

The Independent Panel Report

Jian-Wen Bao (ESRL PSD), Lisa Bengtsson (ESRL PSD),
James Doyle (NRL), Jimmy Dudhia (NCAR)

Outline of the Report

- 1. Executive Summary**
- 2. Background on Suites and Data Used**
- 3. Recommendations on Physics Suites**
- 4. Recommendations for the future**
- 5. Acknowledgements**
- 6. Appendix**

PHYSICS SUITES ASSESSED FOR POSSIBLE GFSv16 IMPLEMENTATION

	<u>Suite 1</u> <u>(GFS v15)</u>	<u>Suite 2</u>	<u>Suite 3</u>	<u>Suite 4</u>
Deep convection	sa-SAS	sa-SAS	sa-CS	sa/aa-GF
Shallow convection	sa-MF	sa-MF	sa-MF	MYNN-EDMF and sa GF
Microphysics	GFDL	GFDL	aa-MG3	aa-Thompson
PBL/Turbulence	K-EDMF	sa-TKE-EDMF	K-EDMF	MYNN-EDMF
Land Surface Model	Noah	Noah	Noah	RUC

*sa = Scale-aware

*aa = aerosol aware

Basis for our recommendation

- Four physics parameterization suites were tested over a large number of cases (every 5 days for 1 year and 16 case studies). A report on the GMTB test setup can be found [here](#).
- The independent panel recommendation is based on a substantial set of verification and evaluation metrics reported from the Environmental Modeling Center (EMC) Model Evaluation Group (MEG), the Global Model Test Bed (GMTB) group in the Developmental Testbed Center (DTC), and EMC's Verification Statistics DataBase (VSDB):
 - **MEG presentation:** https://drive.google.com/file/d/1Fe-lk6egjxQ5ynZ_xF-TfqjQPlp8h119/view?ts=5c93cc43
 - **GMTB report:** https://docs.google.com/document/d/1Vyo1zL5N6GowugDCEWOVFlrK83JcOzWKCCSwGvGe_74/edit?ts=5c89156e#heading=h.8d00ooc534fb
 - **VSDB website:** https://www.emc.ncep.noaa.gov/gmb/wx24fy/NGGPS/phys_cmp/

Executive Summary

- Overall none of the 3 developmental suites were clearly superior to the control GFSv15 physics Suite 1.
- The performance of suite 2 was closest to suite 1 even improving some aspects of the forecasts.
- We encourage work to continue on other suites and new physics packages for future testing.
- We also encourage a similar testing and independent evaluation process on an annual basis in the future.

Recommendations on Physics Suites

- Suite 2, which features a more advanced closure in the PBL scheme, was the closest in performance to Suite 1 and exhibited improvements in some important aspects including somewhat better capturing surface-based inversions and some better precipitation statistics over CONUS.
- Although overall Suite 2 did not perform as well as Suite 1, there are enough positive aspects in the Suite 2 performance to consider further experimentation and tuning in the near term (time permitting) to see if Suite 2 can be implemented in GFSv16. The panel believes the more advanced PBL may ultimately provide improved forecasts of the PBL.
- Suites 3 and 4 both showed promising results in a number of aspects. We strongly encourage and recommend that the developers of both Suites 3 and 4 continue development and testing. We also recommend the developers to consider consolidating the best aspects of all suites, so attention can be focused on a single advanced development suite in the future.

Recommendations for the Future

- We are supportive of a continued annual process in which an independent panel provides analysis and recommendations on the evaluation of parameterization suites considered for future operations.
- We encourage the testing and evaluation of other combinations of physics from the existing four suites, in addition to emerging physical parameterizations.
- Adequate time for tuning and evaluation is needed prior to test phase. The panel is aware of some issues related to the setups of suites 3 and 4 that impacted their results in this round and that these suites would have been improved given adequate pre-testing.
- Our recommendation is that a pre-test period of a few weeks should be built into the schedule using some of this year's initial data (but independent of the data in the next test).
- It is recognized that this year was special because the physics framework was being changed at the same time as new physics were added, but the new framework should make implementations easier in the future.
- Data assimilation cycling was not included in the current test suite and it may be useful in the future to include more testing with the data assimilation as part of the evaluations.

Summary of Systematic Performance of the Physics Suites

Suite 1 (SAS, MF, GFDL MP, K-EDMF, NOAAH)

Positive:

- Overall best anomaly correlation and RMSE of synoptic variables of all suites.
- Overall best day 5 over CONUS precipitation ETS and BIAS scores of all suites.
- Together with suite 2, best track and intensity forecast for Tropical Cyclones.

Room for improvement:

- Increasing cold bias with increasing lead time at 850 hPa and down to the surface. (Addressed with GFDL MP/radiation update?).
- Increased low bias with lead time for 500 hPa heights.
- Precipitation BIAS (too dry) in the Tropics.
- Does not capture warm season low stratus outside the US west coast.
- Struggles with surface temperature inversions
- CERES OLR shows too little high cloud in tropics (as also for suite 2)

Summary of Systematic Performance of the Physics Suites

Suite 2 (SAS, MF, GFDL MP, TKE EDMF, NOAH)

Positive:

- Modest improvement of anomaly correlation for 500 hPa temperature, wind and heights.
- Modest improvements for PBL inversions.
- Improved Tropical Cyclone winds over land (Hurricane Irma).
- Best FSS scores for CONUS precip

Room for improvement:

- Similar to Suite 1, an increased low bias with lead time for 500 hPa heights.
- Degraded RMSE, bias and anomaly correlation for temperature below 850 hPa.
- Cold bias with increasing lead time at 850 hPa and down to the surface, larger than Suite 1.
- Cold bias with increasing lead time near the Tropopause, larger than Suite 1.
- Degradation of precipitation in the Tropics (too dry).
- Small degradation of ETS and BIAS score, and 6h Frequency bias precipitation scores compared with suite 1 over CONUS.
- Colder 2m T BIAS over CONUS than Suite 1 (vs NAM analysis).
- CERES OLR shows too little high cloud in tropics (as suite 1)

Summary of Systematic Performance of the Physics Suites

Suite 3 (CS, MF, MG3 MP, K EDMF, NOAH)

Positive:

- Improved surface inversions compared with suite 1
- Improved 2m T bias over Eastern CONUS (large geographical variations in 2m T and RH).
- CERES OLR shows less tropical bias than suites 1 and 2
- Least wind bias against soundings globally (others have negative bias at high levels)

Room for improvement:

- Degradation of the anomaly correlation of all variables at all verified levels.
- 500 hPa high height bias increasing with forecast lead time.
- Warm temperature bias increasing with forecast lead time in most of the troposphere.
- Very large warm and dry bias in the Tropics (compared with soundings).
- Driest precipitation bias and lowest FSS of all suites in the Tropics.
- Lowest ETS precipitation scores over CONUS compared with all suites.
- Highest bias at heavy rain categories (esp. 6h at 00Z) over CONUS
- Degradation of Tropical Cyclone tracks and intensity compared with suite 1.
- CERES TOA SW too little cloudiness over oceans (more bias than other suites)
- Some unusual soundings noted by MEG

Summary of Systematic Performance of the Physics Suites

Suite 4 (GF, MYNN EDMF/GF, Thompson MP, MYNN EDMF, RUC LSM)

Positive:

- Improved precipitation forecasts in the Tropics.
- Encouraging improvements for PBL inversions. In many cases better than suites 1 and 2 as shown by MEG.
- Improved relationship between surface precipitation and total precipitable water in the tropics.
- Locally much improved 2m T and RH over CONUS, but large geographical variations (sometimes locally much degraded 2m RH).
- Improved 850 hPa temperature BIAS and 500 hPa geopotential height BIAS compared with all other suites.
- Only suite consistently able to capture warm season low stratus along west coast (also seen with CERES SW JJA verification off S. America and S African west coasts and generally less biased cloud cover over oceans) - 2m T at southwest coast much better than other suites possibly because of more clouds
- CERES OLR shows less tropical bias than suites 1 and 2
- From SURFRAD sites SW tends to be cloudier than other suites (often better)

Room for improvement:

- Degradation of the anomaly correlation of all variables at all verified levels.
- 500 hPa high height anomaly correlation scores decreasing with forecast lead time.
- Larger cold bias (compared with suite 1) near the surface and the tropopause.
- Degradation of Tropical Cyclone tracks and intensity.
- Statistically significantly worse precipitation dry bias than other suites for amounts between 10-25 mm/24h over CONUS.
- Extreme precipitation amounts not captured well. Lowest FSS over CONUS.
- Tends to overmix the daytime PBL, not maintaining capping inversions.
- Tends to be drier for 2m Td in central US and more moist on coasts than other suites

Acknowledgements

We would like to thank the following people and organizations for their help, analysis, datasets and collaboration in this process.

- Jack Kain, Ligia Bernardet, Geoff Manikin
- Fanglin Yang, Ratko Vasic, Linus Magnusson (ECMWF)
- EMC, GMTB, MEG, DTC for providing runs, datasets and analyses
- All who contributed to the development of the physics suites

Appendix

Scorecards

00Z and 12Z

Initializations Scored Separately

End Date: 20180110

SUITE 2

00Z RUNS

			N. American							N. Hemisphere							S. Hemisphere							Tropics						
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10				
Anomaly Correlation	Heights	250hPa	▼						▼							▲	▲													
		500hPa														▲	▲	▲												
		700hPa														▼	▼	▲												
		1000hPa	▼	▼						▼	▼					▼	▼	▼	▼	▼	▼									
	Vector Wind	250hPa														▲	▲													
		500hPa	▲							▲	▲					▲	▲	▲	▲											
		850hPa								▲	▲					▲	▲	▲	▲	▼										
	Temp	250hPa								▼							▲	▲	▲											
		500hPa	▲							▲	▲					▲	▲	▲	▲											
		850hPa	▼	▼	▼	▼				▼	▼	▼	▼			▼	▼	▼	▼	▼	▼									
	MSLP	MSL	▼	▼					▼	▼					▼	▼	▼	▼	▼	▼										

EMC Verification Scorecard	
Symbol Legend	
▲	SUITE3 is better than SUITE1 at the 99.9% significance level
△	SUITE3 is better than SUITE1 at the 99% significance level
■	SUITE3 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE3 and SUITE1
■	SUITE3 is worse than SUITE1 at the 95% significance level
▼	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant
Start Date: 20160101	
End Date: 20180110	

SUITE 3

00Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Anomaly Correlation	Heights	250hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		700hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		1000hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Vector Wind	250hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Temp	250hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	MSLP	MSL	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■

EMC Verification Scorecard	
Symbol Legend	
▲	SUITE4 is better than SUITE1 at the 99.9% significance level
▲	SUITE4 is better than SUITE1 at the 99% significance level
■	SUITE4 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE4 and SUITE1
■	SUITE4 is worse than SUITE1 at the 95% significance level
▼	SUITE4 is worse than SUITE1 at the 99% significance level
▼	SUITE4 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant
Start Date: 20160101	
End Date: 20180110	

SUITE 4

00Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Anomaly Correlation	Heights	250hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	■	■	■	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	■	■	■	■	■	■	■	■	■	■	■	■
		700hPa	▼	▼	■	■	■	■	▼	▼	▼	■	■	■	▼	▼	■	■	■	■	■	■	■	■	■	■
		1000hPa	▼	▼	▼	■	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Vector Wind	250hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	■	■	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	■	■	■	▼	■	■	■	■	■	■	■
	Temp	250hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	■	■	■	■	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	■	■	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	MSLP	MSL	▼	▼	▼	■	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■

SUITE 2

▲	SUITE2 is better than SUITE1 at the 99.9% significance level
▲	SUITE2 is better than SUITE1 at the 99% significance level
■	SUITE2 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE2 and SUITE1

	SUITE2 is worse than SUITE1 at the 95% significance level
▼	SUITE2 is worse than SUITE1 at the 99% significance level
▼	SUITE2 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

**00Z
RUNS**

[illegible]

EMC Verification Scorecard	
Symbol Legend	
Start Date: 20160101	
End Date: 20180110	

SUITE 3

▲	SUITE3 is better than SUITE1 at the 99.9% significance level
▲	SUITE3 is better than SUITE1 at the 99% significance level
■	SUITE3 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE3 and SUITE1

■	SUITE3 is worse than SUITE1 at the 95% significance level
▼	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant

00Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
RMSE	Heights	10hPa	▲	■	■	■	■	■	▼	▲	▲	▲	▲	■	■	■	▲	▲	▲	■	■	■	▲	▲	▲	▲
		20hPa	▲	▲	▲	▲	▲	■	▼	▲	▲	▲	▲	▲	■	▲	▲	▲	■	■	▼	▼	■	■	■	■
		50hPa	■	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		100hPa	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		200hPa	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■
		500hPa	■	▼	▼	▼	▼	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		700hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	■	■	■
		1000hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
	Vector Wind	10hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		20hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		50hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		100hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		200hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		500hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼
		700hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼
		850hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		1000hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
	Temp	10hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■
		20hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		50hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		100hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		200hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		500hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		700hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		850hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼
		1000hPa	▼	▼	▼	■	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	■

SUITE 4

	SUITE4 is worse than SUITE1 at the 95% significance level
▼	SUITE4 is worse than SUITE1 at the 99% significance level
▼	SUITE4 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

**00Z
RUNS**

RMSE

			N. American						N. Hemisphere						S. Hemisphere						Tropics						
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	
Bias	Heights	10hPa	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	
		20hPa	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	
		50hPa	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	
		100hPa	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼	
		200hPa	▲	■	■	■	▼	▼	▲	■	▼	▼	▼	▼	▲	■	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼	
		500hPa	▲	▲	■	■	■	■	▲	▲	▲	▲	▲	■	■	▲	▲	■	■	■	■	▲	▲	▲	▲	■	■
		700hPa	▼	▼	■	■	■	■	■	■	■	■	■	■	▲	▲	■	▼	▼	▼	▼	▲	▲	▲	▲	▲	
		850hPa	▼	▼	■	■	■	■	▲	▲	▲	▲	■	■	▼	▼	▼	▼	■	■	▼	▲	▲	▲	▲	▲	
		1000hPa	▲	▼	▼	■	■	■	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▼	■	▲	▲	▲	■	■
	Wind Speed	10hPa	■	■	■	■	■	■	■	▲	▲	▲	■	■	▼	▲	▲	▲	■	■	▼	▼	■	▼	■	■	
		20hPa	■	■	▼	■	■	■	▼	▲	▼	■	■	■	▼	■	■	■	■	■	▼	▼	▼	■	▼	▼	
		50hPa	■	■	■	■	■	■	▼	■	■	■	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	
		100hPa	▲	▲	▲	▲	▲	▲	▲	▲	■	■	■	■	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	
		200hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	
		500hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	
		700hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	
		850hPa	▲	■	■	■	■	■	■	▼	▼	▼	■	■	■	▼	▼	▼	▼	■	■	■	▼	▼	▼	▼	
		1000hPa	▼	■	■	■	■	■	▼	■	▲	■	■	■	■	■	▲	▲	▲	■	■	▼	▼	▲	▲	▲	▲
	Temp	10hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▲	▼	■	■	■	
		20hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	■	▼	▼	▼	▼	
		50hPa	■	▼	■	■	■	■	■	■	▼	▼	▼	▼	■	■	▼	■	■	■	■	▼	▼	▼	▼	▼	
		100hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	■	■	■	■	■	
		200hPa	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▲	■	▼	▼	▼	▼	
		500hPa	▲	▲	■	■	■	■	▲	▲	■	■	■	■	▲	▲	▲	■	■	■	▲	▼	▼	▼	▼	▼	
		700hPa	▲	▲	▲	▲	▲	■	▲	▲	▲	▲	▲	■	▼	■	■	■	■	■	▼	▲	▲	▲	■	■	
		850hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	■	▼	▼	■	■	
		1000hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	

EMC Verification Scorecard							
Symbol Legend							
Start Date: 20160101							
End Date: 20180110							

SUITE 3

▲	SUITE3 is better than SUITE1 at the 99.9% significance level
▲	SUITE3 is better than SUITE1 at the 99% significance level
■	SUITE3 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE3 and SUITE1

■	SUITE3 is worse than SUITE1 at the 95% significance level
▼	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant

00Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10

Bias	Heights	10hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		20hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▲	▲	▲	▲
		50hPa	▲	▼	▲	▲	▲	▲	▼	▼	▼	▲	▲	▲	▼	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		100hPa	▲	▼	▲	▲	▲	▲	▼	▼	▼	▼	▲	▲	▼	▼	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼
		200hPa	▲	▼	▲	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼	▼	▲
		500hPa	▲	▲	▲	▲	▲	▼	▲	▲	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		700hPa	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲
		850hPa	▼	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▼	▼	■	■	■	■	▼	▲	▲	▲	▲	▲
		1000hPa	▼	■	■	■	■	■	▼	▼	▼	■	■	■	▼	▲	▲	▲	▲	▲	▼	▼	■	▲	▲	▲
	Wind Speed	10hPa	■	■	■	▼	▼	■	■	▲	■	▼	▼	▼	▼	■	▼	▼	▼	▼	■	▲	▲	▼	■	■
		20hPa	▲	■	▼	■	■	■	▲	▲	■	▲	▲	▲	■	▲	■	■	■	■	▲	▲	▼	▲	▲	■
		50hPa	▲	■	▲	■	▲	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	■	■	▼	■	▲
		100hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼
		200hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		500hPa	▲	▲	■	■	▼	▼	▲	▲	■	▼	▼	▼	▲	▲	■	■	▼	▼	▲	▲	▲	■	■	▼
		700hPa	▼	■	■	■	▼	▼	■	■	■	▼	▼	▼	▲	■	■	▼	▼	▼	▲	▲	▲	■	■	▼
		850hPa	■	■	■	■	■	▼	▲	▲	■	■	■	▼	▲	■	▼	▼	▼	▼	▼	▼	▼	▼	■	▼
		1000hPa	▼	▼	■	■	■	▲	■	■	▲	▲	▲	▲	▼	■	▲	▲	▼	▲	▼	■	▲	▲	▲	▼
	Temp	10hPa	■	■	▲	▲	▲	▲	■	■	▲	▲	▲	■	■	■	■	■	■	■	▲	▲	▲	▲	▲	▲
		20hPa	■	▼	■	■	■	▼	■	▼	▼	▼	▼	▼	■	■	■	■	■	■	▼	▼	▼	▼	▼	▼
		50hPa	▼	▲	▲	▲	▲	■	■	■	■	■	▼	▼	■	■	■	▲	▲	■	▼	▼	▼	▼	▼	▼
		100hPa	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼
		200hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		700hPa	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		850hPa	▼	▼	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		1000hPa	■	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▼	■	▲	▲	▲	▼	▲	▲	▲	▲	▲

SUITE 4

	SUITE4 is worse than SUITE1 at the 95% significance level
▼	SUITE4 is worse than SUITE1 at the 99% significance level
▼	SUITE4 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

**00Z
RUNS**

Bias

EMC Verification Scorecard

Symbol Legend

▲	SUITE2 is better than SUITE1 at the 99.9% significance level
▲	SUITE2 is better than SUITE1 at the 99% significance level
▲	SUITE2 is better than SUITE1 at the 95% significance level
	No statistically significant difference between SUITE2 and SUITE1
▼	SUITE2 is worse than SUITE1 at the 95% significance level
▼	SUITE2 is worse than SUITE1 at the 99% significance level
▼	SUITE2 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

Start Date: 20160101

End Date: 20180110

SUITE 2

12Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Anomaly Correlation	Heights	250hPa	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▼	▼	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		500hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		700hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		1000hPa	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲
	Vector Wind	250hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		500hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		850hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
	Temp	250hPa	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		500hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		850hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
	MSLP	MSL	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲

EMC Verification Scorecard	
Symbol Legend	
▲	SUITE3 is better than SUITE1 at the 99.9% significance level
▲	SUITE3 is better than SUITE1 at the 99% significance level
■	SUITE3 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE3 and SUITE1
■	SUITE3 is worse than SUITE1 at the 95% significance level
▼	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant
Start Date: 20160101	
End Date: 20180110	

SUITE 3

12Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Anomaly Correlation	Heights	250hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		700hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		1000hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Vector Wind	250hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Temp	250hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	MSLP	MSL	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■

EMC Verification Scorecard	
Symbol Legend	
▲	SUITE4 is better than SUITE1 at the 99.9% significance level
▲	SUITE4 is better than SUITE1 at the 99% significance level
■	SUITE4 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE4 and SUITE1
■	SUITE4 is worse than SUITE1 at the 95% significance level
▼	SUITE4 is worse than SUITE1 at the 99% significance level
▼	SUITE4 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant
Start Date: 20160101	
End Date: 20180110	

SUITE 4

12Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Anomaly Correlation	Heights	250hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	■	■	■	■	■	■
		700hPa	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		1000hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Vector Wind	250hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	Temp	250hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	■	■	■	■	■	■
		850hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■
	MSLP	MSL	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	■	■	■	■

SUITE 2

▲	SUITE2 is better than SUITE1 at the 99.9% significance level
▲	SUITE2 is better than SUITE1 at the 99% significance level
■	SUITE2 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE2 and SUITE1

	SUITE2 is worse than SUITE1 at the 95% significance level
▼	SUITE2 is worse than SUITE1 at the 99% significance level
▼	SUITE2 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

**12Z
RUNS**

[illegible]

SUITE 3

	SUITE3 is worse than SUITE1 at the 95% significance level
▼	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

RMSE

[illegible]

SUITE 4

	SUITE4 is worse than SUITE1 at the 95% significance level
▼	SUITE4 is worse than SUITE1 at the 99% significance level
▼	SUITE4 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

12Z RUNS

RMSE

EMC Verification Scorecard
Symbol Legend
Start Date: 20160101
End Date: 20180110

SUITE 2

▲	SUITE2 is better than SUITE1 at the 99.9% significance level
▲	SUITE2 is better than SUITE1 at the 99% significance level
▲	SUITE2 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE2 and SUITE1

■	SUITE2 is worse than SUITE1 at the 95% significance level
▼	SUITE2 is worse than SUITE1 at the 99% significance level
▼	SUITE2 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant

12Z
RUNS

			N. American							N. Hemisphere							S. Hemisphere							Tropics								
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10						
Bias	Heights	10hPa	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼						
		20hPa	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼						
		50hPa	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼						
		100hPa	▲	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▲	▼	▼	▼						
		200hPa	▲	▲	■	■	■	■	▲	▲	▼	▼	▼	▼	▲	■	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼						
		500hPa	▲	▲	▲	▲	▲	■	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	■	▼	▼	▲	▲	▲	▲	▲	▲					
		700hPa	▼	■	▼	▼	▲	▼	▲	▲	■	■	■	▲	▲	■	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲					
		850hPa	▼	▼	▼	▼	■	▼	■	▲	■	■	■	▲	▼	▼	▼	■	■	▲	■	▲	▲	▲	▲	▲	▲					
		1000hPa	▼	▲	▼	▼	▲	■	▼	▼	▼	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲					
	Wind Speed	10hPa	■	▼	▼	▼	■	■	■	■	■	■	■	▲	■	▲	▲	▲	▲	▲	▼	▼	■	▼	■	■	■					
		20hPa	■	■	▼	■	■	■	▼	■	▼	■	■	▼	▼	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					
		50hPa	▼	▼	■	▼	■	■	▼	▼	■	■	■	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					
		100hPa	▲	▲	▼	▼	▲	▼	▲	▲	▲	▲	■	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼					
		200hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					
		500hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		700hPa	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	
		850hPa	▲	■	■	■	▼	▲	■	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		1000hPa	▼	■	■	■	■	■	▼	■	■	■	▲	▼	▲	▲	▲	▲	▲	▲	▼	▼	▲	▲	▲	▲	▼	▼	▲	▲	▲	▲
	Temp	10hPa	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	■	▼	▼	■	■
		20hPa	■	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		50hPa	■	▼	■	■	▼	▼	■	▼	▼	▼	▼	▼	▼	■	■	■	▼	▼	▼	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼
		100hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	■	■	■	■	■	■	■	■	■	■	■
		200hPa	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	■	▼	▼	▼	▼	▼	▼	▼	▼	▼
		500hPa	▲	▲	▲	■	■	▲	▲	▲	■	▼	▼	■	▲	▲	▲	▲	■	▼	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		700hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	■	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		850hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		1000hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	■	■	▲	▲	■	▲	▲	▲	▲

EMC Verification Scorecard							
Symbol Legend							
Start Date: 20160101							
End Date: 20180110							

SUITE 3

▲	SUITE3 is better than SUITE1 at the 99.9% significance level
▲	SUITE3 is better than SUITE1 at the 99% significance level
■	SUITE3 is better than SUITE1 at the 95% significance level
■	No statistically significant difference between SUITE3 and SUITE1

■	SUITE3 is worse than SUITE1 at the 95% significance level
▼	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant

12Z
RUNS

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Bias	Heights	10hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		20hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		50hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		100hPa	▲	▲	▲	▲	▲	▲	▼	▼	▼	▲	▲	▲	▲	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		200hPa	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▲	▼	▼	▼	▲	▲	▲	▲
		500hPa	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▼	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▼
		700hPa	■	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲
		850hPa	▼	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲
		1000hPa	▼	▲	▼	▼	▼	▼	■	▲	▲	■	■	■	▼	▲	▲	▲	▲	▲	▼	▼	▼	■	■	▲
	Wind Speed	10hPa	▼	▼	▼	▼	▼	▼	▼	■	■	▼	▼	■	▼	■	■	▼	▼	▼	▼	■	▲	▼	■	■
		20hPa	▲	▼	▼	■	■	■	■	▲	▼	▲	■	▲	■	▲	▼	■	■	■	■	▲	▼	▲	▲	■
		50hPa	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▲	■	■	▼	■	▲
		100hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼
		200hPa	▲	▲	▲	■	▲	■	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	■	■
		500hPa	▲	■	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼	▲	▲	■	▼	▼	▼	▲	▲	▲	▲	▼	▼
		700hPa	■	■	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▲	■	▼	▼	▼	▼	▲	▲	▲	▲	▼	▼
		850hPa	▲	■	■	▼	▼	▼	▲	■	■	▼	▼	▼	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■
		1000hPa	■	▲	▲	▲	▲	▲	▼	■	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	■	▲	▲	▲	▲
	Temp	10hPa	▲	■	▲	▲	▲	▲	■	▲	▲	▲	▲	▲	■	▼	▲	■	■	■	▲	▲	▲	▲	▲	▲
		20hPa	▼	■	■	■	▼	▼	▲	▼	▼	▼	▼	▼	▲	▼	▼	■	■	■	▼	▼	▼	▼	▼	▼
		50hPa	▼	▲	▲	▲	▲	■	■	■	▼	▼	▼	▼	■	▲	▲	▲	■	■	▼	▼	▼	▼	▼	▼
		100hPa	■	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼
		200hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼
		500hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		700hPa	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		850hPa	▼	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		1000hPa	▼	▼	■	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▼	■	▲	▲	▲	▼	▼	▼	▼	▼	▼

EMC Verification Scorecard
Symbol Legend
Start Date: 20160101
End Date: 20180110

SUITE 4

▲	SUITE4 is better than SUITE1 at the 99.9% significance level
△	SUITE4 is better than SUITE1 at the 99% significance level
■	SUITE4 is better than SUITE1 at the 95% significance level
□	No statistically significant difference between SUITE4 and SUITE1

■	SUITE4 is worse than SUITE1 at the 95% significance level
▼	SUITE4 is worse than SUITE1 at the 99% significance level
▼	SUITE4 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant

**12Z
RUNS**

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Bias	Heights	10hPa	▼	▼	▲	▲	▲	▲	▼	▼	■	▲	▲	▲	▼	▼	▲	▲	▲	▲	▼	▼	▼	▼	▼	■
		20hPa	▼	▼	▲	▲	▲	▲	▼	■	▲	▲	▲	▲	▼	■	▲	▲	▲	▲	▼	▼	▲	▲	▲	▲
		50hPa	▼	■	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		100hPa	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		200hPa	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		500hPa	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		700hPa	■	▲	▲	▲	▲	▲	■	▲	▲	▲	▲	▲	▼	▲	▲	▼	▼	▼	▼	▲	▲	▲	▲	▲
		850hPa	▼	■	▼	■	▲	△	▼	▼	■	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼	■
		1000hPa	▼	▼	▼	▼	▼	▼	■	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼
	Wind Speed	10hPa	▲	▼	■	■	▼	▼	▲	▼	▼	△	■	△	▲	■	■	■	■	■	▲	▲	▼	▲	▼	▼
		20hPa	▲	▼	■	▼	▼	▼	▲	■	▲	▲	▲	■	▲	■	▲	■	△	■	▲	▼	▲	▼	▼	▼
		50hPa	■	▲	▲	▲	▲	▲	■	△	▲	▼	▼	▼	▼	▲	▲	▲	▲	△	▼	▼	▼	■	▼	▼
		100hPa	▼	■	▼	▼	▲	▼	▼	■	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▼	■	■	▲	▲
		200hPa	▼	■	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▼
		500hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲
		700hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▼	▼
		850hPa	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		1000hPa	▲	■	▼	■	▼	▼	■	▼	▼	▼	▼	▼	▲	▼	▼	▼	▼	▼	▲	▲	▼	▼	▼	▼
	Temp	10hPa	▲	▲	■	■	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	■	▼	▲	▲	▲	▲
		20hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▼	▼	▼	▼	▼
		50hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		100hPa	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	▲	▲
		200hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	▲	▲	▲	▲
		500hPa	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▼	▼	▲	▲	▲	▲
		700hPa	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼
		850hPa	▼	■	△	△	▲	▲	■	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	△	▼	▼	▼	▼	▼
		1000hPa	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	■	■	▼	▼	▼	▼	▼	▼	▲	▲	▲	▲	▲	▲

VSDB output (20160106-20171226 12 UTC inits)
for SUITE2 and SUITE1

2016-01-06 12:00:00 - 2017-12-26 12:00:00

			NEC				APL				MDW				LMV				GMC				SEC			
			Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7
RMSE	2m Temp	SFC		▼	▼	▼		▼	▼		▼	▼	▼	▼	▲		▼	▼	▲	▲			▲	▲	▲	▲
	2m RH	SFC	▲				▲	▲	▲		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
RMSE	10m Vector Wind	SFC	▲	▲		▲	▲	▲			▲	▲	▲	▲	▲	▲		▲	▲	▲	▲		▲	▲		
ME	2m Temp	SFC	▼	▼	▼	▼	▼	▼	▼		▼	▼	▼	▼	▼	▼	▼	▼	▲	▲	▲		▲	▲	▲	
	2m RH	SFC	▲	▲	▲	▲	▼	▼	▼	▼	▲	▼	▼	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Bias	10m Wind Speed	SFC	▲	▲	▲	▲	▲	▼			▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲		▲	▲	▲	▲

▲	SUITE2 is better than SUITE1 at the 99.9% significance level
▲	SUITE2 is better than SUITE1 at the 99% significance level
	SUITE2 is better than SUITE1 at the 95% significance level
	No statistically significant difference between SUITE2 and SUITE1
	SUITE2 is worse than SUITE1 at the 95% significance level
▼	SUITE2 is worse than SUITE1 at the 99% significance level
▼	SUITE2 is worse than SUITE1 at the 99.9% significance level
	Not statistically relevant

Figure SC1. Scorecard documenting performance of Suite 2 as compared to Suite 1 over the subregions across the eastern portions of the CONUS for RMSE and mean error (also referred to as bias) of 2-m temperature, 2-m RH, and 10-m winds by forecast lead time 12 UTC initializations during the entire test period (20160101-20171231). Green (red) marks indicate Suite 2 (Suite 1) is better than Suite 1 (Suite 2). Statistical significance is represented by the type of marks: shading, small arrows, and large arrows indicate 95%, 99%, and 99.9% significance, respectively.

VSDB output (20160101-20171231 00 UTC inits)
for SUITE3 and SUITE1

2016-01-01 00:00:00 - 2017-12-31 00:00:00

		N.Hemisphere							S.Hemisphere							Tropics						
		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
RMSE	Temp	P10																				
		P20																				
		P50																				
		P100																				
		P200																				
		P500																				
		P700																				
		P850																				
		P1000																				
		P500																				
RMSE	Rel. Humidity	P700																				
		P850																				
		P1000																				
RMSE	Vector Wind	P10																				
		P20																				
		P50																				
		P100																				
		P200																				
		P500																				
		P700																				
		P850																				
		P1000																				
Bias	Temp	P10																				
		P20																				
		P50																				
		P100																				
		P200																				
		P500																				
		P700																				
		P850																				
		P1000																				
		P500																				
Bias	Rel. Humidity	P700																				
		P850																				
		P1000																				
Bias	Wind Speed	P10																				
		P20																				
		P50																				
		P100																				
		P200																				
		P500																				
		P700																				
		P850																				
		P1000																				
		P500																				

Figure SC2. Scorecard documenting performance of Suite 3 as compared to Suite 1 over the NH, SH, and Tropics for RMSE and bias of temperature, RH, and winds at various pressure levels by forecast lead time for 00 UTC initializations during the entire test period (20160101-20171231). Green (red) marks indicate Suite 3 (Suite 1) is better than Suite 1 (Suite 3). Statistical significance is represented by the type of marks: shading, open arrows, and filled arrows indicate 95%, 99%, and 99.9% significance, respectively.

▲	SUITE3 is better than SUITE1 at the 99.9% significance level
△	SUITE3 is better than SUITE1 at the 99% significance level
■	SUITE3 is better than SUITE1 at the 95% significance level
■	SUITE3 is worse than SUITE1 at the 95% significance level
▽	SUITE3 is worse than SUITE1 at the 95% significance level
▽	SUITE3 is worse than SUITE1 at the 99% significance level
▼	SUITE3 is worse than SUITE1 at the 99.9% significance level
■	Not statistically relevant

PRECIP SCORECARDS

ETS for NH, SH, & Tropics. All symbols are relative to Suite 1

MET output (20160101-20171231 00 UTC inits)
for GFS_suite2_0p25_FCST and GFS_suite1_0p25_FCST
2016-01-01 00:00:00 - 2017-12-31 00:00:00

			NH								SH								TROP												
			f36	f60	f84	f108	f132	f156	f180	f204	f228	f36	f60	f84	f108	f132	f156	f180	f204	f228	f36	f60	f84	f108	f132	f156	f180	f204	f228		
GSS	24-h Accum Pcp	>=0.254										▲	▲	▲	▲	▲	▲					▲	▲	▲	▲		▲	▲	▲		
		>=0.508		▽									△	△	△	△	△					△	△	△	△	▽	△	△		△	
		>=1.270		▽	▽								▽	▽	▽	▽	▽						▽	▽	▽	▽	▽	▽	▽	▽	▽
		>=2.540	▽	▽									▽	▽									▽	▽	▽	▽	▽	▽	▽	▽	▽
		>=3.810	▽	▽																			▽	▽	▽	▽	▽	▽	▽	▽	▽
		>=6.350		▽																			▽	▽	▽	▽	▽	▽	▽	▽	▽
		>=8.890																					▽	▽	▽	▽	▽	▽	▽	▽	▽
		>=12.700																					▽	▽	▽	▽	▽	▽	▽	▽	▽
		>=25.400													▽								▽	▽	▽	▽	▽	▽	▽	▽	▽

SUITE 2

MET output (20160101-20171231 00 UTC inits)
for GFS_suite3_0p25_FCST and GFS_suite1_0p25_FCST
2016-01-01 00:00:00 - 2017-12-31 00:00:00

			NH								SH								TROP												
			f36	f60	f84	f108	f132	f156	f180	f204	f228	f36	f60	f84	f108	f132	f156	f180	f204	f228	f36	f60	f84	f108	f132	f156	f180	f204	f228		
GSS	24-h Accum Pcp	≥0.254	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	△	△	△	△	△	△	△	△	△	
		≥0.508	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	△	△	△	△	△	△	△	△	△	
		≥1.270	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	
		≥2.540	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	
		≥3.810	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	
		≥6.350	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽
		≥8.890	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	
		≥12.700	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	
		≥25.400	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽	▽

SUITE 3

MET output (20160101-20171231 00 UTC inits)
for GFS_suite4_0p25_FCST and GFS_suite1_0p25_FCST
2016-01-01 00:00:00 - 2017-12-31 00:00:00

		NH										SH								TROP									
		f36	f60	f84	f108	f132	f156	f180	f204	f228	f36	f60	f84	f108	f132	f156	f180	f204	f228	f36	f60	f84	f108	f132	f156	f180	f204	f228	
GSS	24-h Accum Pcp	>=0.254																											
		>=0.508	▽	▽	▽	▽	▽	▽				△	△	△	△	△	△	△	△	△	△	△	△	△	△	△	△	△	
		>=1.270							▽												▽				△				△
		>=2.540	▽	▽	▽	▽	▽	▽	▽			▽	▽	▽	▽	▽	▽	▽	▽	▽		▽		▽					
		>=3.810	▽	▽	▽	▽	▽	▽	▽			▽	▽	▽	▽	▽	▽	▽	▽	▽		▽		▽					
		>=6.350	▽	▽	▽	▽	▽	▽				▽	▽	▽	▽	▽	▽	▽	▽	▽		▽		▽		▽			
		>=8.890	▽	▽	▽	▽	▽	▽					▽	▽	▽	▽	▽	▽	▽	▽		▽		▽		▽			
		>=12.700	▽	▽	▽	▽	▽	▽				▽	▽	▽	▽	▽	▽	▽	▽	▽		▽		▽		▽			
		>=25.400	▽	▽	▽	▽	▽	▽				▽	▽	▽	▽	▽	▽	▽	▽	▽		▽		▽		▽			

SUITE 4

△	GFS_suite4_0p25_FCST is better than GFS_suite1_0p25_FCST at the 99.9% significance level
△	GFS_suite4_0p25_FCST is better than GFS_suite1_0p25_FCST at the 99% significance level
△	GFS_suite4_0p25_FCST is better than GFS_suite1_0p25_FCST at the 95% significance level
▽	GFS_suite4_0p25_FCST is worse than GFS_suite1_0p25_FCST at the 95% significance level
▽	GFS_suite4_0p25_FCST is worse than GFS_suite1_0p25_FCST at the 95% significance level
▽	GFS_suite4_0p25_FCST is worse than GFS_suite1_0p25_FCST at the 99% significance level
▽	GFS_suite4_0p25_FCST is worse than GFS_suite1_0p25_FCST at the 99.9% significance level
	Not statistically relevant

Statistics provided by GMTB

2016-01-01 00:00:00 - 2017-12-31 00:00:00

SUITE 2

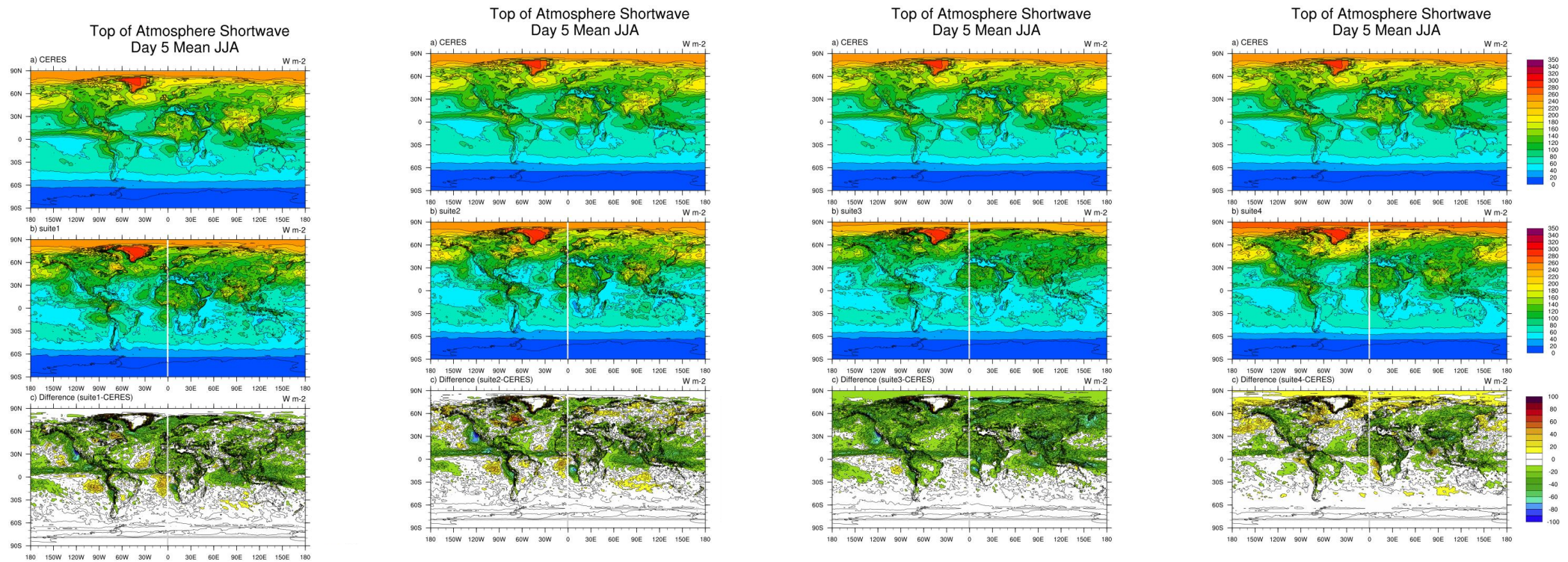
2016-01-01 00:00:00 - 2017-12-31 00:00:00

SUITE 3

2016-01-01 00:00:00 - 2017-12-31 00:00:00

SUITE 4

CERES TOA SW JJA 5-day

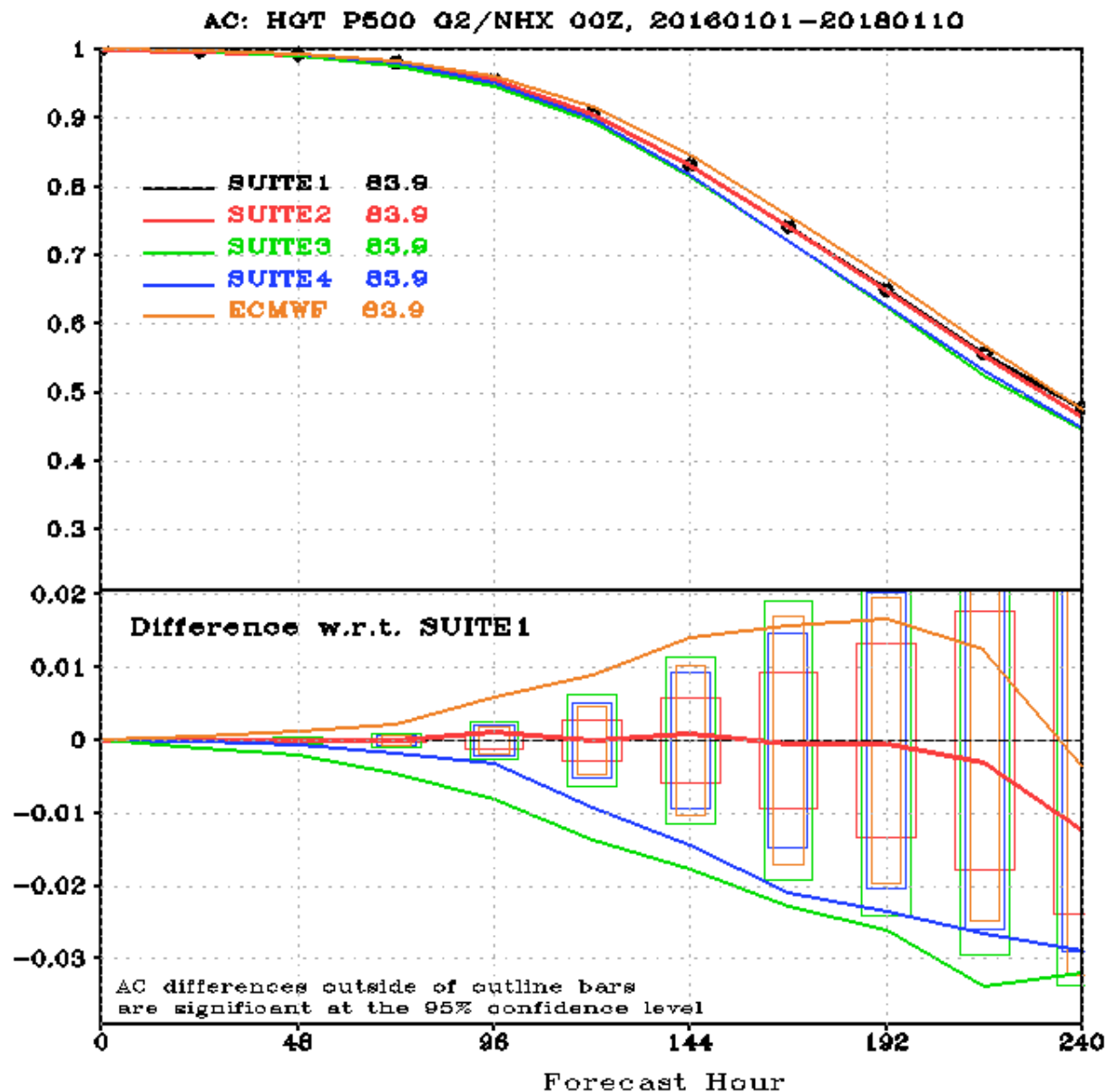


less JJA NH cloud bias over oceans in suite 4 especially off west coasts - similar for SH DJF (not shown)

STATISTICS

(from MEG slides)

500-hPa Geopotential Height Anomaly Correlation (AC) Scores



SUITE 1

SUITE 2

SUITE 3

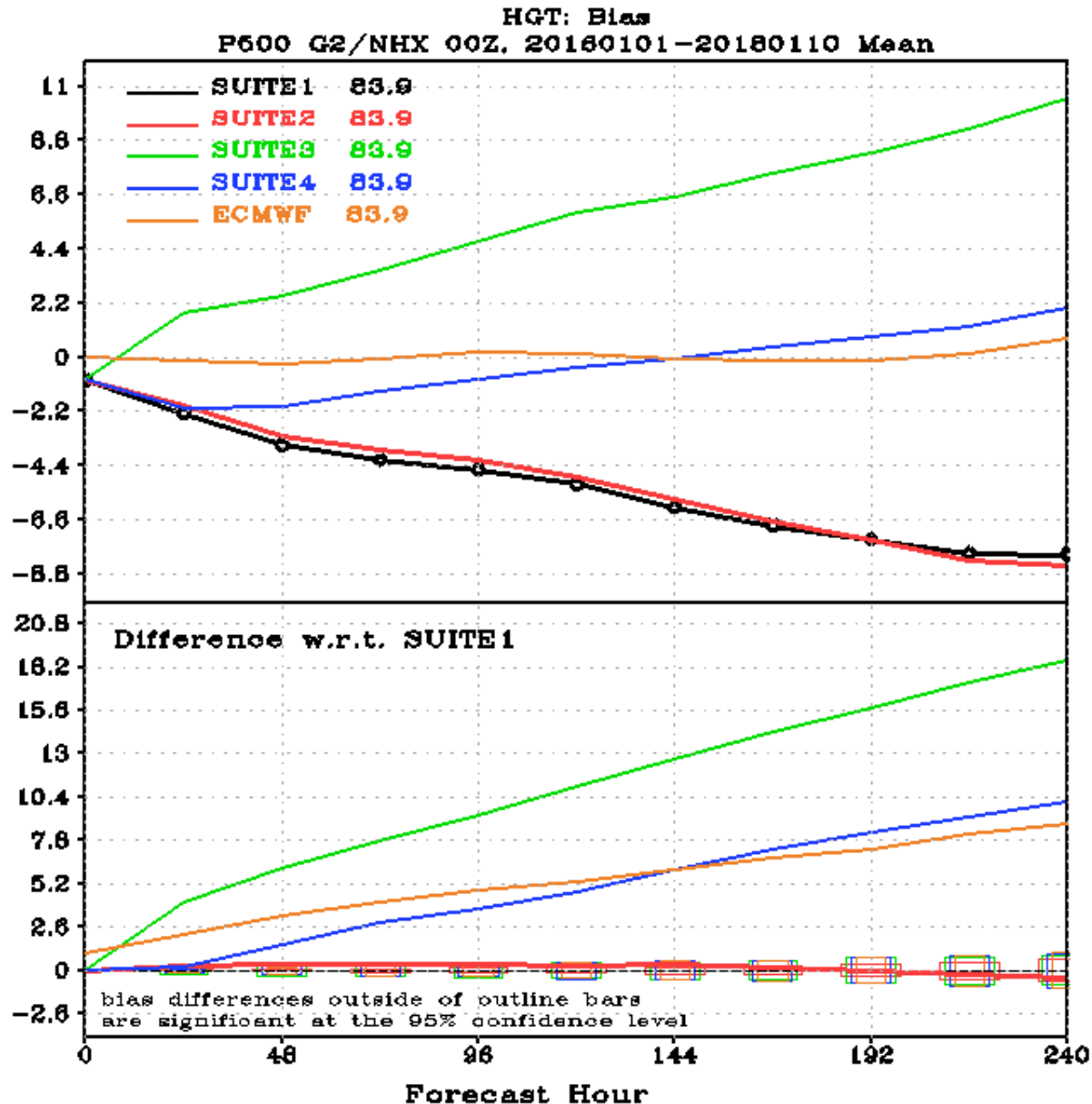
SUITE 4

ECMWF

KEY POINTS:

- Suites 1 and 2 have statistically significantly better AC scores than Suites 3 and 4 well into medium range
- Suites 1 and 2 have nearly identical AC scores
- Suite 4 is slightly better than Suite 3

500-hPa Geopotential Height Bias



SUITE 1
SUITE 2

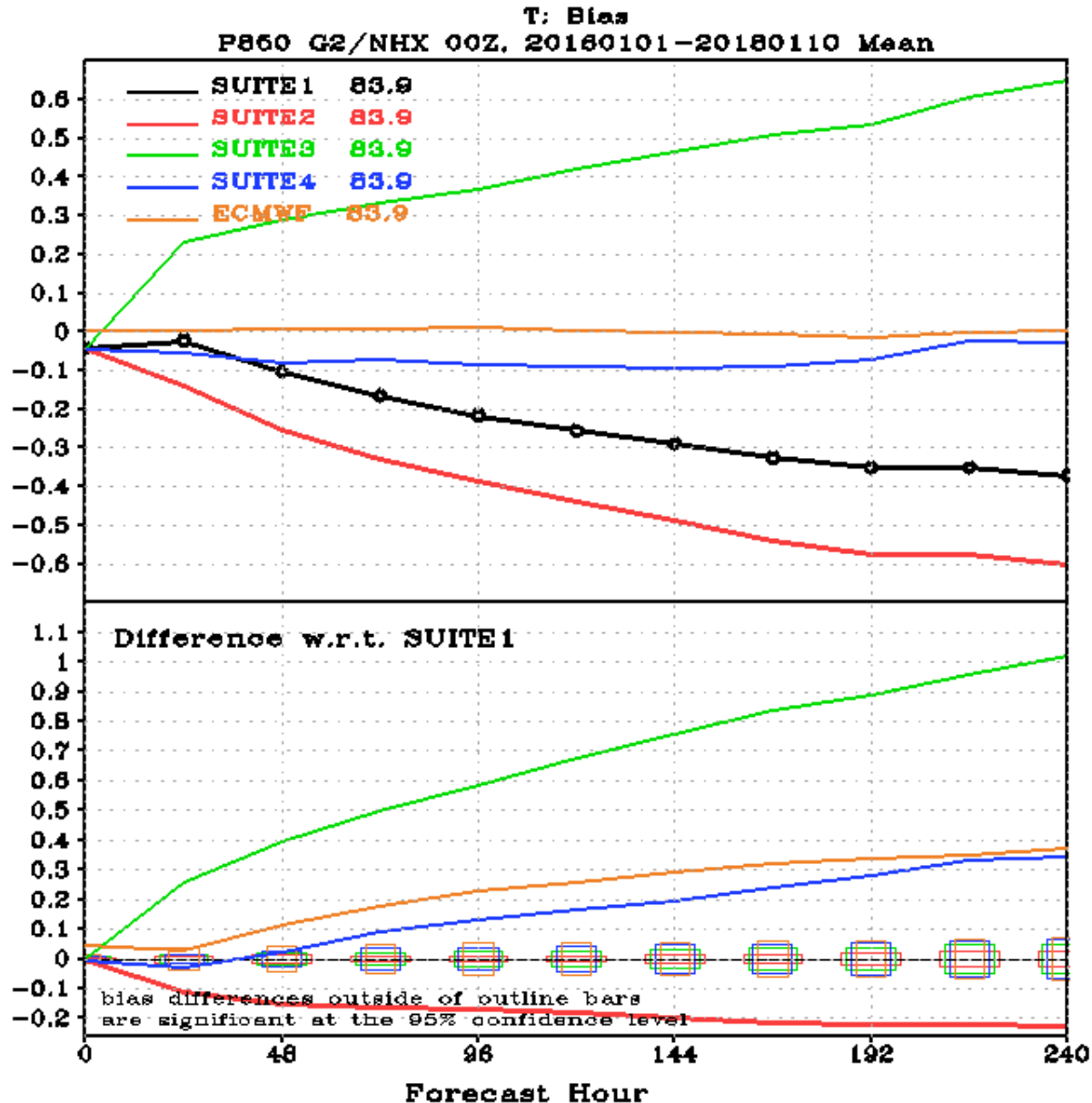
SUITE 3
SUITE 4

ECMWF

KEY POINTS:

- Suites 1 and 2 have a nearly identical low bias that increases with forecast lead time
- Suite 3 has a high bias that increases with forecast lead time
- Suite 4 has the smallest average 500-hPa height bias at almost all forecast hours

850-hPa Temperature Bias



SUITE 1
SUITE 2

SUITE 3
SUITE 4

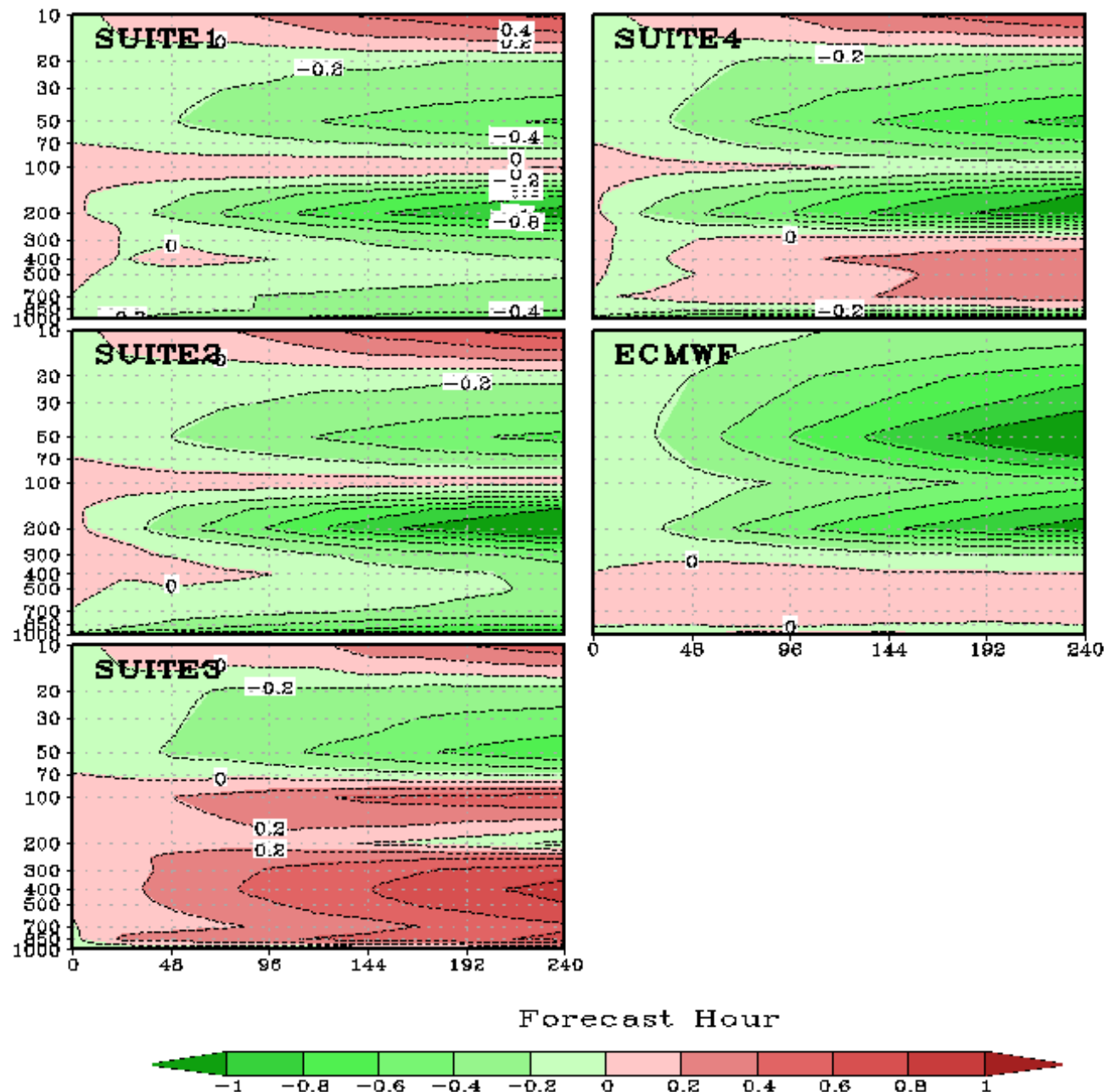
ECMWF

KEY POINTS:

- Suites 1 and 2 have a cold bias that increases with lead time
- Suite 2 has a larger cold bias than Suite 1, potentially due to PBL scheme
- Suites 3 has a warm bias that increases with lead time, while Suite 4 has the smallest temperature bias
- The suites without GFDL microphysics do not show a cold bias increasing w time

Vertical Profile of Temp. Bias

T: Bias
20160101-20180110 Mean, G2/NHX 00Z

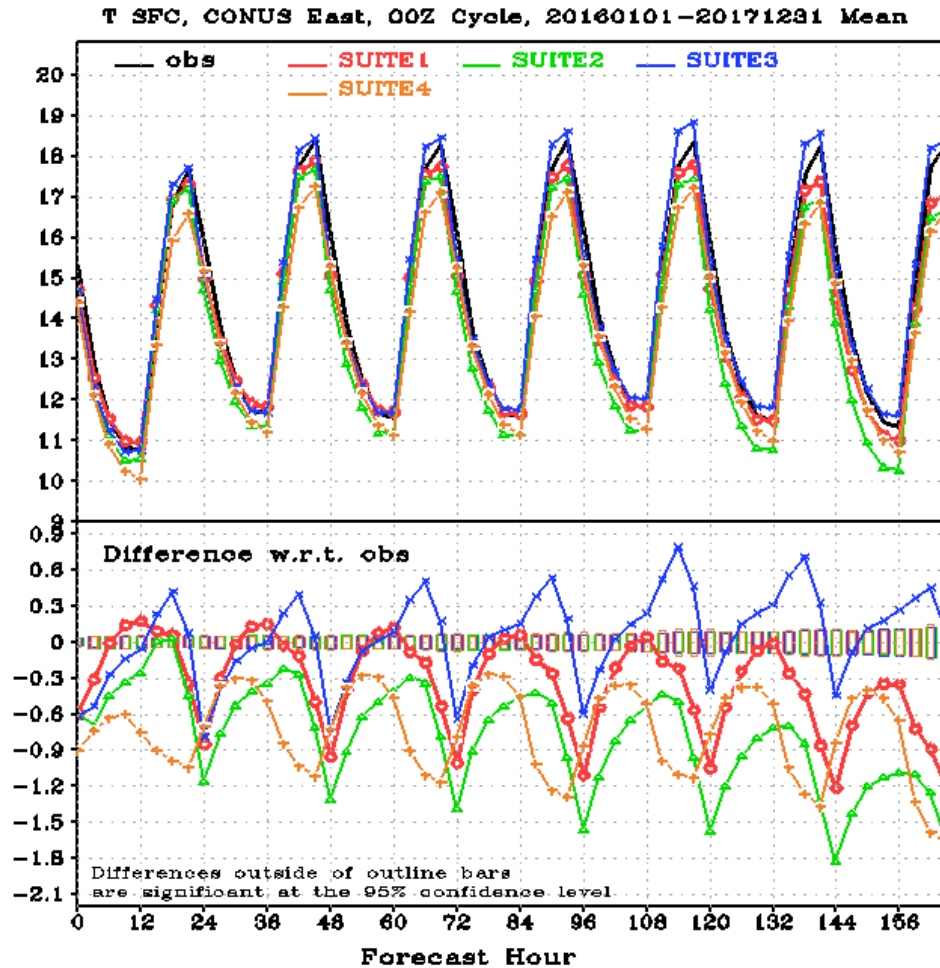


KEY POINTS:

- Suites 1 and 2 have a low-level cold bias that increases with lead time
- Suites 2 and 4 have a worse cold bias than Suite 1 near tropopause and near sfc.
- Suite 3 has a warm bias that increases with lead time throughout most of the troposphere

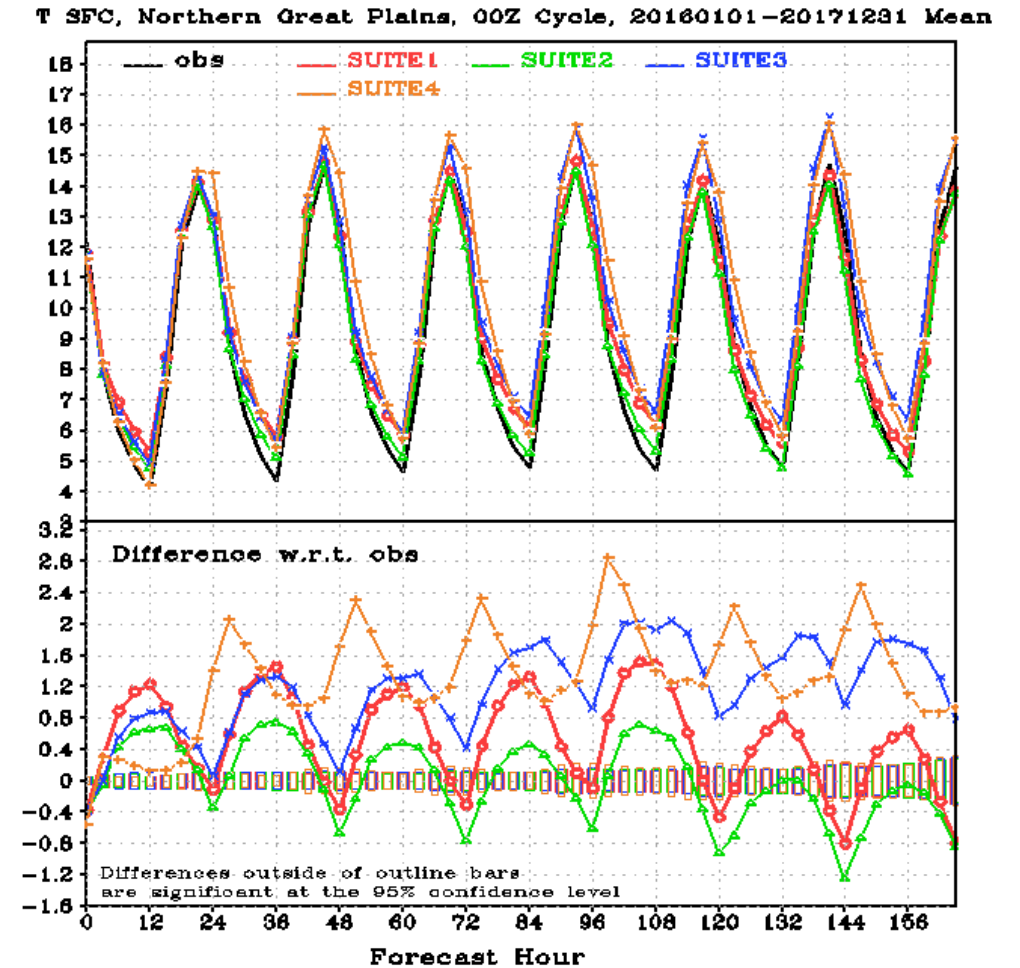
Grid-to-Obs: 2-m Temperatures

CONUS East



- All suites except Suite 3 have a near-surface cold bias that increases with forecast lead time

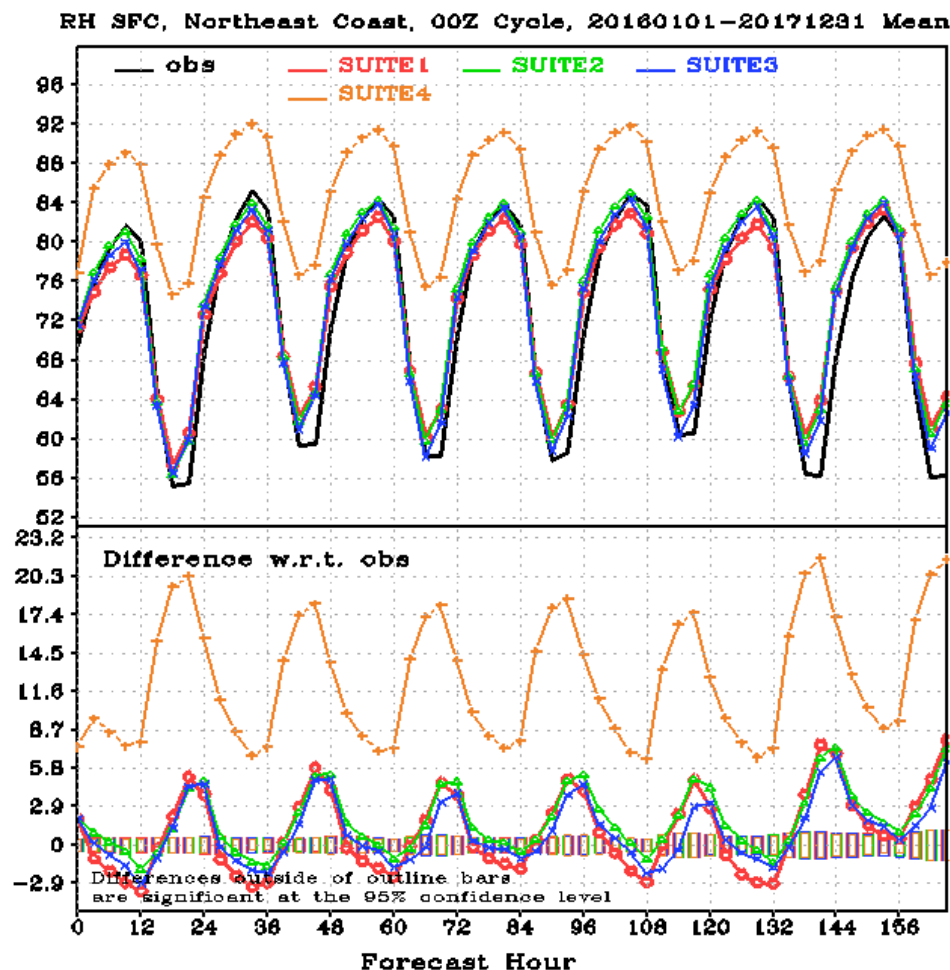
Northern Great Plains



- Suite 4 has the largest near-surface warm bias on the Plains, but all suites are warmer than east

Grid-to-Obs: 2-m Relative Humidity

Northeast Coast

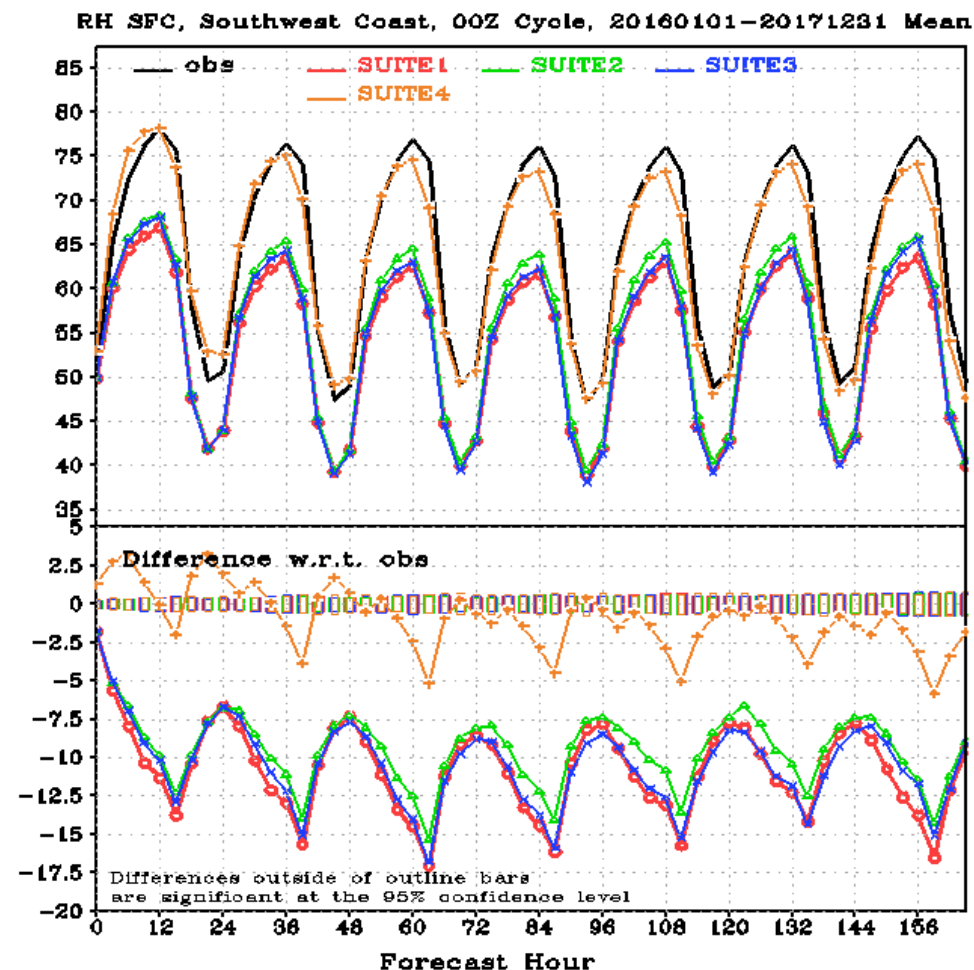


OBS
SUITE 1
SUITE 2
SUITE 3
SUITE 4

lots of regional
variation

- Suite 4 has RH values that are far too high

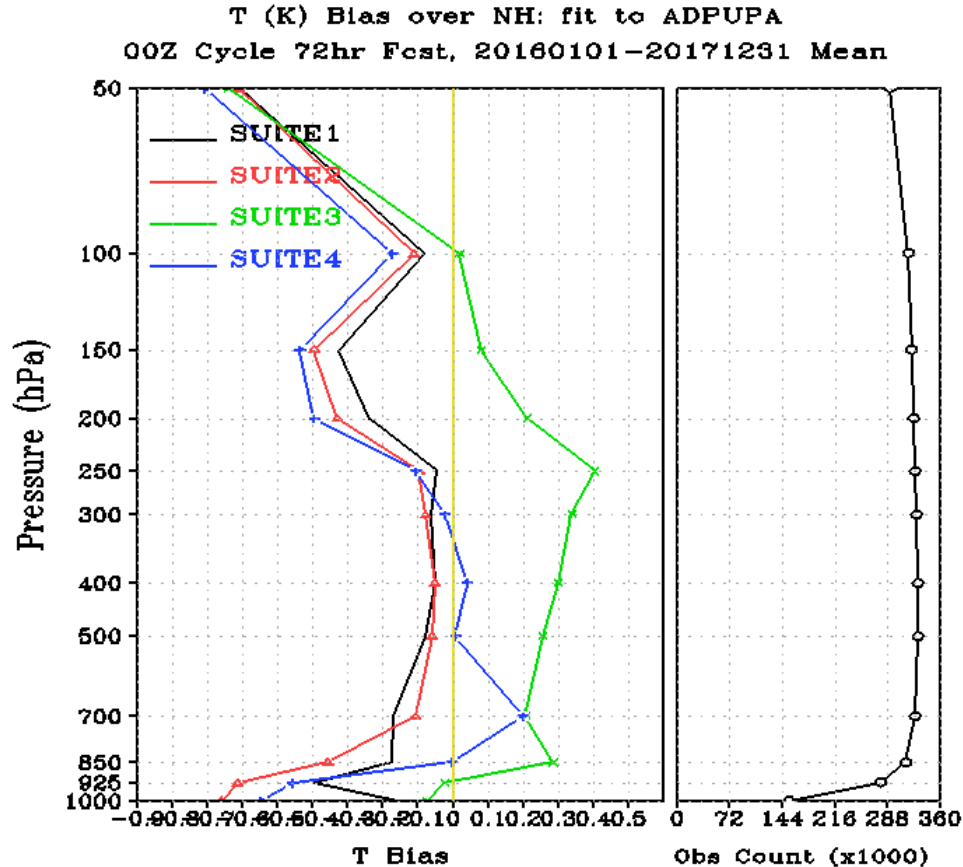
Southwest Coast



- Suite 4 has the best RH values

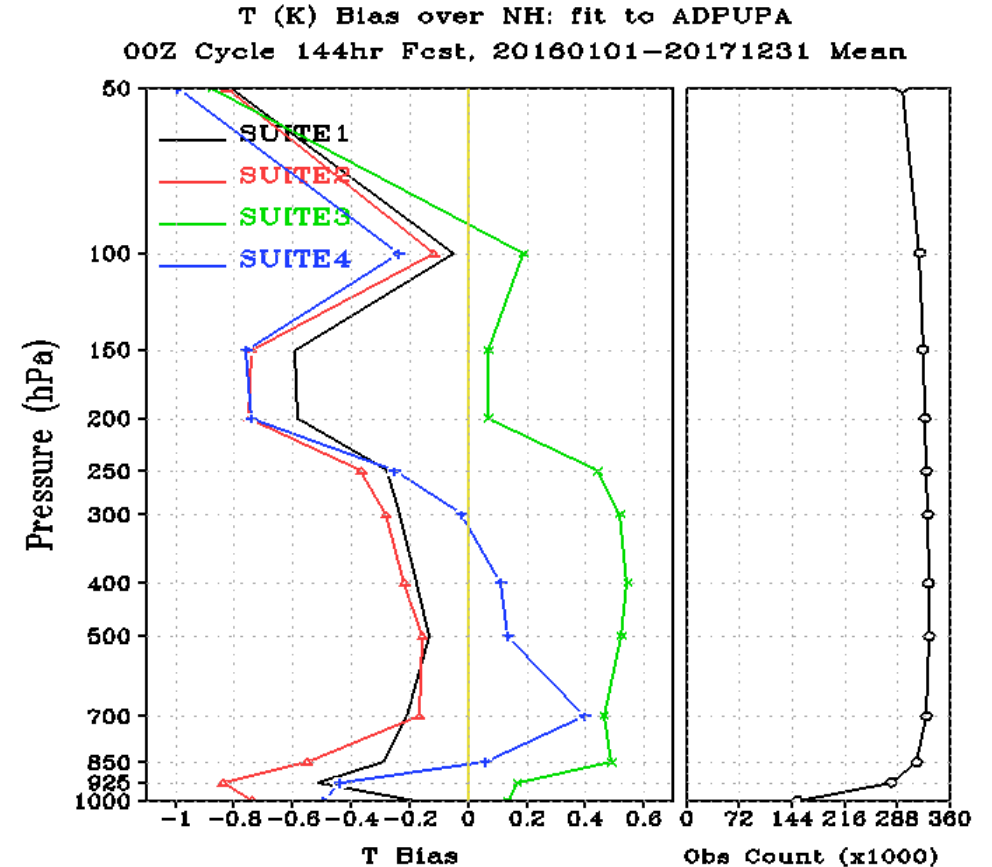
Temperature Profiles compared to RAOBs

F072



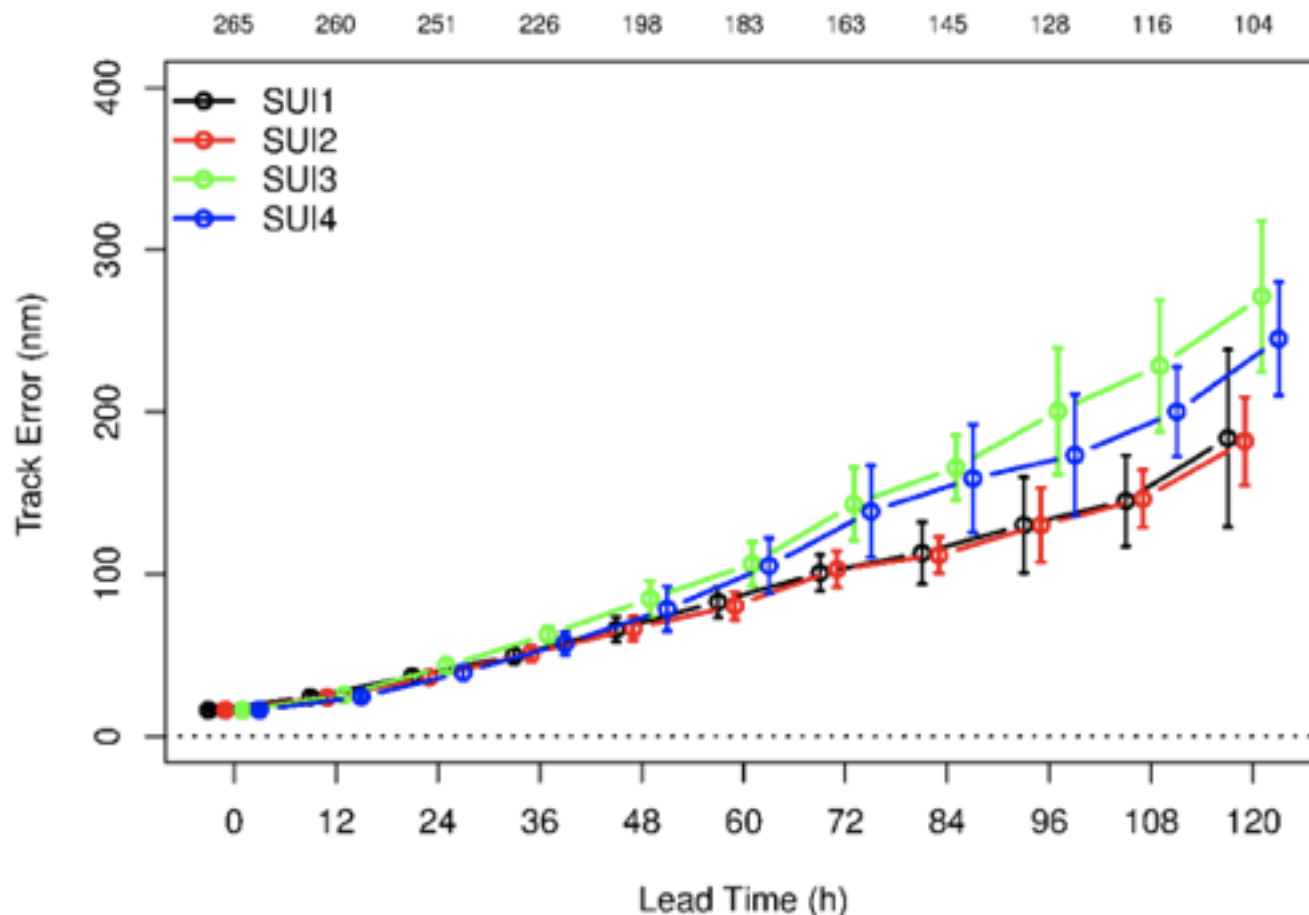
- Suites 1 and 2 have the least bias at F072
- Suite 2 has a larger cold bias at 200 hPa and near surface, while Suite 3 has a warm bias

F144



- Cold bias in Suites 1 and 2 increases with forecast lead time
- Warm bias in Suite 3 increases with lead time

Track Performance – Composite Track Errors

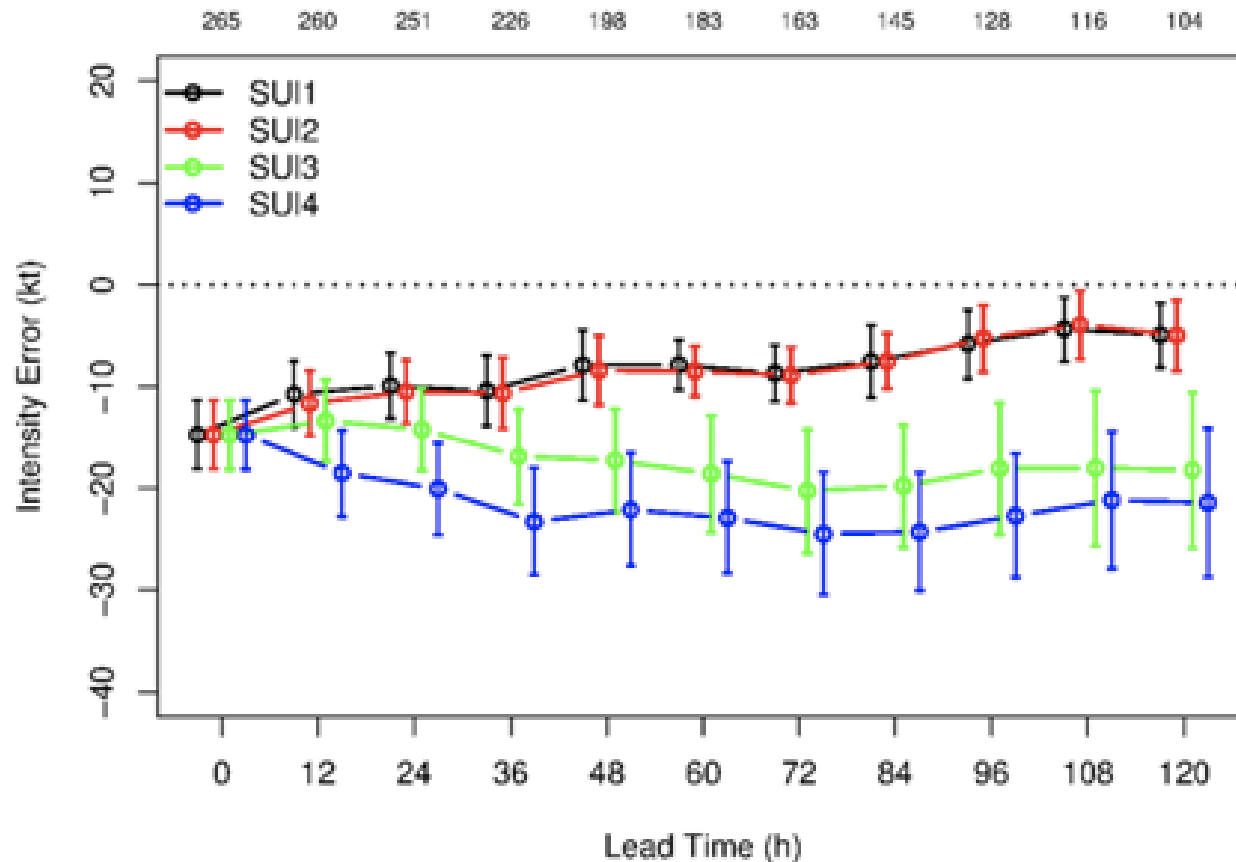


- For many of the 8 TCs examined by the MEG, track forecasts were fairly comparable
- Track errors for Suites 1 and 2 were nearly identical
- Incorrect outlier track solutions were largely limited to forecasts from Suites 3 and 4
- Composite stats for all TCs in all basins show that beyond Day 3, Suites 3 and 4 did have larger track errors than Suites 1 and 2

Statistics provided by GMTB

Suite 1 **Suite 3**
Suite 2 **Suite 4**

Intensity Performance – Composite Vmax Errors



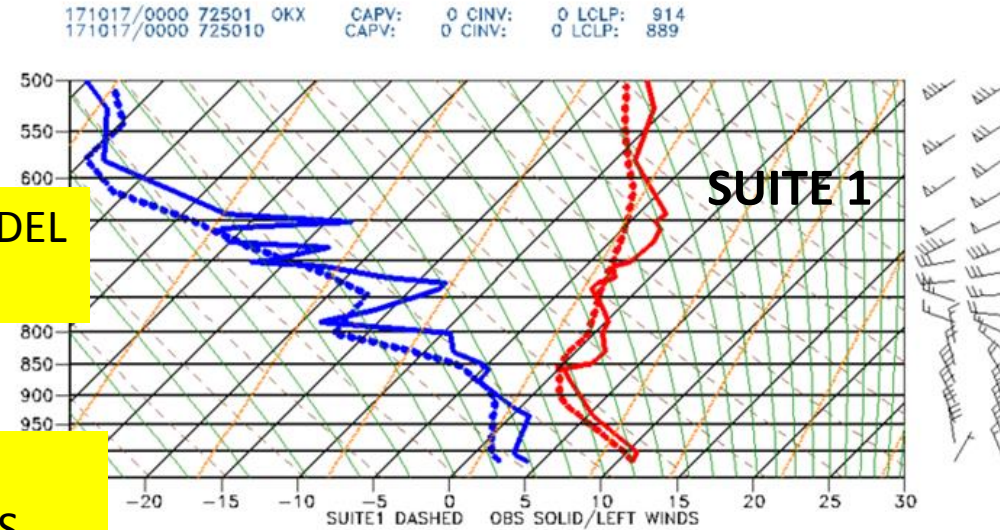
- As expected, all suites had a weak intensity bias compared to Best Track data
- Intensity errors for Suites 1 and 2 were nearly identical
- Intensity forecasts from Suites 3 and 4 were significantly weaker beyond Day 1, while Suites 1 and 2 reduced the intensity errors with time

Statistics provided by GMTB

Suite 1 **Suite 3**
Suite 2 **Suite 4**

Highlighted Comparative Evaluation: PBL Inversion

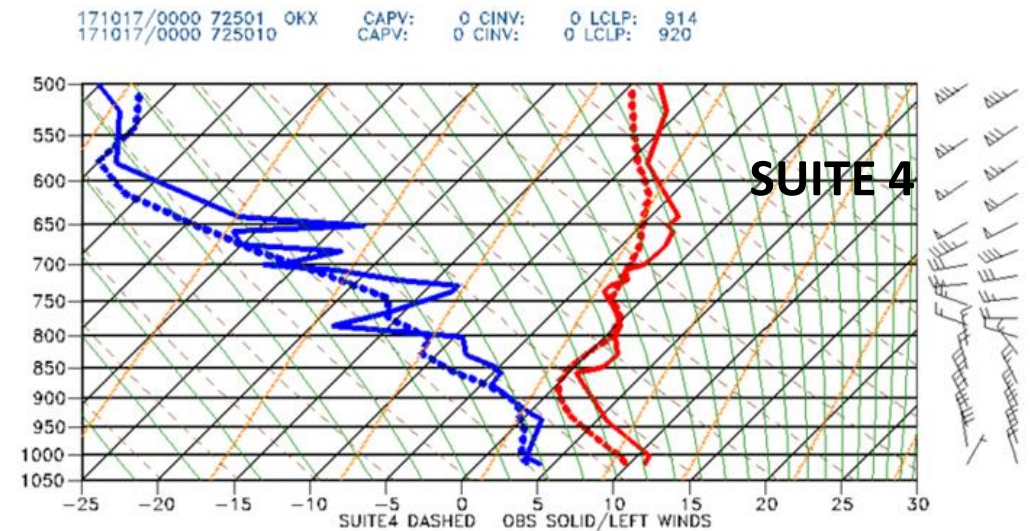
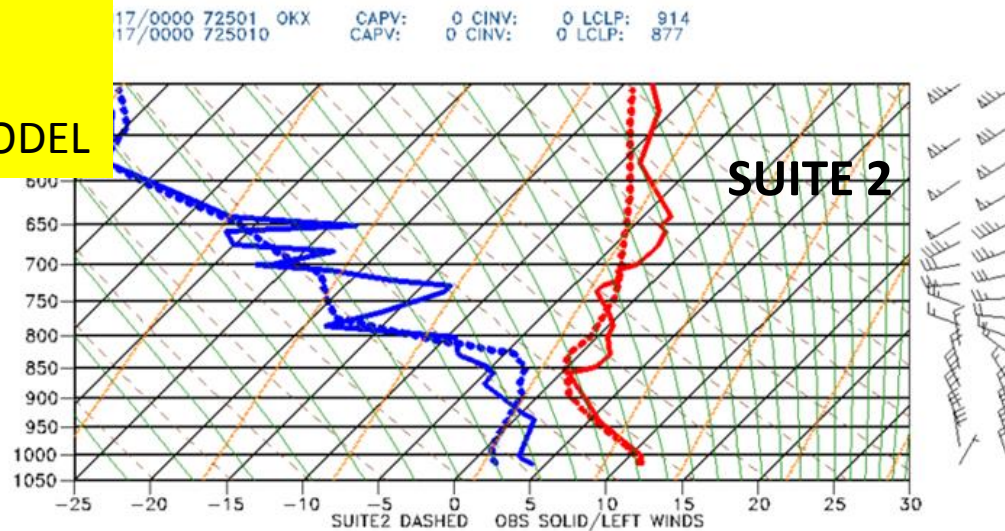
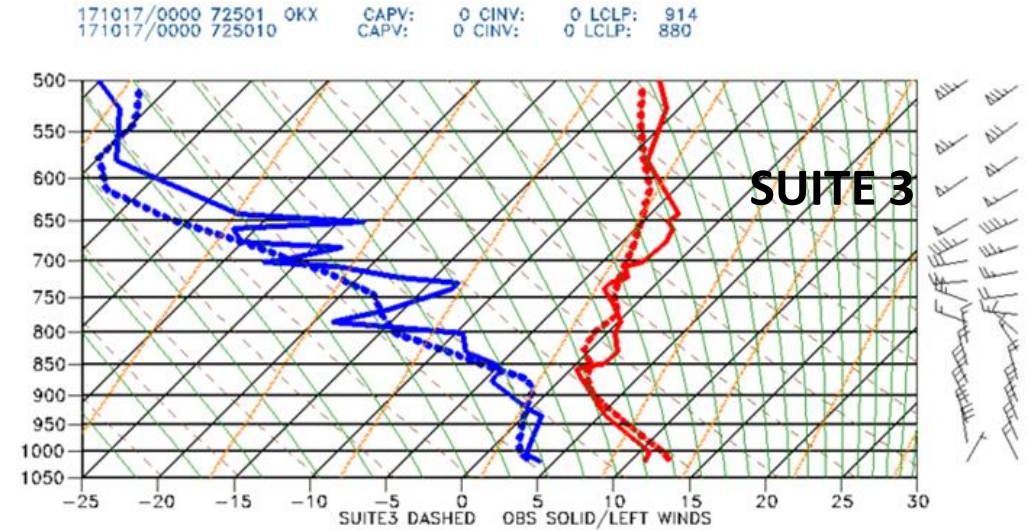
FORMAT of SOUNDINGS SLIDES



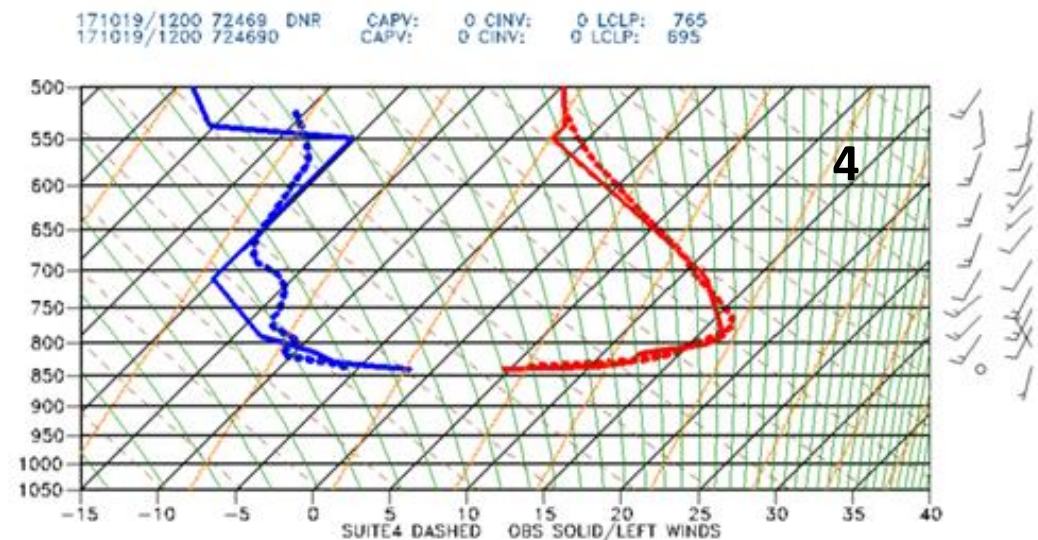
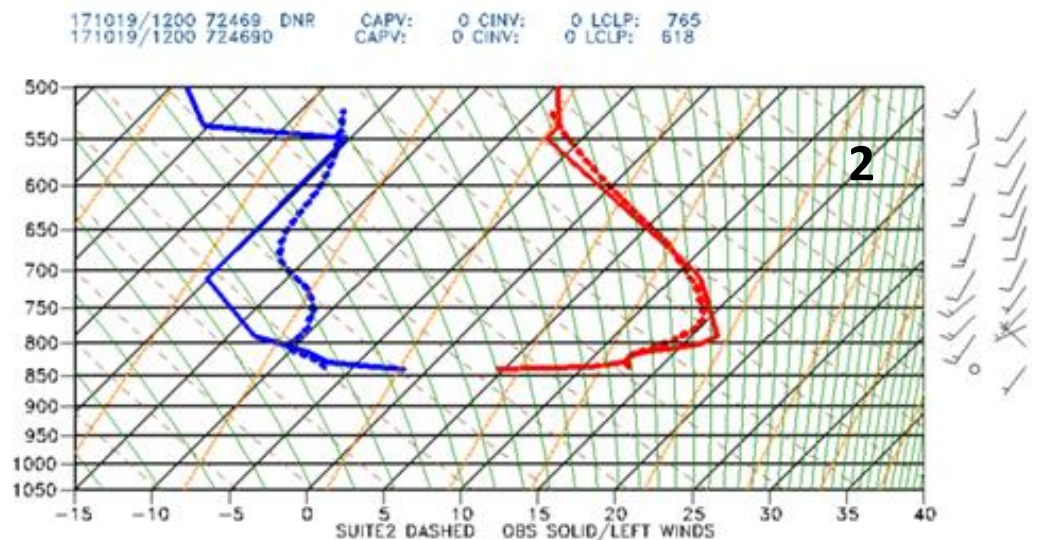
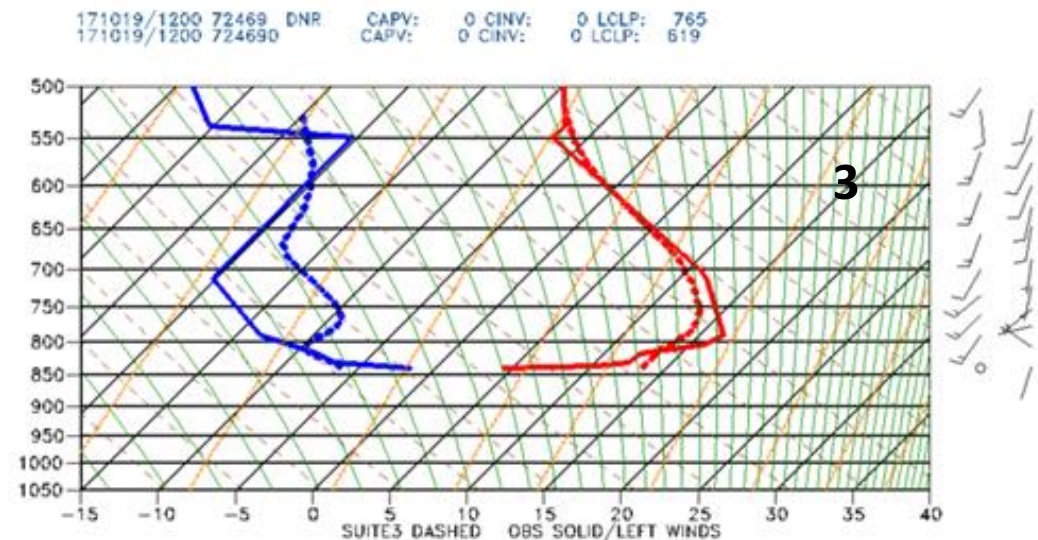
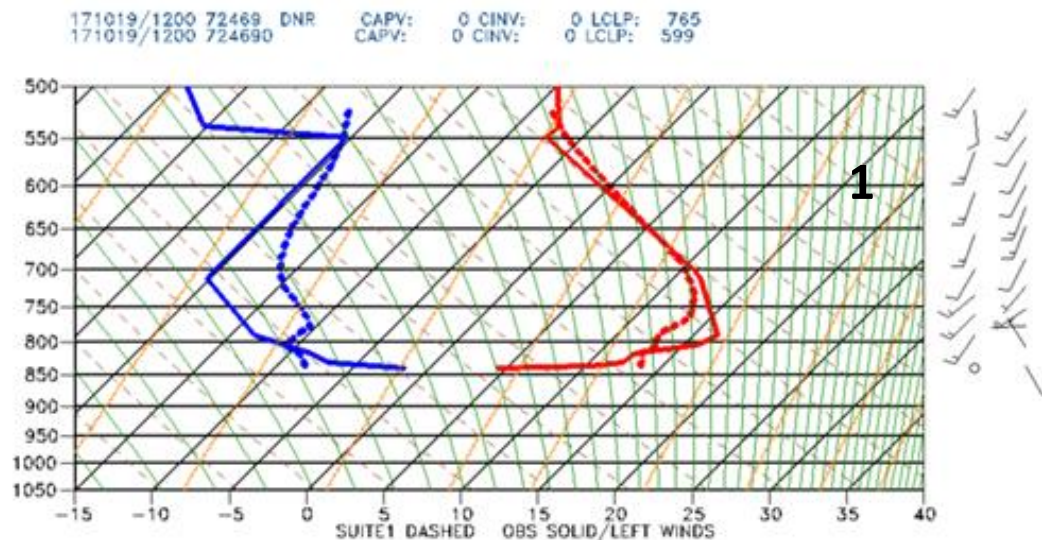
DASHED = MODEL
SOLID = OBS

LEFT WIND
COLUMN = OBS

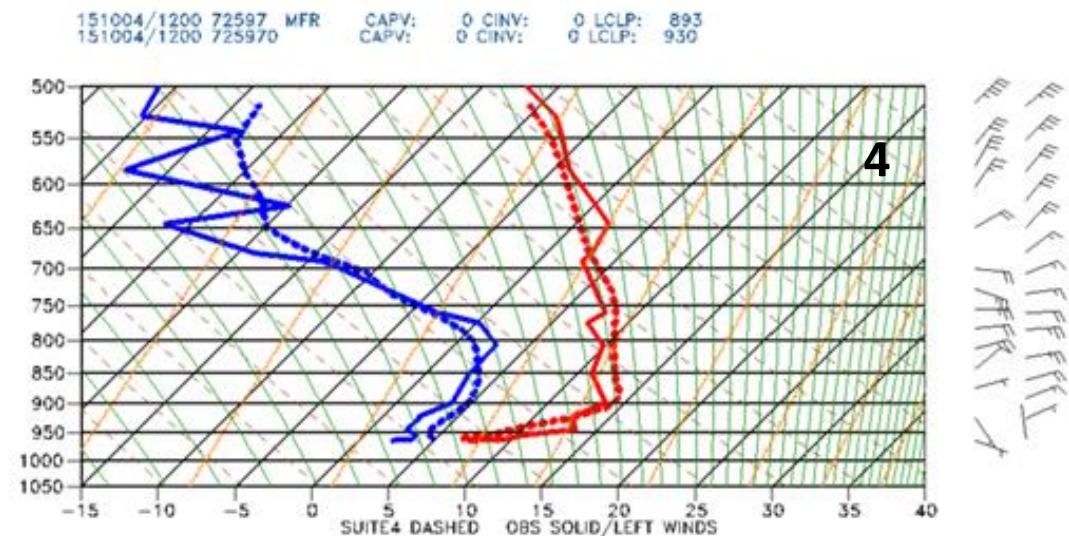
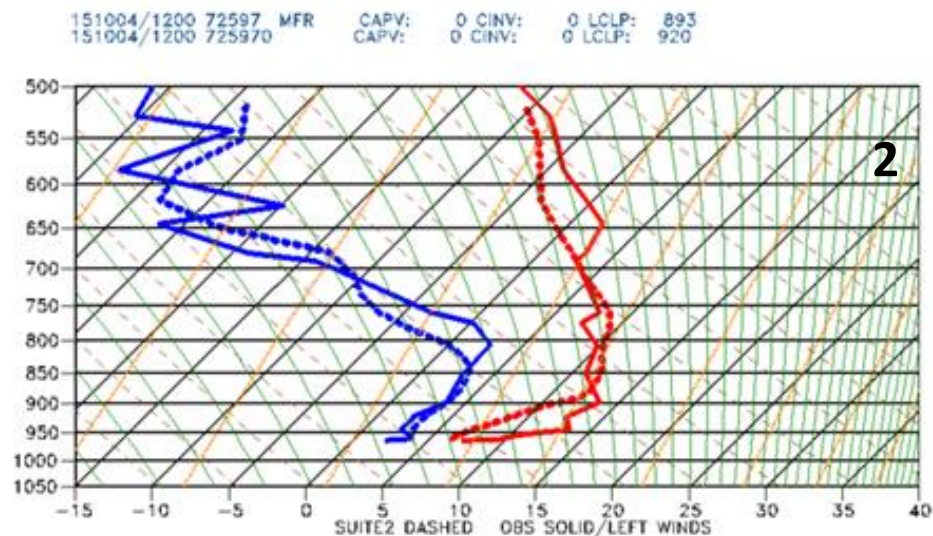
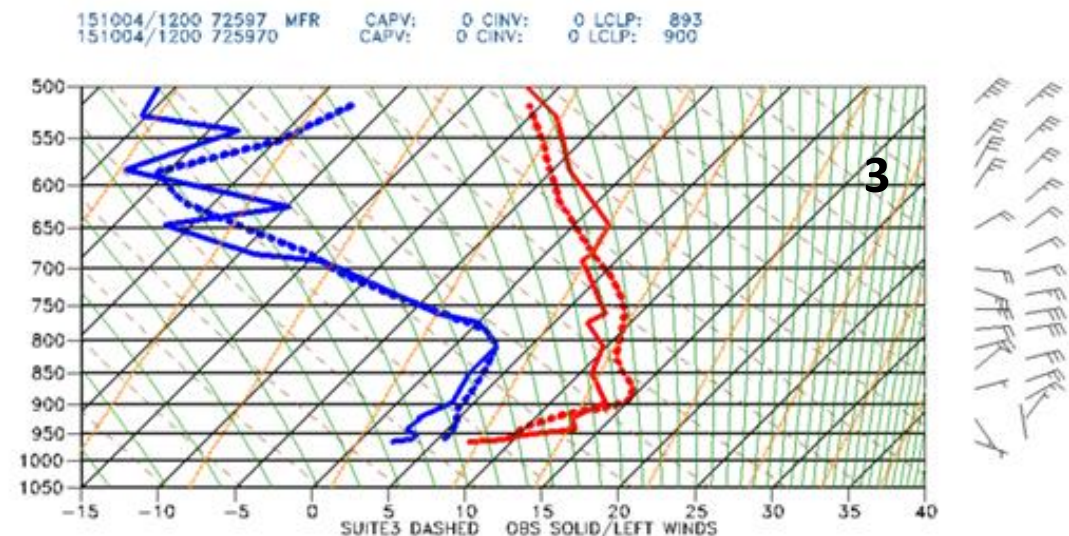
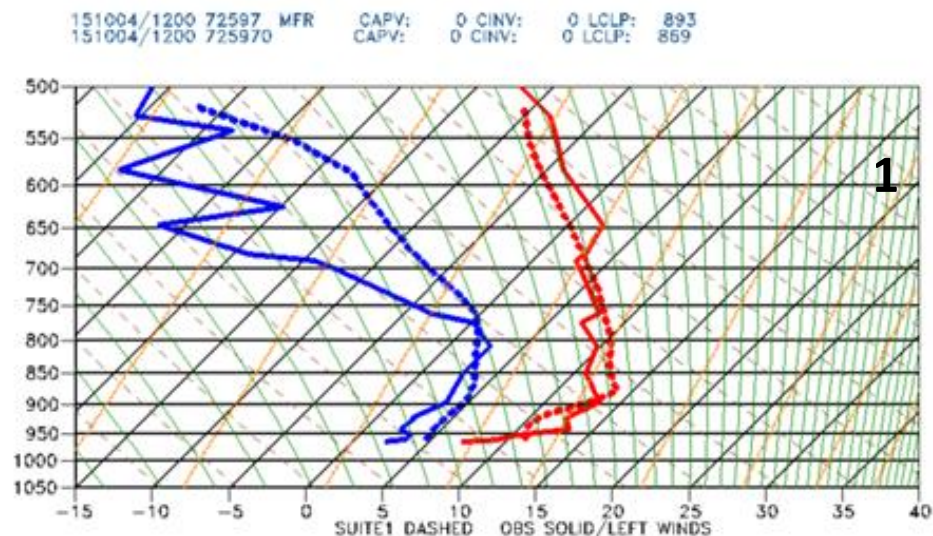
RIGHT WIND
COLUMN = MODEL



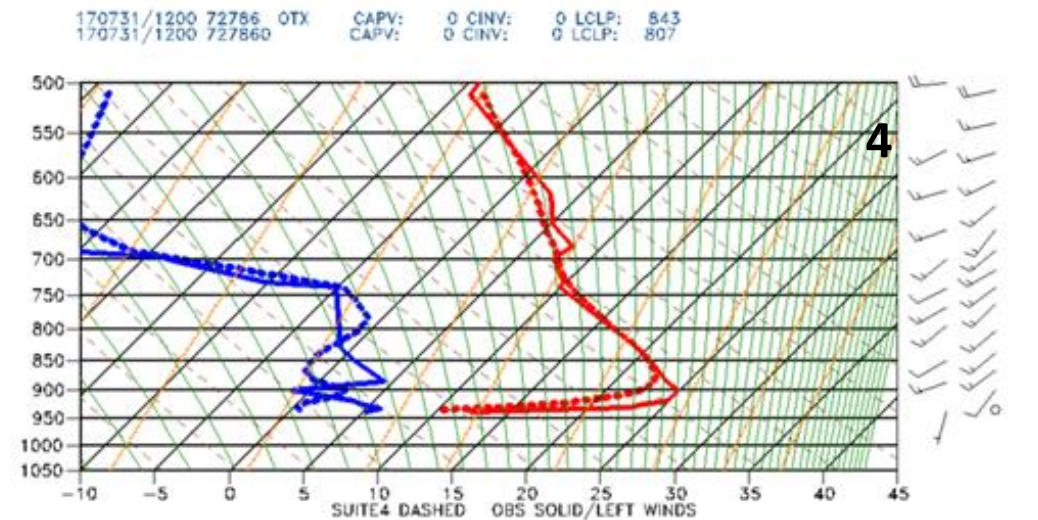
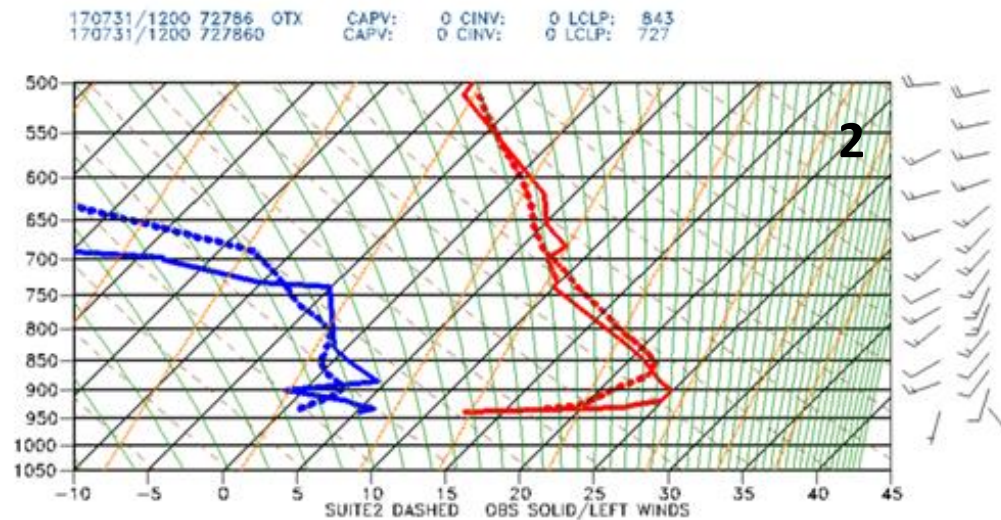
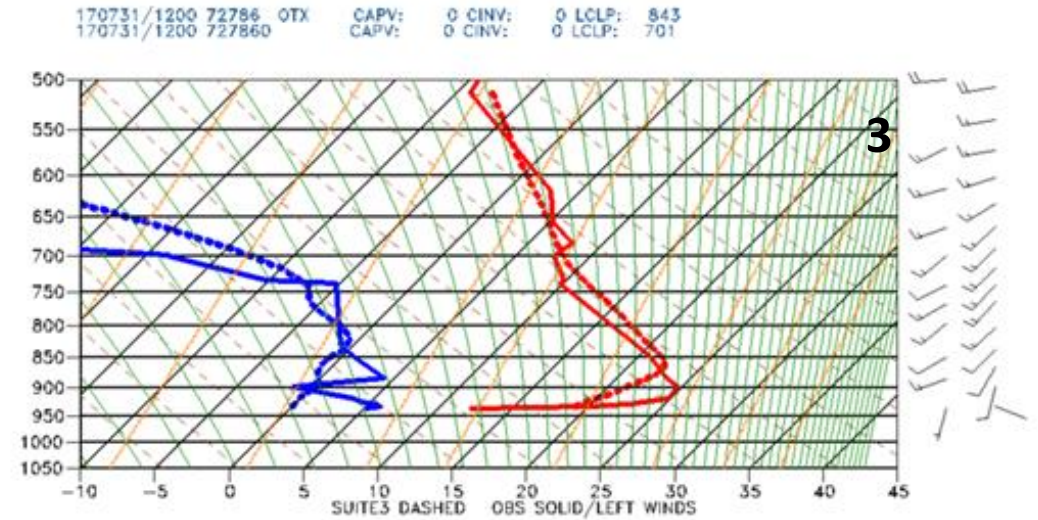
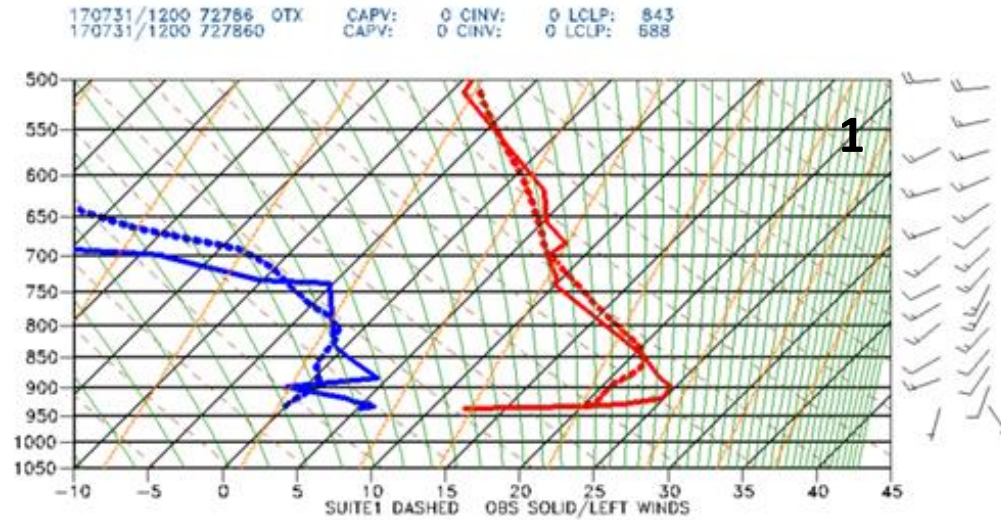
MOST COMMON INVERSION FINDING: Suite 4 handles inversions the best, Suite 1 struggles the most, Suite 2 offers some level of improvement over 1, and Suite 3 sometimes offers modest improvement over 1



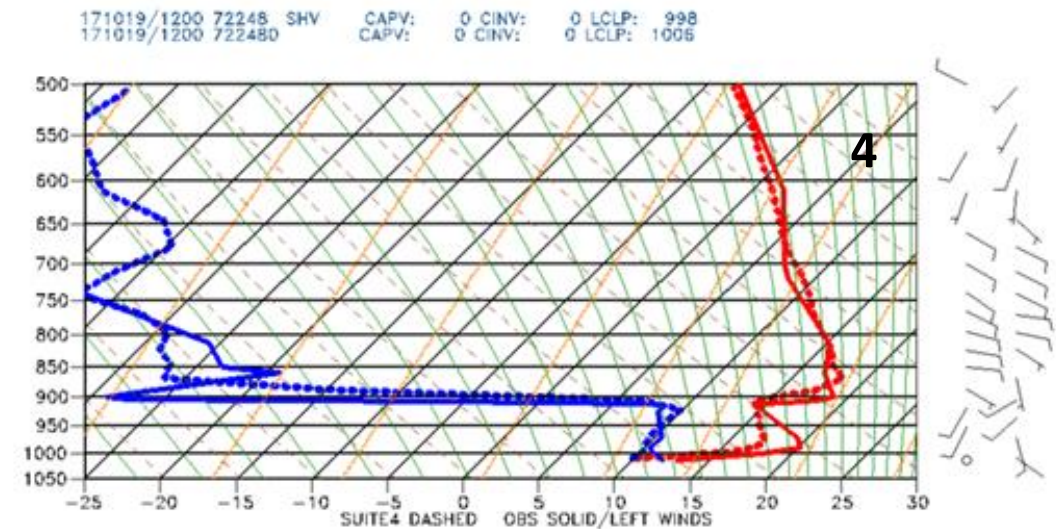
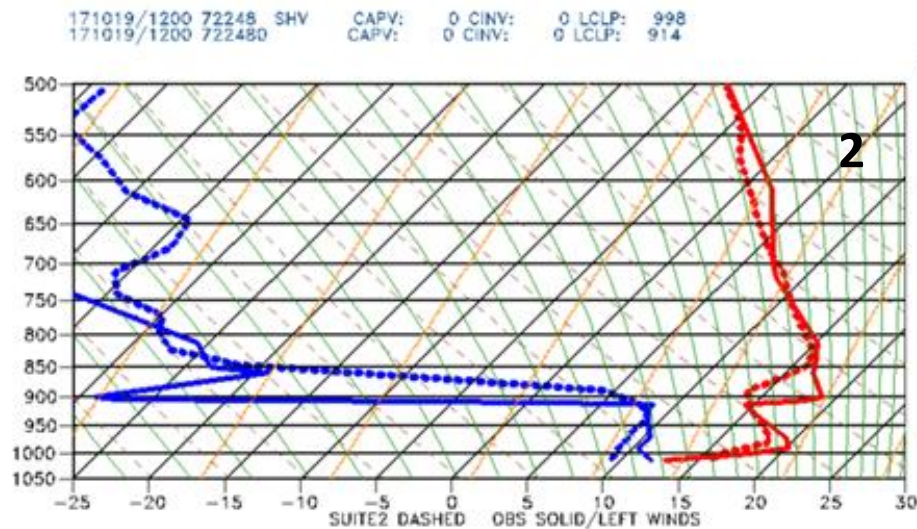
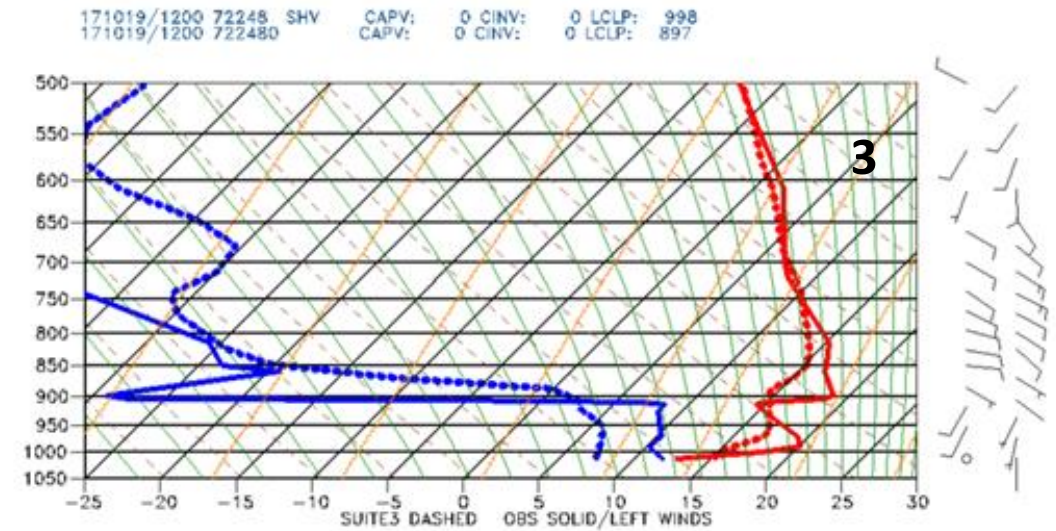
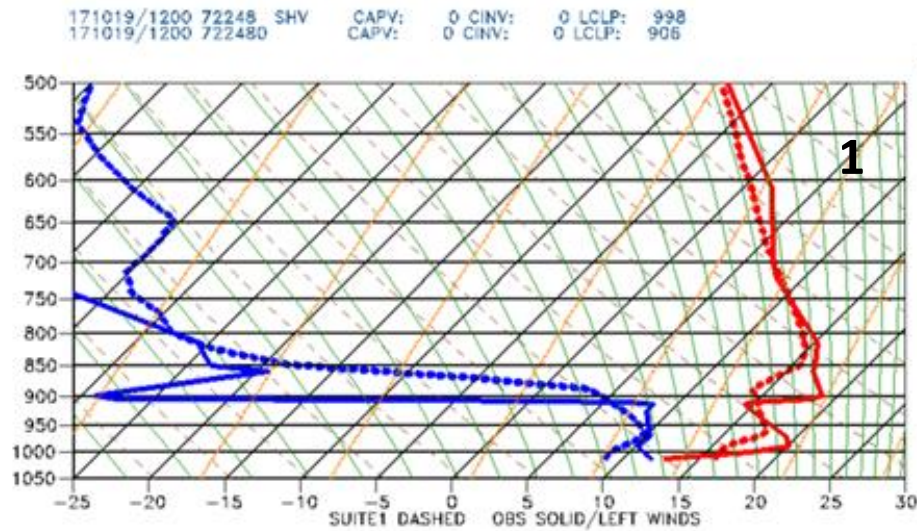
MOST COMMON INVERSION FINDING: Suite 4 handles inversions the best, Suite 1 struggles the most, Suite 2 offers some level of improvement over 1, and Suite 3 sometimes offers modest improvement over 1



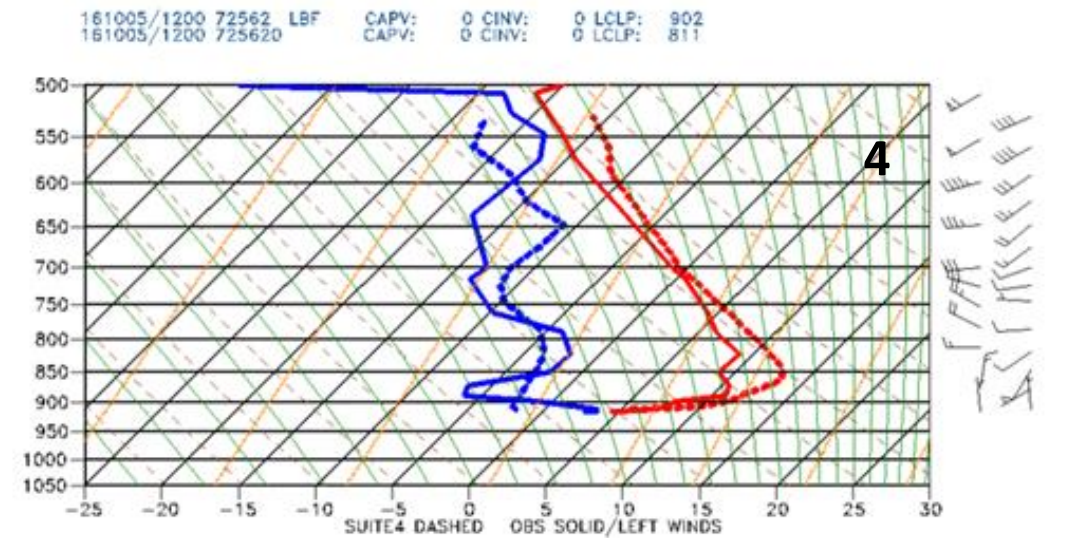
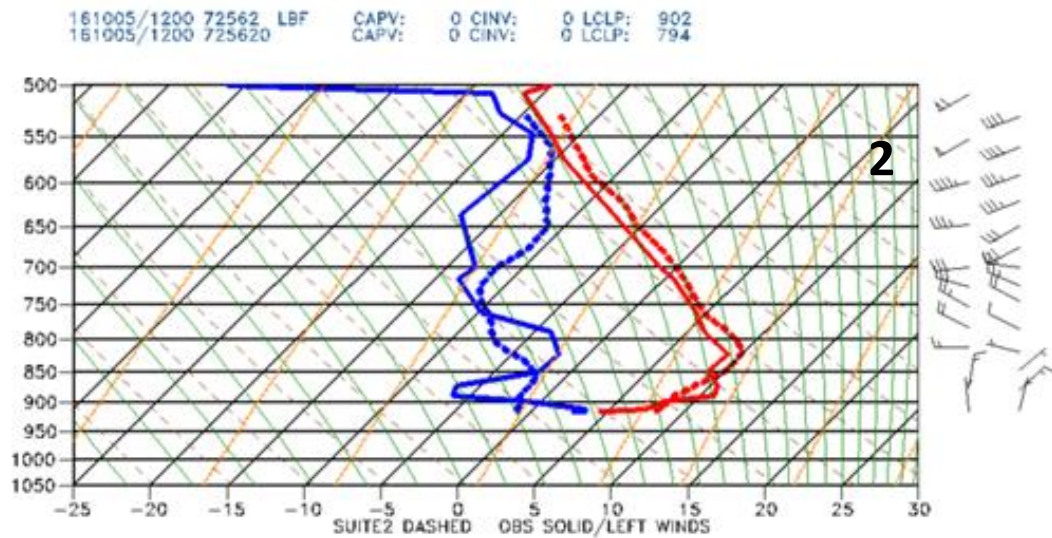
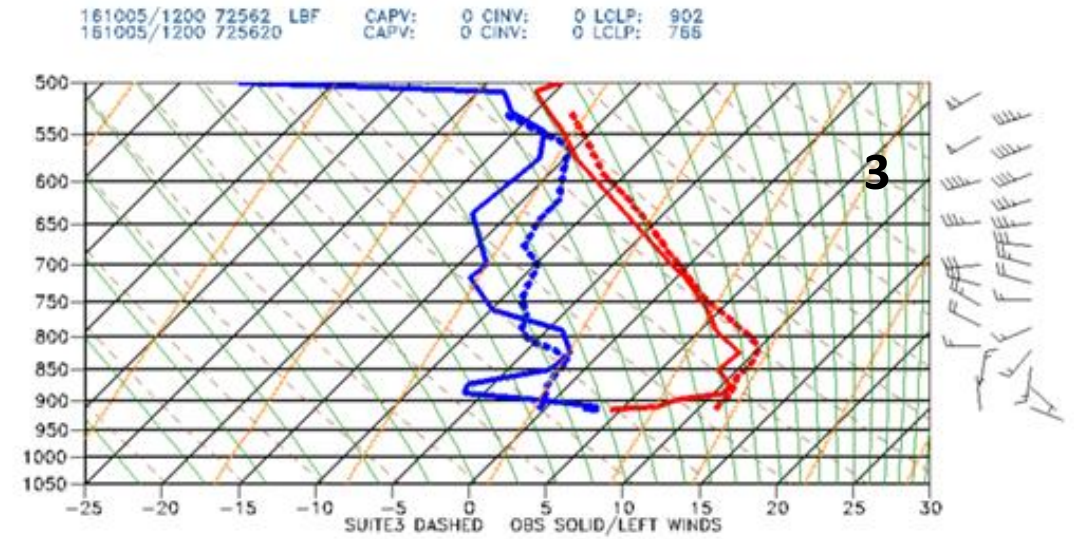
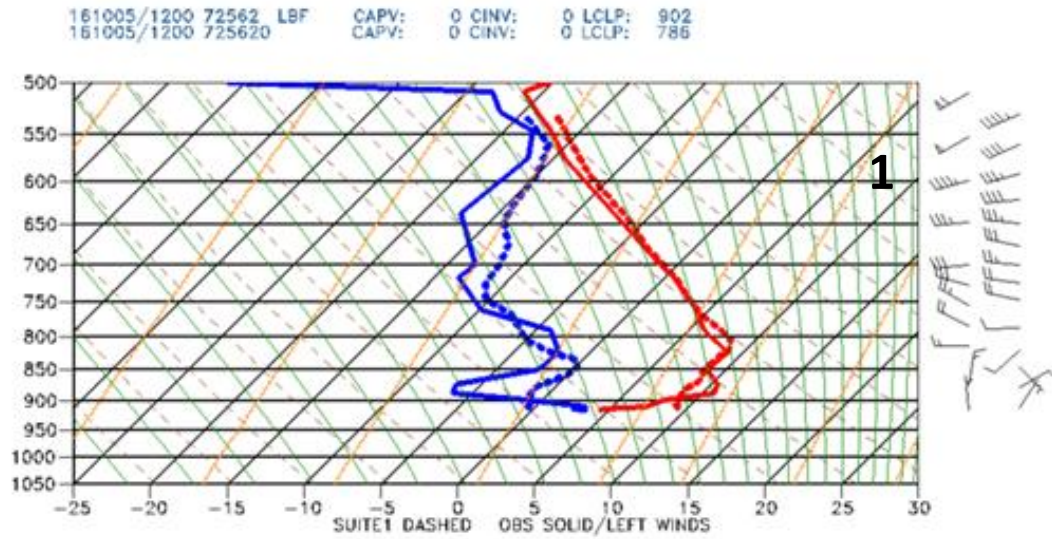
MOST COMMON INVERSION FINDING: Suite 4 handles inversions the best, Suite 1 struggles the most, Suite 2 offers some level of improvement over 1, and Suite 3 sometimes offers modest improvement over 1



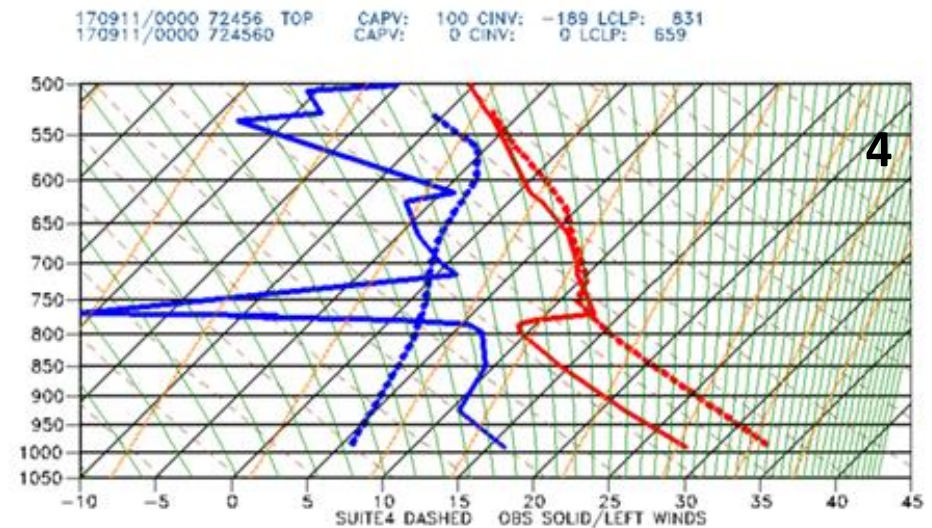
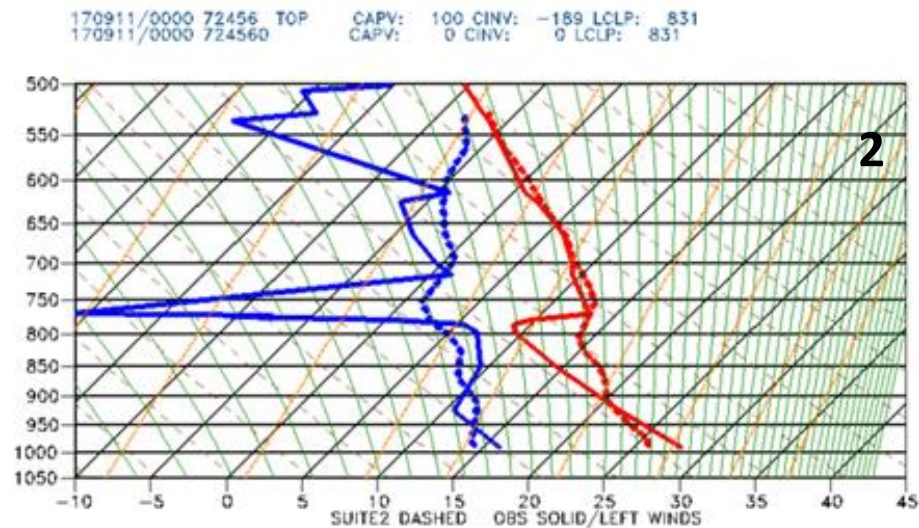
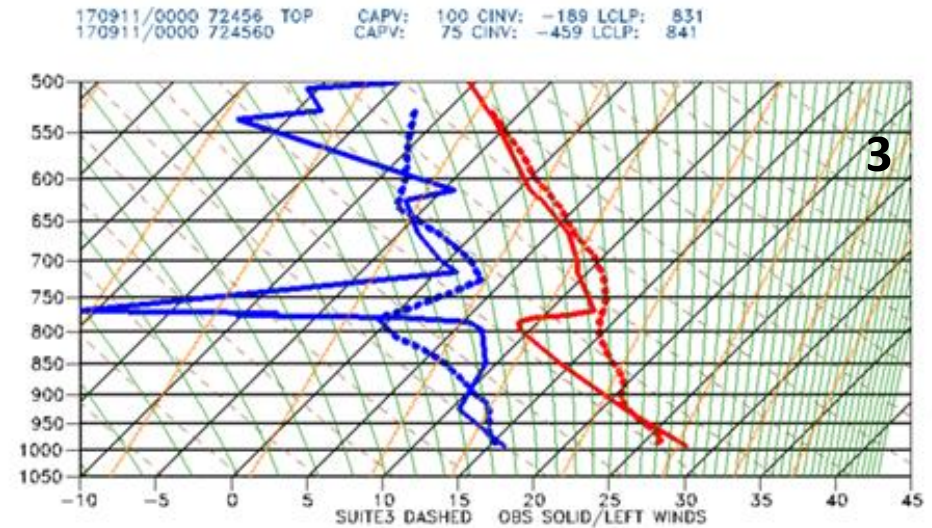
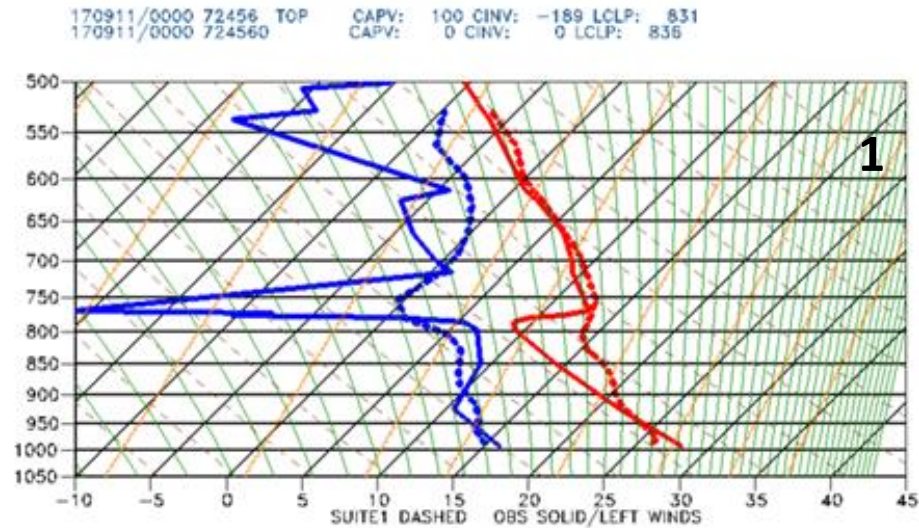
There are examples, however, of Suite 4 predicting low-level inversions that are too strong



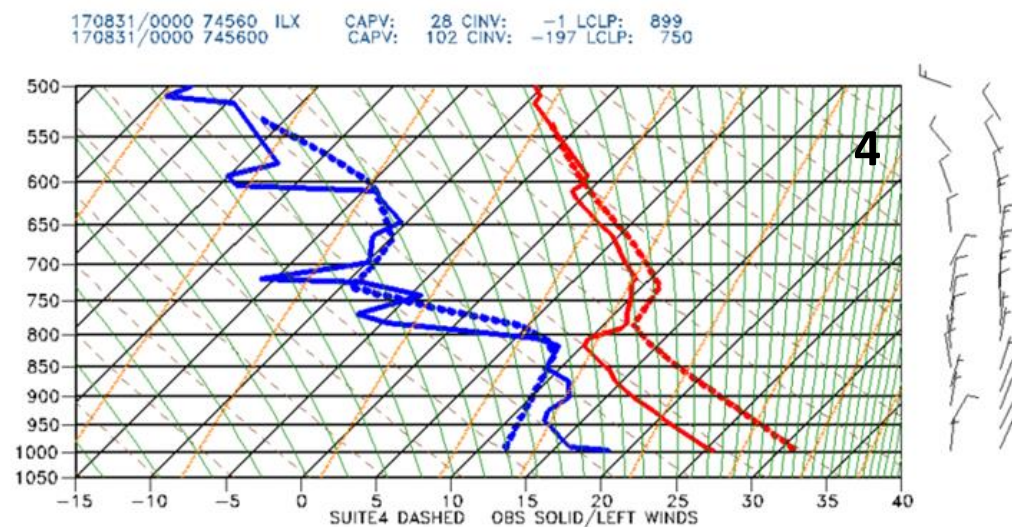
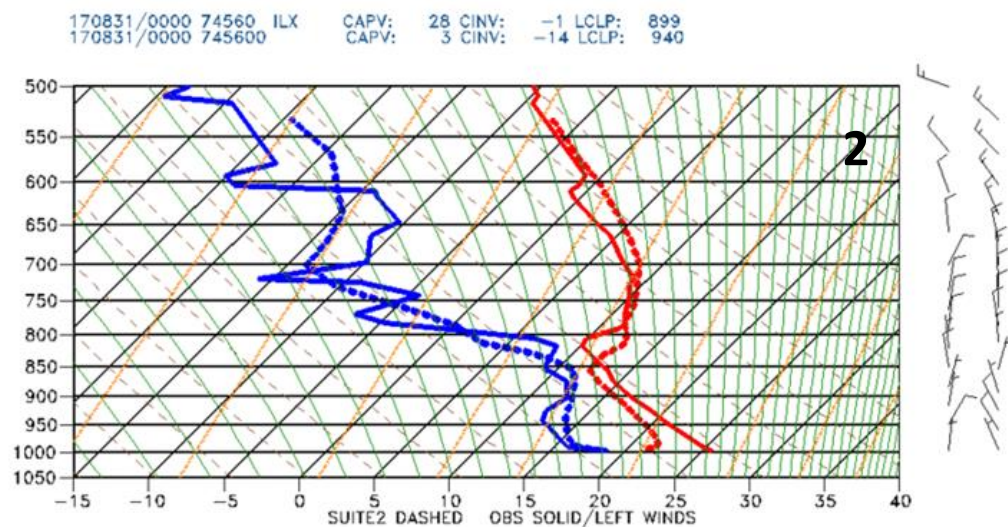
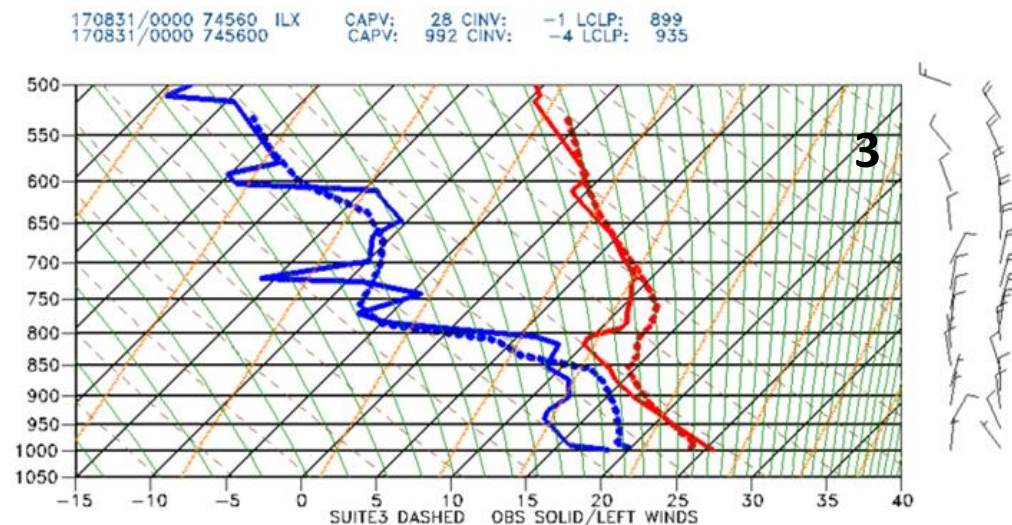
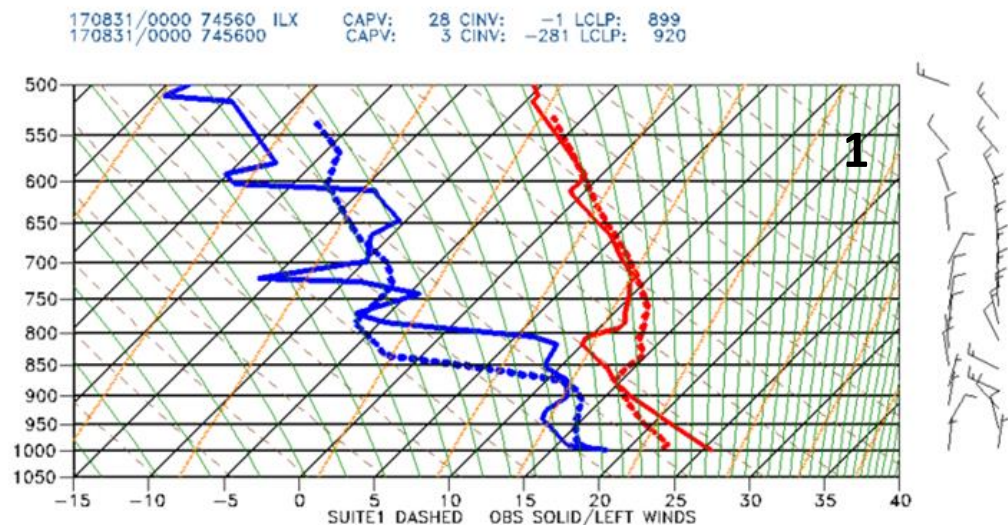
There are also examples in which Suite 4 overall handles an inversion the best but is significantly too warm at the top of the inversion

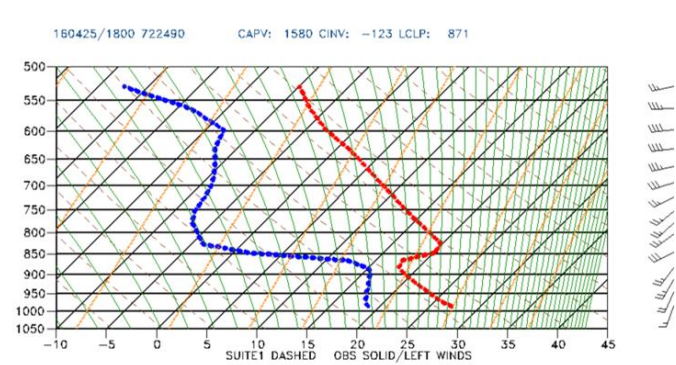


MOST COMMON AFTERNOON FINDING: Suite 4 often overmixes the PBL, leading to the lower levels being too hot and too dry

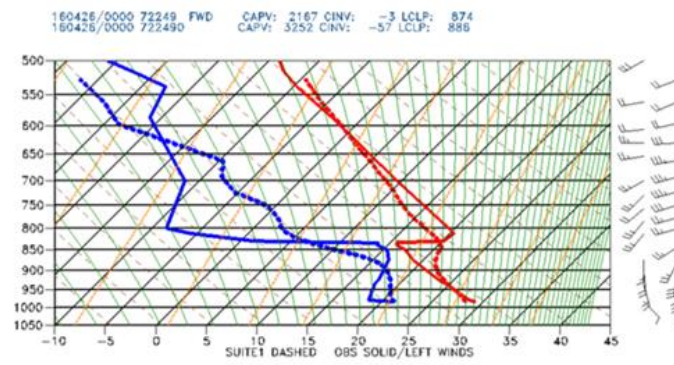


MOST COMMON AFTERNOON FINDING: Suite 4 often overmixes the PBL, leading to the lower levels being too hot and too dry

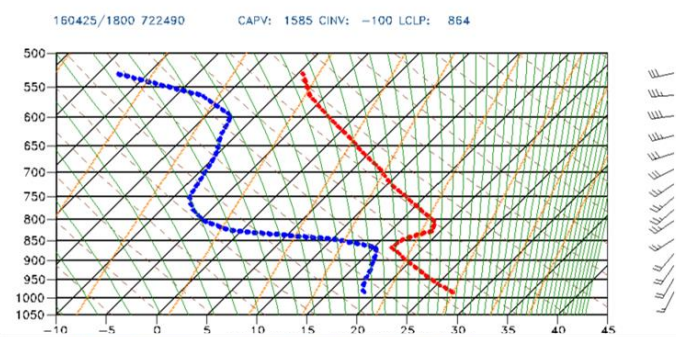




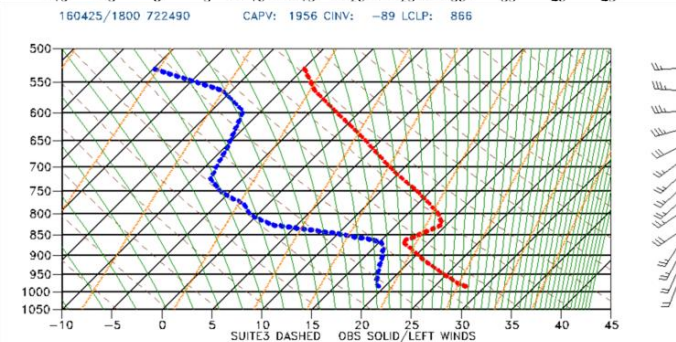
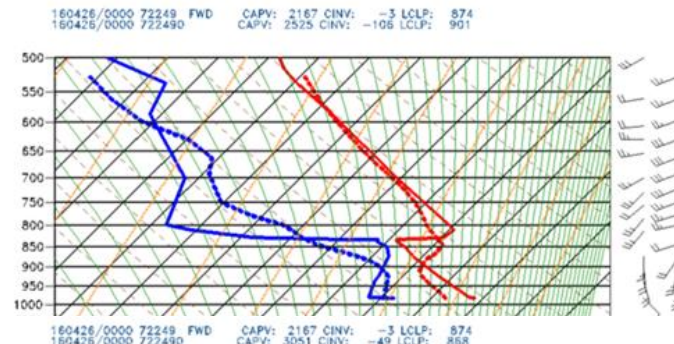
1



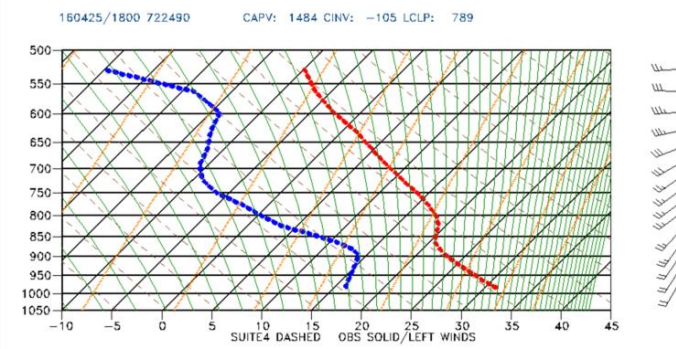
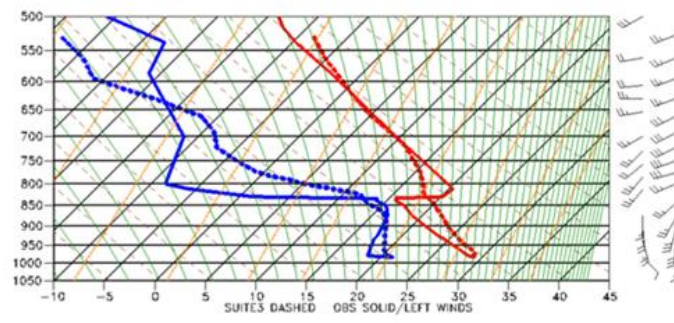
CAPPING INVERSIONS: Suite 4 often mixes out capping inversions which are critical features in pre-convective environments



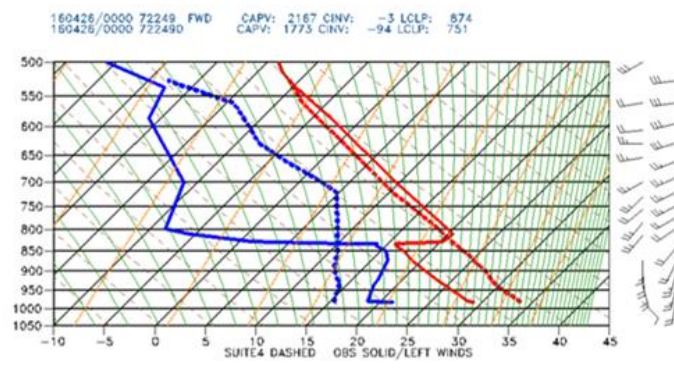
2



3



4

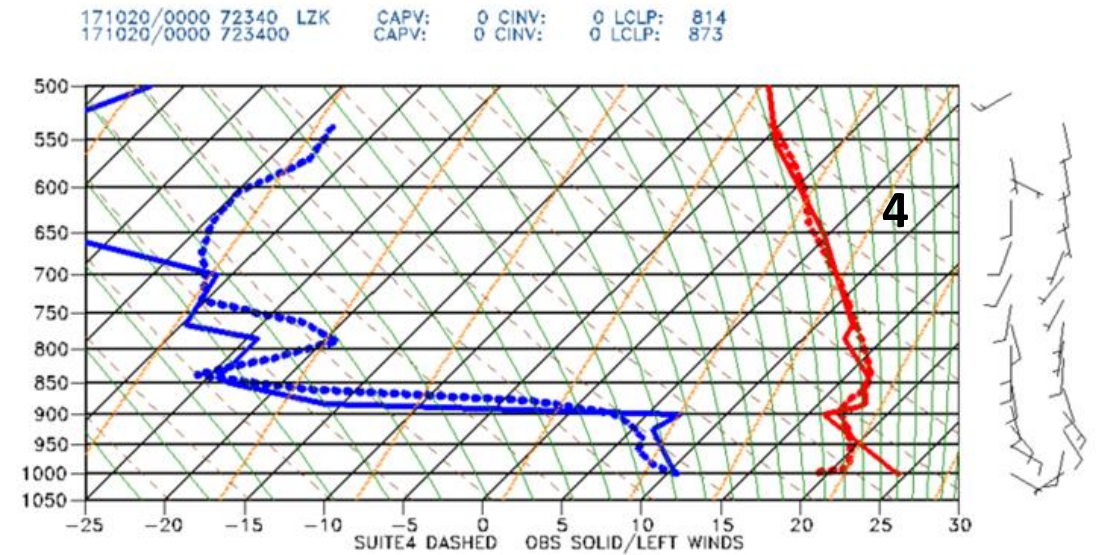
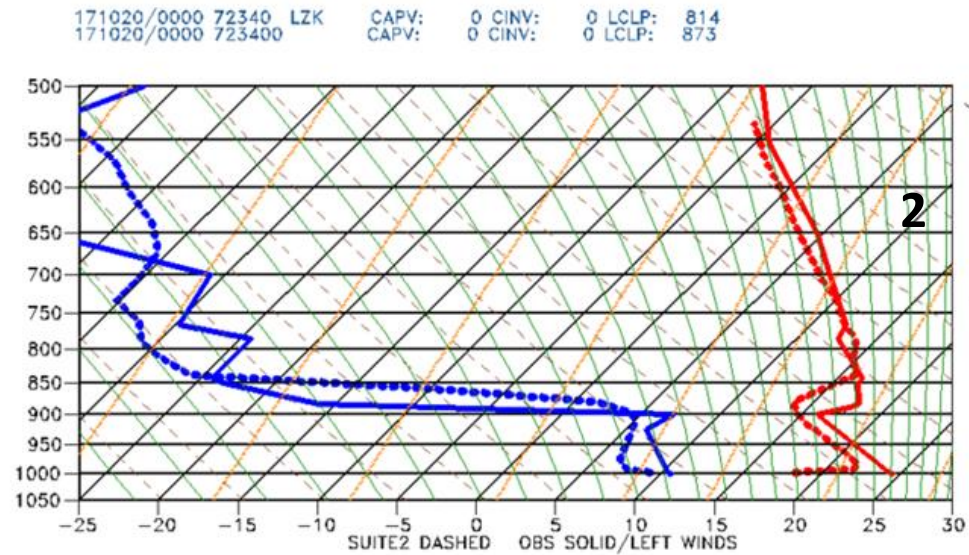
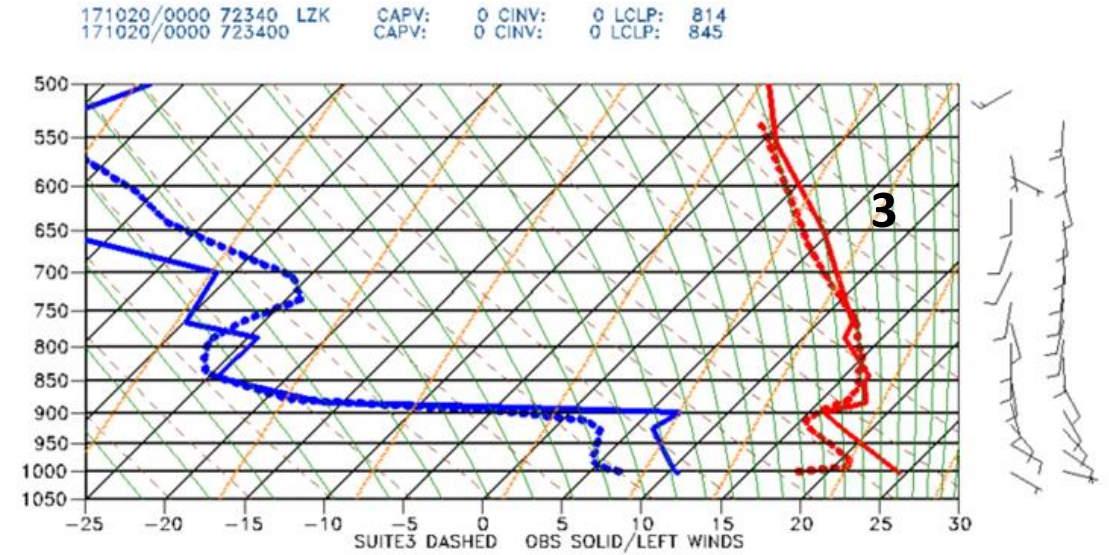
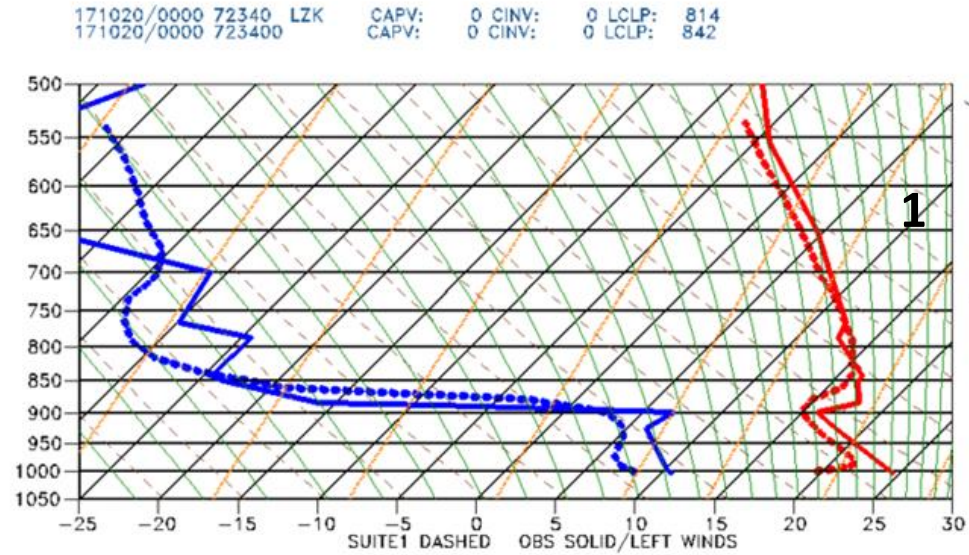


- The process begins during early afternoon

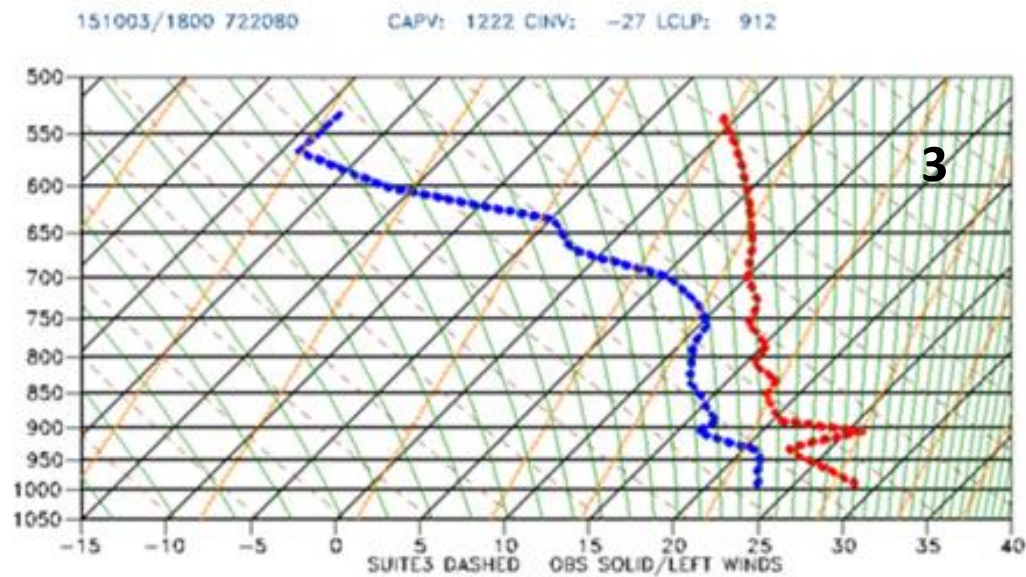
- By early evening, the inversion is too weak in all suites, but it's completely gone in Suite 4

- The temperatures around the inversion level in Suite 3 looks odd by 00z as well

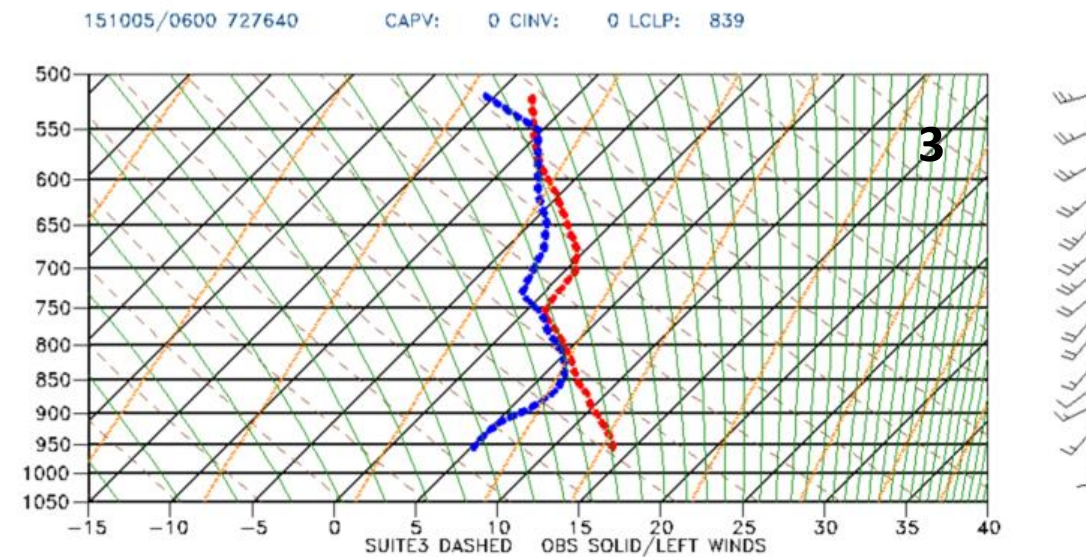
SUITE 2 LOW-LEVEL 00z COLD BIAS: Indications of early decoupling??



SUITE 3 OCCASIONALLY SHOWS SOME ODD FEATURES IN PROFILES



Extremely unrealistic inversion



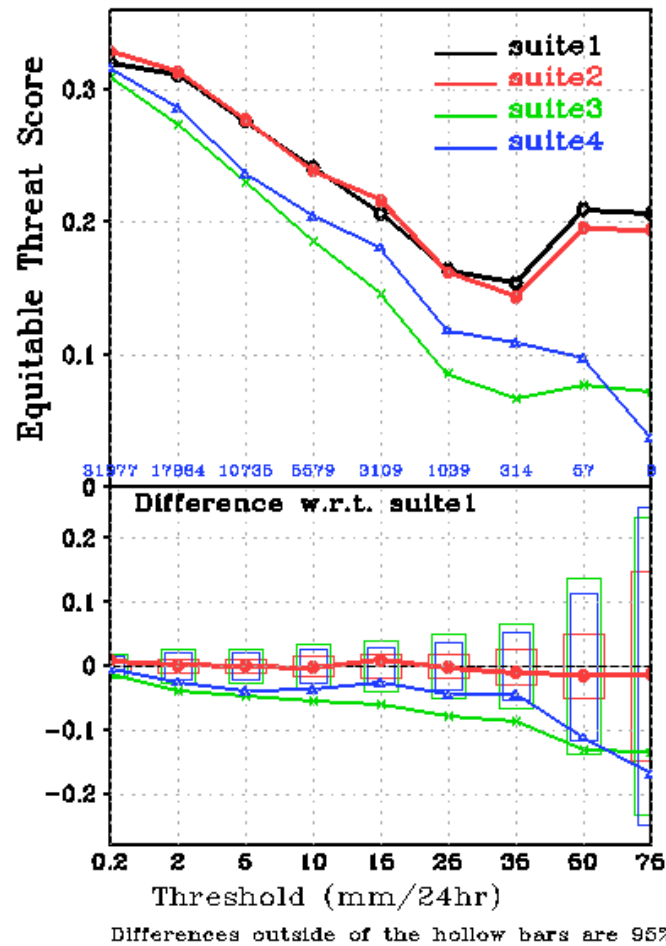
Odd low-level moisture structure

Precipitation

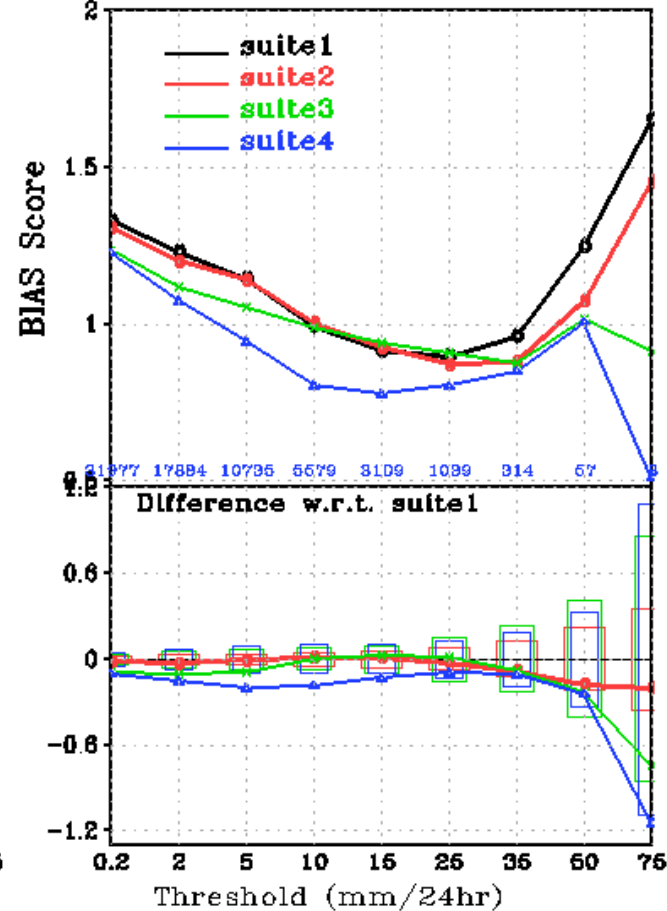
Day-5 CONUS Precipitation ETS and Bias (0000 UTC)

ETS

CONUS Precip Skill Scores, f108-f132, 01Jan2016-10Jan2018 00Z Cycle



Bias



SUITE 1

SUITE 2

SUITE 3

SUITE 4

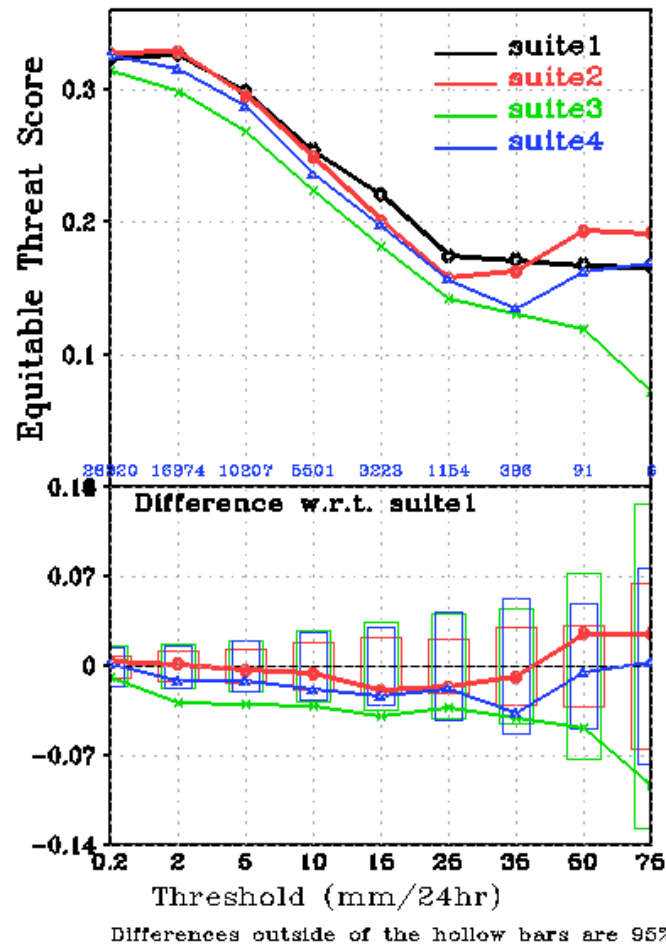
KEY POINTS:

- Suites 1 and 2 have the highest ETSs, while Suite 3 has the lowest ETS
- All four suites have a wet bias at low thresholds & dry bias at mid-thresholds
- Suite 4 has a statistically significantly worse dry bias than other suites for amounts between 10-25 mm/24h

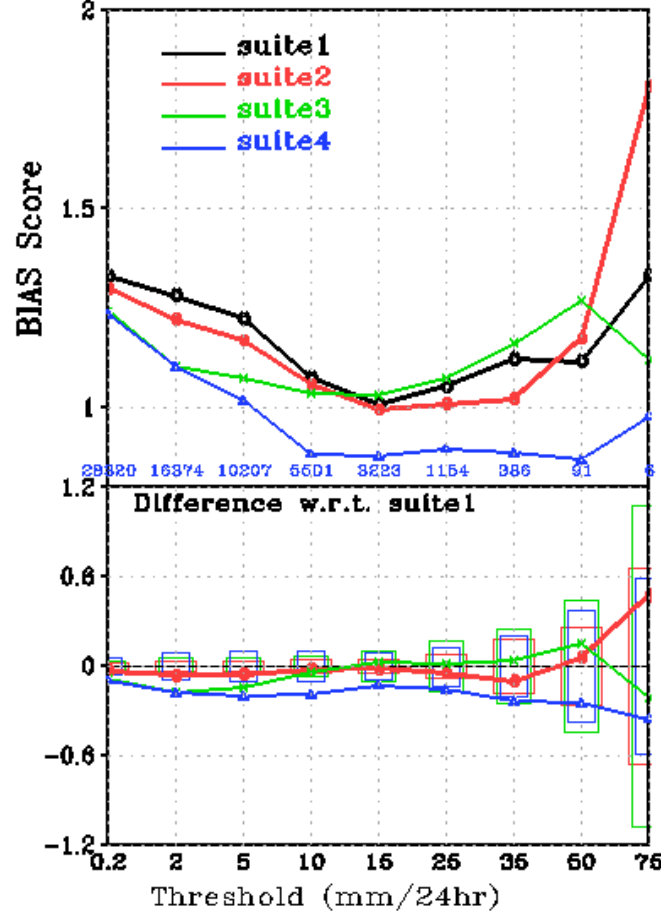
Day-5 CONUS Precipitation ETS and Bias (1200 UTC)

ETS

CONUS Precip Skill Scores, f96-f120, 01jan2016-10jan2018 12Z Cycle



Bias



SUITE 1

SUITE 2

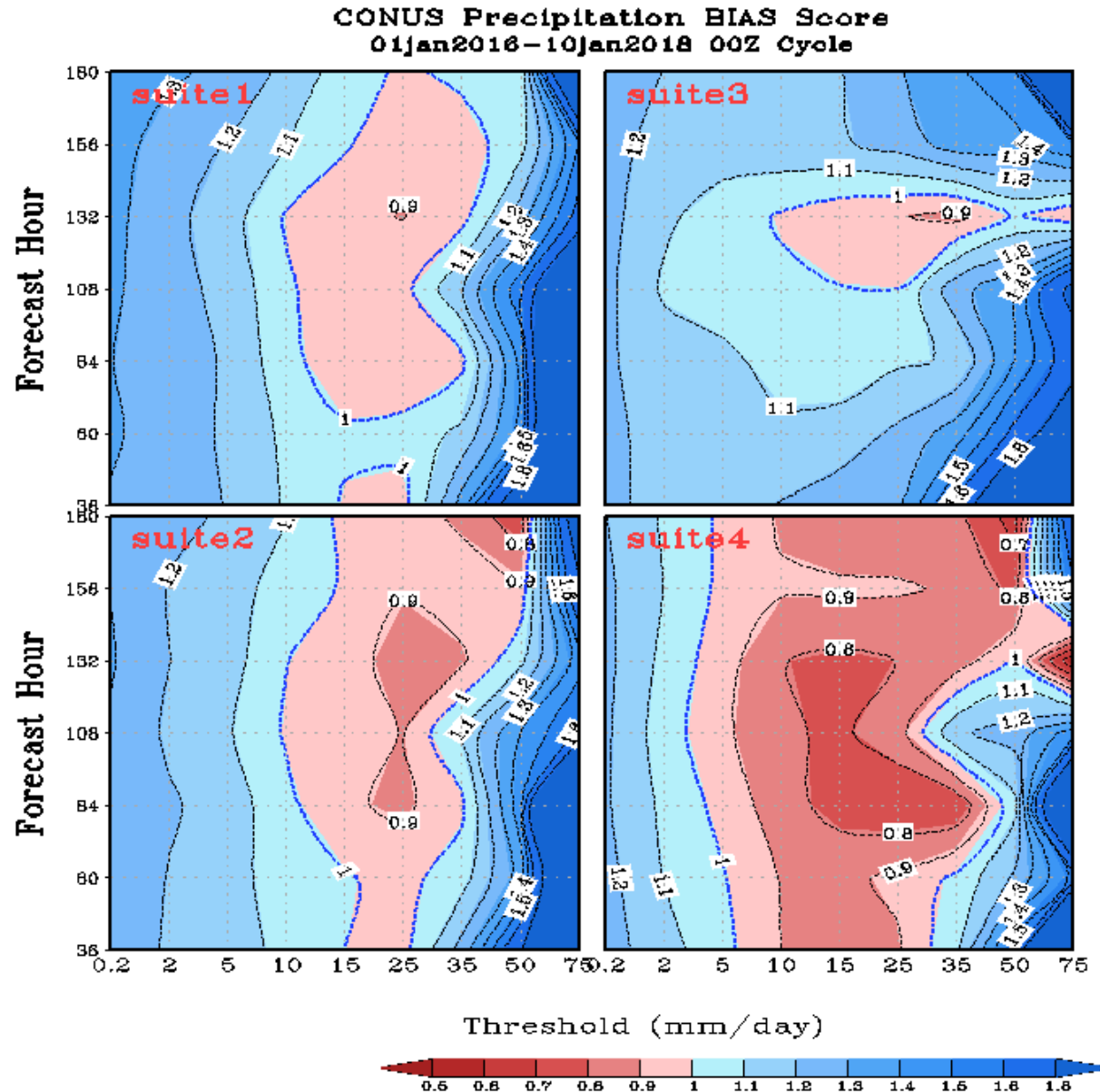
SUITE 3

SUITE 4

KEY POINTS:

- ETSs increase across all four suites
- Suites 1 and 2 have the highest ETSs, while Suite 3 has the lowest ETS
- All four suites have a wet bias at low thresholds that decreases at middle thresholds (Suites 1–3 = no dry bias)

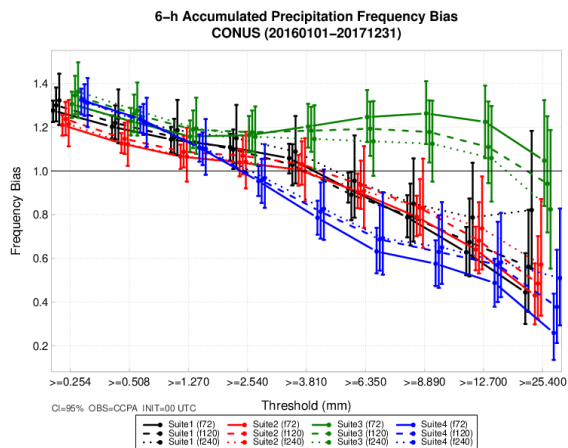
CONUS Precipitation Bias (0000 UTC)



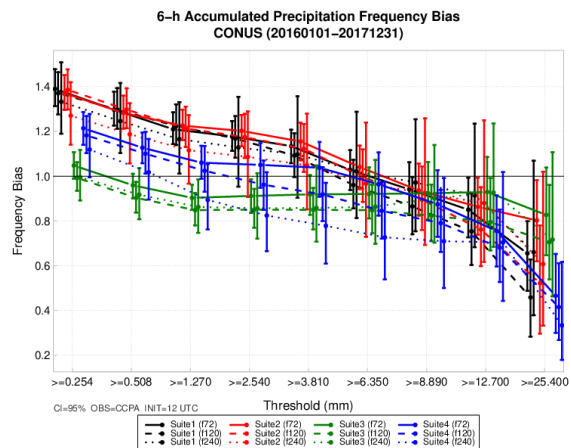
KEY POINTS:

- At lower precipitation thresholds, a wet bias exists at all lead times in all suites
- All four suites have an increasing dry bias with time at middle thresholds
- Suite 4 has the largest dry bias that increases with time at middle and high precipitation thresholds

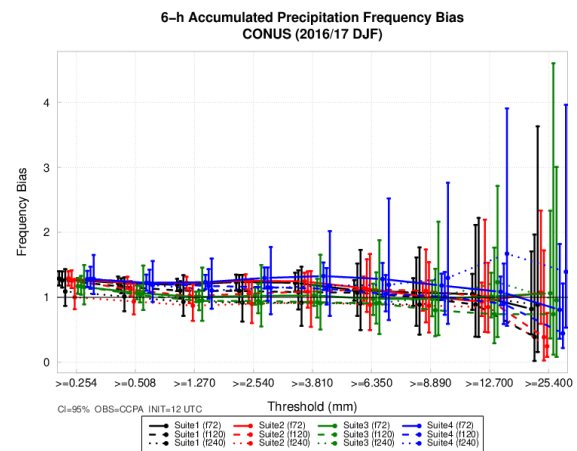
00 UTC



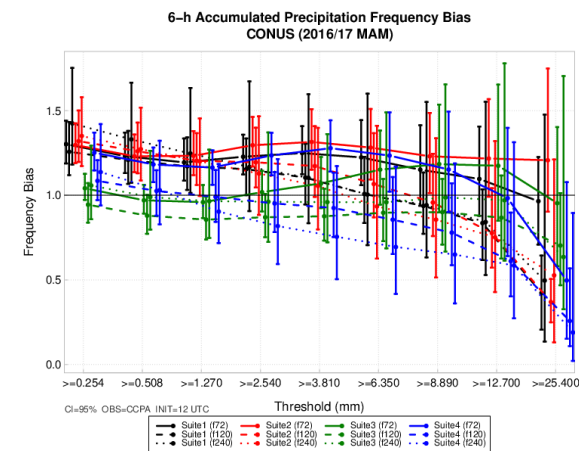
12 UTC



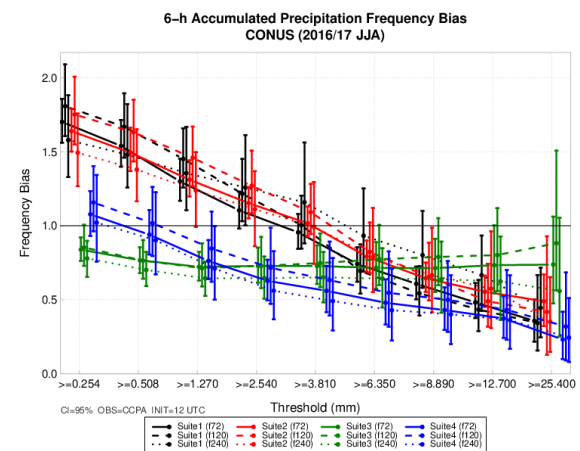
DJF



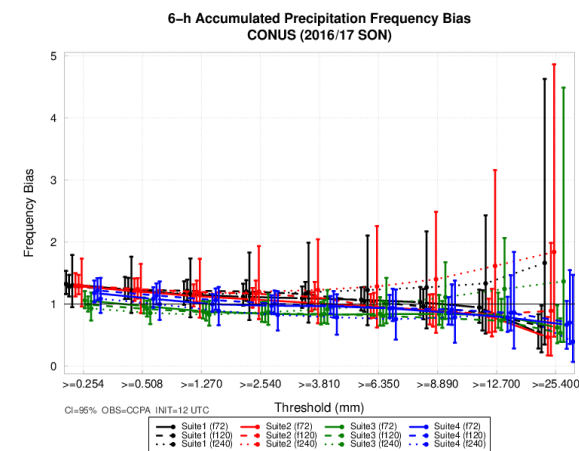
MAM



JJA



SON



https://docs.google.com/document/d/1Vyo1zL5N6GowugDCEWOVFlrK83JcOzWKCCSwGvGe_74/edit?ts=5c89156e#

6 h Frequency bias,
CONUS

6h acc. Fraction Skill Score, CONUS

<https://docs.google.com/document/d/1Vyo1zL5N6GowugDCEWOVFlrK83JcOzWKCCSwGvGe74/edit?ts=5c89156e#>

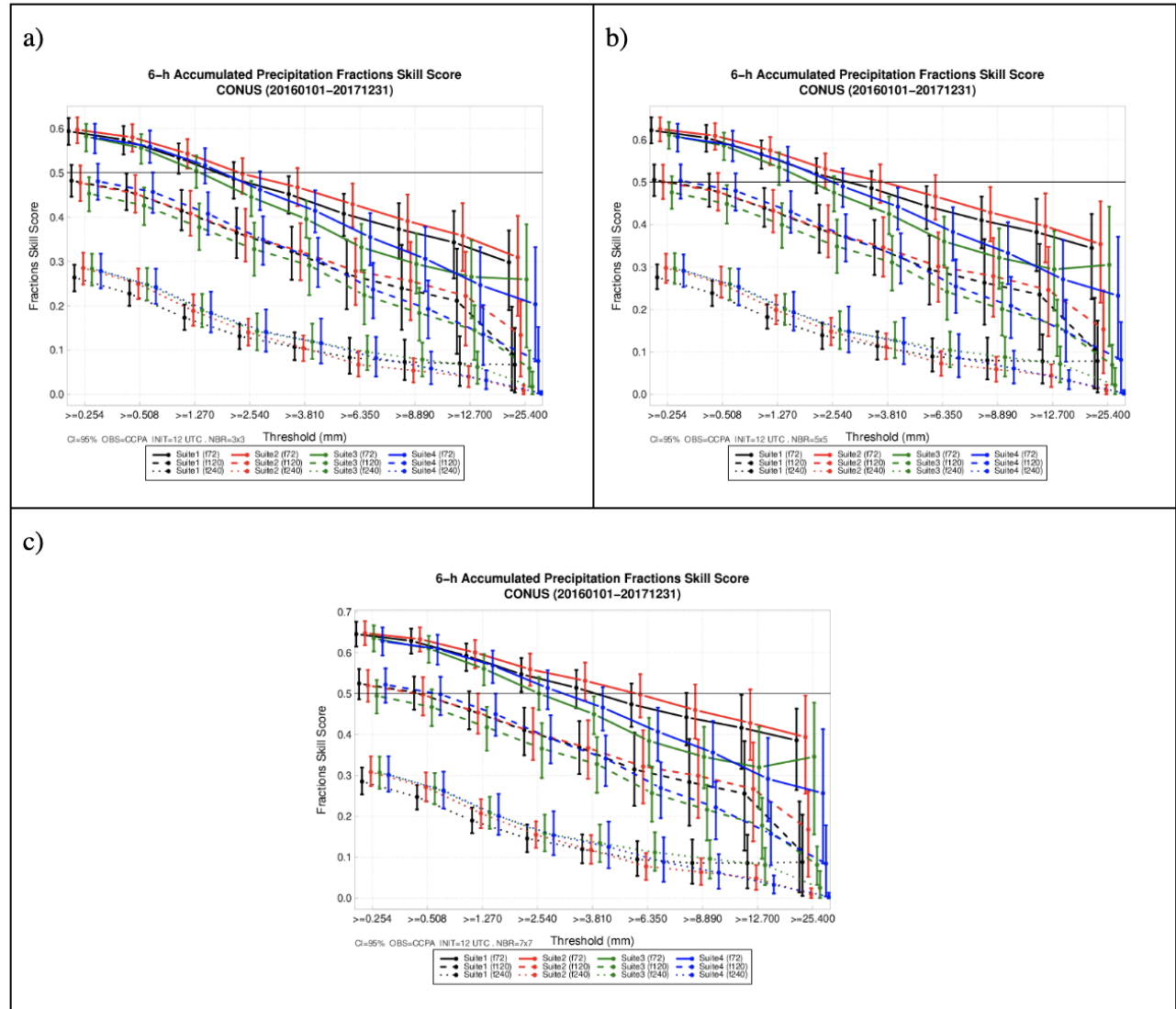
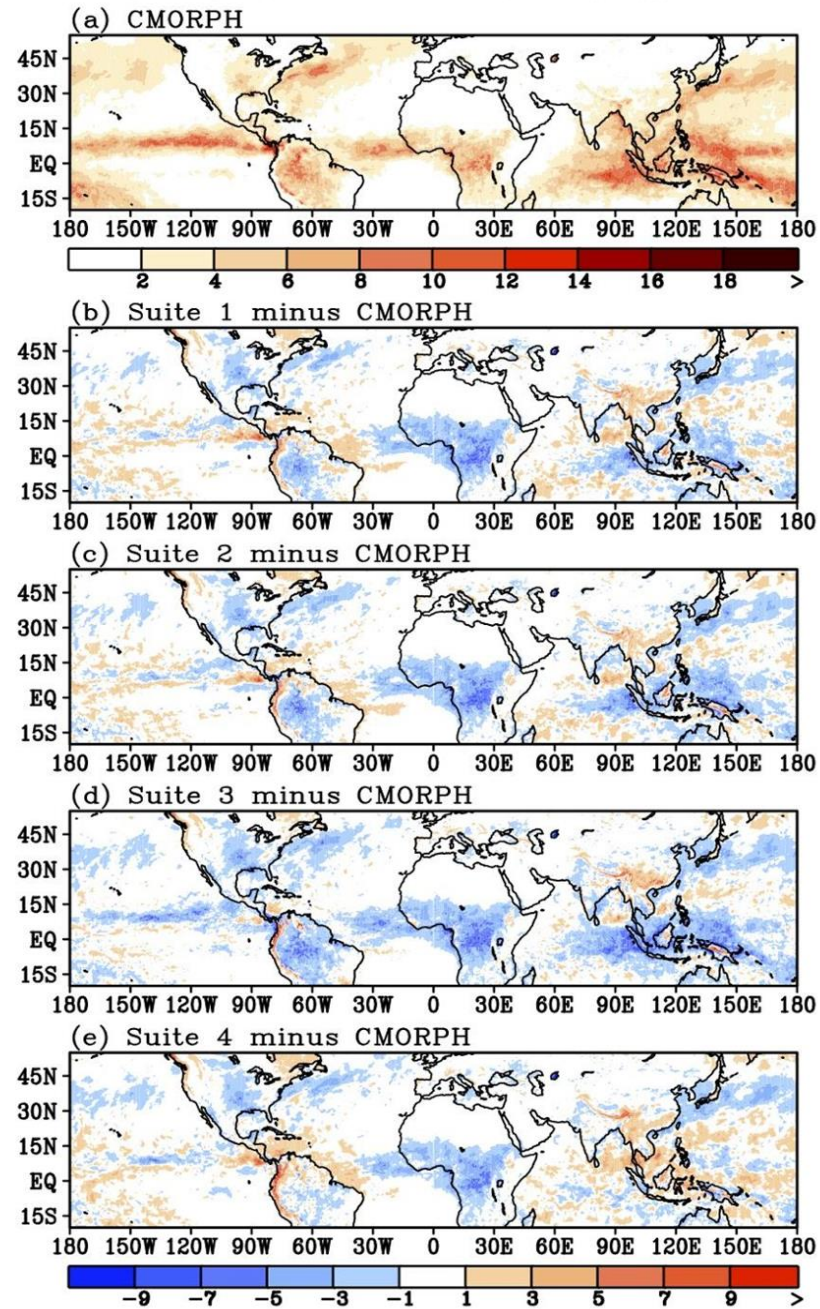


Figure 3. FSS of 6-h accumulated precipitation (mm) for Suite 1 (black), Suite 2 (red), Suite 3 (green), and Suite 4 (blue) aggregated over the CONUS domain for the entire test period (20160101-20171231) for all 12 UTC initializations using a neighborhood size of 49 (7 x 7 grid squares). The 72-h, 120-h, and 240-h forecasts are represented by the solid, dashed, and dotted lines, respectively. The vertical bars surrounding the aggregate value represent the 95% CIs.

Accumulated precip bias (mm day⁻¹; Jan 11–Dec 31, 2016–17)
[84–108h (00z cycle) and 108–132h (12z cycle)] minus CMORPH

MEDIUM RANGE PRECIP



CMORPH OBS

Suite 1 - OBS

Too dry in
tropics

Suite 2 - OBS

Even drier in tropics

Suite 3 - OBS

Driest in tropics

Suite 4 - OBS

Too wet in tropics

Statistics provided by GMTB

Summary from
the Model
Evaluation Group
(MEG).

POSITIVES for SUITE 1	NEGATIVES for SUITE 1
overall the best synoptic scores	really struggles with inversions
overall good tropical cyclone tracks and intensity	underdoes instability
	low-level cold bias increases with time
POSITIVES for SUITE 2	NEGATIVES for SUITE 2
synoptic scores overall match Suite 1	underdoes instability (maybe more than Suite 1)
matches Suite 1 for tropical tracks/intensity	larger low-level cold bias than suite 1
improved handling of inversions	even drier than Suite 1 with tropical precip
POSITIVES for SUITE 3	NEGATIVES for SUITE 3
some improvement on inversions	synoptic scores not as good as Suite 1
more representative instability magnitudes	increasing warm bias with time
	tropical cyclone tracks and intensity are worse
POSITIVES for SUITE 4	NEGATIVES for SUITE 4
shows the most promise for dealing with inversions	worse synoptic scores
shows promise for improving instability	tends to overmix PBL, leading to hot and dry on plains
more tropical precipitation	too light on extreme precipitation events and lowest precip bias scores
has the smallest low-level temperature bias	struggles with tropical cyclone intensity and track

Some Caveats and Thoughts

- All of the runs were with 64 levels; with plans for 96 or 128 levels in GFSv16. How will this change the results?
- Data assimilation cycling was not included and it is not clear how this will impact the results.
- The cause of the cold bias increasing w time in suites with GFDL MP, along with the increased cold bias since the radiation fix, is still is not well understood and the physics testing does not address this issue