Hurricane Matthew Case:
28 Sept. 2016, 12UTC – 09 Oct. 2016, 12UTC*

Christopher Williams and Gerard Ketefian

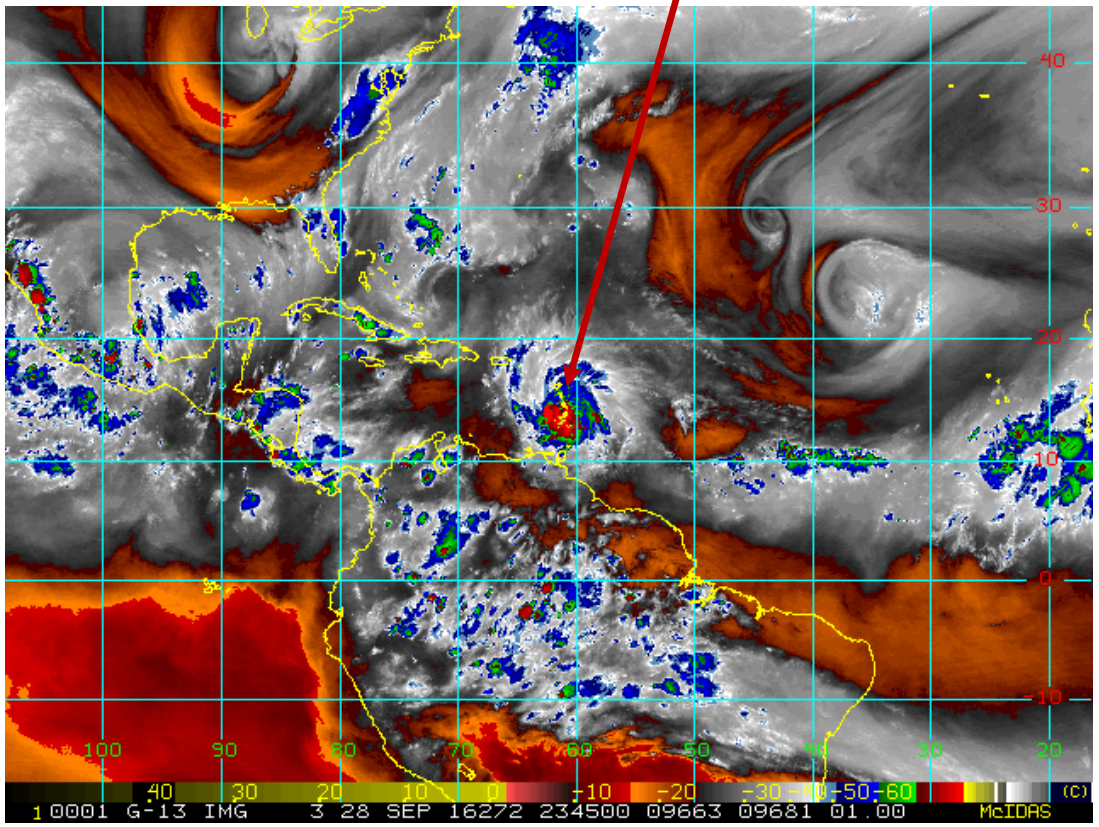
* Model simulations from 29 Sep. 2016, 00UTC – 06 Oct. 2016, 00UTC

Synoptic Discussion

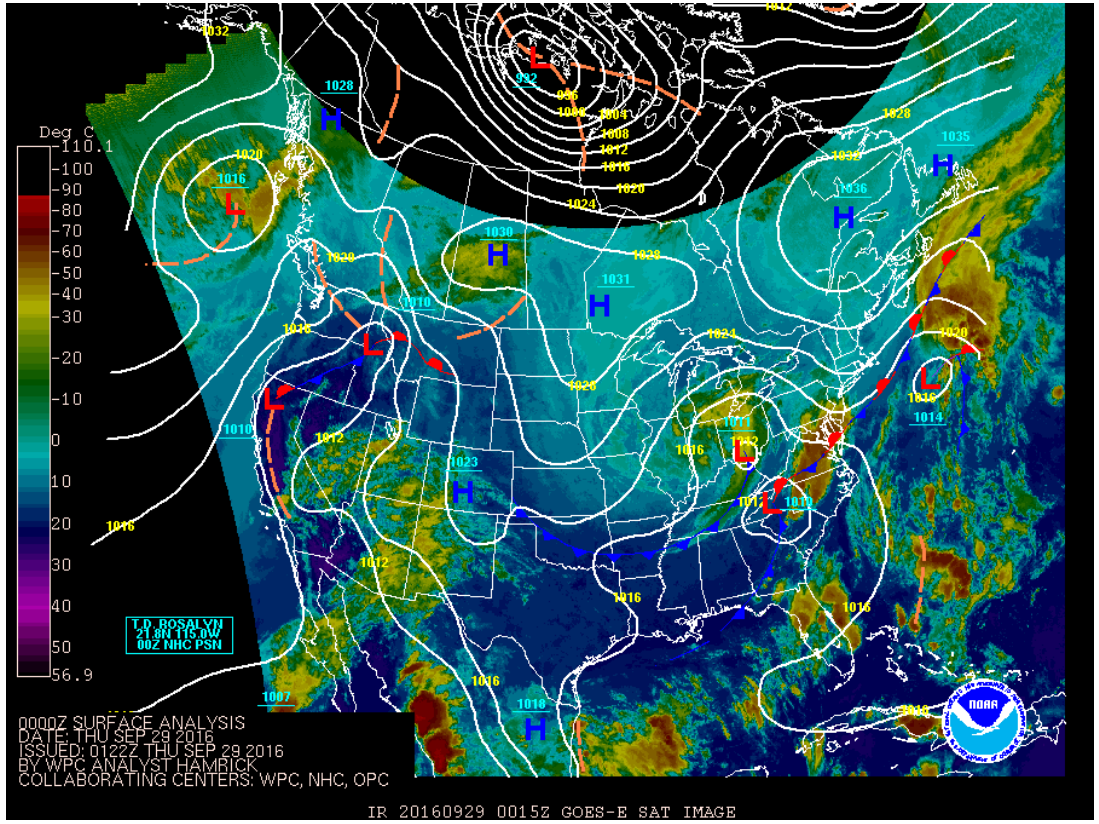
- Hurricane Matthew (AL142016)
- Cat 5 (on the Saffir-Simpson Hurricane Wind Scale)
  - Reached status at lowest latitude ever recorded in AL Basin

29 Sep 2016

Appx 00Z…burgeoning TS Matthew; ridging increases north and west. RI (75 kt intensity increase/30 to 5 n mi eye diameter) soon commences despite moderate vertical wind shear of 15-20 kt



Storm Relative 16 km Geostationary Water Vapor Imagery

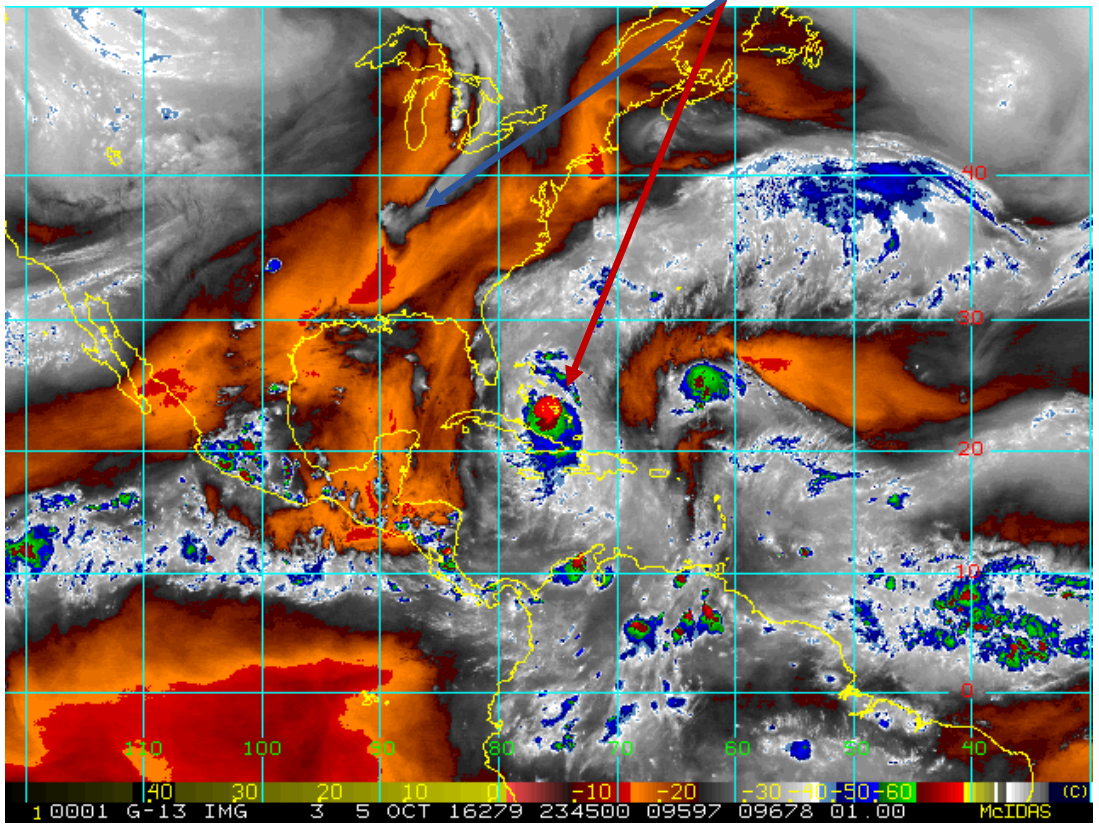


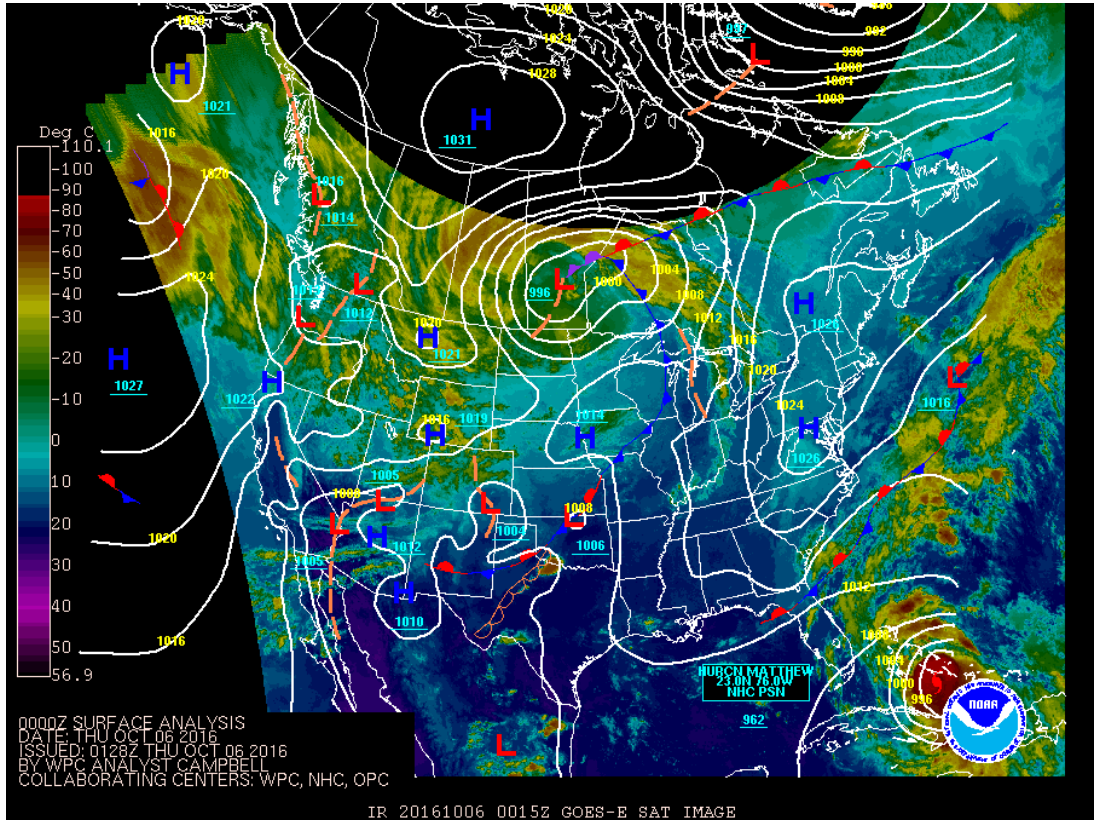
WPC NA Analysis/Satellite Composite

6 Oct 2016

Major Hurricane Matthew; approaching Florida and SE US coast line

Interaction of eastward-moving Central US trough with Hurricane Matthew central to forecast; subtropical ridge over AL necessitated westward peripheral storm motion as well





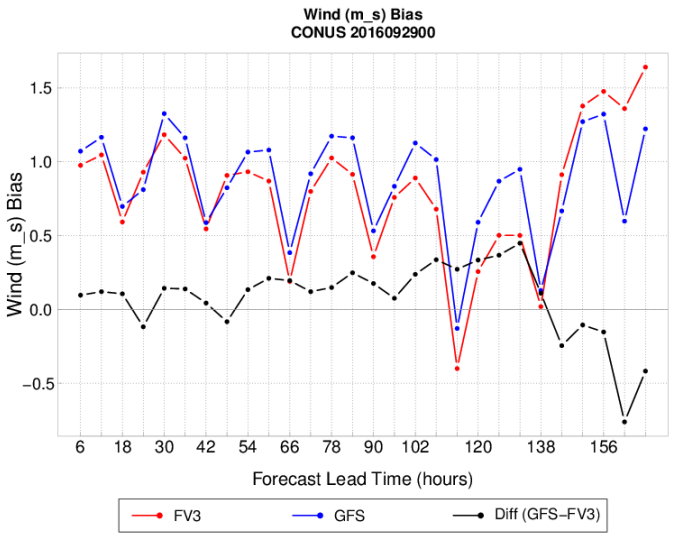
Storm Relative 16 km Geostationary Water Vapor Imagery

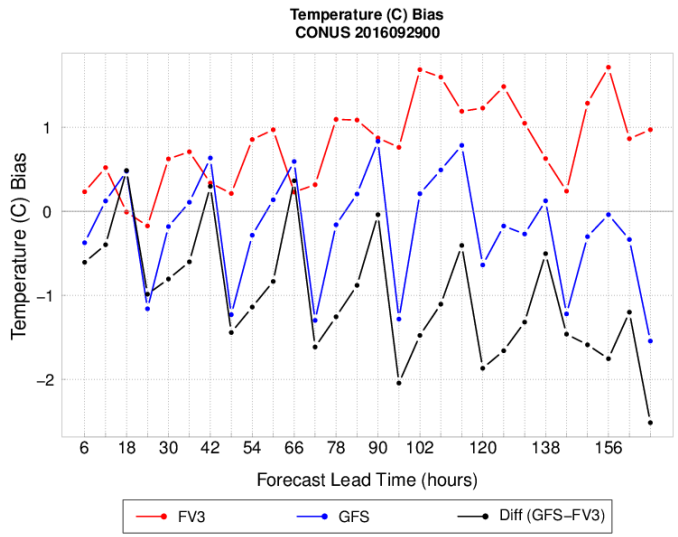
WPC NA Analysis/Satellite Composite

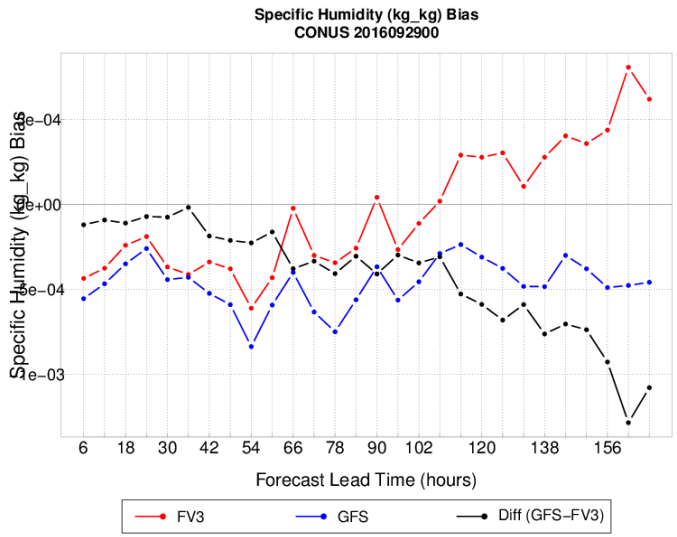
Analysis of Surface Variables:

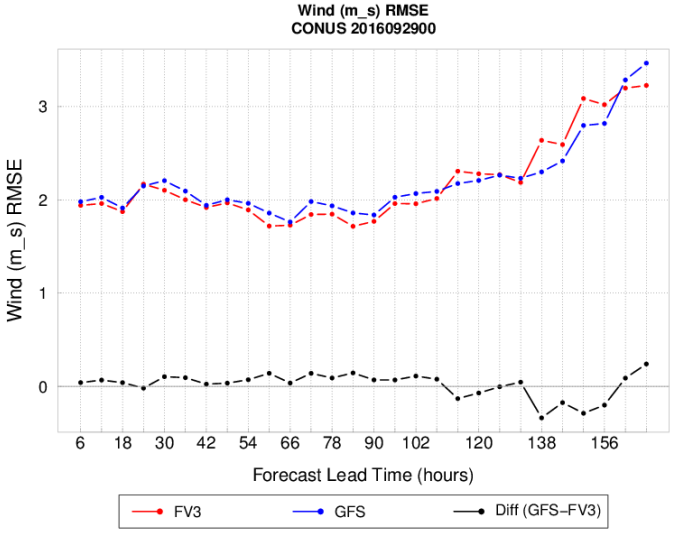
Bias and RMSE vs. Lead Time

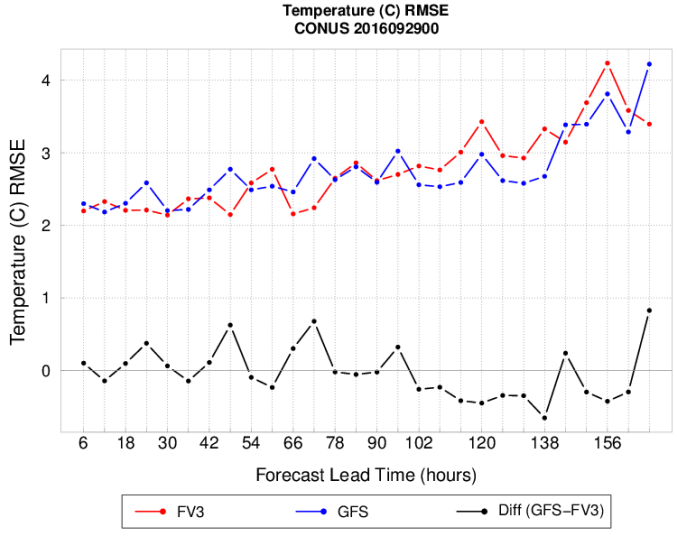
Bias and RMSE of Surface Variables over CONUS

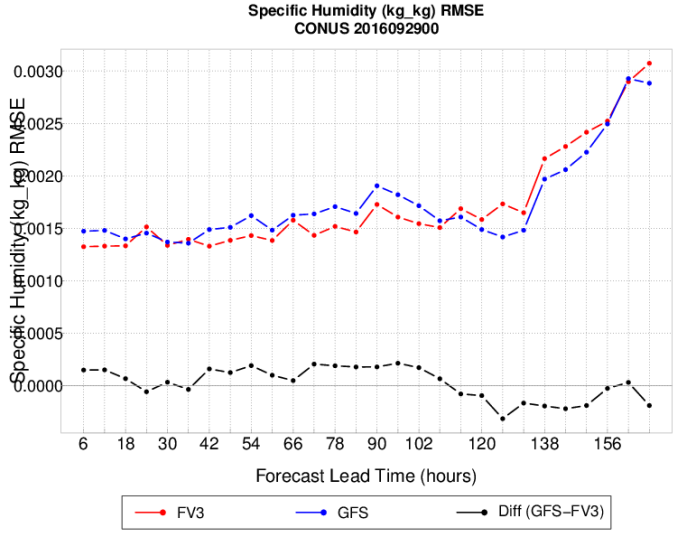












Bias in Surface Variables over CONUS

Wind Speed:

- In both models, the bias is positive at all lead times except at 114 hr (i.e. generally, model winds are faster than observed).
- Both model biases tend to decrease slightly during the first 138 hours. Thereafter, both tend to increase.
- For the first 138 hours, GFS usually has the larger (i.e. more positive) bias. Thereafter, FV3v0 has the larger bias.
- Both models exhibit a distinct diurnal signal. In both models, maxima occur at ~6 UTC (~midnight in the CONUS) while the minima occur at ~18 UTC (~noon in the CONUS).
- Overall (i.e. averaged over all lead times), FV3v0 has slightly smaller absolute bias and thus has slightly better performance.

Temperature:

- FV3v0 has a warm bias at most lead times. GFS bias tends to straddle the zero-line with a diurnal maximum that is positive and a diurnal minimum that is negative.
- GFS bias tends to decrease with lead time while the FV3v0 bias tends to increase.
- FV3v0 bias is usually larger (i.e. more positive) than the GFS bias, i.e. FV3v0 is generally warmer.
- Both models exhibit a distinct diurnal signal. The maxima in FV3v0 occur at ~12 UTC (~6 am in the CONUS) and in GFS at ~18 UTC (~noon in the CONUS). In both models, the minima occur at ~24 UTC (~6 pm in the CONUS).
- Difficult to say which model has overall (i.e. averaged over all lead times) better performance (i.e. smaller absolute bias).

Specific Humidity:

- Both models have a dry bias during the first 102 hours (except for FV3v0 at 90 hr). Thereafter, FV3v0 transitions to a moist bias while GFS maintains a dry bias.
- FV3v0 bias remains approximately constant during the first ~60 hours, then increases with lead time. There is no noticeable trend in the GFS bias with lead time.
- FV3v0 bias is larger (i.e. more positive) than the GFS bias at all lead times, i.e. GFS is always drier.
- In both models, a diurnal signal is readily discernible only between 48 hr and 96 hr (2 to 4 days).
- Overall (i.e. averaged over all lead times), FV3v0 has slightly smaller absolute bias and thus has slightly better performance.

RMSE of Surface Variables over CONUS

Wind Speed:

- RMSEs of both models are approximately steady for the first 90 hours at ~1.9 m/s. Both then gradually increase to ~3.4 m/s at 168 hr.
- Differences between the two models are small ($\leq$ 15% of their absolute RMSEs).
- A diurnal signal is not readily discernible in either model.
- One model does not consistently outperform the other.

Temperature:

- Both models exhibit a gradual increase in RMSE with lead time. Both start at ~2.2°C and end at ~4°C.
- There is no tendency for one model to have larger RMSE than the other. Differences between the models are $\leq$ 30% of their absolute RMSEs.
- GFS usually exhibits a diurnal signal. Such a signal is less apparent in FV3v0.
- One model does not consistently outperform the other.

Specific Humidity:

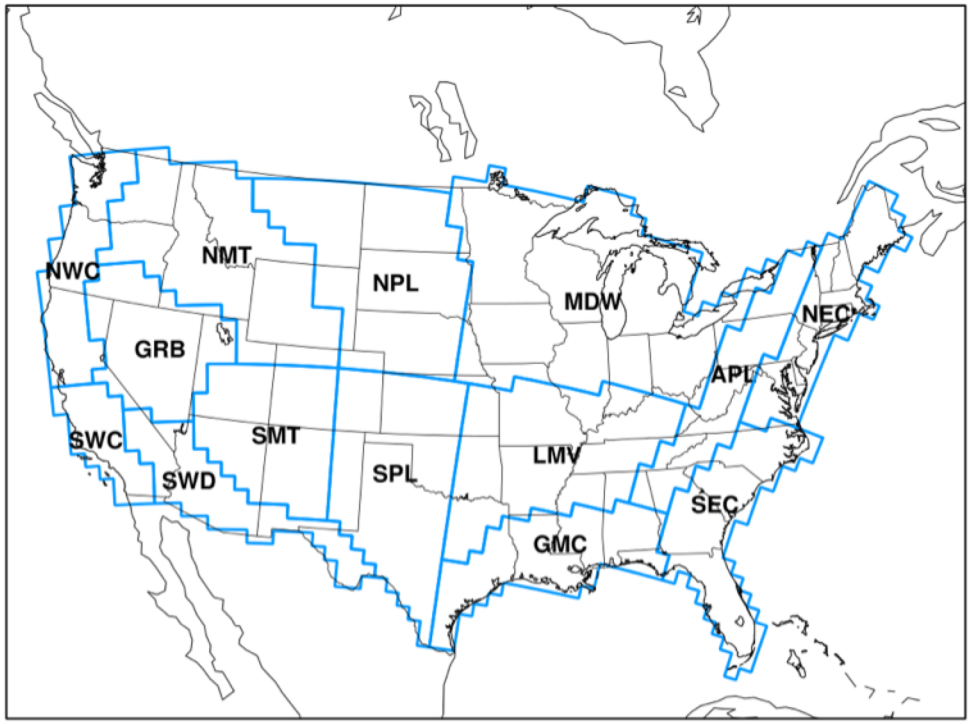
- GFS usually has slightly larger RMSE during the first 108 hours. Thereafter, FV3v0 usually has slightly larger RMSE. Differences between the models are $\leq$ 20% of their absolute RMSEs.
- RMSEs of both models increase only slightly during the first 132 hours, averaging about 1.5×10^{-3} kg/kg. Both then exhibit a noticeably steeper rise in RMSE to about 3×10^{-3} kg/kg at 168 hr.
- A diurnal signal is not readily discernible in either model.
- Overall (i.e. averaged over all lead times), FV3v0 has slightly smaller RMSE and thus has slightly better performance.

Analysis of 24-hour Accum. Precip. Statistics:

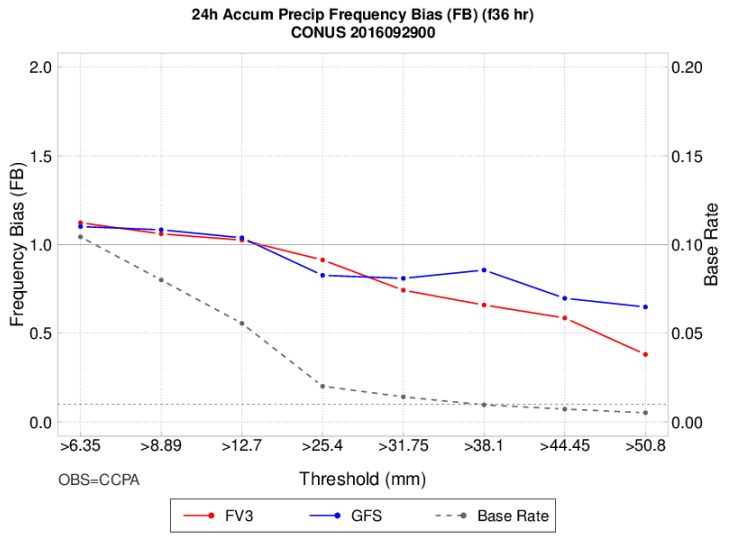
Frequency Bias (FB) and Equitable Threat Score (GSS) vs. Precip. Threshold and Lead Time

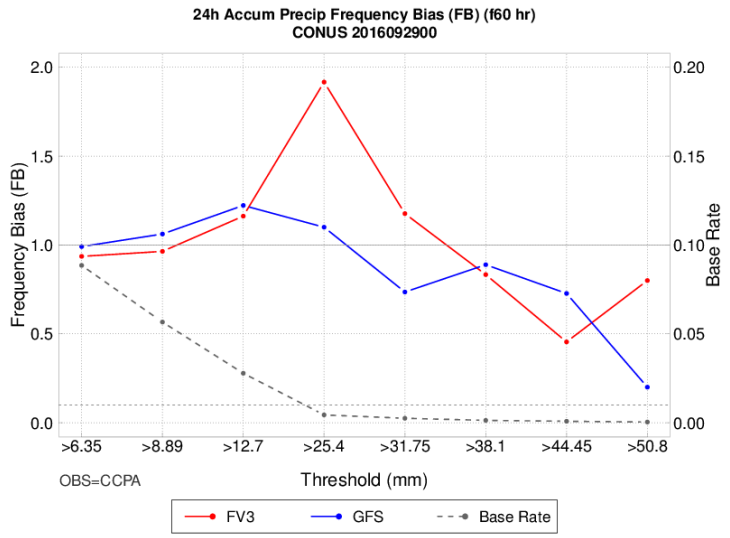
Preliminary Notes

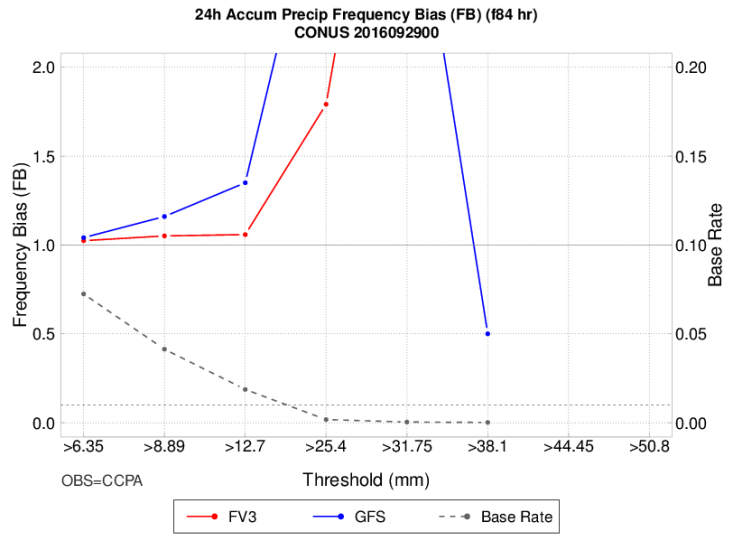
- Consider error statistics (FB and GSS) calculated over CONUS only since global observations are not available to calculate global statistics.
- CONUS domain show in figure to the right.
- FB values can range from 0 to ∞. Ideal value is 1.
- GSS values can range from -1/3 to 1. Ideal value is 1.
- In plots of FB or GSS vs. precip. threshold or lead time (in following slides), base rates are plotted along the right vertical axis and represented by the dashed gray curve.
- Disregard all error statistics for which the base rate is too low, say less than a cutoff of $BR_{cutoff} = 0.01 = 1\%$.
  - BR_{cutoff} indicated in plots with a dashed grey horizontal line.
  - BR always decreases with increasing precip. threshold. Thus, smaller thresholds will have acceptable base rates while larger ones may not.
  - The threshold beyond which we disregard the statistic may change with lead time.

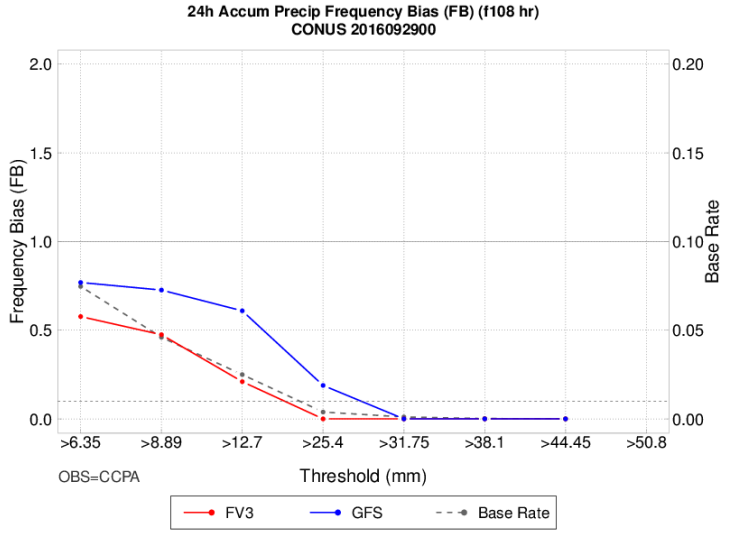


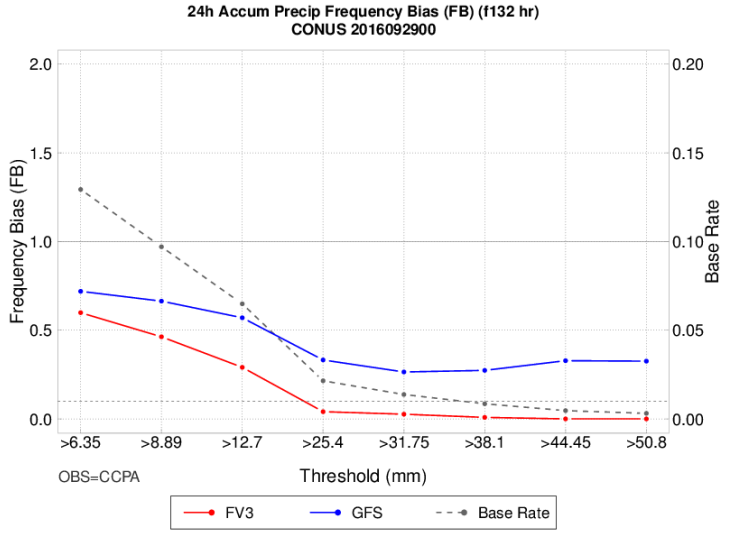
24-Hour Accum. Precip. Frequency Bias (FB) vs. Threshold over CONUS at Various Lead Times

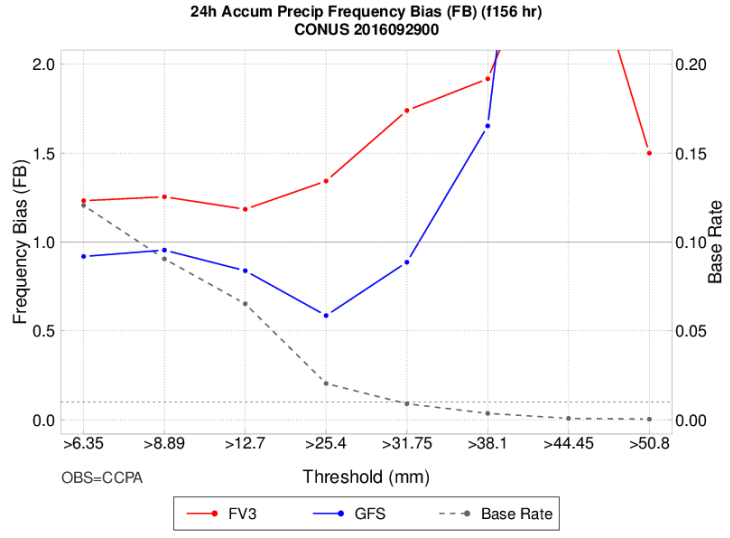








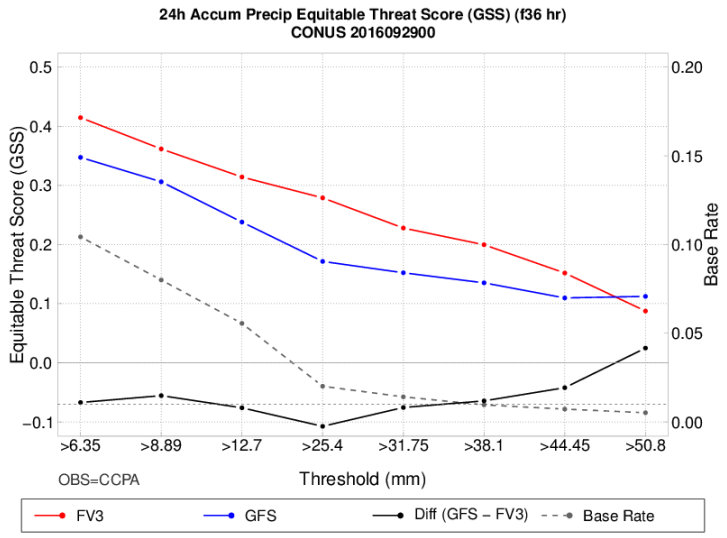


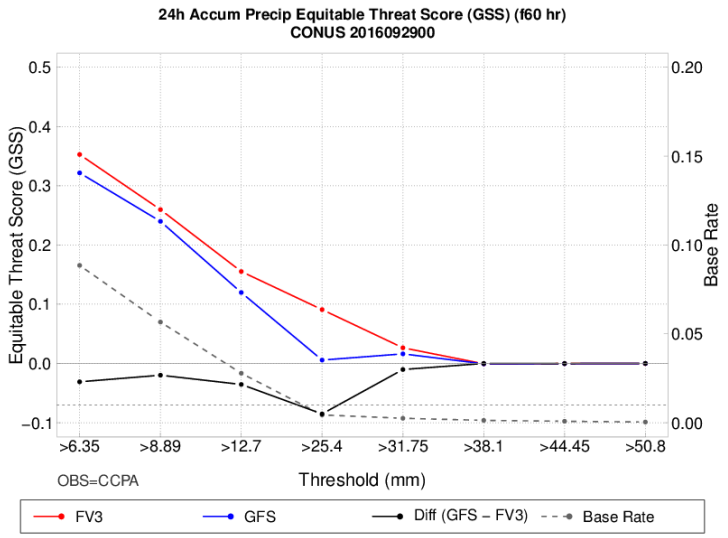


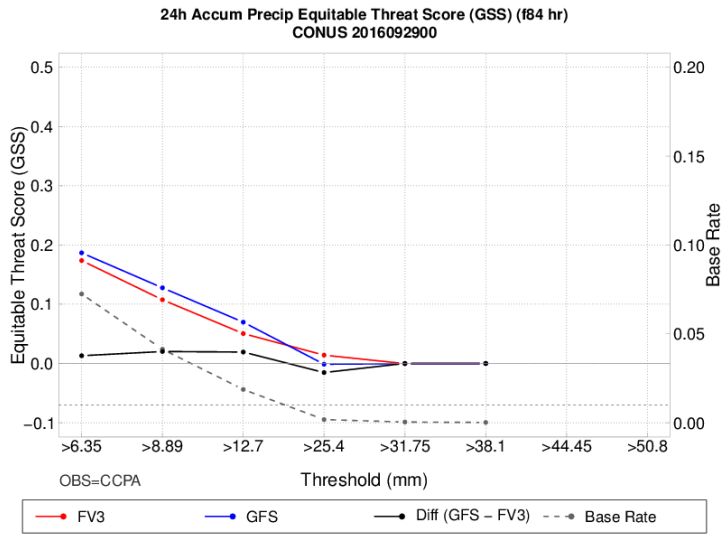
24-Hour Accum. Precip. Frequency Bias (FB) vs. Threshold over CONUS at Various Lead Times

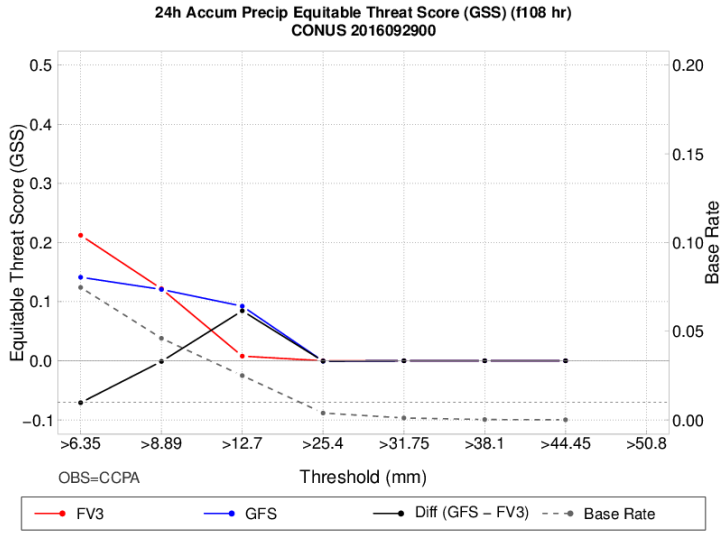
- At a given lead time, disregard FB results beyond the last threshold for which the base rate is above BR_{cutoff} = 0.01. For the lead times shown, i.e. 36 hr, 60 hr, 84 hr, 108 hr, 132 hr, and 156 hr, the largest threshold for which we consider FB values meaningful are, respectively, >31.750 mm, >12.700 mm, >12.700 mm, >12.700 mm, >31.750 mm, and >25.400 mm.
- At the first lead time of 36 hr, both models have approximately the same FB values. For mid lead times (60 hr, 84 hr, 108 hr, and 132 hr), GFS has higher FB values; and at the final lead time of 156 hr, FV3v0 has the higher FB values.
- There does not seem to be a consistent trend in FB with increasing threshold – sometimes FB increases, at other times it decreases or remains about the same.
- In both models, FB values tend to be at or slightly above 1 at 36 hr, 60 hr, and 84 hr and below 1 at 108 hr and 132 hr. At 156 hr, FV3v0 values move above 1 while GFS values remain below (although closer to 1 than at 108 hr and 132 hr).
- Both models exhibit best results (i.e. FB values nearest to 1) at the earlier times, probably at 36 hr. Results tend to degrade with increasing lead time, although there is some recovery at 156 hr.
- For the statistically most meaningful thresholds (i.e. whenever the base rate is above BR_{cutoff} = 0.01), GFS has FB values that are overall (i.e. averaged over thresholds) closer to 1 and thus has the better performance.

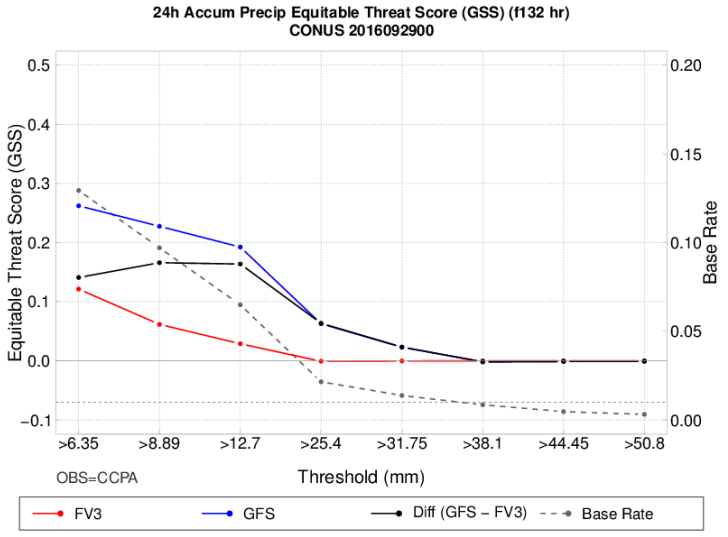
24-Hour Accum. Precip. Equitable Threat Score (GSS) vs. Threshold over CONUS at Various Lead Times

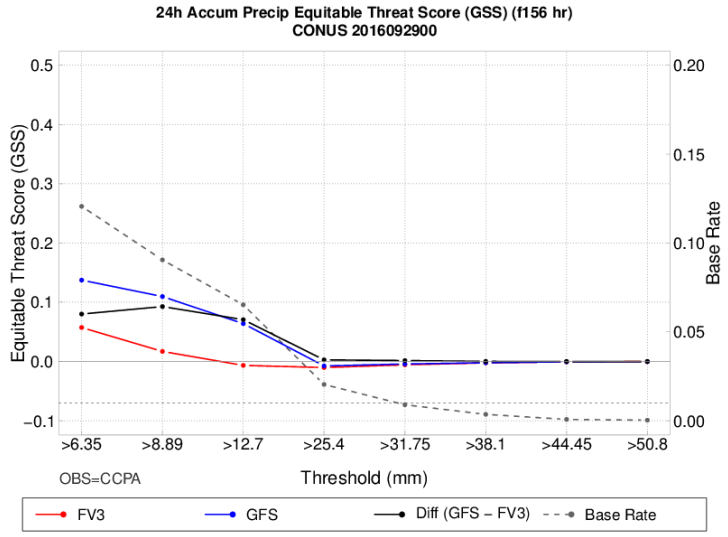








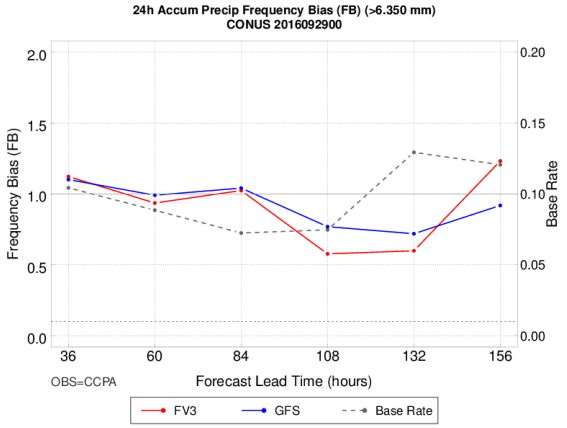


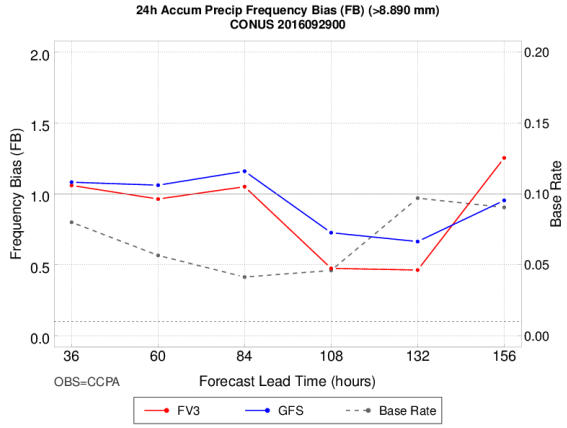


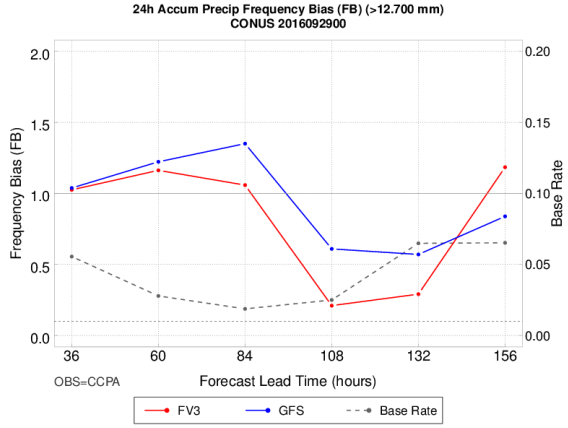
24-Hour Accum. Precip. Equitable Threat Score (GSS) vs. Threshold over CONUS at Various Lead Times

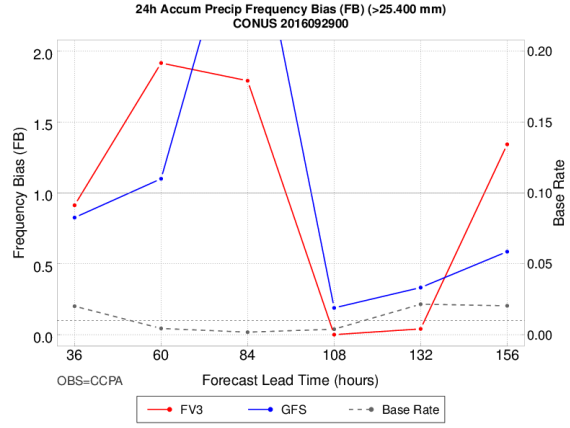
- At a given lead time, disregard GSS results beyond the last threshold for which the base rate is above BR_{cutoff} = 0.01. For the lead times shown, i.e. 36 hr, 60 hr, 84 hr, 108 hr, 132 hr, and 156 hr, the largest threshold for which we consider GSS values meaningful are, respectively, >31.750 mm, >12.700 mm, >12.700 mm, >12.700 mm, >31.750 mm, and >25.400 mm.
- Both models exhibit a generally decreasing trend in GSS with lead time.
- In both models, GSS remains between ~0.4 and ~0.0 at all lead times and for all thresholds.
- Both models exhibit best results (i.e. GSS values nearest to 1) at the earliest lead time of 36 hr. Results tend to degrade with increasing lead time, although there is a slight recovery in GFS going from 108 hr to 132 hr.
- FV3v0 has larger GSS values than GFS (and thus better performance since all reported values are less than the ideal value of 1) at 36 hr and 60 hr but usually has smaller values than GFS (and thus worse performance) at 84 hr and beyond.
- For the statistically most meaningful thresholds (i.e. wherever the base rate is above BR_{cutoff} = 0.01), GFS has GSS values that are overall (i.e. averaged over thresholds) slightly closer to 1 and thus has the slightly better performance.

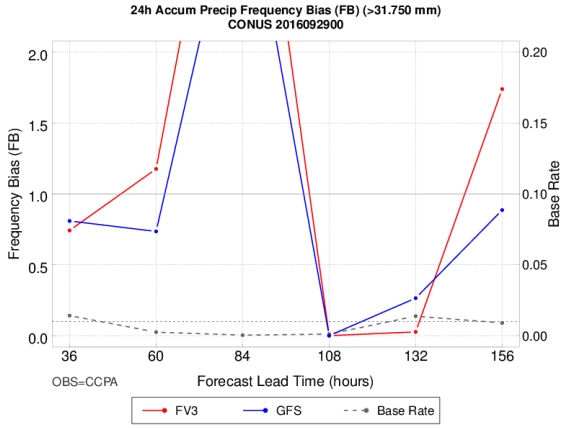
24-Hour Accum. Precip. Frequency Bias (FB) vs. Lead Time over CONUS for Various Precip. Thresholds

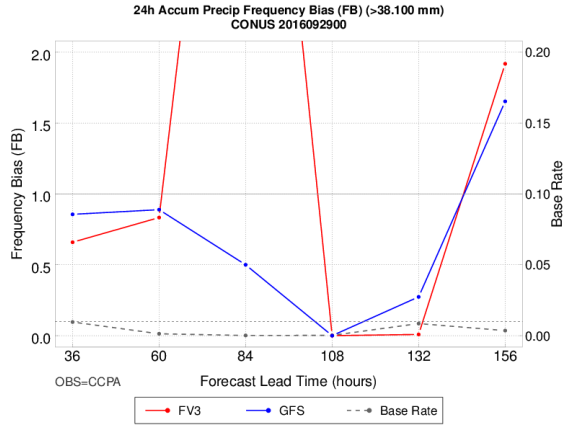


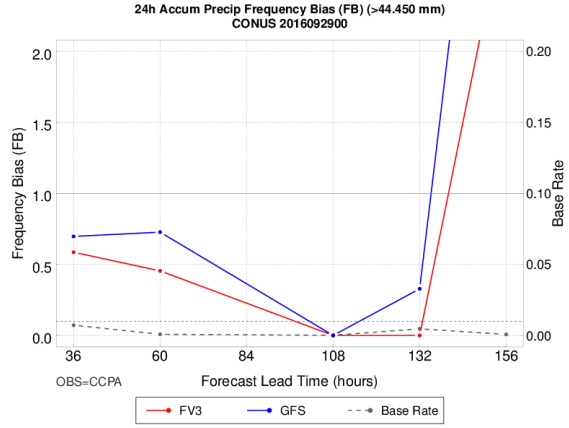


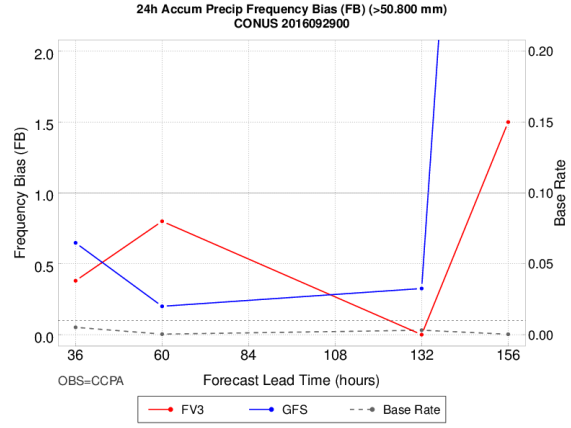








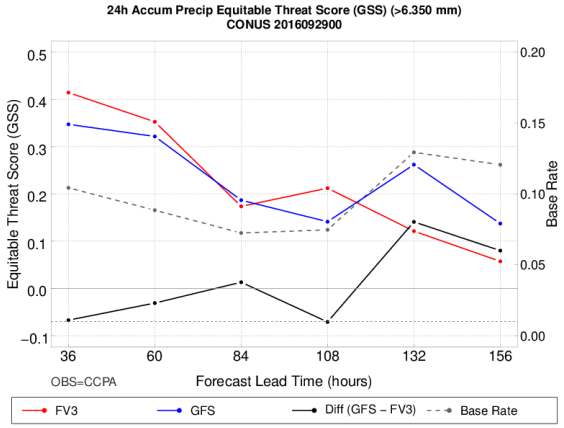


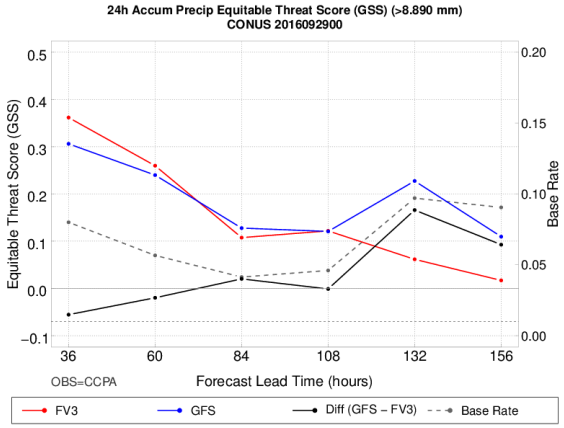


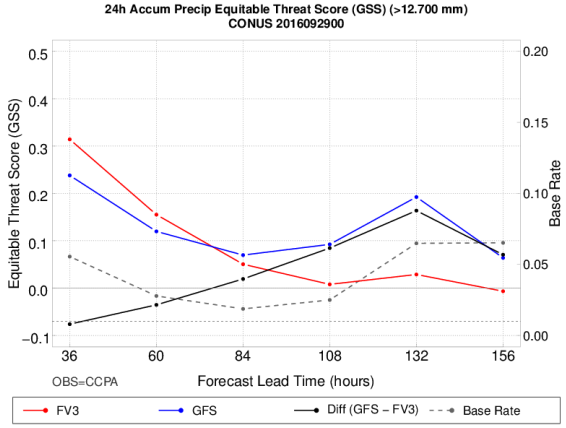
24-Hour Accum. Precip. Frequency Bias (FB) vs. Lead Time over CONUS for Various Precip. Thresholds

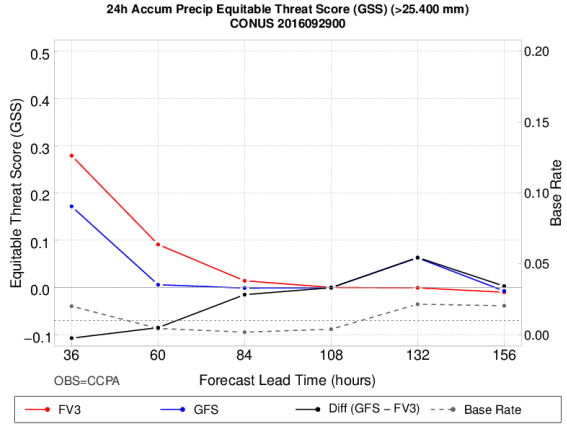
- Disregard FB results for larger thresholds, i.e. >31.750 mm, >38.100 mm, >44.450 mm, and >50.800 mm, since base rate for those is either below or only slightly above BR_{cutoff} = 0.01, i.e. there are not enough samples for FB calculations to be statistically meaningful.
- FB results for >25.400 mm are only marginally acceptable since the base rate at certain lead times is below BR_{cutoff} and at other times is not significantly above (it is at most ~0.02).
- Both models exhibit relatively small changes in FB between 36 hr and 84 hr, a large decrease between 84 hr and 108 hr, a small change between 108 hr and 132 hr, and a large increase between 132 hr and 156 hr.
- Both models exhibit best results (i.e. FB values nearest to 1) for the lowest threshold (>6.350 mm). Results degrade with increasing threshold, i.e. the spread in FB tends to grow going from >6.350 mm to >8.890 mm to >12.70 mm (and possibly to >25.400 mm).
- For the statistically most meaningful thresholds (>6.350 mm, >8.890 mm, and >12.70 mm), GFS exhibits slightly less spread around the ideal FB value of 1 and thus has the slightly better performance.

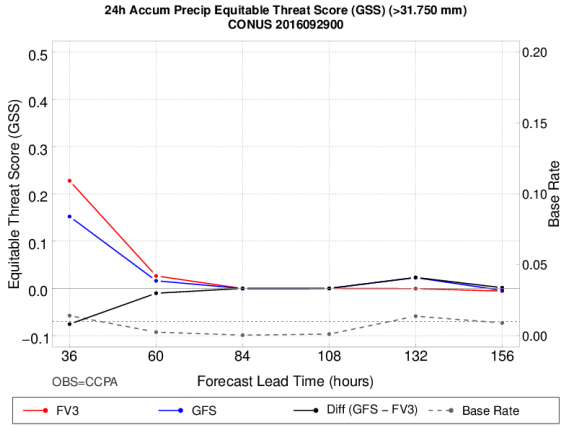
24-Hour Accum. Precip. Equitable Threat Score (GSS) vs. Lead Time over CONUS for Various Precip. Thresholds

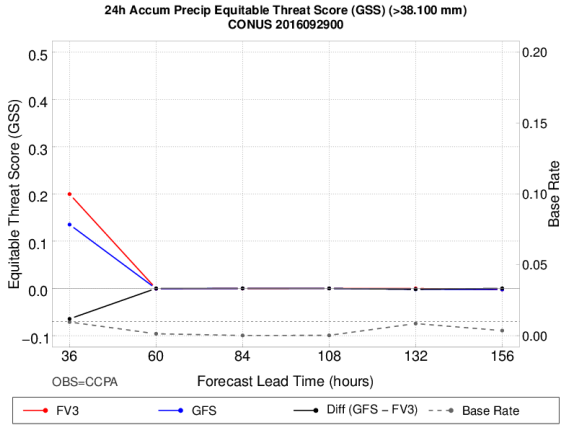


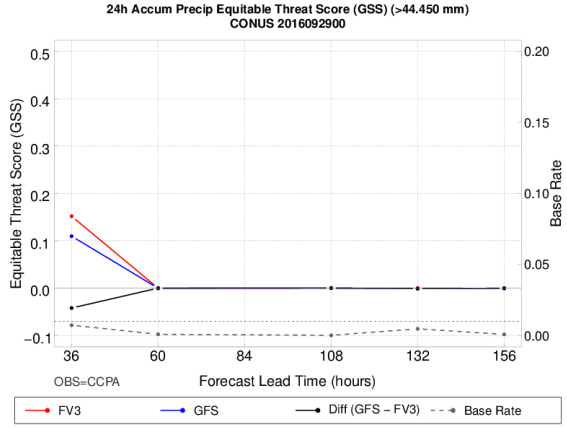


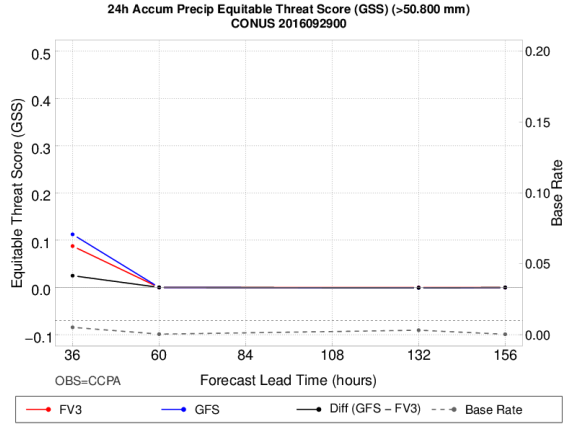












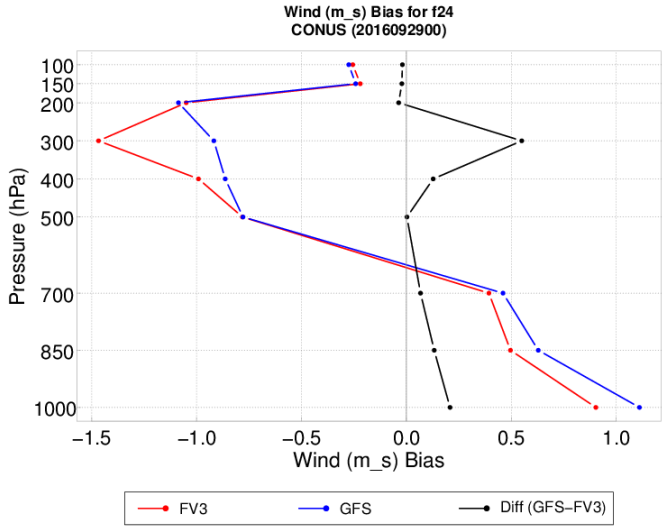
24-Hour Accum. Precip. Equitable Threat Score (GSS) vs. Lead Time over CONUS for Various Precip. Thresholds

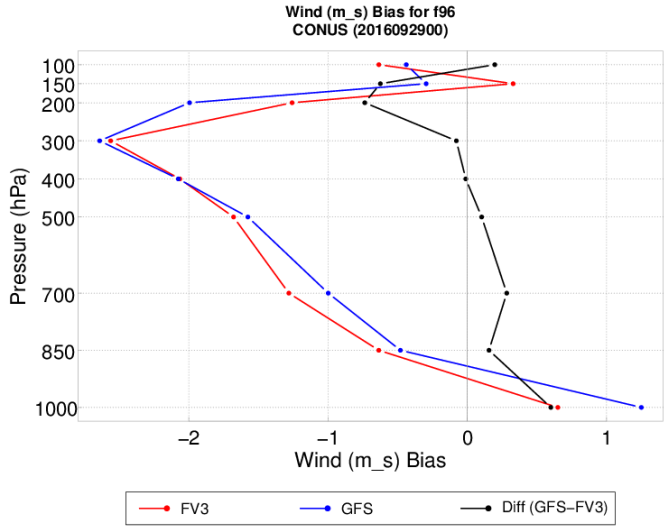
- Disregard GSS results for larger thresholds, i.e. >31.750 mm, >38.100 mm, >44.450 mm, and >50.800 mm, since base rate for those is either below or only slightly above BR_{cutoff} = 0.01, i.e. there are not enough samples for GSS calculations to be statistically meaningful.
- GSS results for >25.400 mm are only marginally acceptable since the base rate at certain lead times is below BR_{cutoff} and at other times is not significantly above (it is at most ~0.02).
- Both models exhibit a generally decreasing trend in GSS with lead time, but GFS tends to decrease more slowly than FV3v0.
- In both models, GSS remains between ~0.4 and ~0.0 at all lead times and for all thresholds.
- Both models exhibit best results (i.e. GSS values nearest to 1) for the lowest threshold (>6.350 mm). Results degrade with increasing threshold, i.e. the GSS curves shift down closer to 0 or become slightly negative going from >6.350 mm to >8.890 mm to >12.70 mm (and possibly to >25.400 mm).
- FV3v0 has larger GSS values than GFS (and thus better performance since all reported values are less than the ideal value of 1) at 36 hr and 60 hr but usually has smaller values than GFS (and thus worse performance) at 84 hr and beyond.
- For the statistically most meaningful thresholds (>6.350 mm, >8.890 mm, and >12.70 mm), GFS exhibits slightly higher time-averaged value of GSS than FV3v0 and thus has the slightly better performance.

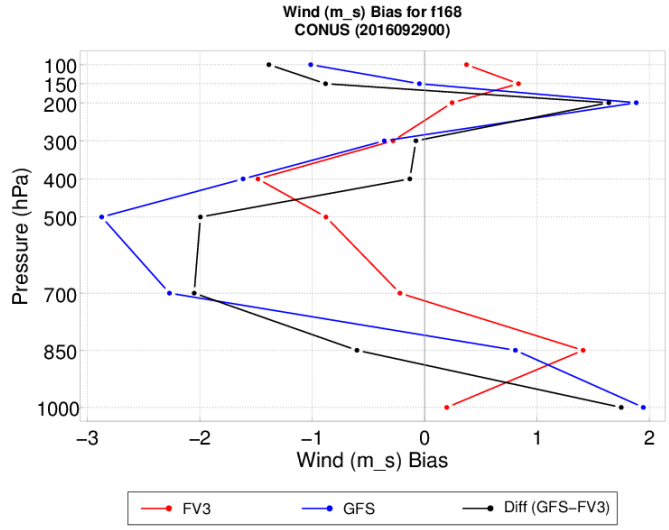
Analysis of Vertical Profiles:

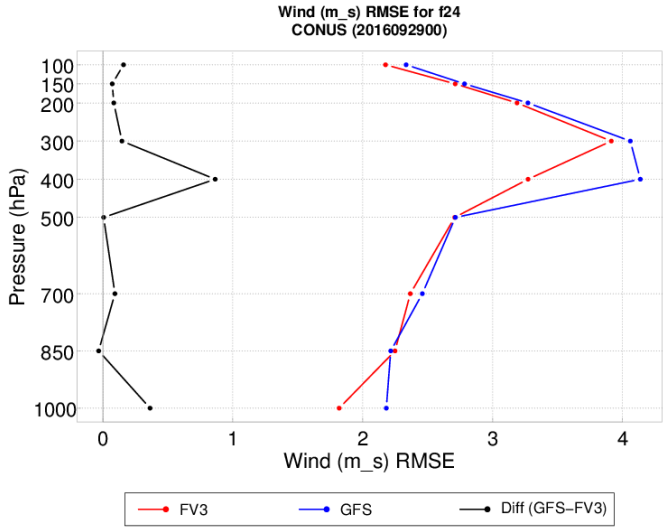
Bias and RMSE vs. Pressure

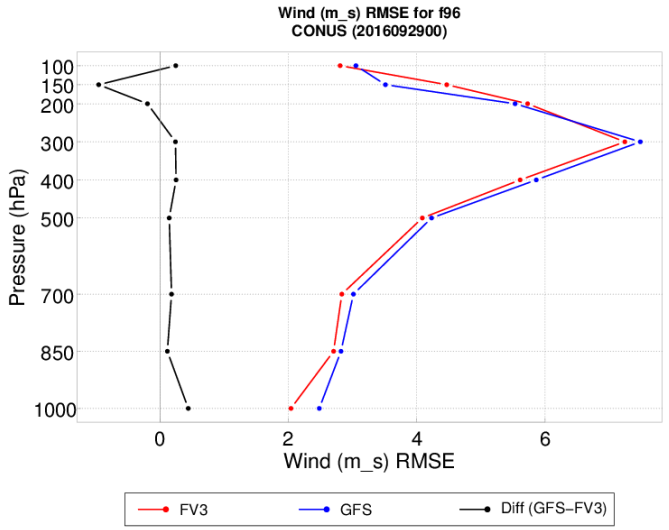
Vertical Profiles of Bias and RMSE of Wind Speed over CONUS at 24 hr, 96 hr, and 168 hr

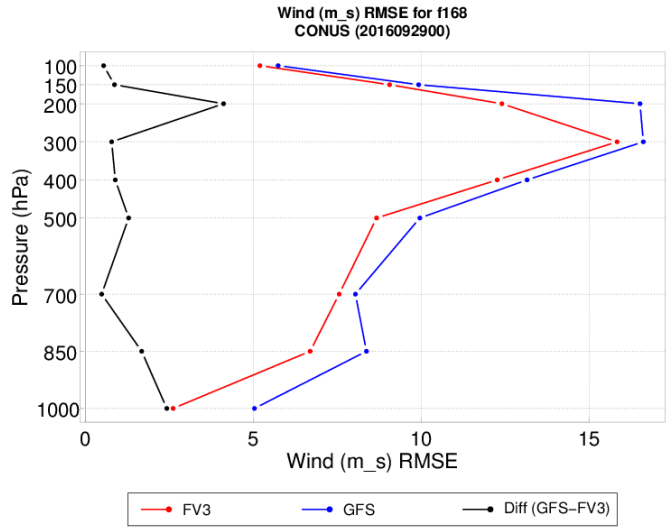












Vertical Profiles of Bias and RMSE of Wind Speed over CONUS

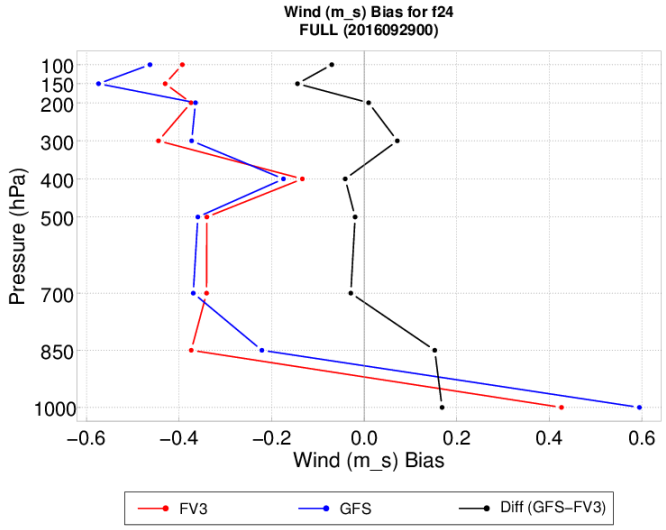
Bias:

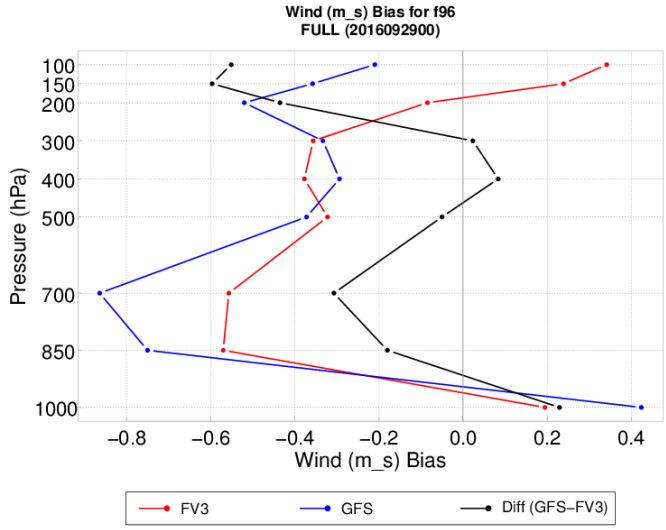
- Both models exhibit a fast (i.e. positive) bias at lower levels and a slow bias with an associated bias minimum at mid-levels.
- Above 300 hPa, behavior is more variable in time – models may have a second region of slow bias or may have a fast bias.
- In both models, magnitude of slow bias at mid-levels grows with time.
- One model does not consistently outperform the other.

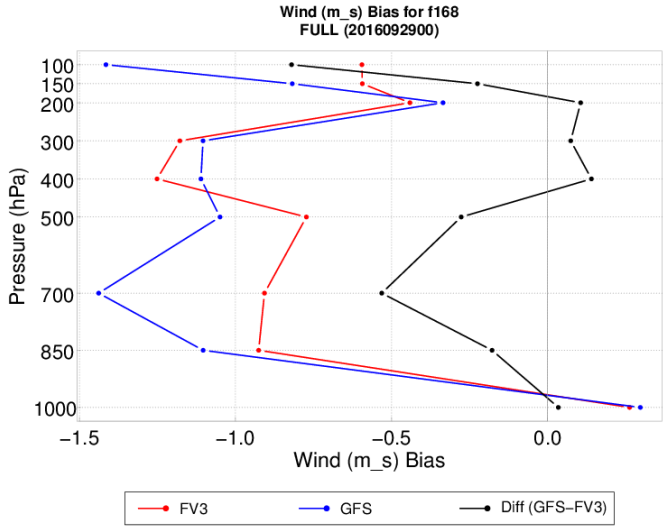
RMSE:

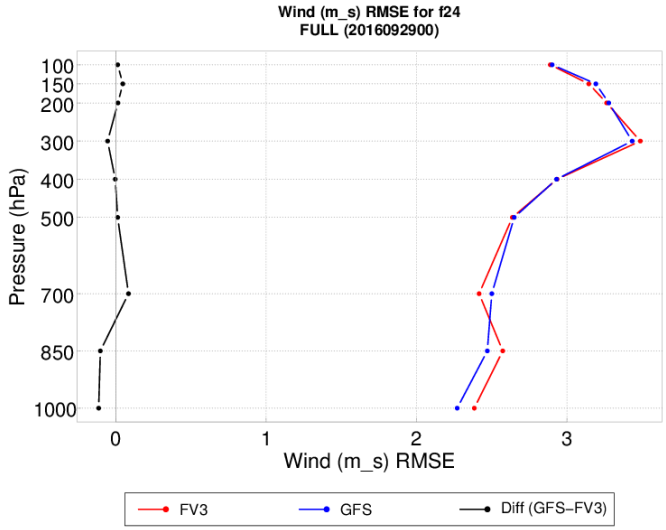
- The two models have similar RMSE profile shapes, i.e. they usually have local minima and maxima at the same levels.
- At early lead times, both models exhibit a slightly increasing RMSE with height near the surface. The rate of increase of RMSE with height near the surface grows with time such that both models exhibit a relatively steep RMSE profile at later times.
- Both models exhibit a persistent maximum at ~300 hPa that grows with time.
- Both models exhibit a growth in overall RMSE magnitude with time.
- One model does not consistently outperform the other.

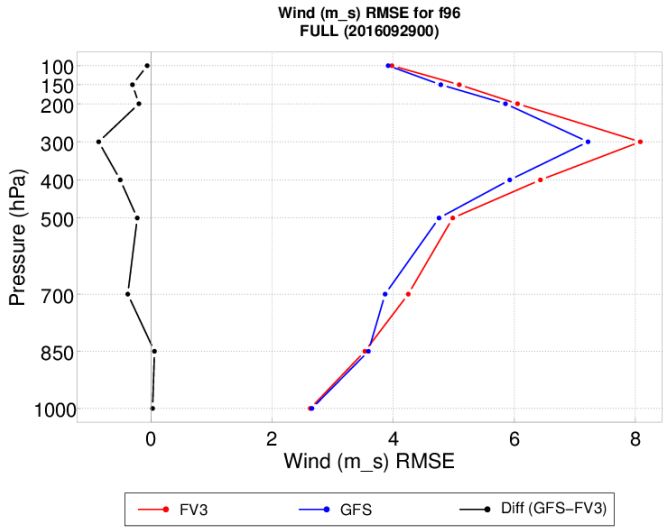
Vertical Profiles of Bias and RMSE of Wind Speed over Global Domain at 24 hr, 96 hr, and 168 hr

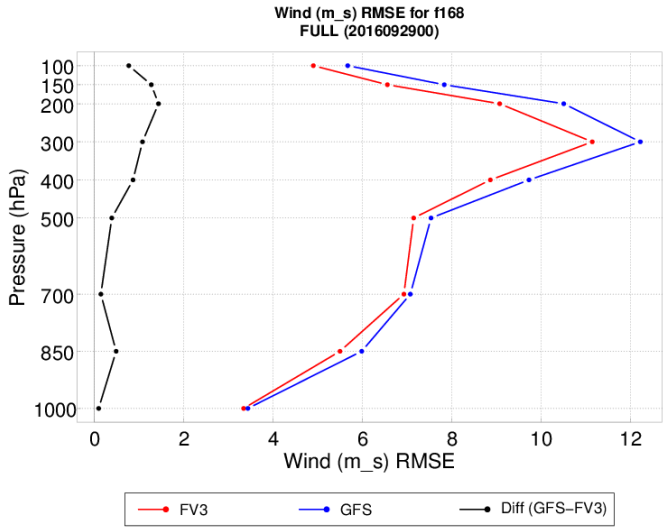












Vertical Profiles of Bias and RMSE of Wind Speed over Global Domain

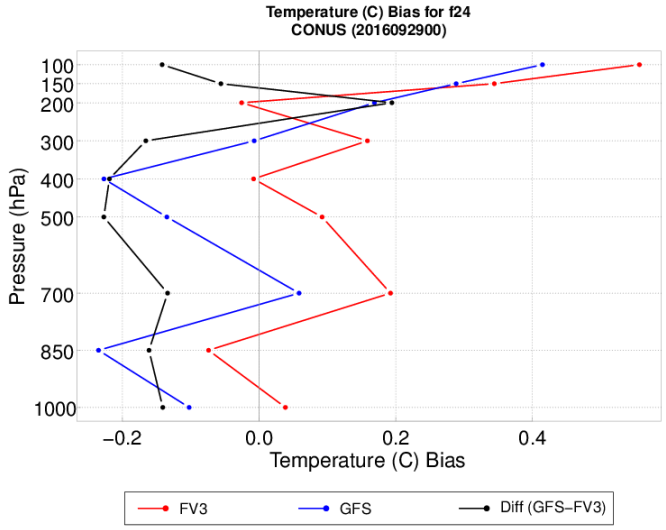
Bias:

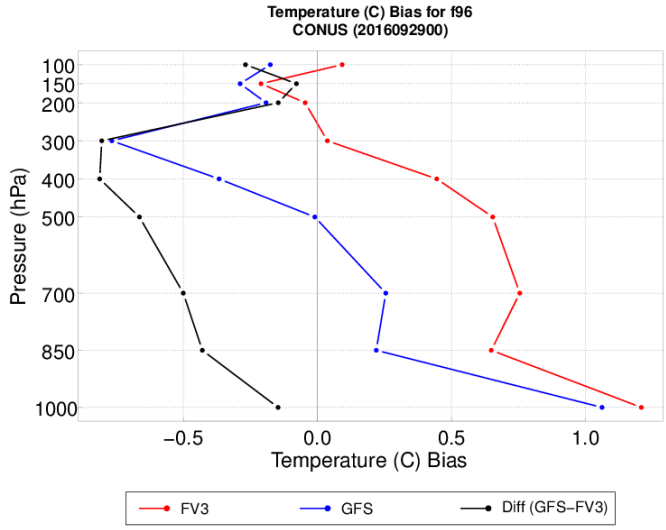
- Both models exhibit a fast (i.e. positive) bias at 1000 mb and a slow bias with an associated bias minimum at mid-levels, from 850 hPa to ~300 hPa.
- Above 300 hPa, behavior is more variable in time – models may have a second region of slow bias or may have a fast bias.
- In both models, magnitude of slow bias at mid-levels grows with time.
- One model does not consistently outperform the other.

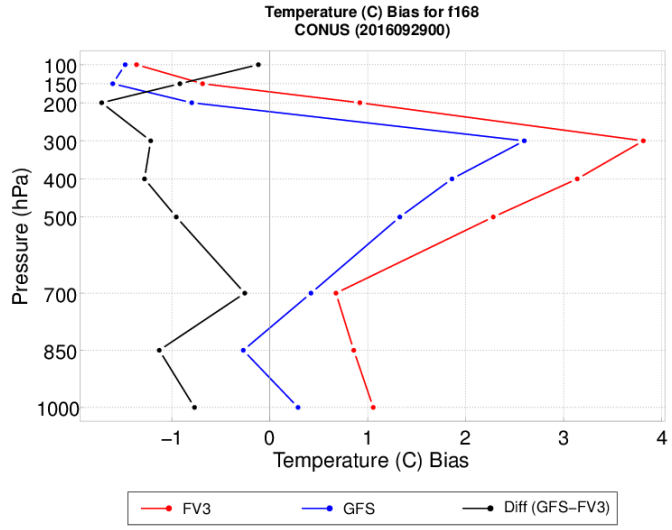
RMSE:

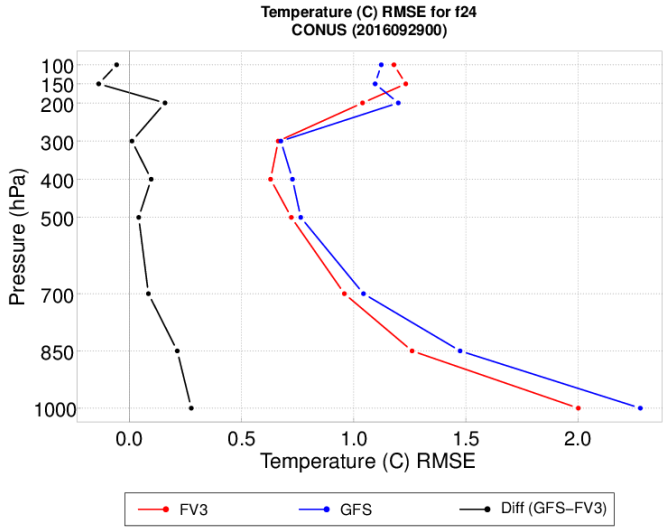
- The two models have similar RMSE profile shapes, i.e. they usually have local minima and maxima at the same levels.
- At early lead times, both models exhibit a slightly increasing RMSE with height near the surface. The rate of increase of RMSE with height near the surface grows with time such that both models exhibit a relatively steep RMSE profile at later times.
- Both models exhibit a persistent maximum at ~300 hPa that grows with time.
- Both models exhibit a growth in overall RMSE magnitude with time.
- One model does not consistently outperform the other.

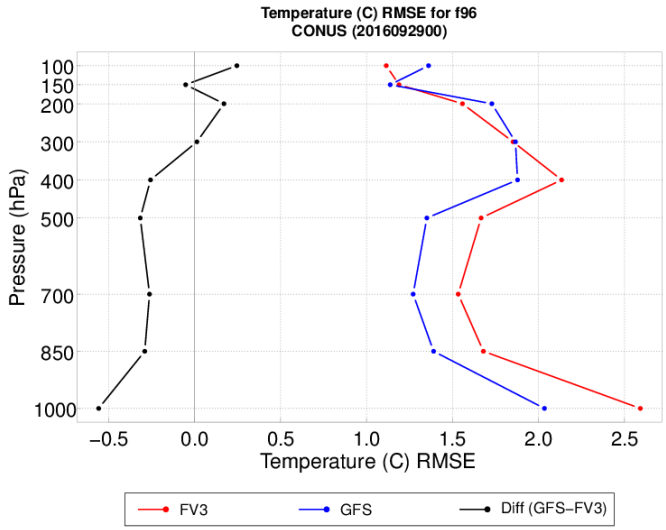
Vertical Profiles of Bias and RMSE of Temperature over CONUS at 24 hr, 96 hr, and 168 hr

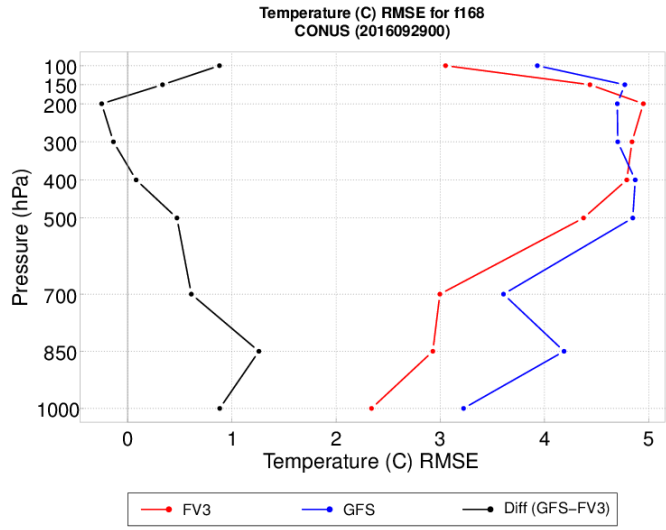












Vertical Profiles of Bias and RMSE of Temperature over CONUS

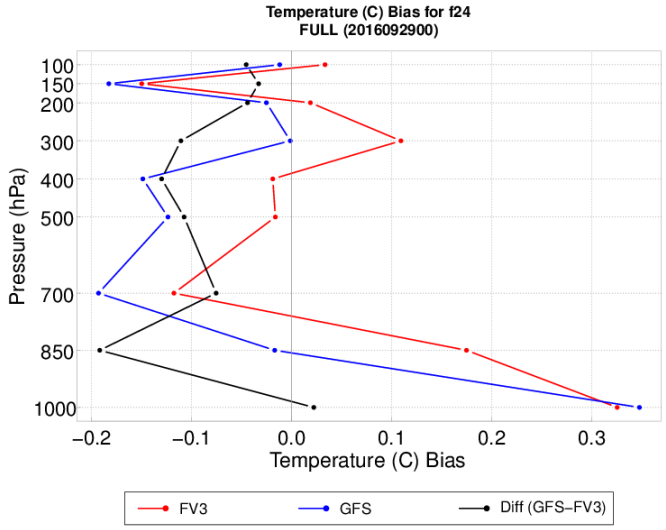
Bias:

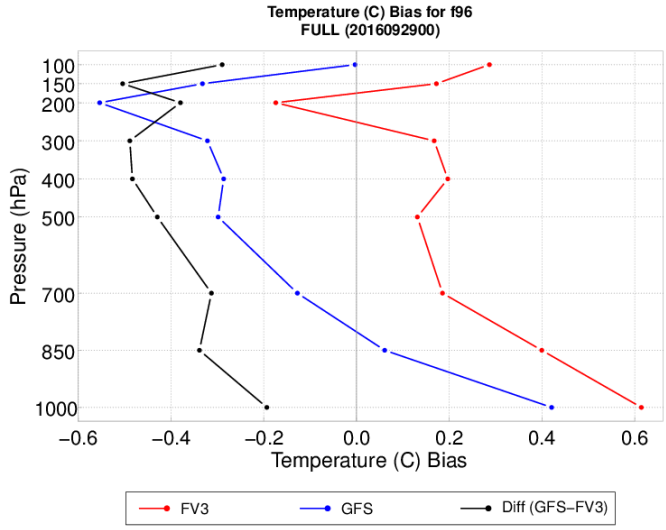
- After 72 hr, the two models tend to have similar profile shapes, i.e. they usually have local minima and maxima at the same levels.
- GFS is usually cooler than FV3v0 (i.e. its profile is shifted to the left relative to FV3v0); maximum difference between the models is $< 2°C$.
- Both models usually exhibit a decreasing bias with height near the surface. The rate of decrease with height becomes smaller with time (i.e. profiles tend to become more vertical near the surface at later times).
- Both models become warmer (i.e. profiles shift to the right) with increasing lead time (but with the GFS remaining to the left of FV3v0), so it is difficult to say which one is “better”.

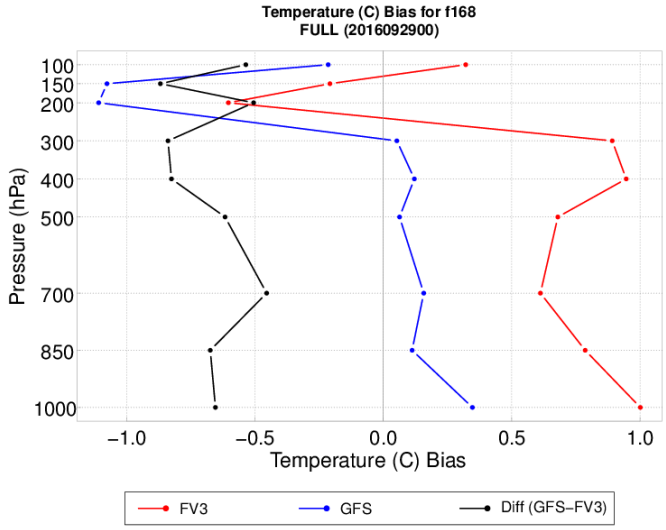
RMSE:

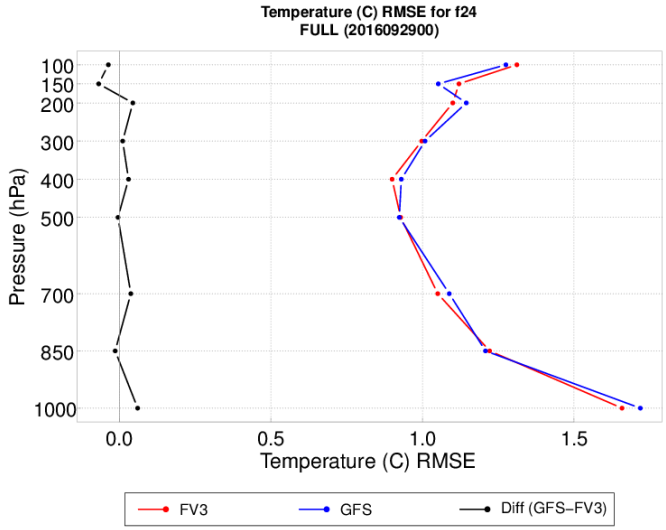
- The two models often have similar profile shapes, i.e. they usually have local minima and maxima at the same levels.
- At early lead times, both models exhibit a decreasing RMSE with height near the surface. This gradually shifts in time to an increasing RMSE with height near the surface.
- Both models exhibit an intermittent local minimum/maximum pair at upper levels (above 400 hPa).
- Both models exhibit a growth in overall RMSE magnitude with time.
- One model does not consistently outperform the other.

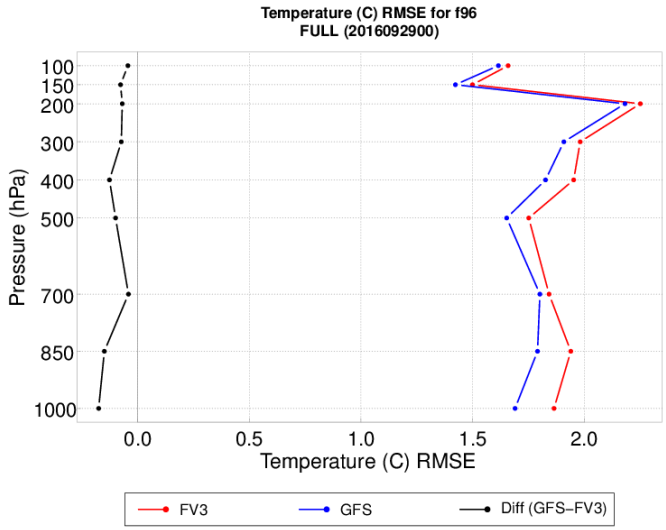
Vertical Profiles of Bias and RMSE of Temperature over Global Domain at 24 hr, 96 hr, and 168 hr

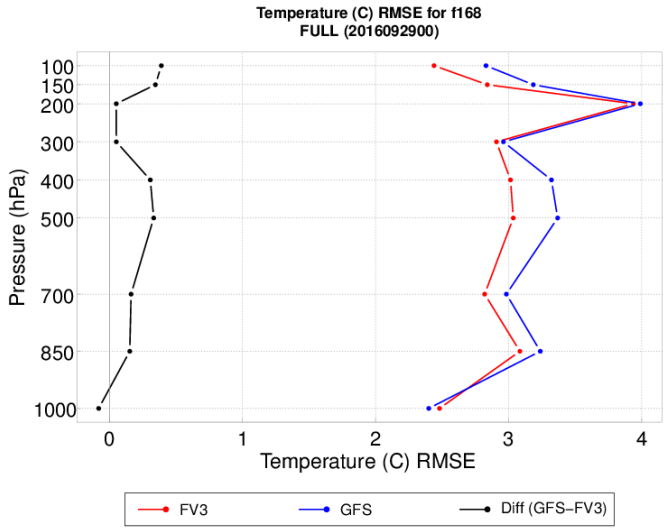












Vertical Profiles of Bias and RMSE of Temperature over Global Domain

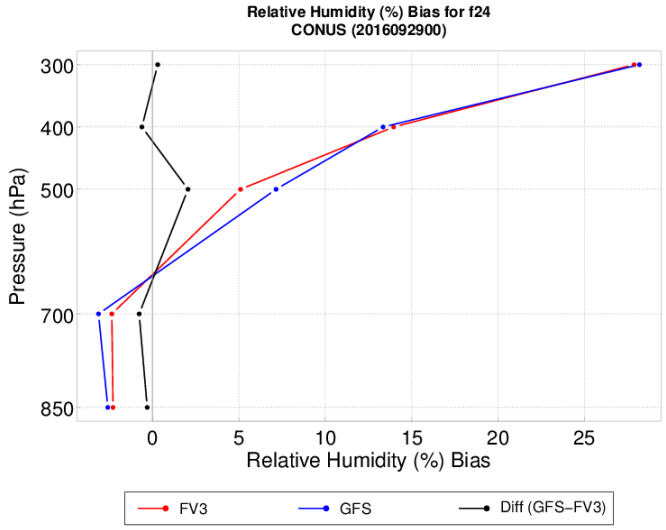
Bias:

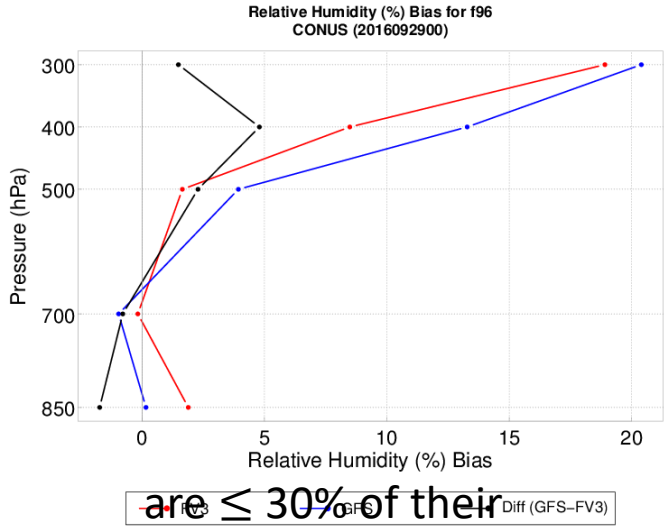
- The two models have similar profile shapes, i.e. they usually have local minima and maxima at the same levels.
- GFS is almost always cooler than FV3v0 (i.e. its profile is shifted to the left relative to FV3v0); maximum difference between the models is $< 1°C$.
- Both models usually exhibit a decreasing bias with height near the surface. The rate of decrease with height becomes smaller with time (i.e. profiles tend to become more vertical near the surface at later times).
- Both models exhibit a persistent local minimum at 200-150 hPa that grows in magnitude (i.e. becomes more negative) with time.
- Both models become warmer (i.e. profiles shift to the right) with increasing lead time (but with the GFS remaining to the left of FV3v0), so it is difficult to say which one is “better”.

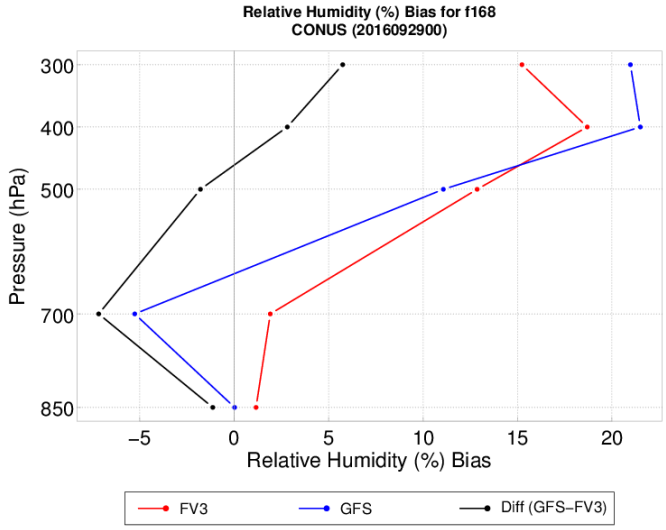
RMSE:

- The two models have similar profile shapes, i.e. they usually have local minima and maxima at the same levels.
- At early lead times, both models exhibit a decreasing RMSE with height near the surface. This gradually shifts in time to an increasing RMSE with height near the surface.
- Both models exhibit a persistent local minimum/maximum pair at upper levels (above 400 hPa). The maximum of this pair grows in magnitude and is a prominent feature at 200 hPa at later lead times.
- Both models exhibit a growth in overall RMSE magnitude with time.
- One model does not consistently outperform the other.

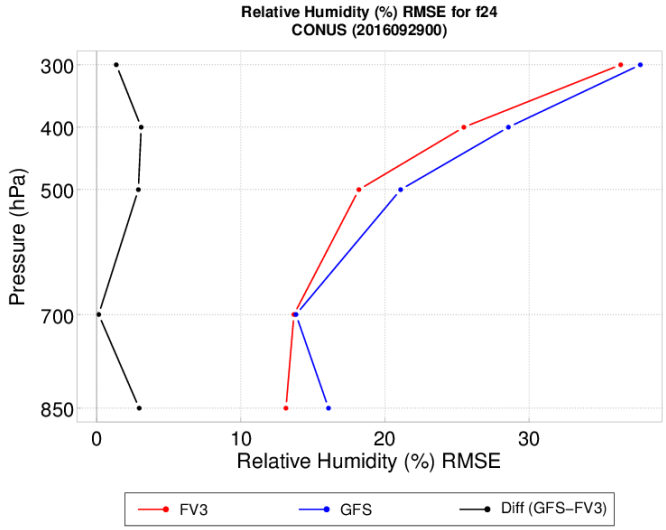
Vertical Profiles of Bias and RMSE of Relative Humidity over CONUS at 24 hr, 96 hr, and 168 hr

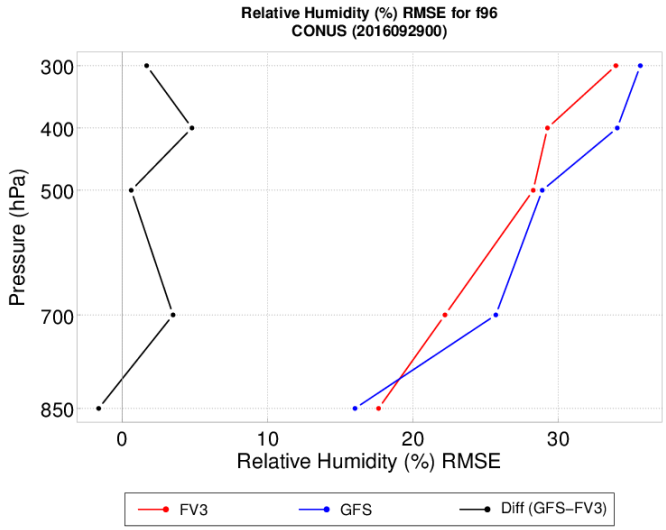


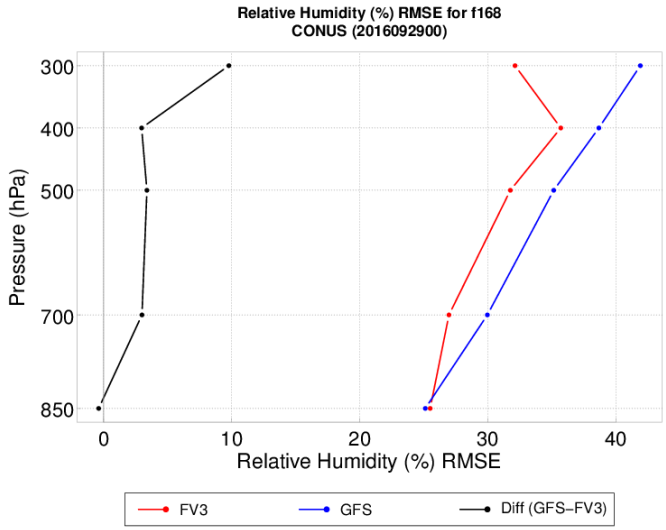




are ≤ 30% of their







Vertical Profiles of Bias and RMSE of Relative Humidity over CONUS

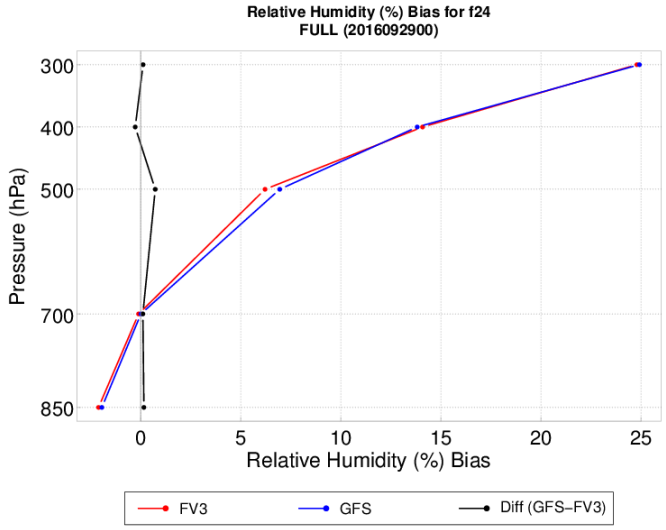
Bias:

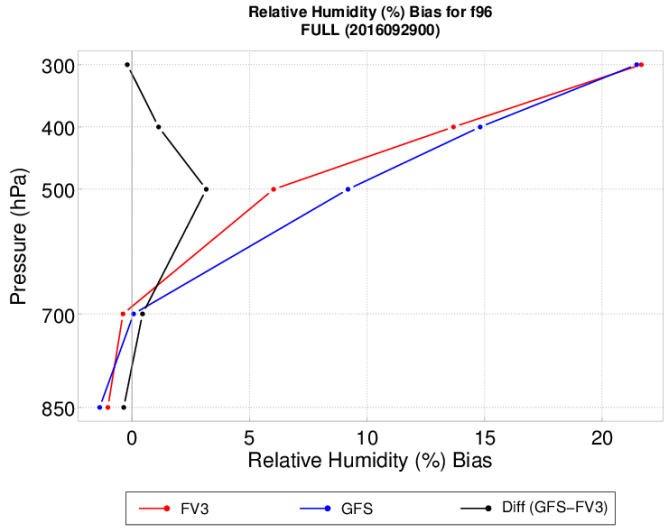
- The two models usually have similar profile shapes, i.e. at a given lead time, both are concave, linear, or convex.
- Both models start close to zero at 850 hPa (bias magnitude $< 5\%$) but generally grow with height to between 15% and 30% at 300 hPa.
- Both models are too moist at upper levels.
- One model does not consistently outperform the other.

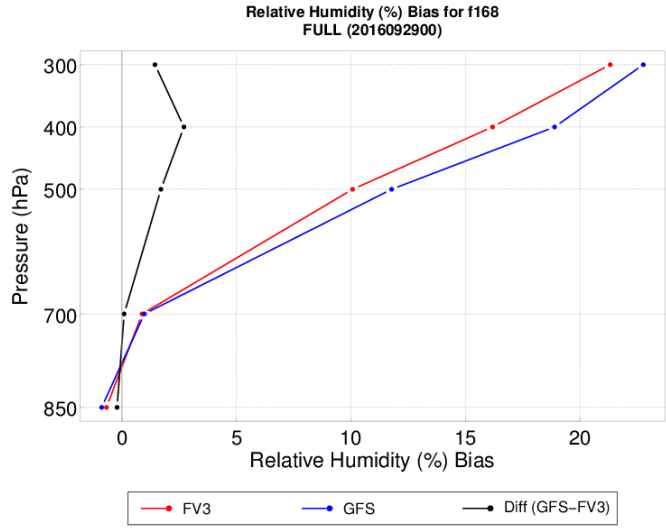
RMSE:

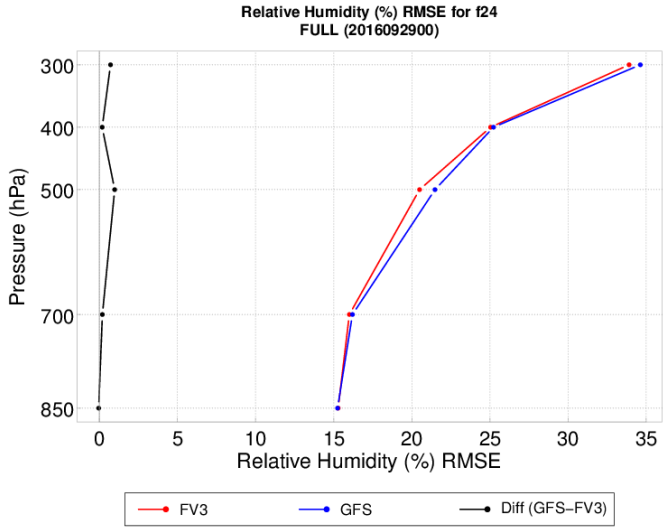
- The two models usually have similar profile shapes, i.e. at a given lead time, both are concave, linear, or convex.
- Both models usually exhibit a growth in RMSE with height.
- Both models exhibit an increase in the slope of the RMSE profile at lower levels with time, i.e. the profiles are initially both convex and become more linear with time.
- Overall, FV3v0 has slightly smaller RMSE.

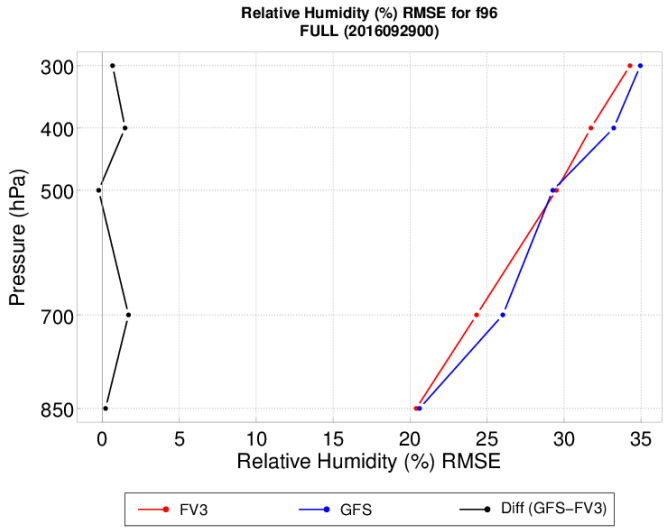
Vertical Profiles of Bias and RMSE of Relative Humidity over Global Domain at 24 hr, 96 hr, and 168 hr

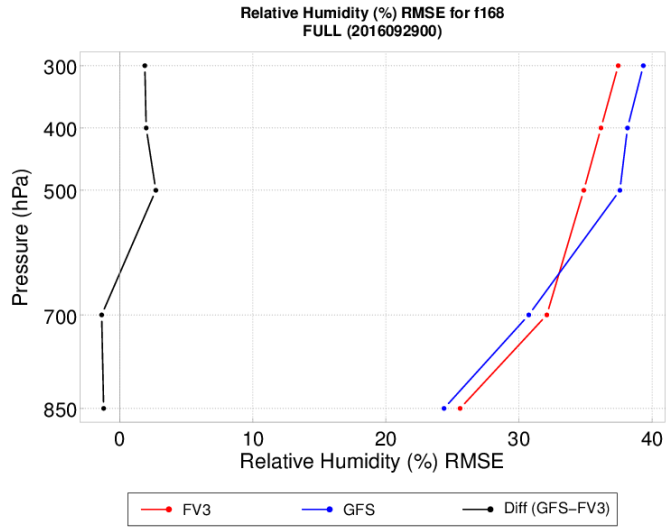












Vertical Profiles of Bias and RMSE of Relative Humidity over Global Domain

Bias:

- The two models have similar profile shapes, i.e. at a given lead time, both are concave, linear, or convex.
- Both models start close to zero at 850 hPa (bias magnitude $< 5\%$) but grow with height to between 18% and 26% at 300 hPa.
- Both models are too moist at upper levels, with GFS usually slightly moister than FV3v0.
- Overall, FV3v0 has slightly smaller bias magnitudes.

RMSE:

- The two models have similar profile shapes, i.e. at a given lead time, both are concave, linear, or convex.
- Both models exhibit a growth in RMSE with height.
- Both models exhibit an increase in the slope of the RMSE profile at lower levels with time, i.e. the profiles are initially both convex, become linear at mid-lead times, and become slightly concave at later lead times.
- Overall, FV3v0 has slightly smaller RMSE.

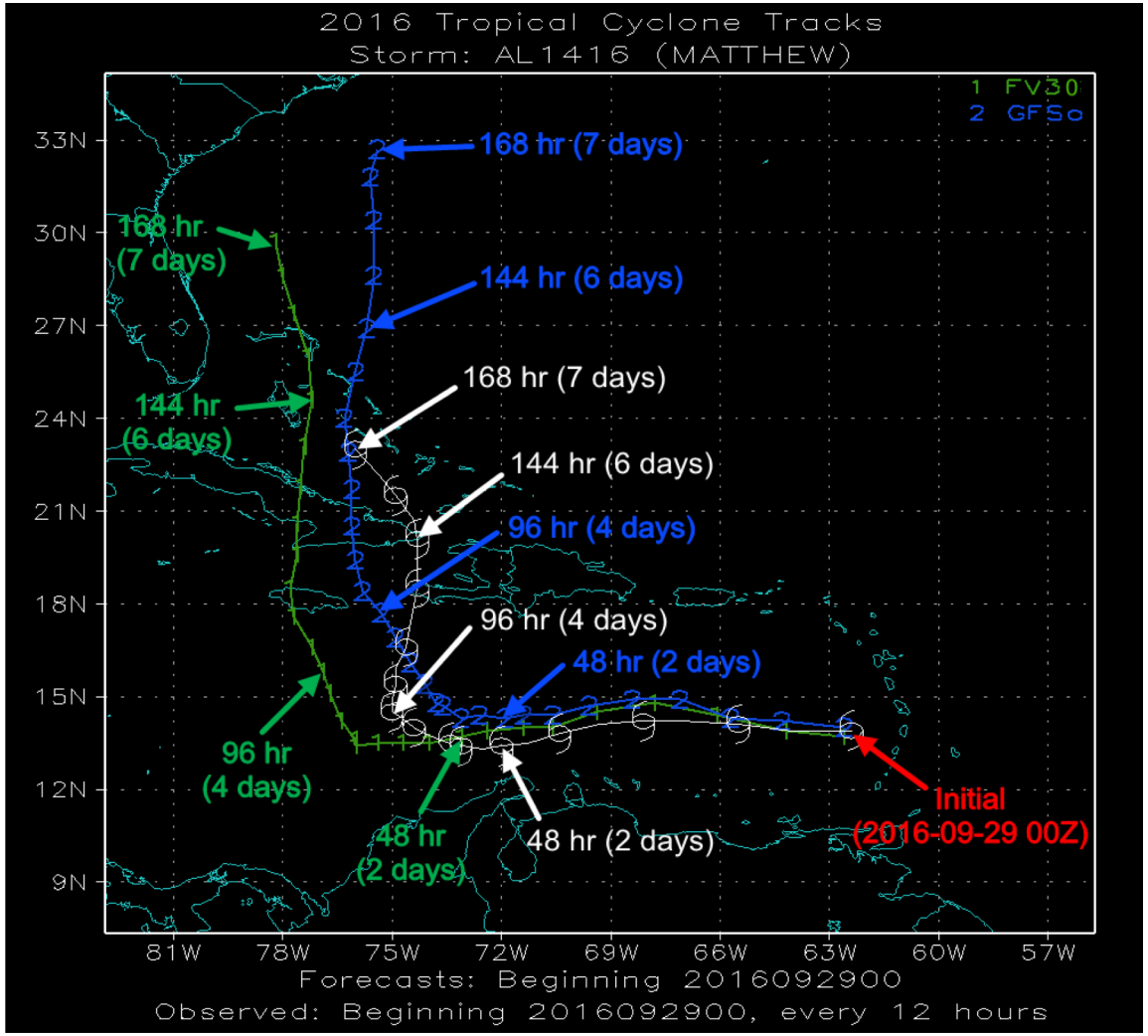
Hurricane Track and Intensity

Hurricane Track

- Hurricane track plot information:
  - Actual track in white, marked every 12 hours.
  - GFS track in blue, marked every 6 hours.
  - FV3v0 track in green, marked every 6 hours.

- Both GFS and FV3v0 move hurricane too fast – center is too far north after 7 days, with GFS moving hurricane faster than FV3v0.
- FV3v0 takes hurricane too far west while GFS follows actual track more closely in space (but not in time since it moves the hurricane north too quickly).
- At final forecast time (168 hr), FV3v0 center is closer to actual than GFS center.

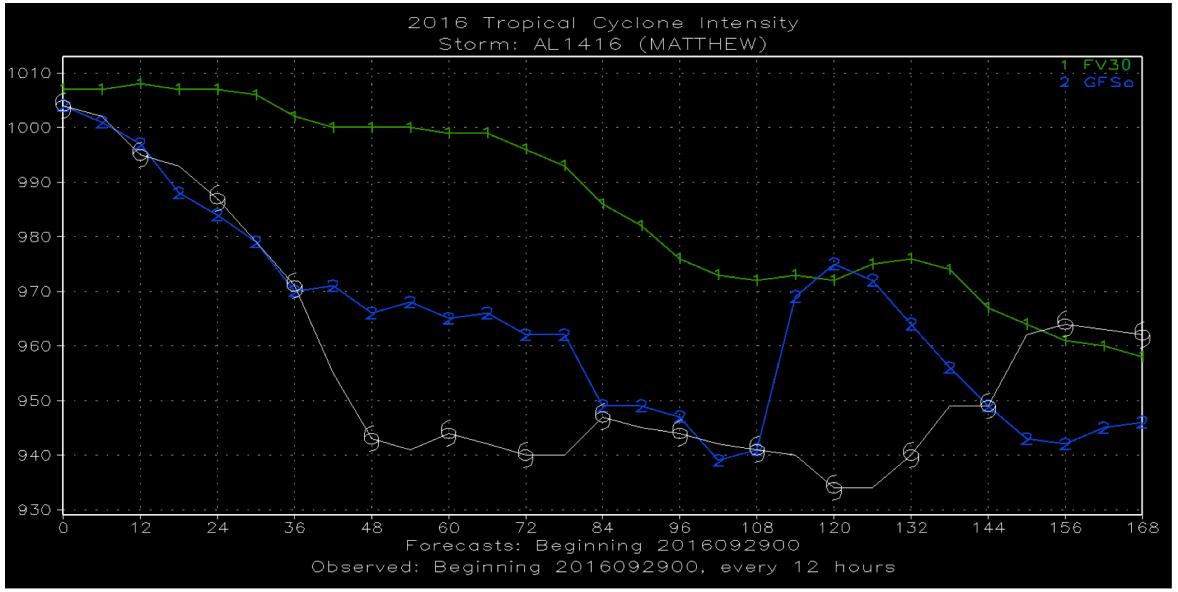
➢ GFS track has better spatial accuracy while FV3v0 has better temporal accuracy. Before ~96 hr, GFS center is closer to actual, while after ~96 hr, FV3v0 center is closer to actual.

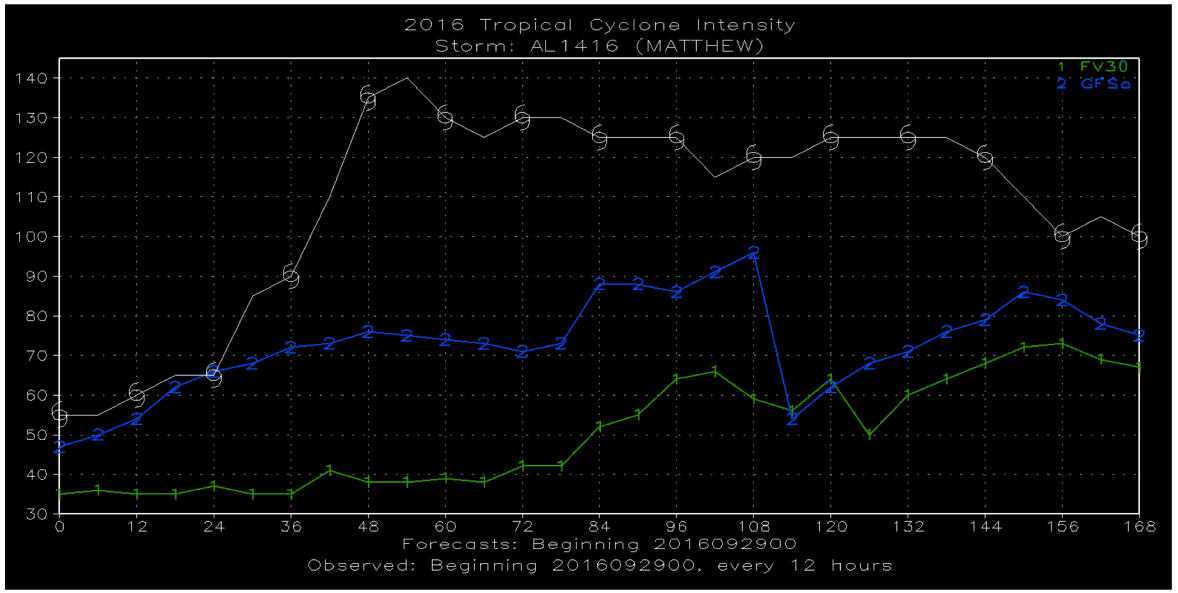


Hurricane Intensity

- Hurricane intensity plot information:
  - Actual track in white, marked every 12 hours.
  - GFS track in blue, marked every 6 hours.
  - FV3v0 track in green, marked every 6 hours.
  - Top plot is mean sea level pressure (MSLP) of center (in mb), bottom plot is maximum wind (knots).

- Both models underpredict intensity, i.e. MSLP too high and max wind too slow.
- GFS intensity (both MSLP and max wind) is almost always closer to observed than FV3v0 intensity.





Summary

- For surface variables (wind speed, temperature, and specific humidity), FV3v0 yields a slightly better forecast.
- For precipitation metrics (frequency bias and equitable threat score), GFS yields a slightly better forecast.
- For vertical profiles (wind speed, temperature, and relative humidity), FV3v0 yields a slightly better forecast for relative humidity (while forecasts of the other two variables are about the same).
- For hurricane track, GFS yields a slightly better forecast before ~96 hr while FV3v0 yields a slightly better forecast after.
- For hurricane intensity, GFS usually yields a better forecast.

➢ Overall, GFS may have the slightly better forecast due to its more accurate prediction of hurricane intensity, but differences between the models are small enough to warrant further study.