# Global Observing System Simulation Experiments (OSSEs) for Error-Added Geostationary Hyperspectral Infrared Sounder (Geo-HSS) Infrared Atmospheric Sounding Interferometer (IASI) Observations[†]

*Sean P.F. Casey[1‡], Robert Atlas[2], Ross N. Hoffman[1], Lidia Cucurull[3], and Narges Shahroudi[4]*

*[1]University of Miami/Cooperative Institute for Marine and Atmospheric Sciences (CIMAS) and National Oceanic and Atmospheric Administration (NOAA)/Atlantic Oceanic and Meteorological Laboratory (AOML)/Hurricane Research Division (HRD), Miami, FL*

*[2]NOAA/AOML, Miami, FL*

*[3]NOAA/AOML, Miami, FL, and NOAA/Earth System Research Laboratory (ESRL), Boulder, CO*

*[4]NOAA/National Environmental Satellite, Data, and Information Service (NESDIS)/Center for Satellite Applications and Research (STAR), College Park, MD*

## Motivation

This project looks at the forecast impacts of a Geo-HSS constellation, using multiple metrics. This project also looks at significance; can we show that these forecast impacts are statistically significant, at the 95% level?

## Geo-Hyper IR (IASI) Observations

Figure 1 shows simulated 12.8-micron (window channel) brightness temperatures for five IASI-like instruments simulated from the Goddard Earth Observing System Model, Version 5 (GEOS-5) 7-km Nature Run (G5NR), and located in geostationary orbit at 0°E, 60°E, 140°E, 225°E, and 285°E without added errors. Control observations, based on instruments that were operational in August 2014, were simulated as well (Casey et al. 2017b). The Global Forecast System (GFS) operational 3-Dimensional Ensemble Variation (3DEnVar) version from 2015 is used in these experiments. More information on the OSSE system is also given by Casey et al. (2017b) and on the simulation, calibration and validation of test Geo-HSS observations by Zhou et al. (2017).

## Description of Experiments

A data drop-out scenario is used to identify the impact and significance of Geo-HSS assimilation. First, a two-week (2006080100-2006081418) spin-up period assimilates all control observations, plus the 5 Geo-HSS (IASI) observations. Next, experiment perhss is run from 2006081500-2006083118, assimilating the same observations as the spin-up period. Concurrently experiment pernohss that denies the 5 Geo-HSS observations runs for the same 2006081500-2006083118 period. So far, all observations are error-free. Explicit observation error biases and variances are then calculated for all

[†] Extended abstract for the poster presentation of Casey et al. (2017a).

[‡] Sean.Casey@noaa.gov

control instruments following Errico et al. (2013) and Casey et al. (2017b). The explicit ob error magnitudes for IASI_Metop-B are then applied to each Geo-HSS platform. Finally, experiments errhss and errnohss are set up similar to perhss and pernohss, respectively, with errhss run from 2006080100-2006083118 and errnohss run from 2006081500-2006083118. For all experiments, Global Spectral Model (GSM) forecasts are run every 12 hours at 00Z and 12Z.

Many differences in standard metrics for these experiments are small in magnitude and to highlight differences that do exist the specialized plotting convention used for many results shown in this paper is explained in Figure 2. Experiment perhss will be shown in blue, pernohss in red, errhss in cyan, and errnohss in magenta. Error bars denote 95% confidence intervals for the metric. Circles around a given value (blue arrow) denote significant differences based on a paired t-test, where a difference is considered statistically significant if the error bar from one experiment does not contain the mean value of another experiment (blue oval, comparison between errhss and errnohss). If the error bars overlap and include the value of another experiment (red circle), the difference is not significant.

**Hurricane Case Study**

The first metric used for assessing forecast impact is the impact on forecast tracks of AL01, an Atlantic Basin TC that was first identified on 2006081818, reached a maximum 10-m wind speed of 107 knots on 2006082700, and made landfall in Labrador, Newfoundland, Canada on 2006083012. Figure 3 shows 00Z forecast tracks for AL01 for all four experiments. Small differences are noted in these maps, but nothing stands out as a major effect of assimilating Geo-HSS observations on the track forecasts (with the possible exception of greater variability noted in errnohss).

Figure 4 shows the mean forecast track error as a function of forecast hour, along with confidence intervals (as described in Figure 2) for the comparison between (top) perhss and errhss, and (bottom) pernohss and errnohss. Both of these comparisons show nonsignificant differences, though a visual inspection shows that the impact of adding errors is stronger in the no-Geo-HSS case than in the Geo-HSS case. As Figure 5 shows, this changes the observed impact that Geo-HSS has in the error-added case. While Geo-HSS has a mixed impact in the perfect case, with no forecast hour differences being significant, it has a clear positive impact in the error-added case that is significant at forecast hours 54-66.

Figure 6 is similar to Figure 4, only for mean maximum-wind error. Once again, the no-Geo-HSS case appears more sensitive to these added errors than the with-Geo-HSS case. Figure 7, however, is unlike Figure 5 in that none of the forecast error differences are considered statistically significant, given high variability in maximum-wind errors in the OSSE system.

All mean minimum-surface-pressure errors are negative, as shown in Figure 8; this is likely due to the lower resolution of GFS compared to the 7-km G5NR. While the addition of errors does not have a significant impact on the Geo-HSS case, the no-Geo-HSS case with added errors sees significantly larger (in magnitude) minimum-surface-pressure errors at forecast hours 144-162. Given the large variability at these forecast times, however, Figure 9 shows that these forecast impacts of Geo-HSS are insignificant in both cases.

Overall, the addition of observation errors increases the forecast errors as expected, but adding errors to the control configuration has a greater impact on hurricane forecast errors than for the Geo-HSS configuration. This leads to different conclusions in the case of hurricane track error, where the differences are deemed not significant in the perfect-ob comparisons but significant in the error-added comparisons. This does not, however, change the conclusions with respect to hurricane intensity errors (both maximum-wind and minimum-surface-pressure errors) because in these comparisons the case-by-case variability is large.

**Global-Scale Comparisons**

The second metric used for assessing forecast impact is global-scale, using latitude bands around the globe. This will demonstrate whether there are significant impacts from Geo-HSS assimilation on the primary metrics used to predict hemispheric and tropical meteorology on a synoptic scale. These latitude bands, along with sample differences between the perhss and pernohss cases, are shown in Figure 10. For Northern and Southern Hemisphere, 500 hPa geopotential height anomaly correlation is analyzed, and for the tropics, 200 hPa vector wind RMSE is used. (Note Figure 10 shows scalar wind differences for clarity.)

Figure 11 uses the same plotting style described in Figure 2, now for Northern Hemisphere 500 hPa Geopotential Height Anomaly Correlation. In both Control and Geo-HSS configurations, the error-added observations yield significantly lower AC scores. Experiment errhss yields statistically significant differences from perhss for forecast hours 0-60 and 132-168, whereas errnohss is significantly different from pernohss for fewer forecast hours (0-54 and 156-168). This leads to a small difference in Geo-HSS impact results (Figure 12), where no comparison is statistically significant for the perfect-ob case, but Geo-HSS significantly reduces AC at the analysis time (remember that these are verified against the nature run, not against self-analyses, so AC forecast hour 0 is not 100% by default).

Figure 13 shows similar behavior in the Southern Hemisphere, where error addition lowers AC scores significantly for 0-102 hr in the Geo-HSS case and 0-90 hr in the no-Geo-HSS case. As in Figure 12, Figure 14 shows a negative significant impact for Geo-HSS in the error-added case, while no significant difference is noted for the perfect-observation case.

The impact on Geo-HSS observations compared to control observations is clear in Figure 15, which shows RMSE of 200 hPa vector wind in the tropics. For the Geo-HSS case, the error-added observations yield significantly higher RMSE at all forecast hours. For the no-Geo-HSS case, though, this increase in RMSE is only significant at hours 0-90 and 162-168. As such, Figure 16 shows that, while the perfect-ob case shows no significant impact for Geo-HSS, the error-added case shows significant degradation caused by Geo-HSS at forecast hours 0-66!

Unlike with the hurricane case study, the impact of added error appears to be greater when Geo-HSS is included compared to the control (no-Geo-HSS) case. As such, a slight improvement in forecast skill (for perfect data) becomes a somewhat negative forecast impact for error-added observations.

**Overall Forecast Score**

While these primary metrics are for the most part not statistically significant, our final metric tests whether the combination of these and other pressure-level metrics into an overall forecast score (OFS, Boukabara et al. 2016) is statistically significant. In the OFS, an overall forecast quality is assessed based on normalization of AC and RMSE statistics for multiple forecast hours, verification hours, variables, and pressure levels. Individual overall forecast scores $O_{AC}$ and $O_{RMSE}$ are calculated as:

$$O_{AC} = (1/n)\sum_{i=1}^{np}\sum_{j=1}^{nlev}\sum_{k=1}^{nhr}\sum_{m=1}^{nday}\frac{(AC_{i,j,k,m} - min_{i,j,k})}{(max_{i,j,k} - min_{i,j,k})}$$

$$O_{RMSE} = (1/n)\sum_{i=1}^{np}\sum_{j=1}^{nlev}\sum_{k=1}^{nhr}\sum_{m=1}^{nday}\left[1 - \frac{(RMSE_{i,j,k,m} - min_{i,j,k})}{(max_{i,j,k} - min_{i,j,k})}\right]$$

Note that each normalized score within the summations is in the range [0,1]. An overall metric combining these two, $O_{combined}$, is calculated using the normalization scores for both $O_{AC}$ and $O_{RMSE}$. Only metrics that were numerically uncorrelated ($R^2 < 0.25$) were used; this meant using forecasts for HGT, RH (RMSE only), T, and WIND initialized every 12 hours, verified every 18 hours, and calculated at three pressure levels (250, 500 and 850 hPa).

Figures 17 through 20 show the OFS scores for all four experiments calculated across the full globe (Figure 17), Northern Hemisphere (Figure 18), Southern Hemisphere (Figure 19), and Tropics (Figure 20). All four of these figures show the same relation among the four experiments. For perfect observations, Geo-HSS has a small but significant positive impact compared to the no-Geo-HSS case. However, for error-added observations, Geo-HSS has a clear *negative* impact

**Discussion**

The addition of explicit ob errors does, in this case, change the OSSE results for five geostationary Geo-IASI satellites. In comparison to perfect-observation results, the error-added results show a slightly more positive impact for Geo-HSS in the hurricane case study, but more negative results for global primary metrics and OFS. The fact positive results with perfect obs turn into negative results with error-added observations suggests an issue with the error-addition applied to at least the Geo-HSS instruments, if not some of the control instruments as well.

One potential clue is found in Figure 3. AL01, the storm used in the hurricane study, is effectively what could be called a “fish storm.” It stays hundreds of nautical miles away from land until it approaches Nova Scotia and transitions into an extratropical storm. In going from perfect to error added data, we see positive change in forecasts for this ocean-only storm, but negative impacts when areas including land are included. This suggests that land vs. ocean differences, and how they are applied to simulated Geo-HSS observations, could be to blame.

For errhss, the same explicit ob bias and variance magnitudes are applied to the Geo-HSS platforms as are calculated for IASI_Metop-B. If we run an additional iteration, an “errV2”, we would then have different bias/variance magnitudes added to IASI_Metop-B and the Geo-HSS observations. Differences

between IASI_Metop-B and the Geo-HSS instruments for this errV2 case could explain the discrepancy for the "error version 1" (errhss) analysis/forecast degradation.

Figure 21 shows the difference between explicit observation biases and variances for Metop-B and each Geo-HSS instrument for oceanic observations. These differences are rather small, no more than ~0.1 K for bias and ~0.15K for variance. For land observations, however, Figure 22 shows much larger differences for surface sensitive channels, up to ~0.8 K for bias and ~1 K for variance. There are also very clear differences in the individual Geo-HSS instruments at these wavenumbers, with GOES 13 and GOES 15 showing greater explicit ob errors than JMA, ISAT, and M10.

Would an additional separation by land surface type for radiance explicit observation errors remedy this situation, yielding a control error-added dataset whose explicit ob errors could be applied to simulated test radiances without encountering this perfect vs. error-added result discrepancy? Figure 23 tries to answer this by showing the Implicit Observation Bias (top) and Variance (bottom) for IASI channel 1271 (962.5 $cm^{-1}$) from experiment perhss, comparing IASI_Metop-B implicit observation errors (x-axis) with IASI_G13 implicit observation errors (y-axis). Each data point represents a given land surface type (9 total classifications). Linear fit and correlation ($R^2$) are provided as well.

Implicit observation errors, rather than explicit observation errors, show what the observation innovation mean/variance are solely due to potential errors in simulation, assimilation, and differences between the background and the observation. By using perhss, we can isolate implicit observation errors, since no explicit observation error is added. Channel 1271 is chosen because it is a surface-sensitive channel, and it is associated with the largest differences between Metop-B and GOES 13 and 15 in Figure 22. Figure 23 suggests that a separation by land type may be helpful in terms of variance, with the individual implicit observation biases/variances being moderately correlated and somewhat resembling a 1:1 relation. For biases, however, this does not appear to be the case, with the exception of land type 1 (water), which shows biases for both instruments of around -0.1 K. For other land types, IASI_G13 appears to have identified biases around 0.4 K colder than IASI_Metop-B. (This could be due to differences in sun angle, as IASI_Metop-B passes over at two set local times each day, whereas IASI_G13 views the same area continuously.) Figure 23 suggests that separating out explicit observation error calculation by land type for radiances could be helpful when testing additional radiance experiments in terms of matching the observed variance of a control experiment, though the observed bias may need to be monitored closely.

**References**

Boukabara, S.-A., K. Garrett, and V. K. Kumar, 2016: Potential gaps in the satellite observing system coverage: Assessment of impact on NOAA's numerical weather prediction overall skills. *Mon. Wea. Rev.*, 15 July 2016. In press. doi:10.1175/MWR-D-16-0271.1.

Casey, S. P. F., R. Atlas, R. N. Hoffman, L. Cucurull, and N. Shahroudi, 2017a: Global OSSEs for Error-Added Geo-Hyper IR (IASI) Observations. *21st Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface. American Meteorological*

*Society 97th Annual Meeting,* Seattle, WA, January 2017. https://ams.confex.com/ams/97Annual/webprogram/Paper307087.html.

Casey, S. P. F., R. Atlas, R. N. Hoffman, L. Cucurull, J. S. Woollen, I. Moradi, N. Shahroudi, S. A. Boukabara, K. Ide, R. Li, N. Privé, and F. Yang, 2017b: Creation of a Control Dataset and Forecast System for Global OSSEs. *21st Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface. American Meteorological Society 97th Annual Meeting,* Seattle, WA, January 2017. https://ams.confex.com/ams/97Annual/webprogram/Paper307082.html.

Errico, R. M., R. Yang, N. C. Privé, K.-S. Tai, R. Todling, M. E. Sienkiewicz, and J. Guo, 2013: Development and validation of observing-system simulation experiments at NASA's Global Modeling and Assimilation Office. *Quart. J. Roy. Meteor. Soc.*, **139**, 1162-1178, doi:10.1002/qj.2027.

Zhou, Y., N. Shahroudi, K. Ide, and S. A. Boukabara, 2017a: Assessment of the cubesat infrared atmospheric sounder impact through the community global OSSE package. *21st Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, American Meteorological Society, Boston, MA, Seattle, Washington, poster 671. The poster presented actually matches Zhou et al. 2017b Available online at https://ams.confex.com/ams/97Annual/webprogram/Paper315199.html.

Zhou, Y., N. Shahroudi, K. Ide, and S. A. Boukabara, 2017b: Assessment of the geostationary hyperspectral infrared sounders impact through the community global OSSE package. *21st Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, American Meteorological Society, Boston, MA, Seattle, Washington, paper 3.4. The talk given actually matches Zhou et al. 2017a. Available online at https://ams.confex.com/ams/97Annual/webprogram/Paper302427.html.

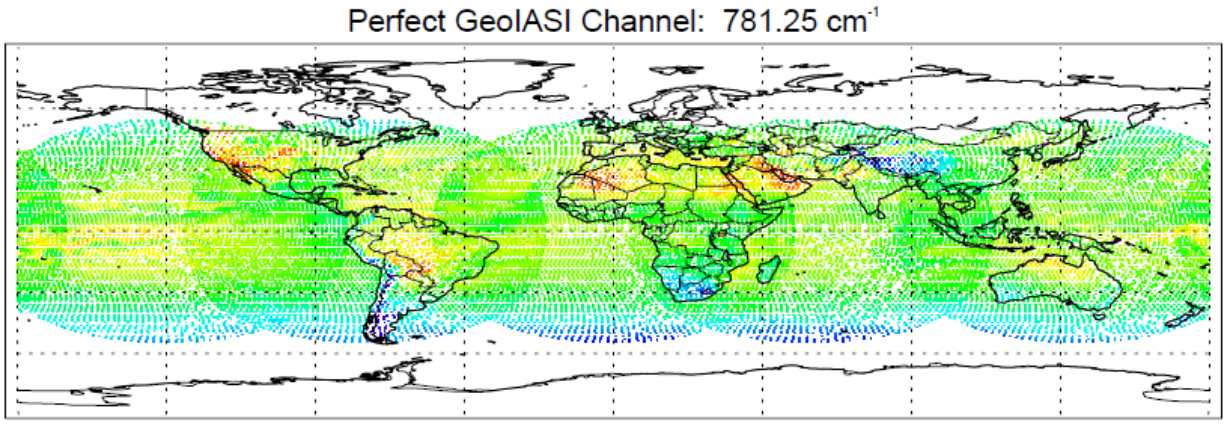



Figure 1: Simulated 12.8-micron observations from five Geostationary Hyper-Spectral Sounder (Geo-HSS) satellites seated at 0°E, 60°E, 140°E, 225°E, and 285°E.

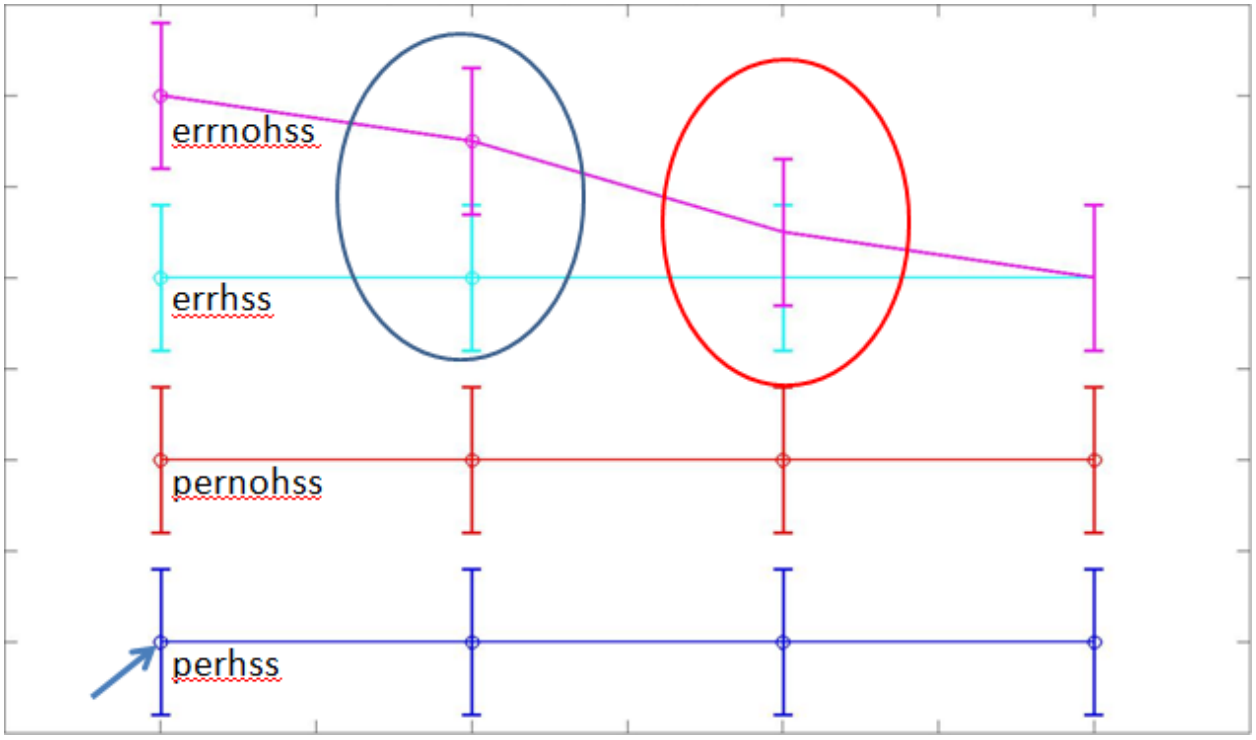



Figure 2: Explanation for plots used in Figures 4-15. Experiment perhss is shown in blue, pernohss in red, errhss in cyan, and errnohss in magenta. Error bars denote 95% confidence interval. Circles around a given value (blue arrow) denote significant differences. If the error bars overlap, but the error bar from one experiment does not contain the value of another instrument (blue oval), the difference is considered statistically significant. If the error bars overlap and include the value of another instrument (red circle), they are not significant.

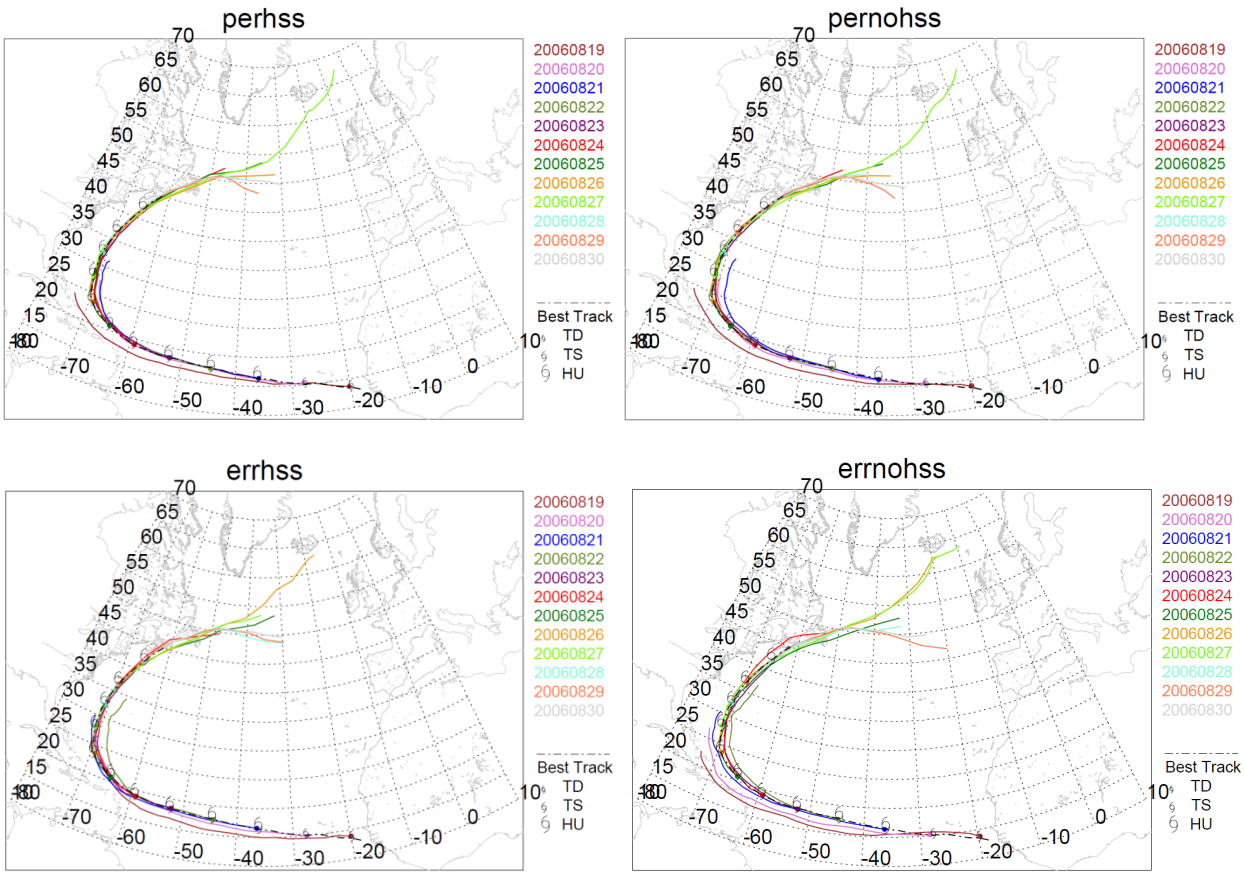



Figure 3: G5NR AL01 00Z forecasts for (top left) perhss, (top right) pernohss, (bottom left) errhss, and (bottom right) errnohss, with best-track information included in black.

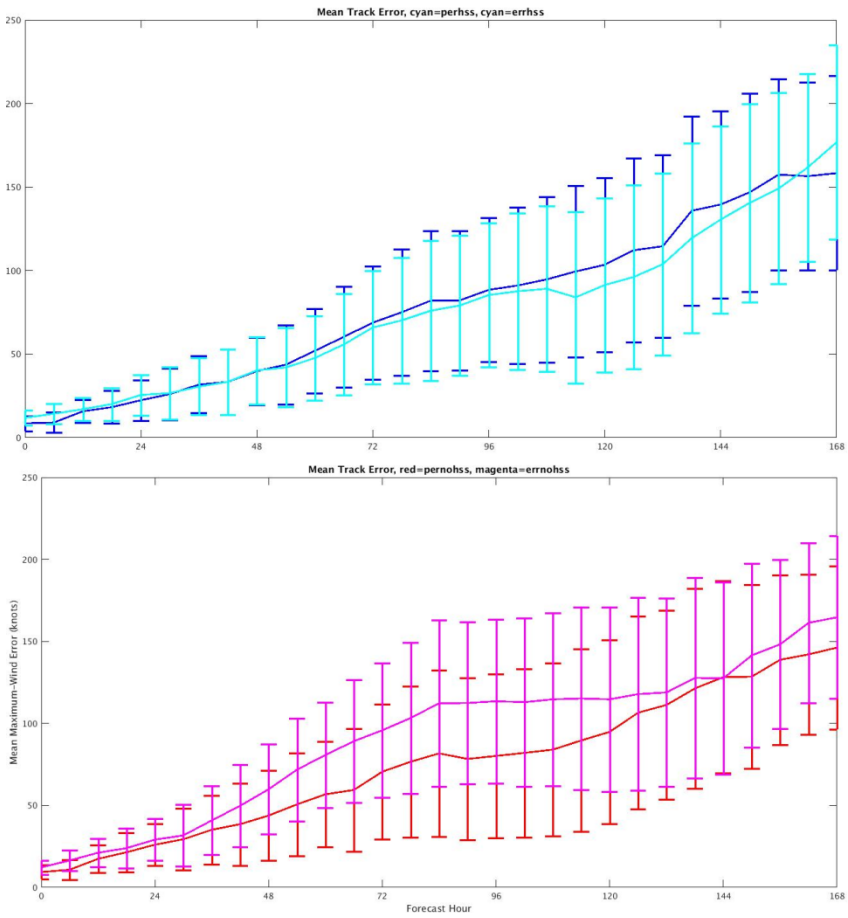



Figure 4: The effect of errors on TC AL01 forecast track error, comparing the differences between (top) experiments perhss (blue) and errhss (cyan), and (bottom) experiments pernohss (red) and errnohss (magenta). Figure description given in Figure 2.

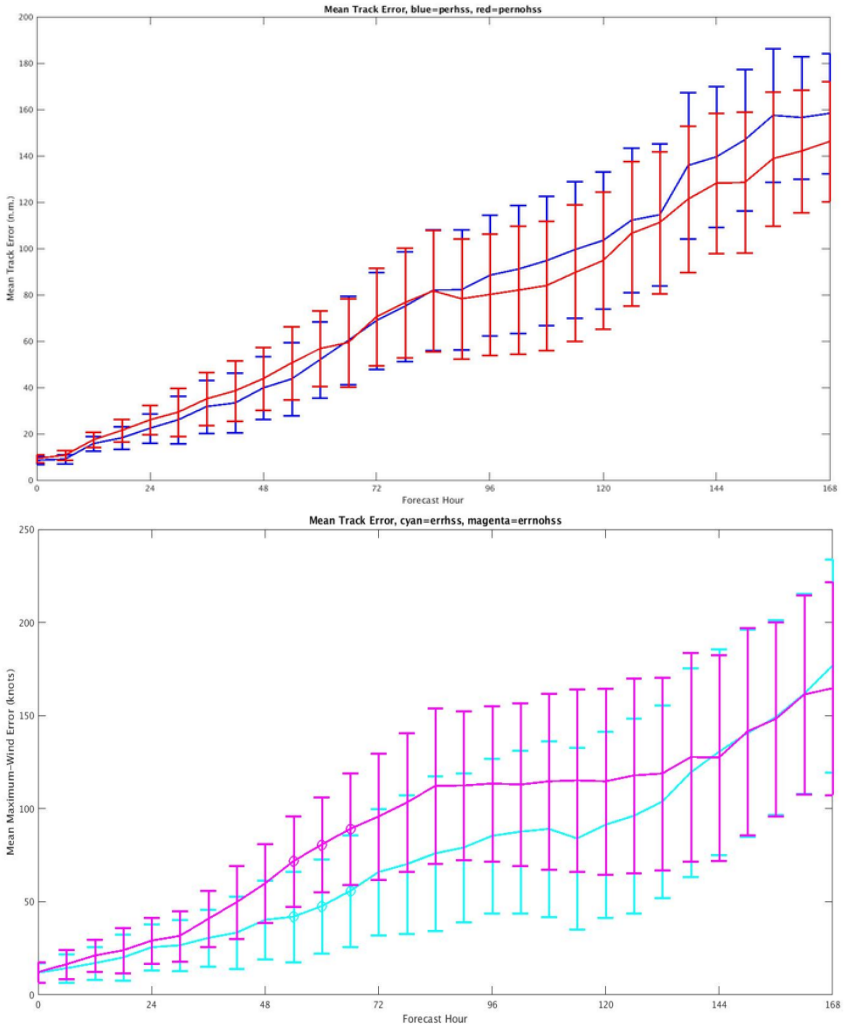



Figure 5: The effect of adding Geo-HSS radiance observations on forecast track error. As in Figure 4, but comparing (top) perhss (blue) and pernohss (red), and (bottom) errhss (cyan) and errnohss (magenta).

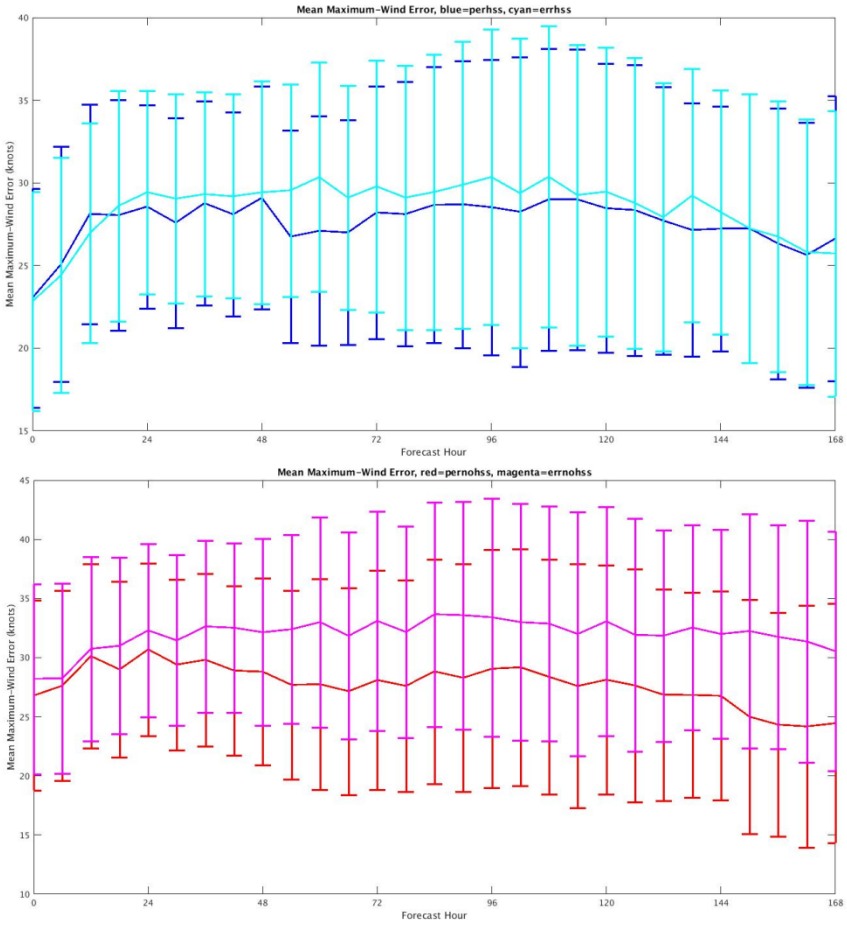



Figure 6: The effect of errors on TC AL01 maximum-wind (intensity) error. As in Figure 4.

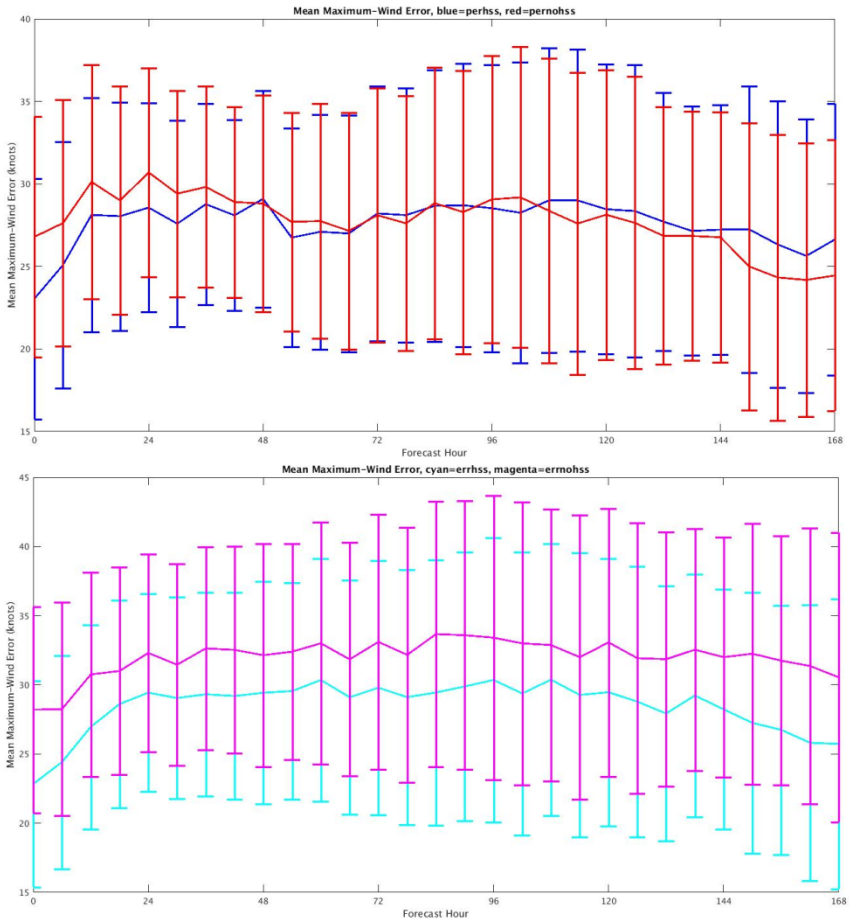



Figure 7: The effect of adding Geo-HSS radiance observations on maximum-wind (intensity) error. As in Figure 5.

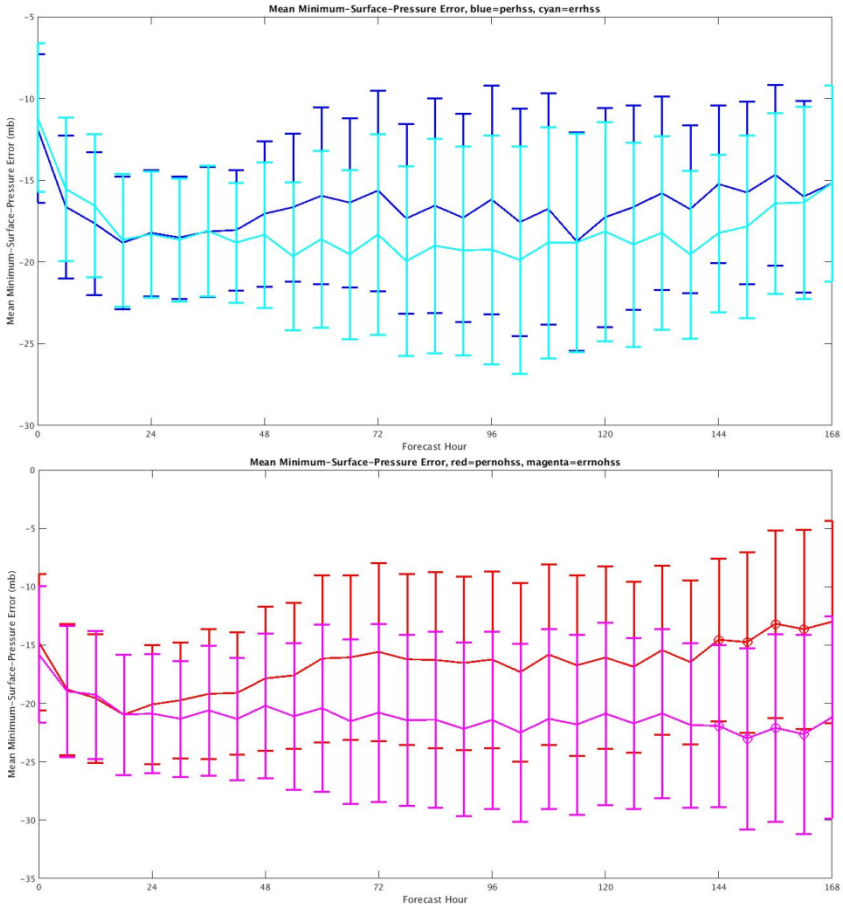



Figure 8: The effect of errors on TC AL01 minimum-central-pressure (intensity) error. As in Figure 4.

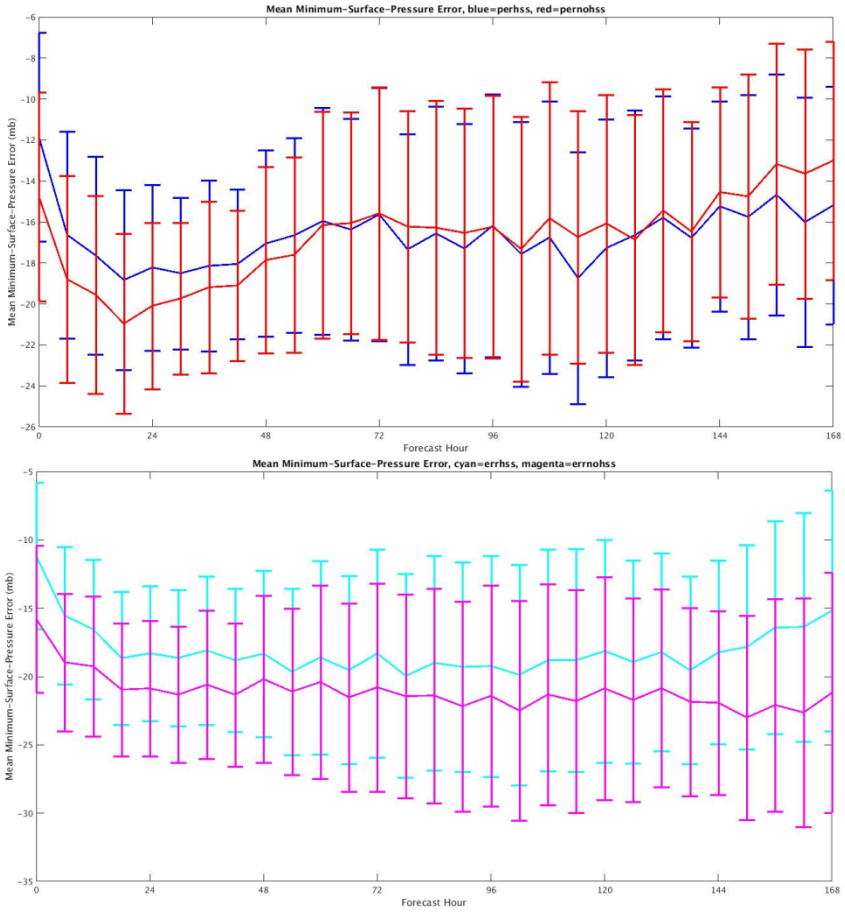



Figure 9: The effect of adding Geo-HSS radiance observations on minimum-central-pressure (intensity) error. As in Figure 5.

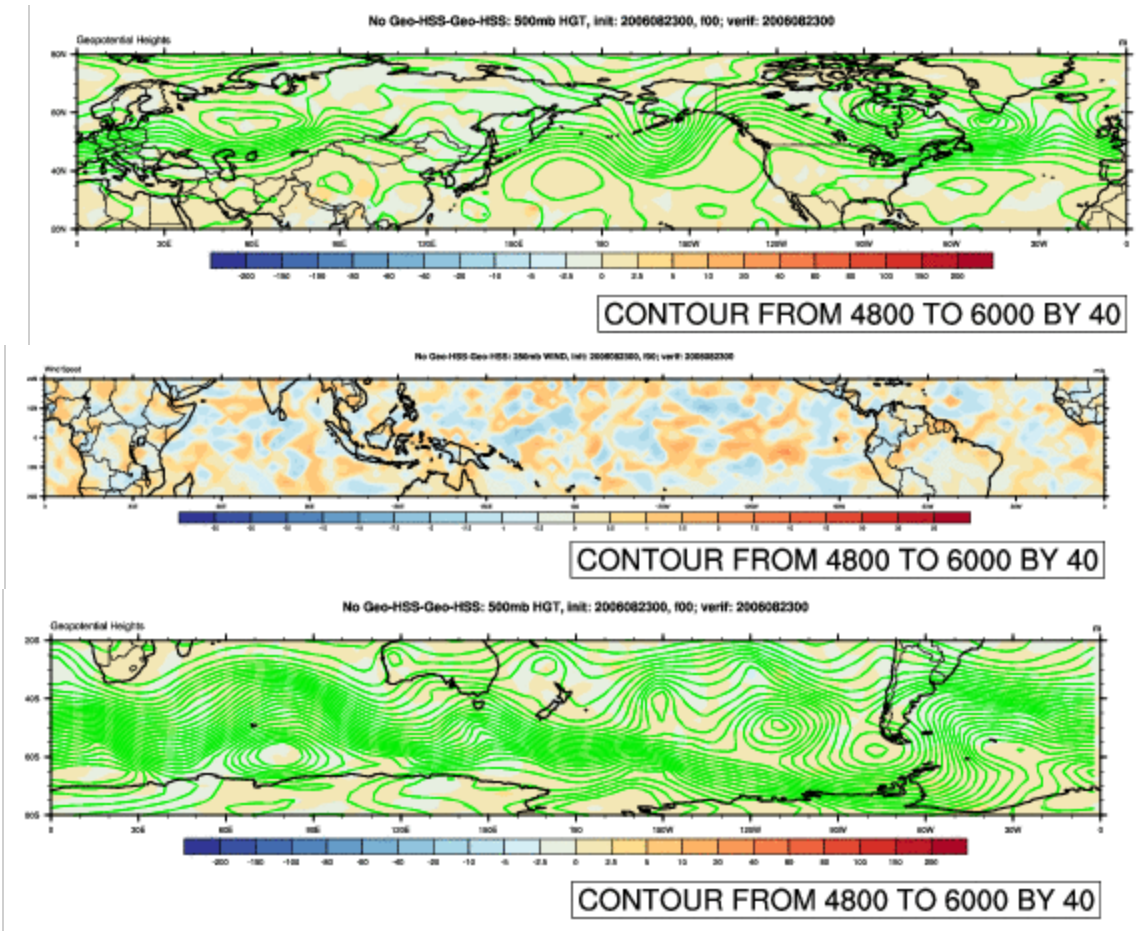



Figure 10: Latitude-band global scale comparison range for Northern Hemisphere (20°-80°N, top), Tropics (20°S-20°N, middle), and Southern Hemisphere (80°-20°S, bottom). Northern and Southern Hemisphere plots include contoured 500 hPa Geopotential Heights for the Geo-HSS case on 2006082300 (green), as well as the difference (no Geo-HSS minus Geo-HSS, red/blue contours). Red/blue contours on Tropical plot shows difference (no Geo-HSS minus Geo-HSS) between 200 hPa wind speed.

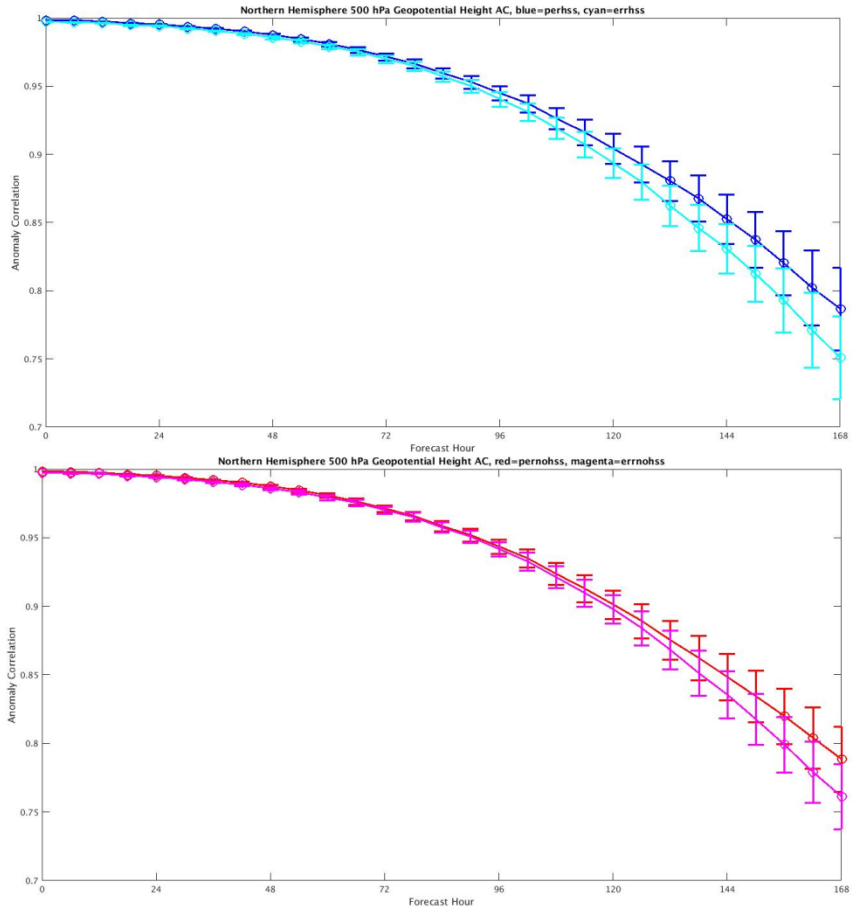



Figure 11: The effect of errors on the Northern Hemisphere 500 hPa Geopotential Height Anomaly Correlation. As in Figure 4.

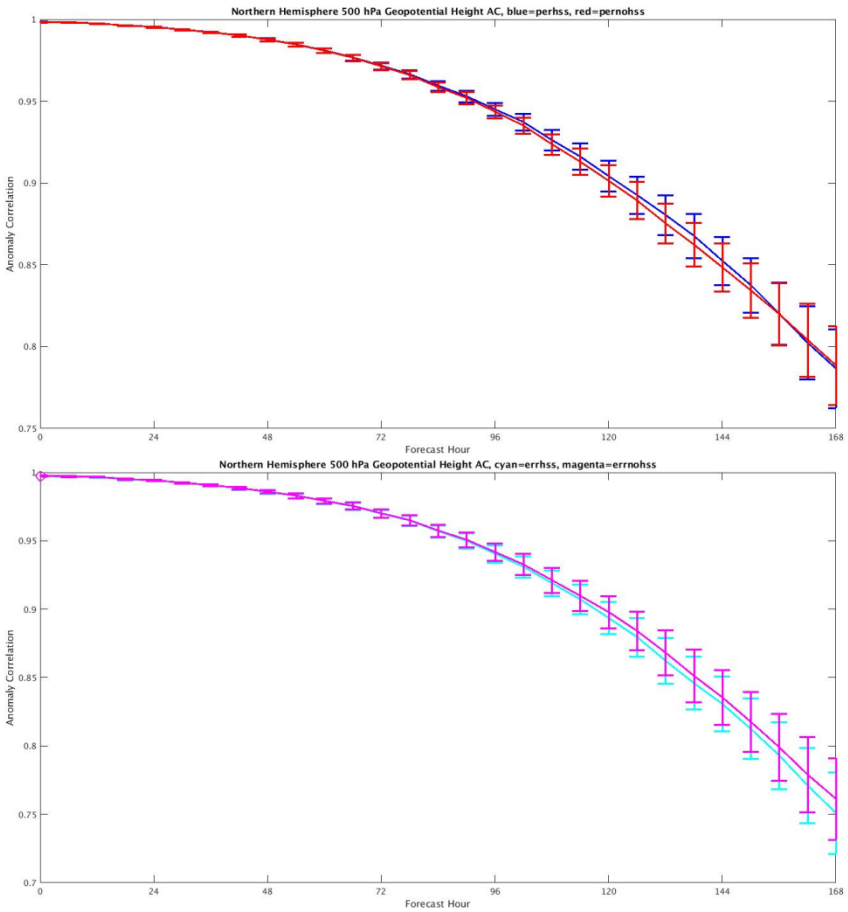



Figure 12: The effect of adding Geo-HSS radiance observations on Northern Hemisphere 500 hPa Geopotential Height Anomaly Correlation. As in Figure 5.

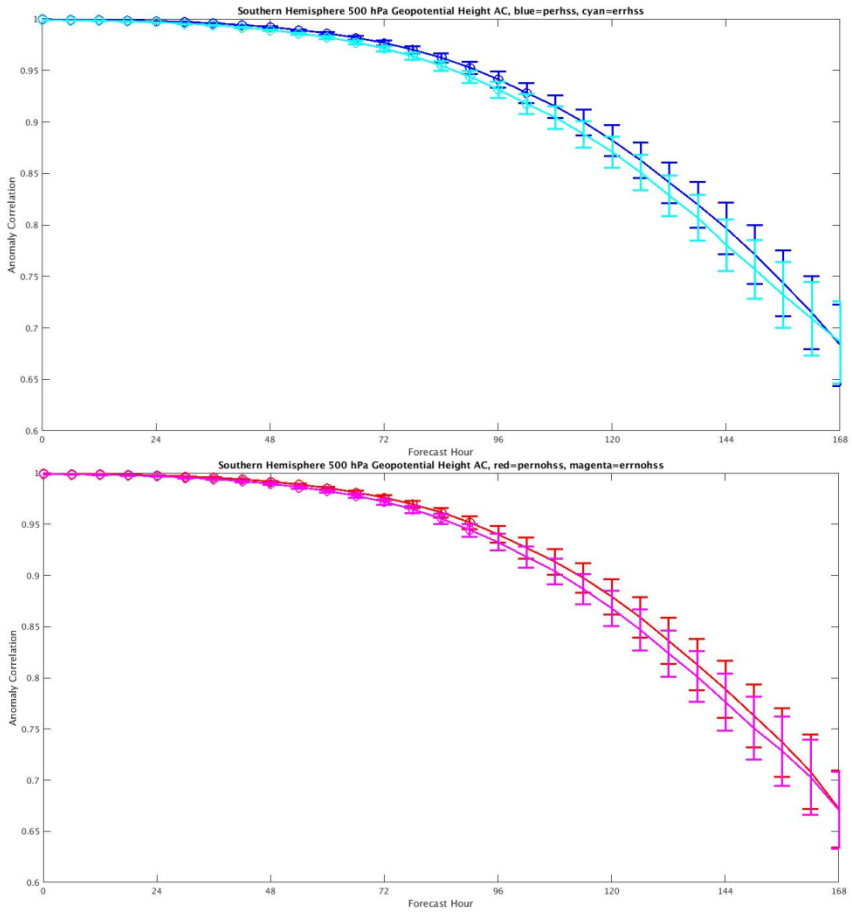



Figure 13: The effect of errors on the Southern Hemisphere 500 hPa Geopotential Height Anomaly Correlation. As in Figure 4.

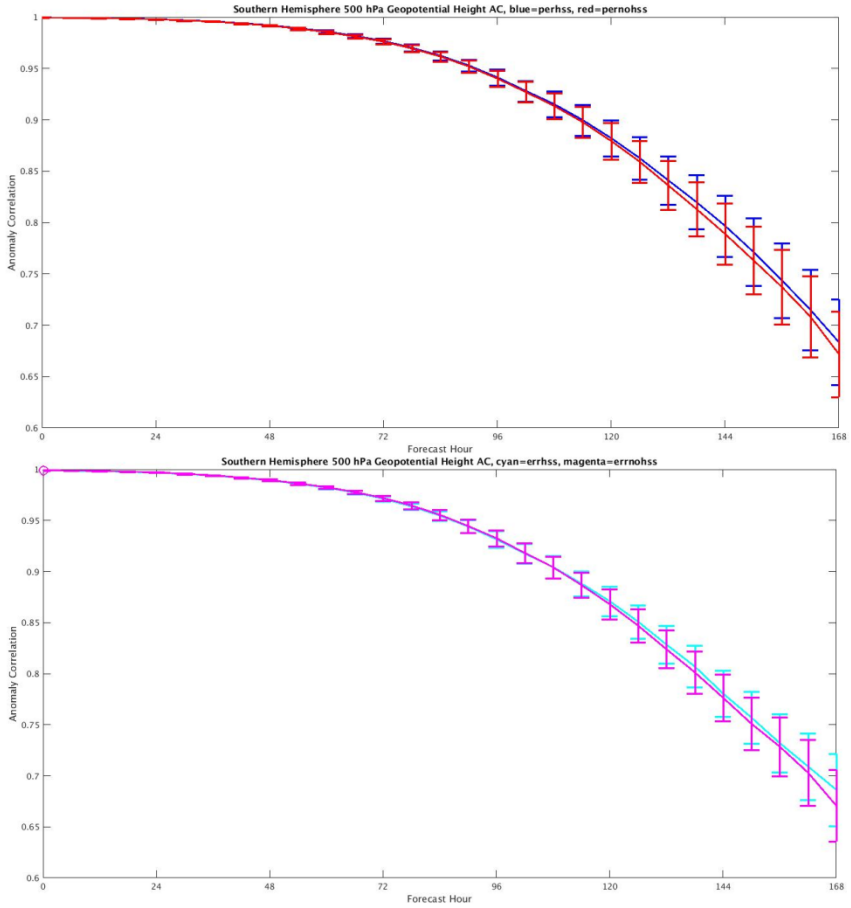



Figure 14: The effect of adding Geo-HSS radiance observations on Southern Hemisphere 500 hPa Geopotential Height Anomaly Correlation. As in Figure 5.

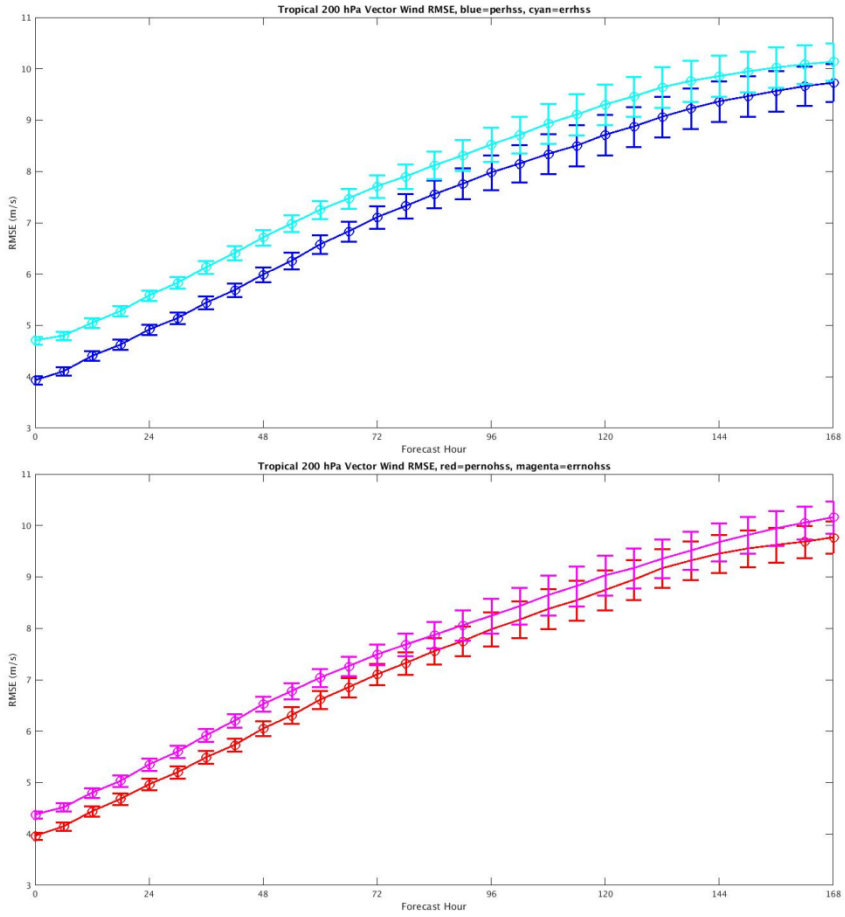



Figure 15: The effect of errors on the 200 hPa Tropical Vector Wind RMSE. As in Figure 4.

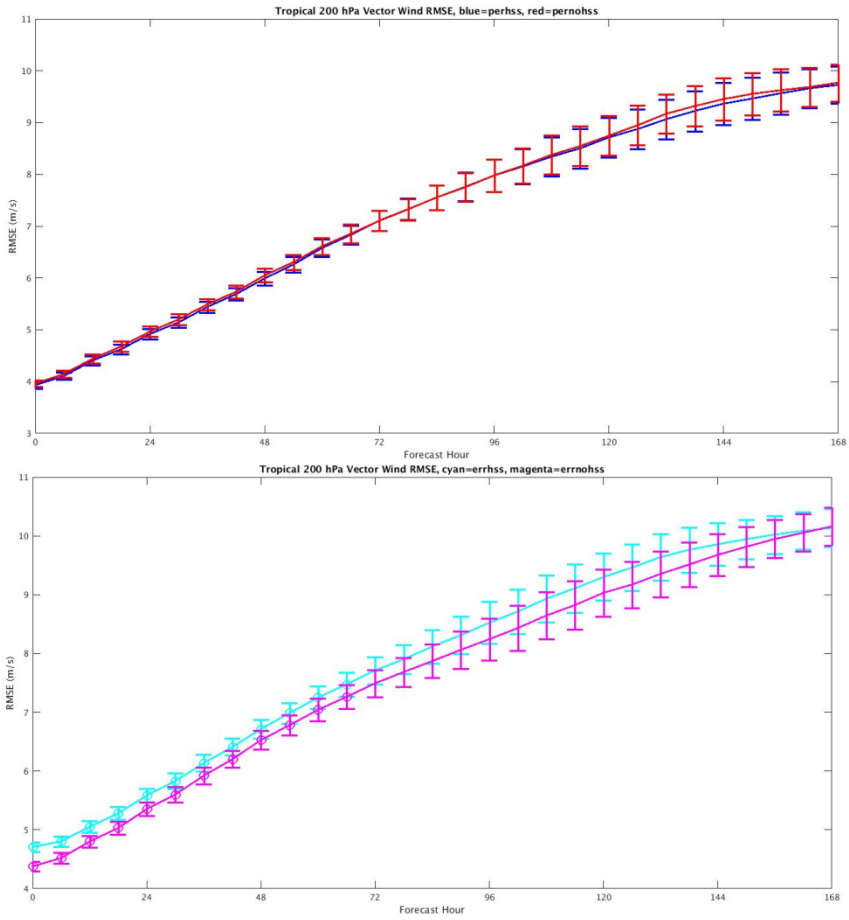



Figure 16: The effect of adding Geo-HSS radiance observations on 200 hPa Tropical Vector Wind RMSE. As in Figure 5.

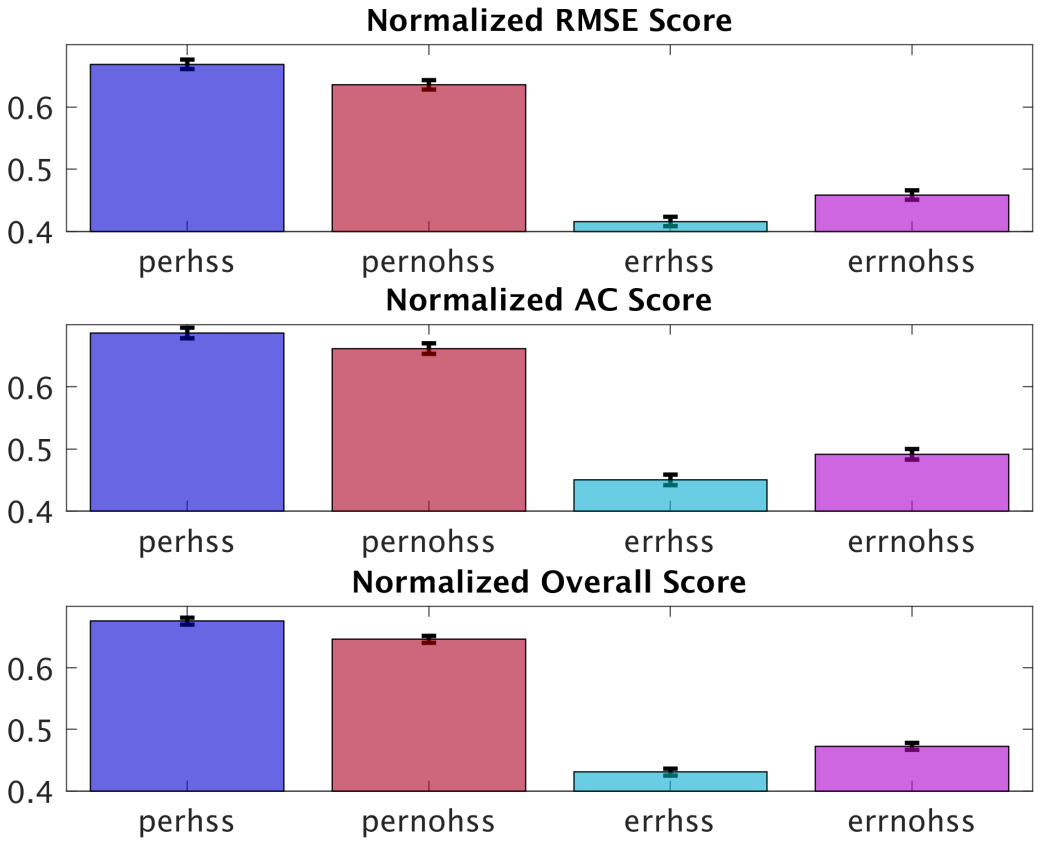



Figure 17: Normalized overall forecast score (OFS) for perhss (blue), pernohss (red), errhss (cyan), and errnohss (purple), in terms of RMSE (top), AC (middle), and combined (bottom) for the global domain. Black bars denote 95% confidence interval.

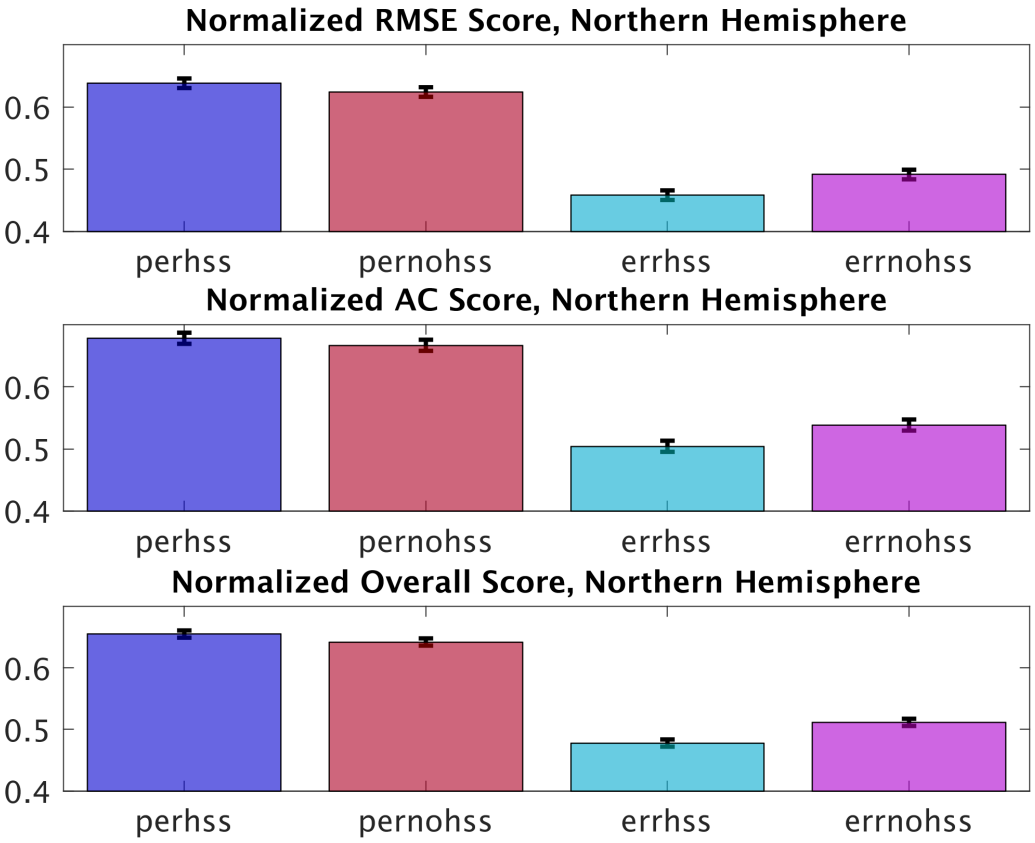



Figure 18: As in Figure 17, for the Northern Hemisphere.

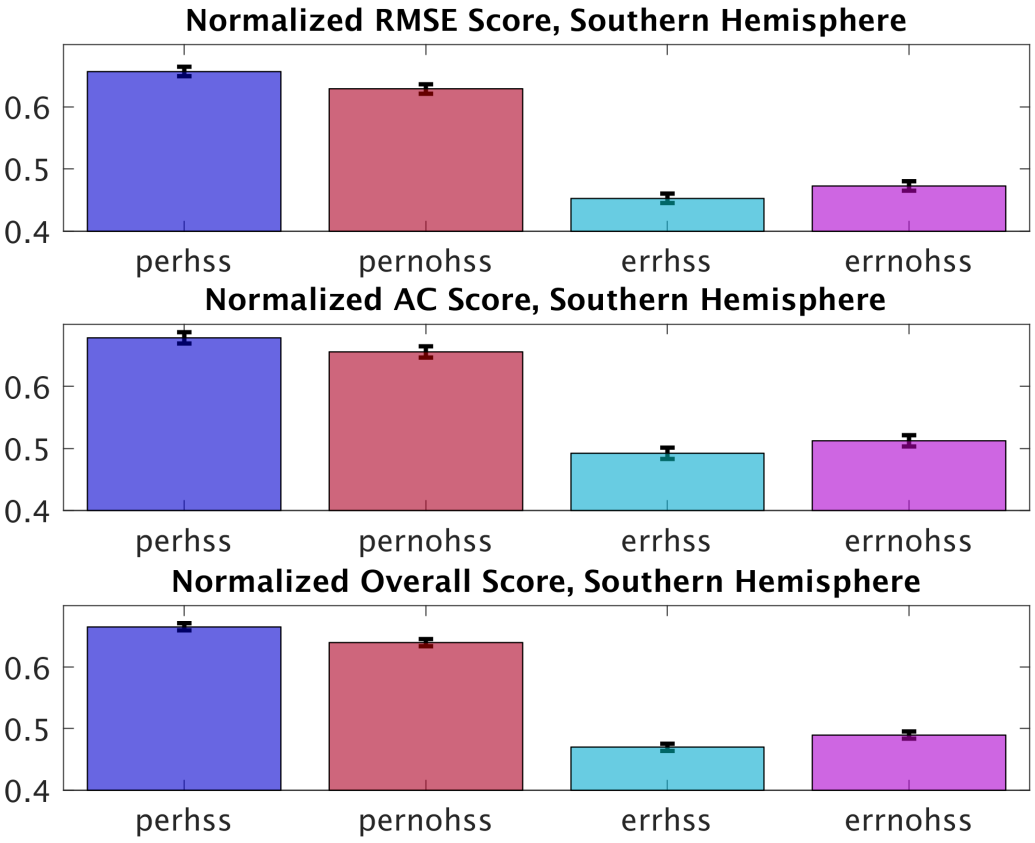



Figure 19: As in Figure 17, for the Southern Hemisphere.

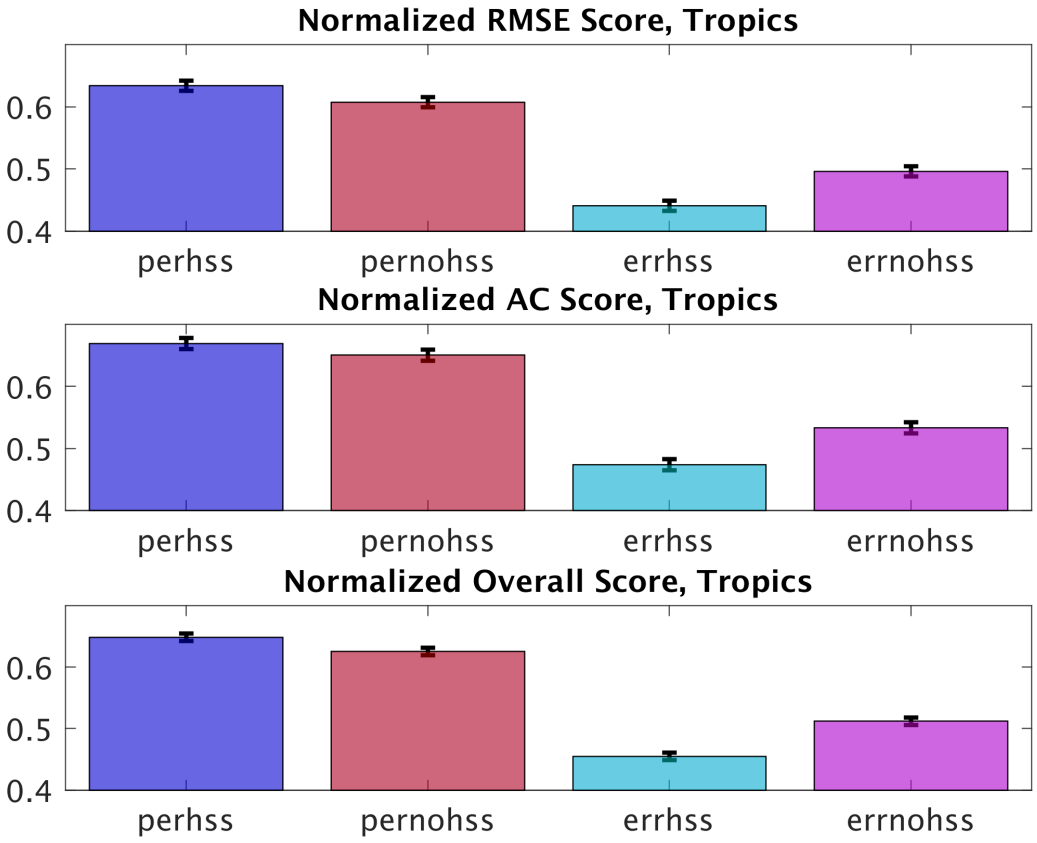



Figure 20: As in Figure 17, for the Tropics.

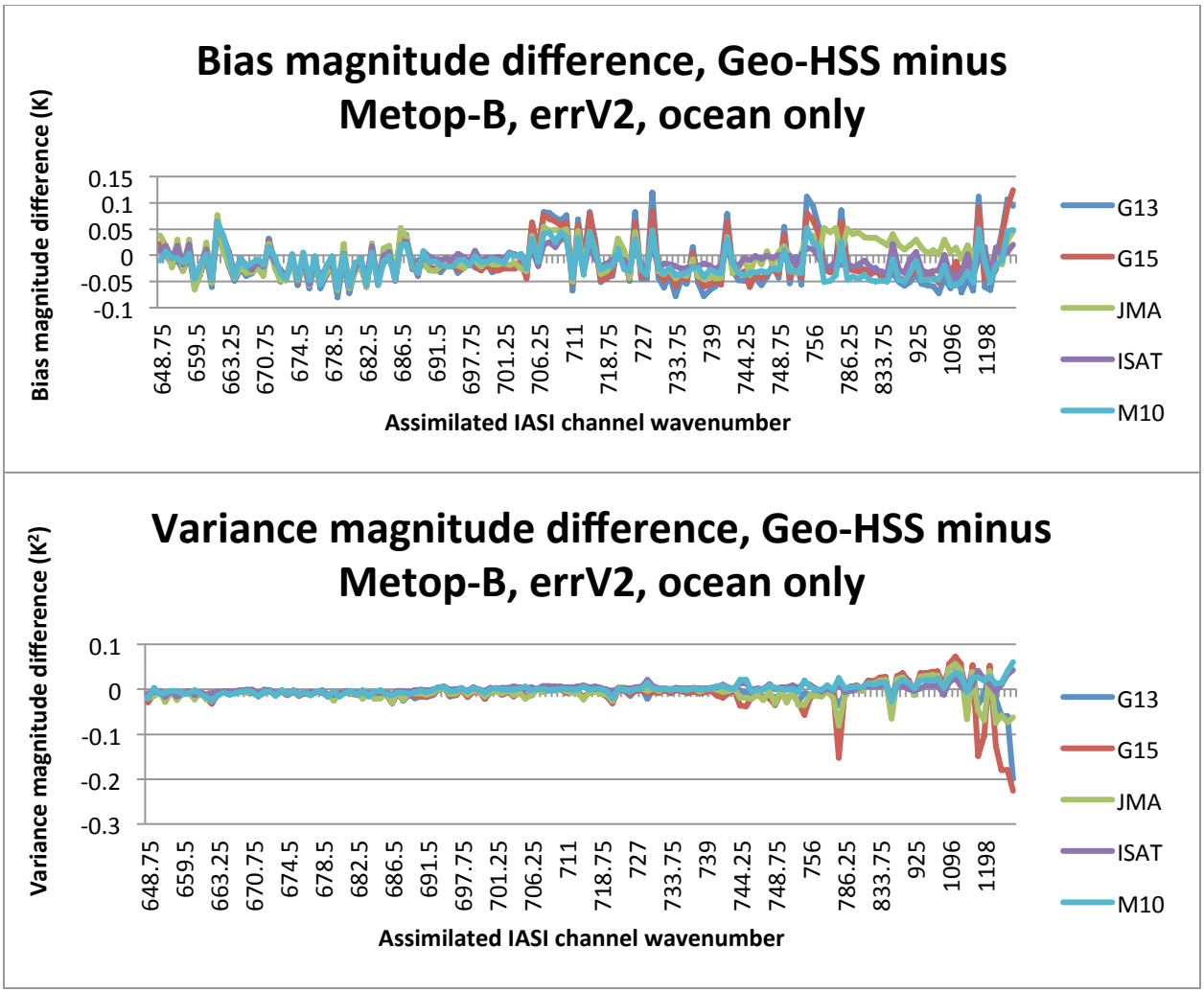



Figure 21: Difference between errV2 explicit ob biases (top) and variances (bottom) for Metop-B versus channel for each Geo-HSS satellite for oceanic observations.

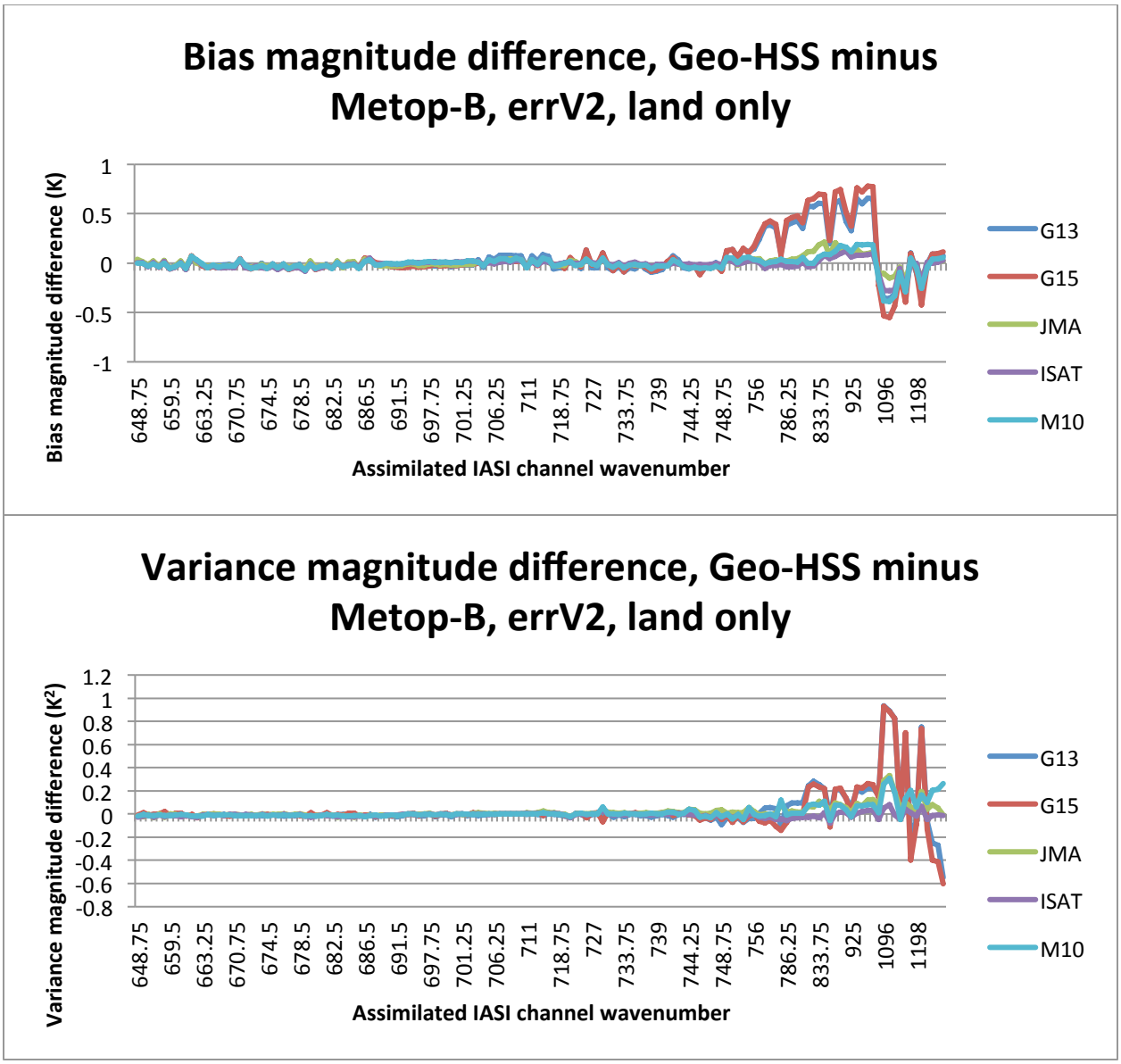
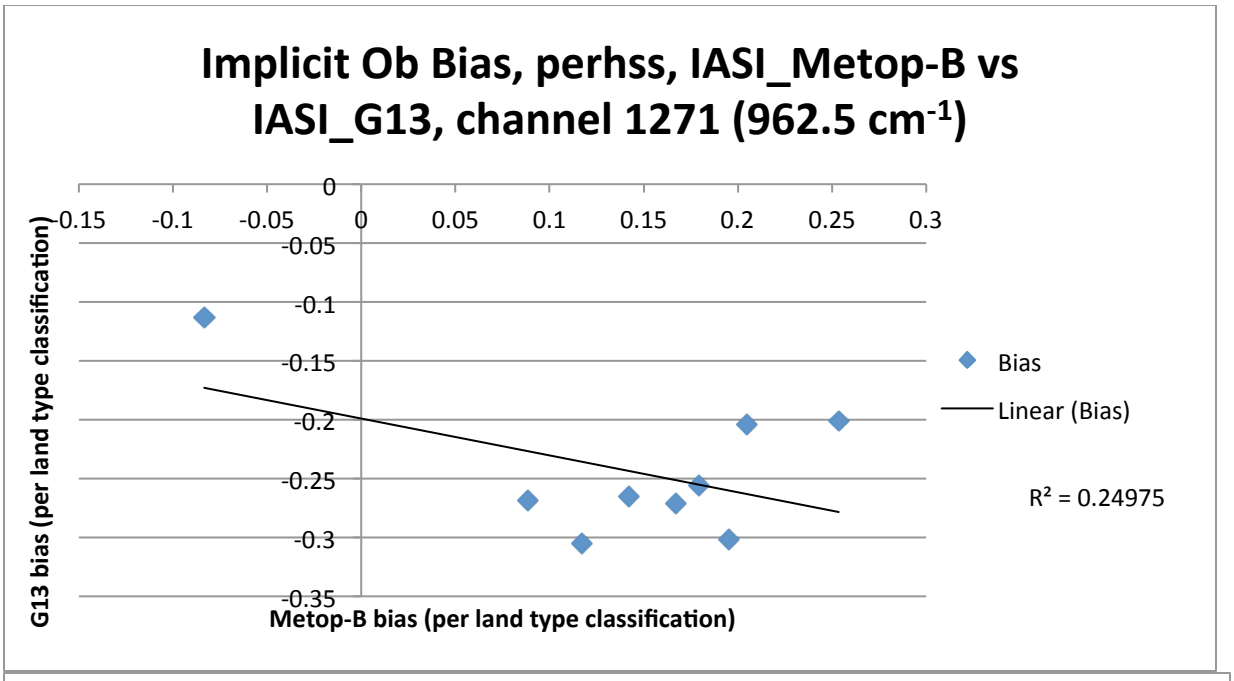



Figure 22: As in Figure 21, for land observations.



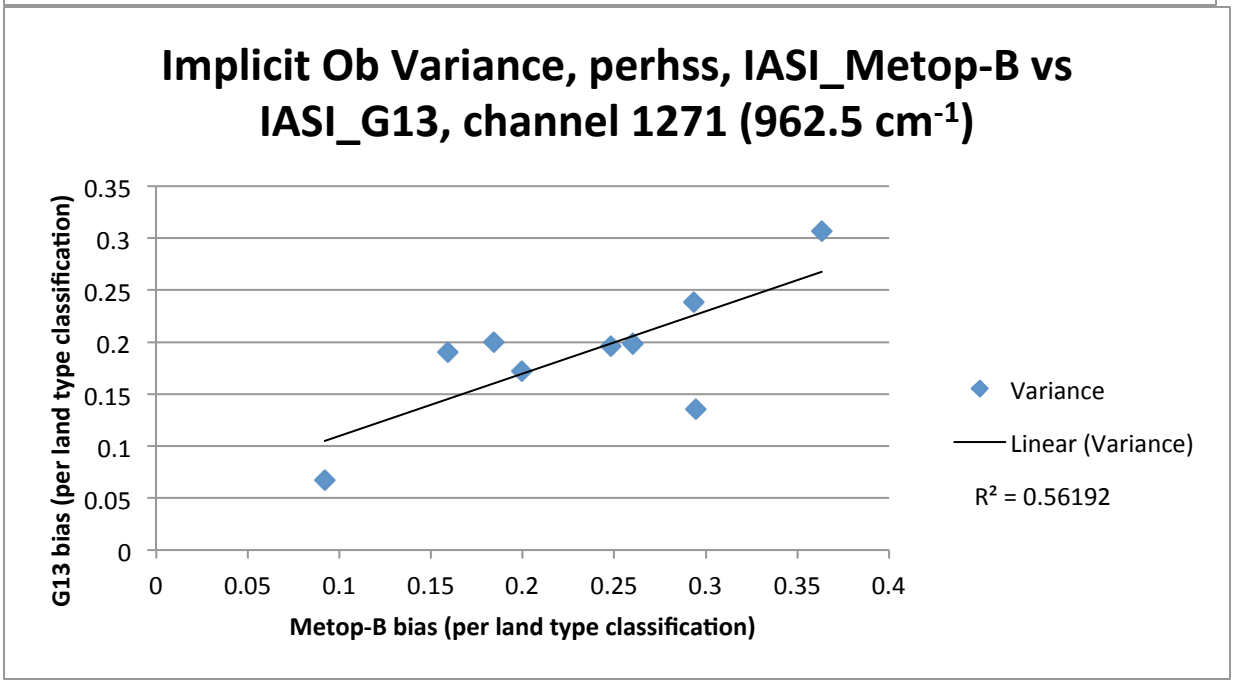




Figure 23: Implicit Observation Bias (top) and Variance (bottom) for IASI channel 1271 (962.5 $cm^{-1}$) from experiment perhss, comparing IASI_Metop-B implicit observation errors (x-axis) with IASI_G13 implicit observation errors (y-axis). Each data point represents a given land surface type (9 total classifications). Linear fit and correlation ($R^2$) are provided as well.