Forecast Impact Studies from Data Addition Experiments Using the NCEP Global Forecast System

James A. Jung [1] and Mitchell D. Goldberg [2]

[1] Cooperative Institute for Meteorological Satellite Studies (CIMSS), Madison Wisconsin

[2]NOAA/NESDIS/Joint Polar Satellite System Office, Lanham Maryland

Introduction

Addition to baseline Observing System Experiments (OSEs) are used to quantify the contributions made to forecast skill by remotely sensed satellite data. The impact is measured by comparing the analysis and forecast results of an assimilation–forecast system using a minimum of data, adding a particular observing system then comparing it to the full suite of observations. The case studies chosen consist of the time period of December 2014 – January 2015.

The National Centers for Environmental Prediction (NCEP) Global Data assimilation System / Global Forecast System (GDAS/GFS) is used for the data assimilation system and forecast model. The baseline experiment uses all of the operational conventional data available plus the Global Positioning System – Radio Occultation (GPS-RO). The experimental runs individually add data from infrared sensors; Atmospheric Infrared Sounder (AIRS) from Aqua, the Cross-track Infrared Sounder (CrIS) from Suomi NPOES Preparatory Project (SNPP) and the Infrared Atmospheric Sounding Interferometer (IASI) from Metop-b, and microwave sensors; Advanced Microwave Sounding Unit / Microwave Humidity Sensor (AMSU/MHS) from NOAA-19, the Advanced Technology Microwave Sounder (ATMS) on SNPP and the Special Sensor Microwave Imager/Sounder (SSMIS) from F18. The Control simulation uses almost all available operational satellite and conventional data.

The impact of each observing system is assessed by comparing the analyses and forecast results over extended periods. Analysis differences, anomaly correlations, and Root-Mean Square Error (RMSE) are evaluated for all experimental runs. Analysis differences of geopotential height, relative humidity and temperature are shown along with the anomaly correlation die off curves and histograms from geopotential heights during January.

Assimilation System

For these experiments, the January 2015 version of the semi-lagrangian GDAS/GFS was used at a reduced resolution from operations. A horizontal resolution of 670 spectral triangular waves (T670) was used, with a Gaussian grid of 1344 X 672, which corresponds to approximately 27 km horizontal resolution. The vertical domain ranges from the surface to 0.27 hPa and is divided into 64 unequally spaced sigma/pressure layers with enhanced resolution near the bottom and top of the model domain. There are 15 layers below 800 hPa and 24 layers above 100 hPa. Comprehensive documentation of the GFS, including any recent changes, can be found online at http://www.emc.ncep.noaa.gov/GFS/doc.php. The current Gridpoint Statistical Interpolation (GSI) analysis scheme is a three-dimensional ensemble-variational hybrid (3DEnVAR) scheme that provides the initial conditions for the GDAS/GFS from a blend of a first guess and both conventional and satellite observations (Parrish and Derber 1992, Kleist et al. 2009a, Kleist et al. 2009b). The GSI ensemble is composed of 80 members running at a reduced resolution of T256 or approximately 50 km. An Ensemble Kalman Filter (EnKF) generates the flow dependent error covariance estimates and hybrid algorithm. The GSI, with subsequent changes is documented online at http://www.emc.ncep.noaa.gov/gmb/gdas/.

The NCEP global data assimilation system consists of a first or early cycle with a T-3.0 to T+2.5 hour data cut-off window for all observations available by T+2.5 hours after synoptic time, where T indicates the analysis time, typically at the synoptic times of 00, 06, 12 and 18 UTC. In operational practice, an extended range forecast is issued from each analysis. For this study, only the 00 UTC forecasts are used out to 168 hours. The analysis process is repeated 6 hours later to provide the final analysis for the six hour forecast for the next early cycle first guess. This final analysis includes observational data that arrived after the cut-off for the early analysis. The final analysis is our best estimate of the atmosphere and in this study it was used as truth for the analysis and forecast quality assessment.

Experimental Design

Diagnostics presented here include statistics commonly used by NCEP and other NWP centers world-wide. The computation of Anomaly Correlations (AC) for forecasts produced from the GDAS/GFS are completed using code developed and maintained at NCEP. NCEP (NWS 2006) provides a description of the method of computation while Lahoz (1999) presents an overall description of what the anomaly correlation is typically used for. The fields being evaluated, which are truncated to only include spectral wave numbers 1 through 20, are limited to the zonal bands 20°-80° of each Hemisphere.

All diagnostics exclude the first 14 days of the time period. This delay in evaluating the statistics allows for the impact of the new data to be acclimated into the model initial conditions. The diagnostics presented here are for 1 – 31 January 2015. The forecast diagnostics for this paper were also terminated at 168 hours to concentrate on the shorter term forecast impacts.

Results

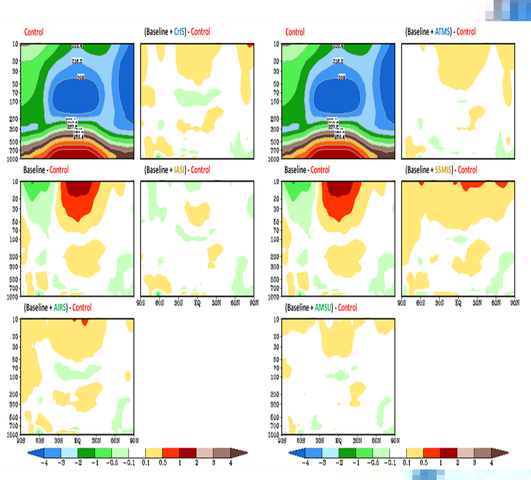

Figure 1. Average 00Z analysis latitude-height plots of temperature difference with respect to the control for the infrared sensors (AIRS, CrIS, and IASI) on the left and the microwave instruments (AMSU, ATMS and SSMIS) on the right. Larger differences with respect to the control suggest greater analysis deviations and a less accurate model initialization. Units are [m]

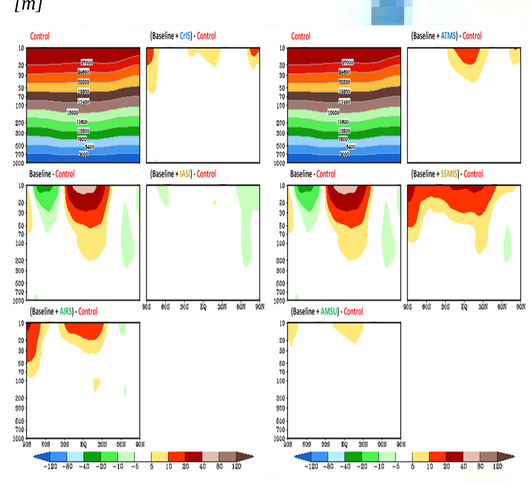

Figure 2. Average 00Z analysis latitude-height plots of geopotential height difference with respect to the control for the infrared sensors (AIRS, CrIS, and IASI) on the left and the microwave instruments (AMSU, ATMS and SSMIS) on the right. Larger differences with respect to the control suggest greater analysis deviations and a less accurate model initialization. Units are [K]

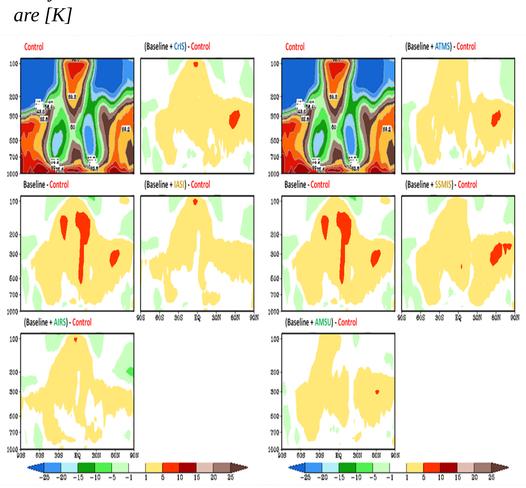

Figure 3. Average 00Z analysis latitude-height plots of relative humidity difference with respect to the control for the infrared sensors (AIRS, CrIS, and IASI) on the left and the microwave instruments (AMSU, ATMS and SSMIS) on the right. Larger differences with respect to the control suggest greater analysis deviations and a less accurate model initialization. Units are [%]

In general, the analyses will change when data are added or removed from the data assimilation system. In a cycling system such as the GDAS/GFS, these changes will evolve and often magnify over time. These changes may eventually lead to systematic biases in various fields when compared to the control fields (generated by using all available observations). Regions with minimal bias of the perturbation analysis with respect to the control analysis indicate either that the forecast model has little bias here, or that other observations are able to keep the fields from drifting away from the control.

Zonal mean analysis differences of temperature, geopotential height and relative humidity are shown in figures 1-3 respectively. These plots are the average differences spanning 1 – 31 January 2015. The experiments are divided into two groups, infrared (left) and microwave (right) only for comparison to similar instruments. The upper left is the average field of the control, the center left is the difference between the baseline experiment and the control, and are identical in each group.

The baseline analysis differences are greatest with respect to the control suggesting each of the sensors in this study improves the analysis. The average temperature is generally warmer when only the baseline data are used and the various single satellite sensors are added as shown in Figure 1. This is consistent with the height differences as shown in Figure 2. These results suggest the GDAS/GFS has a warm temperature and a higher geopotential height bias. In the infrared (left) panels, adding the IASI sensor generates the closest analysis of temperature and geopotential height fields to the control. The CrIS sensor shows less improvement with the AIRS sensor improving the analysis the least, especially in the stratosphere. In the microwave (right) panels, the addition of the AMSU sensor generates temperature and geopotential height analysis fields closest to the control while the ATMS and SSMIS show less improvement respectively. The GDAS/GFS also has a "wet" moisture bias which is indicated by the difference between the baseline experiment and the control shown in Figure 3. None of the infrared (left) panels show any significant changes in the humidity analysis. This is most likely due to the fact that AIRS is the only infrared instrument where only a few water vapor channels are used. The CrIS and IASI water vapor radiances are not used operationally in this version of the GDAS/GFS. The microwave (right) panels also show minimal improvements in the relative humidity bias. The MHS, coupled with AMSU and the water vapor channels on ATMS are being used. The water vapor channels on SSMIS are not used in this version of the GDAS/GFS.

The anomaly correlation die-off curves presented in figures 4-6 are for the Southern Hemisphere during 1-31 January 2015 and are for 250, 500 and 1000 hPa respectively. They consist of spectral waves 1-20 and are computed according to WMO standards. Each experiment is verified with respect to the control analysis. All of the panels have the Control and Baseline experiments. The left panels have the addition to the baseline of the infrared instruments (AIRS, CrIS, and IASI), the right panels have the addition to the baseline of the microwave instruments (AMSU, ATMS, SSMIS). The greater the difference between the anomaly correlation scores of the various experiments and the Baseline, the larger the impact the single satellite-instrument has on the quality of the forecast. The lower portion of each panel shows the significance tests which are computed with respect to the differences between the baseline and the various experiments. Values outside (above or below) bars of corresponding color are significant at the 95% confidence level.

The Control simulation has the highest and the baseline experiment has the lowest average geopotential height anomaly correlation scores at all forecast ranges for this season (1-31 January 2015) in the Southern Hemisphere. In general, all of the instruments show improvements in forecast skill over the baseline. The infrared instruments are about equal through the day 4 forecast. IASI tends to show the greatest improvements out to day 7. This is consistent in the three levels shown here. The ATMS addition experiment shows the greatest improvements in forecast skill through day 7 with AMSU and SSMIS providing less improvements. Also note that IASI, AIRS, and ATMS produce similar forecast improvements during this time period in the Southern Hemisphere.

Another and more in-depth look at the forecast impact of each sensor is to review the statistics of the anomaly correlations. The shape of the histogram of the anomaly correlations as shown in figure 7 can reveal; entire distribution shifts, bi-modal distributions or is influenced by a few forecasts being better/worse than the control.

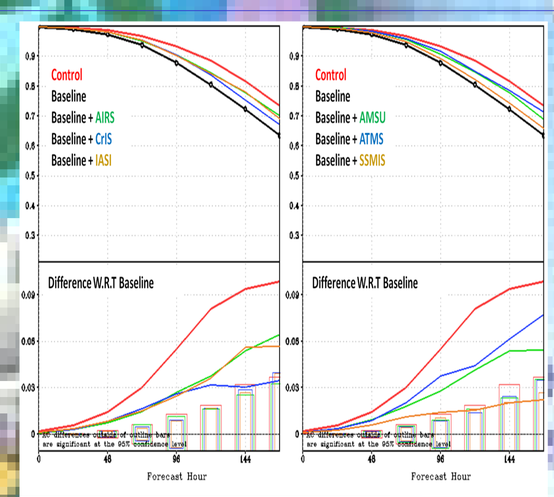

Figure 4. 250 hPa geopotential height anomaly correlations through day 7 for the Southern Hemisphere. The left panel contains the infrared sensors, the right panel contains the microwave sensors. The bottom portion of both panels is the difference between the baseline and each experiment with the statistical significance test. Lines outside (above or below) the corresponding color box are significant at the 95% confidence level.

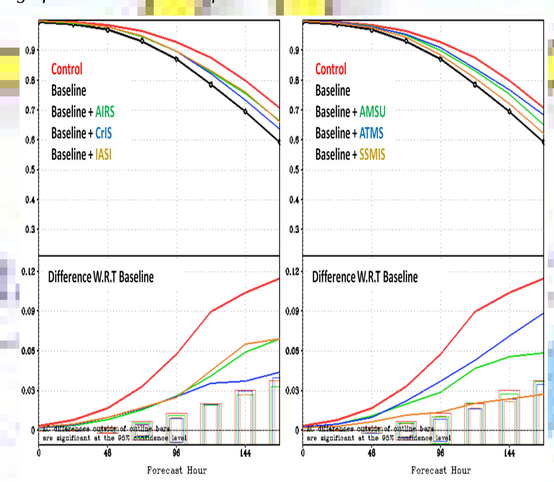

Figure 5. 500 hPa geopotential height anomaly correlations through day 7 for the Southern Hemisphere. The left panel contains the infrared sensors, the right panel contains the microwave sensors. The bottom portion of both panels is the difference between the baseline and each experiment with the statistical significance test. Lines outside (above or below) the corresponding color box are significant at the 95% confidence level.

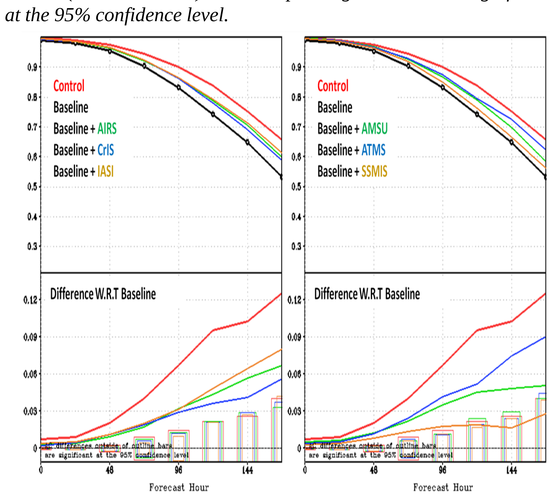

Figure 6. 1000 hPa geopotential height anomaly correlations through day 7 for the Southern Hemisphere. The left panel contains the infrared sensors, the right panel contains the microwave sensors. The bottom portion of both panels is the difference between the control and each experiment with the statistical significance test. Lines outside (above or below) the corresponding color box are significant at the 95% confidence level.

Figure 7 is the histogram of the 500 hPa day 5 Southern Hemisphere anomaly correlations for the infrared (left) and microwave (right) instruments. The infrared sensors generally have a similar histogram to the control but with a longer lower scores tail, indicating similar variability in the forecasts but with more lower scoring forecasts. For the microwave, the AMSU and ATMS have broader peaks than the control, indicating more forecast variability and with less higher scoring forecasts.

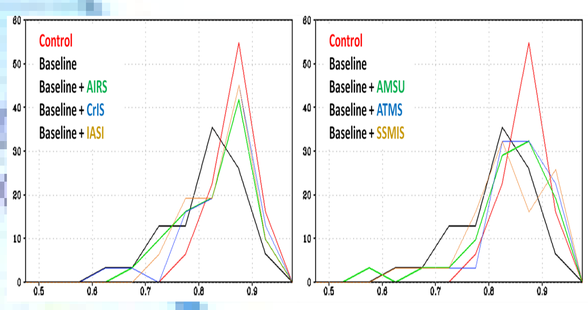

Figure 7. Histogram of 500 hPa Southern Hemisphere anomaly correlations. The left panel contains the infrared experiment, the right panel contains the microwave experiments.

Summary

Results from an addition to baseline Observing System Experiments are summarized here for the Southern Hemisphere summer, 1-31 January 2015. The baseline experiment consists of all of the conventional observations (rawinsondes, aircraft, synoptic surface, ship, buoy, and etc.) and GPS-RO. Single instruments were then added to the baseline experiment and their impact on the analysis and forecasts were quantified. Three hyperspectral infrared instruments (AIRS, CrIS and IASI) and three microwave instruments (AMSU/MHS, ATMS and SSMIS) addition experiments were conducted. Verification and forecast performance was determined with respect to a Control experiment, which contains almost all of the operational data.

The baseline experiment analysis differences are greatest with respect to the control suggesting each of the sensors improves the baseline analysis. The average temperature is generally warmer when only the baseline data are used and the various single satellite sensors are added. These results suggest the GDAS/GFS has a warm temperature and a higher geopotential height bias. Individually adding the IASI and AMSU sensor generates the closest analysis to the Control for temperature and geopotential height. The GDAS/GFS also has a "wet" moisture bias and is indicated by the difference between the baseline experiment and the Control. None of the sensors show any significant improvements to the moisture field.

The control simulation has the highest and the baseline experiment has the lowest average anomaly correlation at all forecast ranges. The addition of the IASI and AIRS sensors seem to add the most improvement to forecast skill of the infrared sensors. The ATMS generally added the most skill to the forecast base from the microwave sensors. The IASI, AIRS and AMSU add about equal skill to the forecast in the Southern Hemisphere in this study.

Acknowledgements

The authors wish to thank Andrew Collard, Daryl Kleist and Russ Treadon of NOAA/NWS/NCEP/EMC for their scientific input and assistance with the verification statistics. The study was undertaken with the Joint Center for Satellite Data Assimilation (JCSDA) computing resources. The authors also wish to thank Dave Kemeza of NASA and Scott Nolin and Jesse Stroik of UW-Madison for providing hardware/software support and maintenance of the JCSDA computers. This work was supported under NOAA grant NA10NES4400013.