CSTAR Final Report

Adaptive, High Resolution Modeling for the Arctic Test Bed at NWS Alaska

Reporting Period: 01 May 2016 - 30 April 2019

Grant number: NA16NWS4680006

PI: John Pace (University of Alaska Fairbanks)
Co-I's: Don Morton (Boreal Scientific Computing, LLC);
Jiang Zhu (UAF)

Highlights

- Completed work in development of a robust system for launching WRF clusters in the cloud and directing WRF simulation setup and execution from local machines
- Conducted two studies evaluating impact of assimilating VIIRS wind data in a numerical weather model.
- Evaluated impact of assimilating VIIRS wind data using two methods.
- Results show a noticeable impact on model analyses but a smaller impact on model forecasts.
- Documented work and prepared final reports

Satellite wind data assimilation in CSTAR final report

Overview: One of the integrated objectives of the CSTAR project is to investigate the impact of assimilation of satellite data in a weather model for short-term forecasts of Alaska region. One operational NCEP weather prediction model is the HRRR-Alaska (HA) model. In its current configuration, the HA model does not assimilate satellite VIIRS wind data. In this study, we intended to study assimilation of the satellite polar wind data in the HA model. Because the operational HA model could not run in our test domain, we built a HRRR-Alaska-Like (HAL) model for the study. The HAL model was built as close as possible to the HA model, and has the same domain and parameters as the HA model, but uses lower resolution than the operational HA model. For the purposes of this study, we did not configure the HAL model to assimilate the same observations as the operational HA model. In our experiment, we ran the HAL model 4 times a day over two monthly periods. At each analysis time, we ran the model in three modes: (1) control run: running model without any data assimilation; (2) experiment run1: running model with GDAS observational data assimilation; and (3) experiment run2: running model with GDAS observation and VIIRS wind data assimilation. The performance of data assimilation is evaluated in terms of Root Mean Square Error (RMSE) and Correlation Coefficient (CC) between model output and the RAOB data. We used two methods to compare model grid output and observations: (a) pairing model outputs with RAOB observation points with using the MET tool (Brown et al); and (b) interpolating profiles from model output to the pressure levels of RAOB profiles. We used two methods to ingest VIIRS polar wind data in the GSI data assimilation system (Hu et al). Method 1: treating the wind data as upper air observations, we appended the data into the conventional observation gdas.prepbufr file, and assimilated the files. Method 2: directly assimilating the wind data stored in gdas.satwnd.bufr_d files in the GSI system. In this project, we performed two one-month retrospective studies. First, in the study of 2015091600 to 2015101518, we used Method 1 to assimilate GDAS and VIIRS wind data, and we used point match to evaluate the performance. Second, in the study of 2018100100 to 2018103118, we assimilated GDAS and VIIRS wind data directly into the GSI (Method 2), and used profile comparison to compare performance of two data assimilation runs.

Study I: Assimilate GDAS prepbufr with appended VIIRS wind data (Method 1) and compare performance of GDAS-only and GDAS+VIIRS using RMSE of matching points

Download the wind data

We downloaded the VIIRS polar wind data (in netCDF format) files from the NASA CLASS website (https://www.avl.class.noaa.gov/saa/products/). We treated polar wind data in these files as conventional upper air wind observations (APDUPA, data type=120 in prepbufr).

Process the data

Data processing includes filtering and appending. We use the time window [analysis time-1.5 hours, analysis time+1.5 hours] to filter the files and read data from these files, applying quality controls to select valid "good" data, and appended the data to a gdas.prepbufr file. In a polar wind data file, there are three quality controls: Flag, QI, and EE (Daniels et al). Flag indicates which quality check was failed. FLAG=0 means the data passed all quality checks. QI is an index of the evaluated quality of the data. Its range is 0 to 100, with 100 being the best. EE is estimated error of the data. After evaluating various combinations of control parameters, we determined the best combination was QI>80 or EE<3.5 m/s and not using the FLAG. The python scripts to do this process include process_viirswind2gdas.py and netcdf2bufr.py (APPENDIX A 1-2).

Assimilate the data

We used the GSI system to assimilate the modified gdas.prepbufr file. In GSI system, we modify three files: run.sh; comgsi_namelist.sh; and convinfo. In the run.sh file, we link the gdas.prepbufr to ob.bufr and PREPBUFR=ob.bufr. The comgsi_namelist.sh file is for setting GSI parameters edited following the definition in section OBS_INPUT:

prepbufr	ps	null	ps	1.0	0	0
prepbufr	t	null	t	1.0	0	0
prepbufr	q	null	q	1.0	0	0
prepbufr	pw	null	pw	1.0	0	0
prepbufr	uv	null	uv	1.0	0	0
prepbufr	spd	null	spd	1.0	0	0
prepbufr	dw	null	dw	1.0	0	0
prepbufr	sst	null	sst	1.0	0	0

The convinfo file controls which types of data are assimilated in the GSI system. We set rows type=120 and iuse=1 and type=220 and iuse=1 to assimilate RAOB control variables.

Results

First, we verify that VIIRS wind data are appended into the GDAS observation files. Figure 1 shows the points included in the gdas.prepbufr and the gdas+VIIRS prepbufr files at the analysis time of 2015101300.

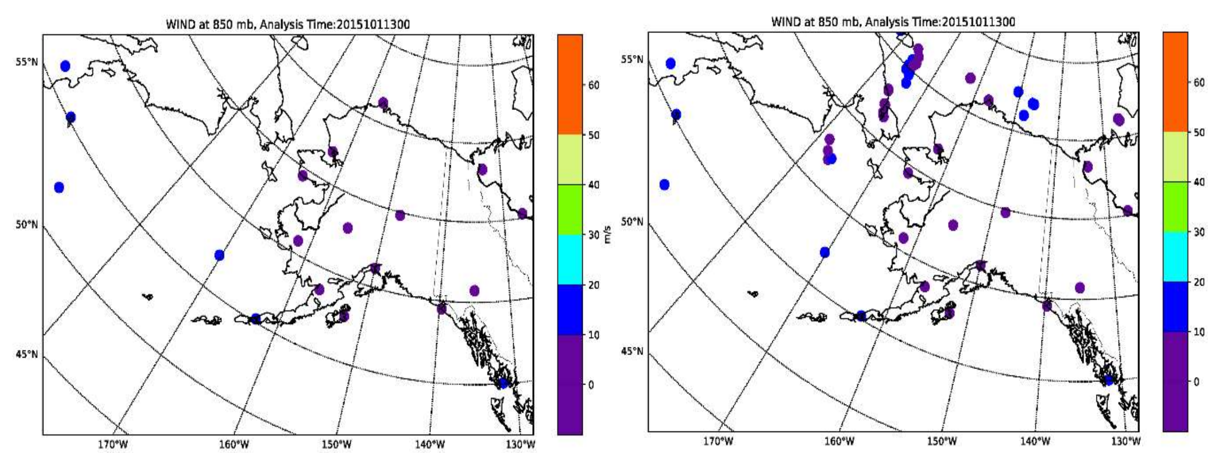

Figure 1. GDAS and GDAS+VIIRS wind observations at 850-mbar level for analysis time 2015101300. Left panel shows the wind observations in the GDAS file, and right panel shows the wind observations in GDAS+VIIRS file.

There are fewer data points in the GDAS-only set (left-hand panel) than in the GDAS+VIIRS set (right-hand panel), which verifies that VIIRS polar wind data are appended to the GDAS+VIIRS prepbufr file. These files were used in the run1 and run2 experiments, respectively.

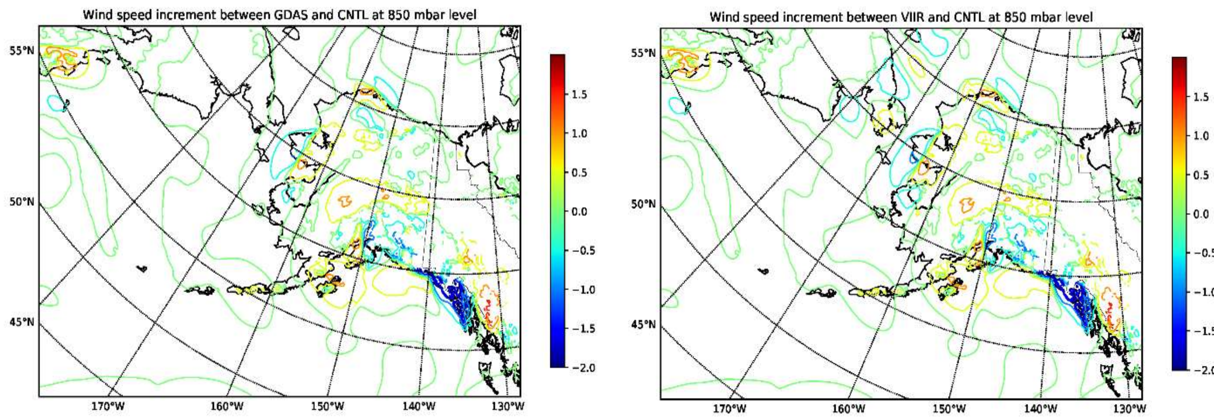

Figure 2. Analysis increments of wind at 850-mbar pressure level from experiment 1 and experiment 2.

Figure 2 shows the differences between the run1 and run2 results, compared to the control run. The left-hand panel in Figure 2 shows results with assimilation of GDAS observations, and the right-hand panel shows results with assimilation of GDAS+VIIRS observations. The increments are similar. There is a noticeable difference at the west side of the upper Bering Sea. Comparison of Figures 1 and 2 shows this difference occurred in the area where more VIIRS wind data were used. This indicates that assimilation of VIIRS wind data can indeed affect the analysis.

We paired the model output with observation points at several standard pressure levels (1000, 850, 750, 500, and 300 mbar) using the MET tool, and calculated the values of RMSE for the three data assimilation (DA) modes (control – no data assimilation; run1 – GDAS DA; run 2 – GDAS+VIIRS DA). For clarity, we calculated the <u>RMSE ratio</u> as RMSE of a DA run divided by RMSE of the control run, converted to percent. A RMSE ratio less than 100% means the performance of the DA run is better than the control run.

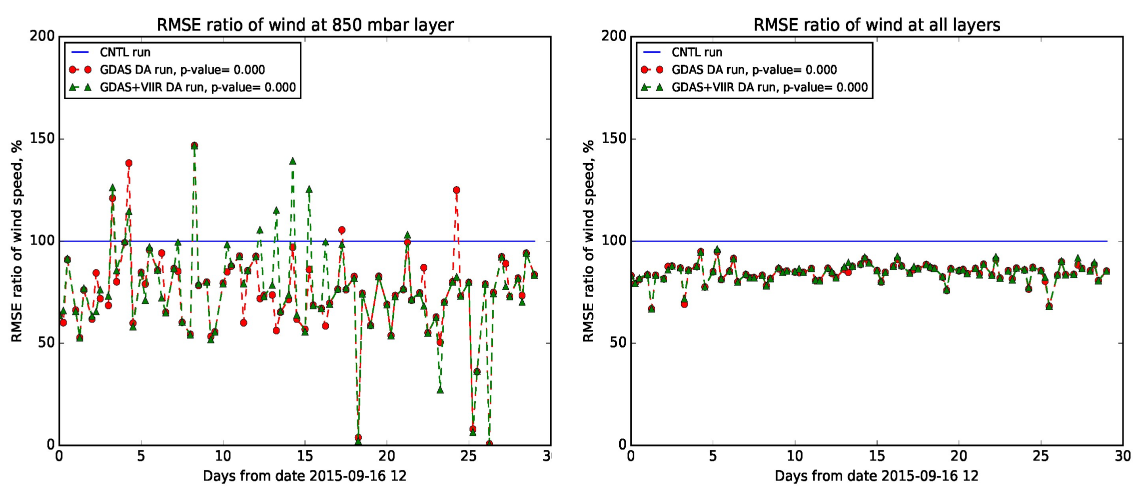

Figure 3. RMSE ratios of monthly wind analyses from control, GDAS, and GDAS+VIIRS DA runs

Results from one month of RMSE ratios of wind speed analyses produced by the three modes are shown in Figure 3. The left-hand panel shows RMSE ratios at the 850 mbar level. Some analyses from the GDAS+VIIRS DA run were much better than the GDAS DA run, and some others were worse. The majority of analyses had similar performance. The right-hand panel compares the RMSE ratios from the two DA runs over all pressure levels. The similarity of the two ratios indicates the addition of VIIRS data had only a small overall effect.

Integrating results over the entire one-month period, Figures 4 and 5 show RMSE ratios for temperature, relative humidity, and wind speed, at the 850 mb level and all pressure levels, .

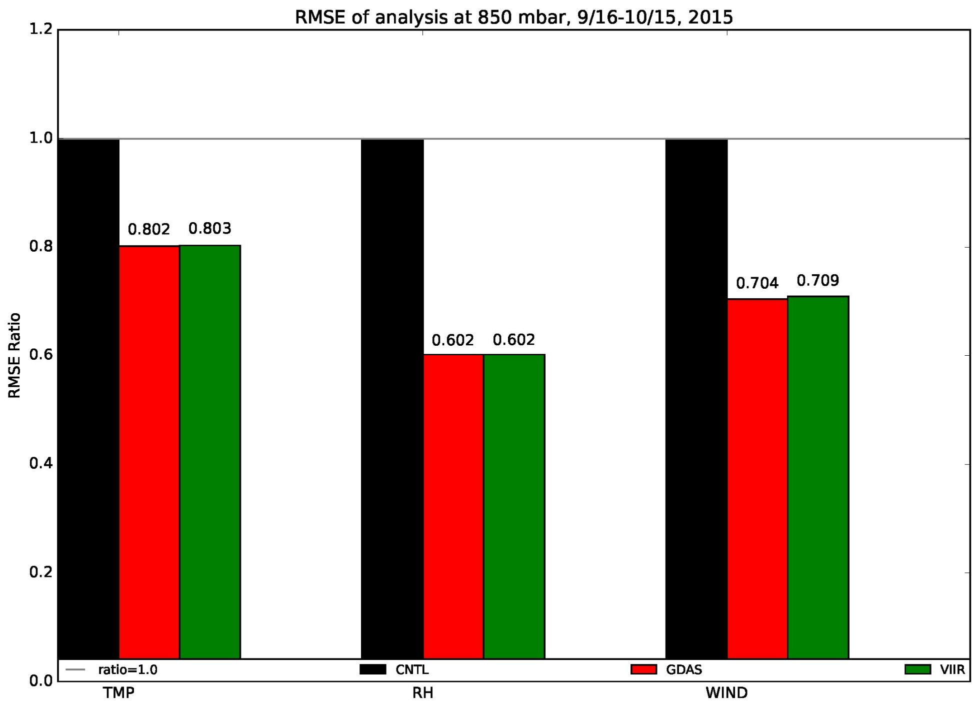

Figure 4. Monthly RMSE ratios of temperature, relative humidity, and wind speed at 850-mbar level

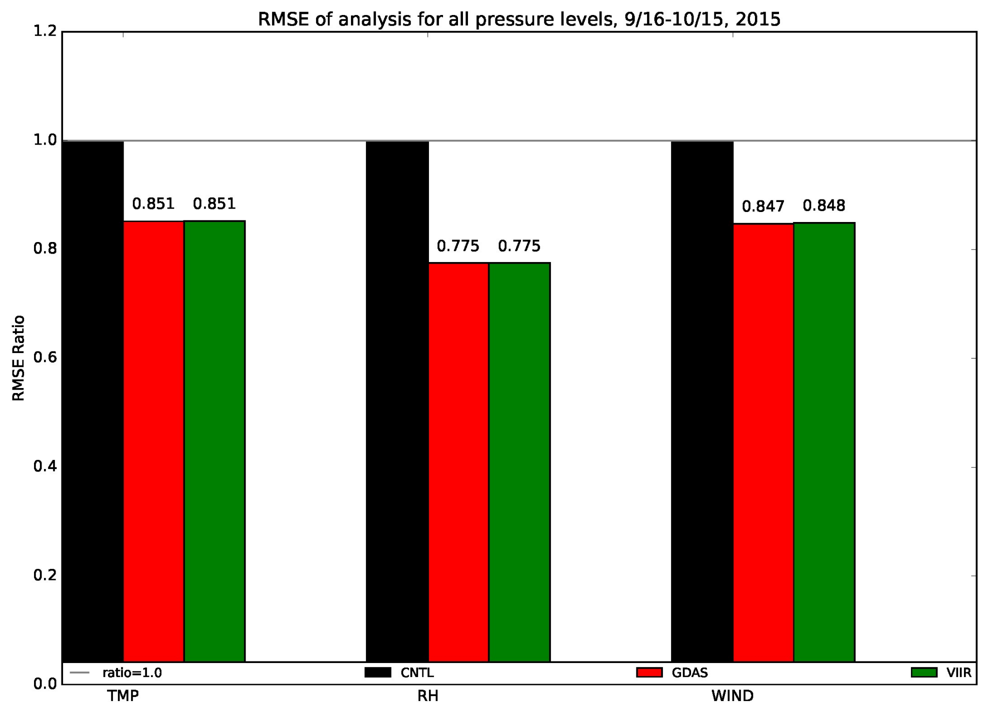

Figure 5. Monthly RMSE ratios of temperature, relative humidity, and wind speed over all pressure levels

In both figures, the black bar represents RMSE ratio of the control run, which is always equal to 1, the red bars represent RMSE ratio of experiment run1 (GDAS only), and the green bars represent RMSE ratio of experiment run2 (GDAS+VIIRS). Values of RMSE ratio of temperature, relative humidity, and wind speed from both data assimilation runs are lower than the control run in Figures 4 and 5. In Figure 4, the GDAS DA run produced a slightly better analysis at 850 mbar than the GDAS+VIIRS wind data DA run. In Figure 5, both DA runs have similar RMSE ratios.

The conclusions from Study I are that both the GDAS and GDAS+VIIRS wind DA runs improved the analysis, and that the GDAS+VIIRS wind data assimilation did not have significant improvement compared to the GDAS-only data assimilation.

Study II: Assimilate GDAS and VIIRS wind data directly in GSI (Method 2) and compare performance of GDAS-only and GDAS+VIIRS using profile comparison

Download the wind data

VIIRS wind data are included in the gdas.satwnd.bufr_d files. We downloaded the gdas.satwnd.bufr_d files from http://nomads.ncdc.noaa.gov

Process the data

Inside the file, there are one quality mark (QM) and two quality index (QI) flags associated with the data. qifn is QI values without forecast considered, and qinf is QI values with forecast considered. We modified source code read satwnd.f90 in the GSI system to read only data with qifn >= 80 into GSI, then we filtered actually assimilated data with different quality control criteria by modifying setupw.f90 in the GSI system.

Assimilate the data

In order to assimilate both gdas.prepbufr and gda.satwnd.bufr_d directly, we configured the run.sh, comgsi_namelist.sh, and convinfo files. In run.sh, we linked the gdas.prepbufr to ob.bufr and PREPBUFR=ob.bufr. In comgsi_namelist.sh, we defined observation in the OBS_INPUT part as following:

```
prepbufr     ps     null   ps      1.0   0   0
prepbufr     t      null   t       1.0   0   0
prepbufr     q      null   q       1.0   0   0
prepbufr     pw     null   pw      1.0   0   0
prepbufr     uv     null   uv      1.0   0   0
prepbufr     spd    null   spd     1.0   0   0
prepbufr     dw     null   dw      1.0   0   0
prepbufr     sst    null   sst     1.0   0   0
satwndbufr   uv     null   uv      1.0   0   0
```

We edited the coinvinfo file to include the following:

```
uv  260   0   1  3.0  0  0  0  2.5  20.1  1.4  2.5  0.000500  0  0.  0.  0  0.  0.
uv  260 224   1  3.0  0  0  0  2.5  20.1  1.4  2.5  0.000500  0  0.  0.  0  0.  0.
uv  260 225   1  3.0  0  0  0  2.5  20.1  1.4  2.5  0.000500  0  0.  0.  0  0.  0.
```

The workflow of the program to do comparison of variable profiles is as following: define two vectors to hold processed forecast and observation profiles, respectively, for each observation profile within the domain. For each observation profile, find the corresponding forecast profile that is closest to the observation location, interpolate this profile to the pressure levels of the observation profile, and append the observation profile and corresponding forecast profile respectively into two vectors, respectively. Processing all observation profiles in the domain in this way generates the observation vector and the forecast vector. We calculate the RMSE and CC between these vectors. Appendix A 3 lists the profile comparison program: read_profile_ncepbufr.py.

Results

We used this assimilation method to run one month (20181001-20181031) of 12-hour forecasts. We evaluated two data assimilation runs in terms of RMSE and CC between model output profiles and observation profiles. We compared the temperature, relative humidity, and wind speed profiles for analyses and 12-hr forecasts in this study.

The first set of results compares RMSE and CC at analysis time. Figures 6 shows the RMSE ratios (left-hand panel) and CC (right-hand panel) of temperature, relative humidity, and wind speed analysis fields over the entire one-month period.

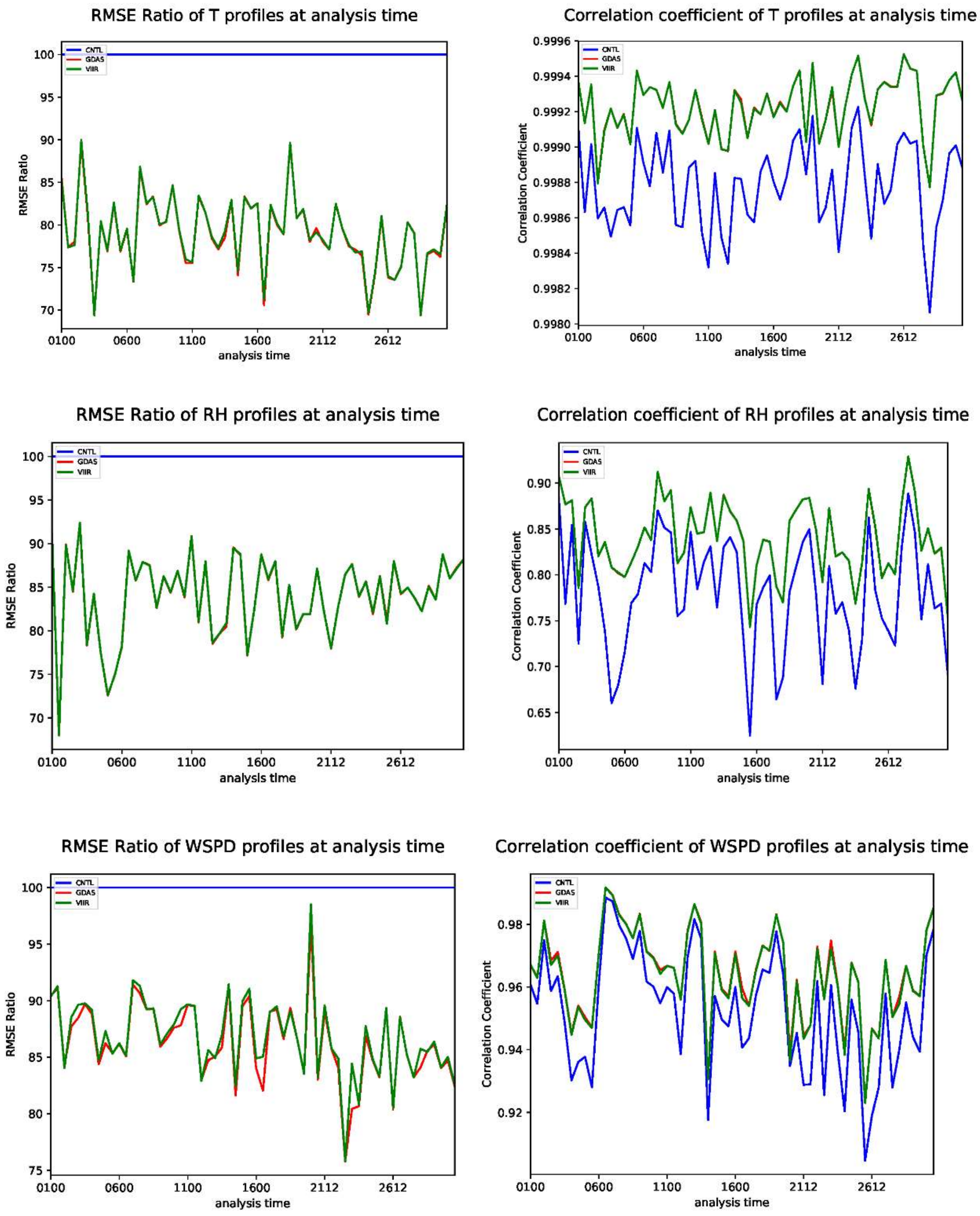

Figure 6. RMSE ratios and CC of temperature, relative humidity, and wind speed at analysis time over the month of Oct. 2018. Orange values are for run1 (GDAS alone) and green lines are for run2 (GDAS + VIIRS).

The results for run1 (orange, GDAS only) are nearly identical to results for run2 (green, GDAS+VIIRS). RMSE ratios of all three variables are less than 100% indicating both experiment runs improved the analyses. For the temperature and relative humidity results, the two data assimilation runs achieved essentially identical results. For wind speed, experiment run1 achieved a slightly better analysis than experiment run2. The right-hand panels in Figure 6 show the CC values of three variables of the analysis fields over the experiment period. The CC values of all three variables from both experiment runs are larger than those values from the control run, indicating better analysis results. There was very little difference between the two data assimilation runs.

The second set of results compares RMSE and CC of 12-hour forecasts from three runs. Figures 7 to 9 show the RMSE ratio and CC values of 12-hour forecasts for one month (2018100100 to 2018103118).

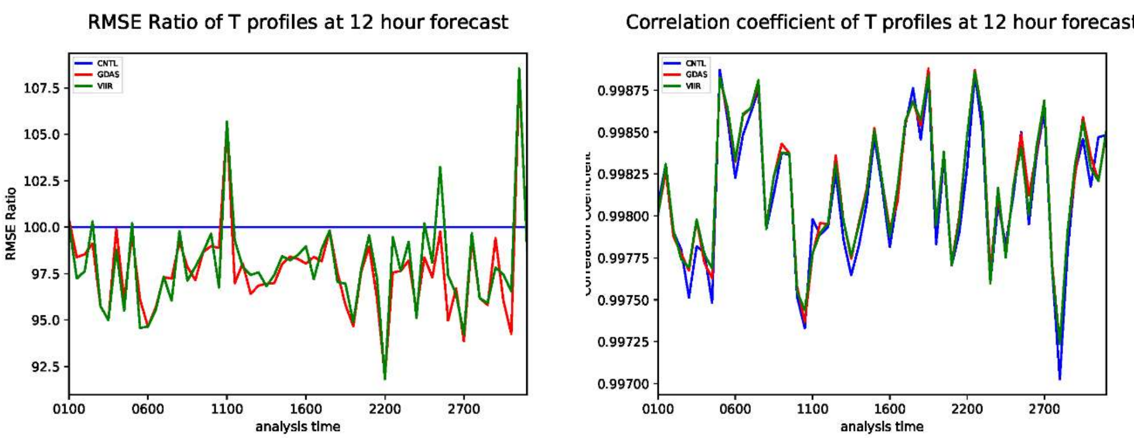

Figure 7. RMSE ratio and cc values of temperature of 12-hour forecasts from Oct.1 to 31, 2018

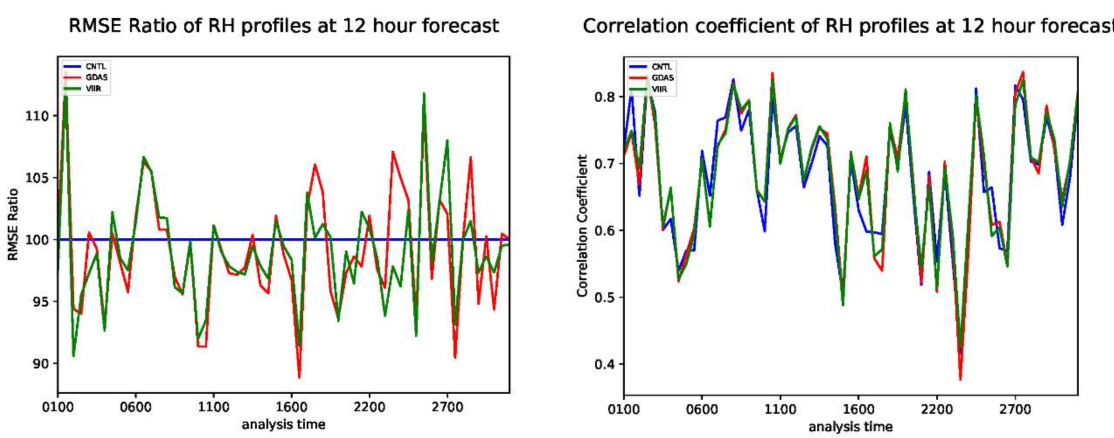

Figure 8. RMSE ratio and CC values of relative humidity of 12-hour forecasts from Oct. 1 to 31, 2018

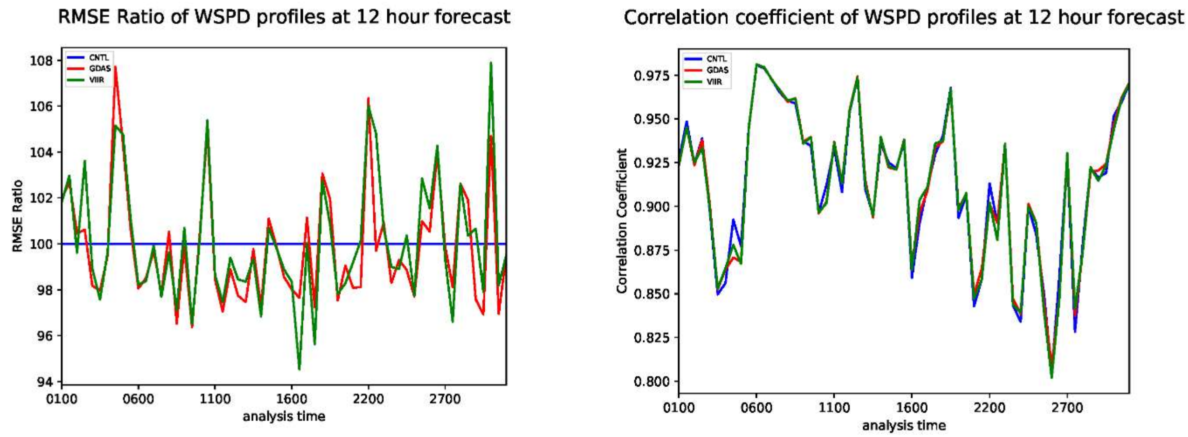

Figure 9. RMSE ratio and CC values of wind speed of 12-hour forecasts from Oct. 1 to 31, 2018

Figures 7 to 9 show most, but not all, of the RMSE ratio values of both data assimilation runs are smaller than those of the control run. The CC values of the three runs do not show much difference, indicating the better CC values from the assimilation runs noted in the analysis (Figure 6) are lost by the 12-hour forecast time. Therefore, both data assimilation runs slightly improved the 12-hour forecasts but not as much as they improved the analysis.

Figures 10 and 11 show the monthly averaged RMSE ratios and CC values at analysis time and 12-hour forecast, respectively. In both figures, the blue color represents results from the control run, the red color represents results from the GDAS observation data assimilation run (experiment run1), and the green color presents the results from GDAS observation and VIIRS wind data assimilation run (experiment run2).

Figure 10 shows the RMSE ratio values of temperature, relative humidity, and wind speed, and CC values of relative humidity and wind speed, at analysis time, are all significantly better than the control run. Figure 11 shows that most of the benefit of the two data assimilation runs is lost by the time of the 12-hr forecast.

Consistent with results in previous figures, there is little difference between the results from the two experimental results. Addition of the VIIRS wind data with GDAS data did not have much effect compared to GDAS alone.

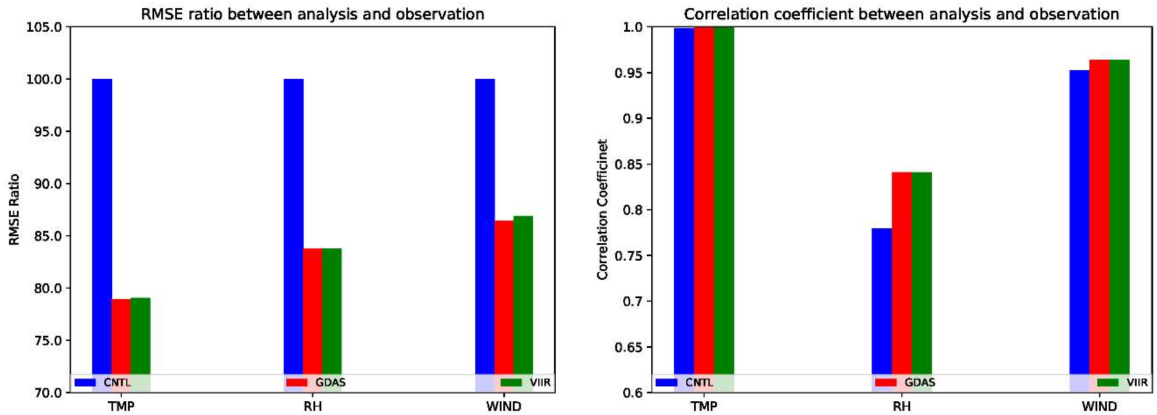

Figure 10. Averaged RMSE ratios and CC values of analysis for month of Oct., 2018

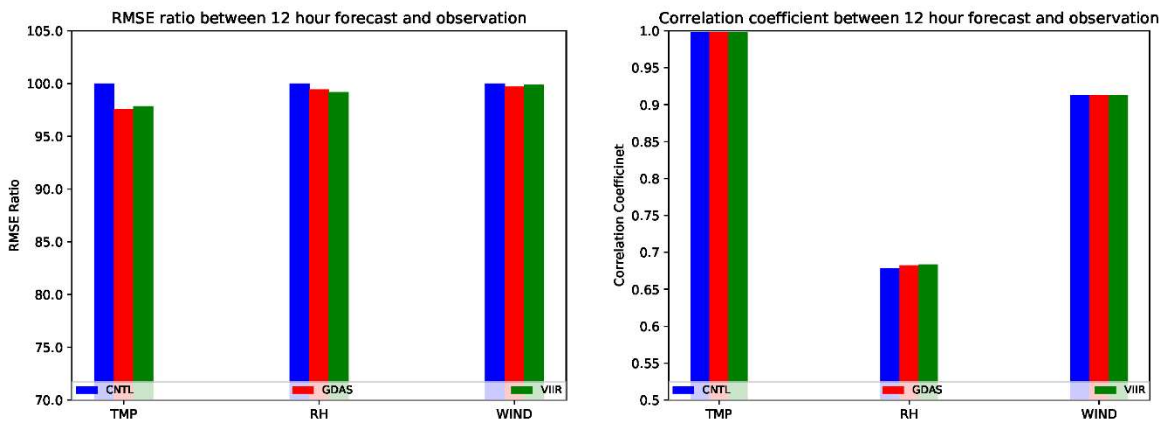

Figure 11. Averaged RMSE ratio and CC values of 12-hour forecast for month Oct., 2018.

Summary: In this project, we used two methods to assimilate VIIRS wind data into a HRRRAK-like model. We did two month-long experiments. We used the points-matching method to evaluate forecasts from the 20150916 to 20151015 period, and we used the profiles-matching method to evaluate forecasts from the 20181001 to 20181031 period. Results indicate that assimilation of conventional observation data and assimilation of observation plus VIIRS wind data both improve the analysis fields of the model, but this improvement in short-term forecasts is smaller. Assimilation of conventional observation plus VIIRS wind data assimilation does not provide significant improvement over assimilation of conventional observation only.

References

Jaime Daniels, Wayne Bresky, Jeff Key, Steve Wanzong, Andrew Bailey. NOAA VIIRS Polar Winds Product System (NVPWPS) Algorithm Theoretical Basis Document Version 1.0, June 18, 2014 (https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_VIIRS_PolarWinds_v1.0.pdf)

Brown B., R. Bullock, T. Fowler, J.H. Gotway, K. Newman, and J. Jensen, Model Evaluation Tools Version 6.0 (METv6.0) User's Guide

Daniels, J., W. Bresky, J. Key, S. Wanzong, and A. Bailey, 2014, NOAA VIIRS Polar Winds Product System(NVPWPS) Algorithm Theoretical Basis Document, ver1.0. 2014

Hu, M., G. Ge, C. Zhou, D. Stark, H. Shao, K. Newman, Jeff Beck, and X. Zhang, 2018, Grid-point Statistical Interpolation (GSI) User's Guide Version 3.7. Developmental Testbed Center. Available at http://www.dtcenter.org/com-GSI/users/docs/index.php, 149 pp.

Kim J.H. and H. Park, 2016, ANV Quality Control Method for the Next Generation Geostationary Satellite of Korea, 13th International Winds Workshop (IWW13), June 27-July 1, 2016. Monterey, CA, USA

Jiang Zhu, Don Morton, M. Hu and T. Alcott and J. Pace. 2019. Assimilation of Polar Winds data in a HRRRAK-like model, 23rd Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans and Land Surface, January 6-10, 2019, Phoenix, AZ.

Appendix A. List of program

Program name	Description
1.process_viirswind2gdas.py	Main program to call netcdf2bufr.py to combine netcdf wind data to gdas data
2.netcdf2bufr.py	Functions to implement the filtering and append of viirs wind data to gdas data
3.read_profiles_ncepbufr.py	Interpolate wrfout data according to observation profiles and calculate RMSE and CC
4.draw_profile_comparison_rmse.py	Draw line plot to show monthly RMSE ratio
5.draw_profile_comparison_cc.py	Draw line plot to show monthly CC values
6.draw_prfiles_rmse_9bars.py	Draw bar plot to show averaged RMSE ratios
7.draw_profiles_cc_9bars.py	Draw bar plot to show averaged CC