

DTC visiting program final report

**Advanced Data Assimilation Techniques Applied to a
Regional High Resolution Rapid Refresh Model in
Alaska (HRRR-Alaska)**

Jiang Zhu

Geographic Information Network of Alaska

University of Alaska Fairbanks

Reported on 08/24/2018

Introduction

Jiang Zhu accomplished his DTC visitor project with title “ Advanced Data Assimilation Techniques Applied to a Regional High Resolution Rapid Refresh Model in Alaska (HRRR-Alaska) “, which started May 2017. In this program, Jiang investigated the assimilation implementation of satellite-derived wind data into the initial field of High-Resolution Rapid Refresh for Alaska (HRRR-Alaska)-like model, evaluated the wind data assimilation impact on Alaska region weather model forecast by case study and statistical analysis. The tasks associated with Jiang’s working process are described below. Case study and statistical analysis results of wind data assimilation within HRRR-Alaska-like model were presented. Conclusions of polar satellite wind data assimilation in Alaska region model was obtained. The implementation of satellite wind assimilation and codes are presented.

Tasks

1. First DTC visit in May 2017

Jiang visited the DTC for 2 weeks in May 2017. With the supervision of Dr. Ming Hu and the help of Drs. Trevor Alcott and Chunhua Zhou, Jiang familiarized himself with the NCAR high performance computing system “Cheyenne”. He built an experimental HRRR-Alaska-like model in retrospective run mode with partial success. The model itself can successfully run without data assimilation. Due to the complexity, he did not make the data assimilation part run successfully.

2. Implemented polar orbit satellite wind data assimilation at University of Alaska

After returning to Alaska, Jiang and his colleague Dr. Don Morton continued trying to build a complete HRRR-Alaska experiment on Cheyenne. They spent extensive time, and got a lot of help from DTC personnel. Finally, they built an HRRR-Alaska-like experimental model on Cheyenne, on Amazon Web Services (AWS) cloud computing environment, and on local machine in University of Alaska. HRRR-Alaska-like model used options and namelist files similar to those used for the NOAA operational HRRR-Alaska model. The only exception was it does not assimilate as many kinds of datasets as HRRR-Alaska model.

3. Report preliminary results at 2018 AMS annual meeting (poster)

Jiang started to use customized HRRR-Alaska-like model to run 24-hour forecasts for the 30-day period 2015091600-2015101518. Jiang reported the comparison of control run and experiment run 1 (assimilation of Global Data Assimilation System (GDAS) observation plus Visible infrared Imaging Radiometer (VIIRS) wind data) results at the 2018 AMS annual meeting in the form of poster (1).

4. Case study and statistical analysis

Jiang conducted a case study with HRRR-Alaska-like model on AWS cloud computing system. Analysis date time 2018060600 was picked as the case study.

5. Second DTC visit, discuss and form the final research report.

Jiang visited DTC the week of August 20-24, 2018. He presented this work in a seminar on August 23, 2018 at NCAR's Foothills Laboratory, and delivered the final research report.

Methodology

In order to investigate satellite wind data assimilation in the HRRR-Alaska model, a HRRR-Alaska-like model is necessary. HRRR-Alaska-like model was built with options and namelist similar to that used for the operational HRRR-Alaska model. The difference is HRRR-Alaska-like model does not assimilate all of the observational data ingested into the operational HRRR-Alaska model, rather, it only assimilates satellite wind data in order to evaluate its impact on HRRR-Alaska-like model short term forecasts. The model runs in three modes. Control run does not assimilate any observational data.

Experiment run 1 assimilates GDAS observation plus VIIRS wind data, and experiment run 2 assimilates VIIRS wind data only. The model runs 4 times a day, with each run generating a 24-hour forecast. VIIRS wind data are used in the experiment. They are available at NOAA CLASS website

(<https://www.class.ncdc.noaa.gov>, JPSS VIIRS Product JPSS_NGRN level 3 data).

Data files are in NetCDF format. In order to be assimilated into model initial field, these files have to be transferred into PrepBUFR format observation files.

In GDAS PrepBUFR files, wind data are recorded as Upper-Air (RAOB, PIBAL, RECCO, DROPS) Reports (ADPUPA). A pre-process program was developed to accomplish the data format conversion. For simplicity, GSI 3D-Var analysis is

used in the data assimilation. MET tools are used to pair the observation points and grid outputs. We calculate RMSEs for the three run modes. RMSE of control run is calculated between control run output and observation data, RMSE of experiment run 1 is calculated between experiment run 1 and observation data, and RMSE of experiment run 2 is calculated between experiment run 2 output and observation data. We get the ratios of RMSEs for two experiment runs. Ratio of RMSEs of an experiment run is defined as the experiment run RMSE divided by control run RMSE. If the RMSE ratio of an experiment run is smaller than 1, it means the output of the experiment run is better than the control run. The smaller the RMSE ratio is, the better the improvement of the experiment run. By comparison of RMSE ratios of the two experiment runs, we can evaluate the performance of two different data assimilation experiments.

The key part in the method is pre-process of VIIRS wind data. The logic of the pre-process is: read GDAS PrepBUFR data, look for the ADPUPA report, get the subset name and date and time of the record. Open VIIRS wind data NetCDF file, go through each record, check if the wind data is valid by using condition of “good data”: data quality >80%, and both pressure and height values are valid; only “good” data is appended to the GDAS PrepBUFR file or write into a new PrepBUFR file. The complete Fortran code of the pre-process is included in the archived files.

Results

A case study is used to illustrate the VIIRS data distribution in the model domain. While the Alaska spatial and temporal satellite overpass coverage is greater than that for the lower 48 states, the high quality VIIRS wind data is very coarse in the domain. One of the reasons is satellite winds are mainly derived from motion of clouds, and days of cloud coverage over vast interior region is not as many as coastal region. Another possible reason is high quality requirement (quality >80%) filters out many data.

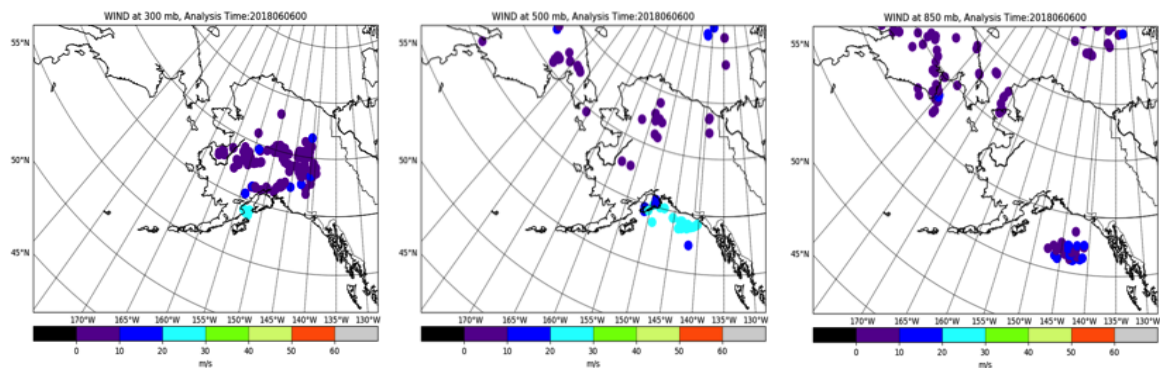


Figure 1. VIIRS wind data at three pressure levels in the experimental model domain. From left to right, panels display wind data at 300, 500, and 850 mbar pressure levels, respectively.

Figure 1 illustrates all wind data at three pressure levels. Actual numbers of data points are 395, 2121, and 1365 at 300, 500, and 850 mbar pressure levels.

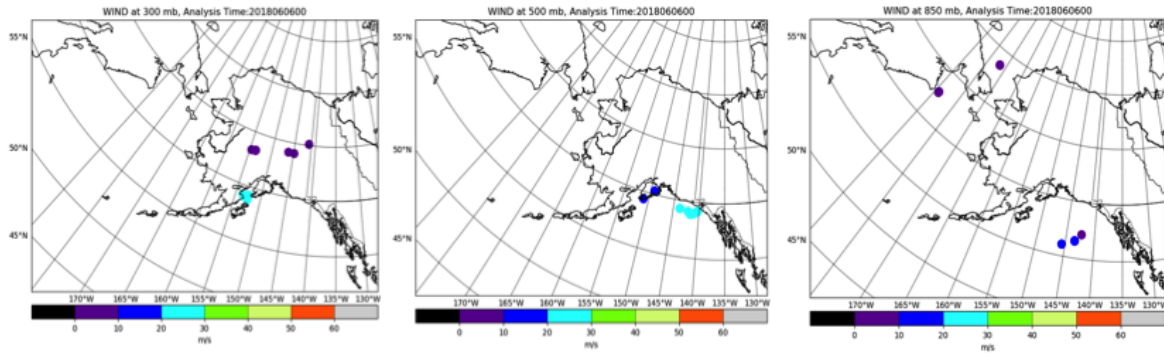


Figure 2. Distribution of VIIRS wind data with quality $\geq 80\%$ in the model domain. From left to right panels, there are wind data distributions at 300, 500, and 850 mbar pressure levels, respectively.

Figure 2 illustrates the “good” data distribution in three pressure levels. The distribution of “good” wind data is coarse. The actual number of points at 300, 500, and 850 mbar are 201, 528, and 119, respectively. The ratios of “good” data to all data at 300, 500, and 850 mbar of pressure levels are 0.51, 0.25, and 0.09. Data quality increases with height.

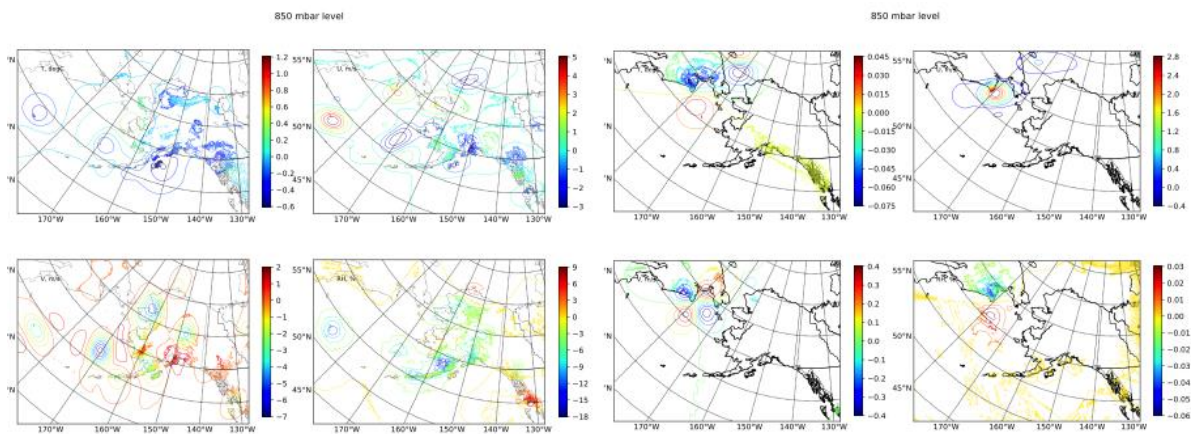


Fig.3. Analysis increments for two experiment runs at 850 mbar pressure level. The group of 4 panels in left two columns show increments from experiment run 1, and the group of 4

panels in right two columns are increments from experiment run 2. In each group, starting from upper left and going clockwise, there are T, U, V, and Q increments at 850 mbar pressure level.

Figure 3 shows the increments of two experiment runs. Left group shows that modification of initial fields by experiment run 1 is much more intensive than experiment run 2, because experiment run 1 assimilates not only wind data, but also other data from conventional observations. Right group shows the wind data above Bering Sea make the increment change. From the case study, we know that VIIRS wind data impact the initial fields. Because the VIIRS wind data are too coarse, the impact is very limited. If wind data are combined with GDAS observation, they play a positive role in improvement of initial fields.

Observation points and grid data over one month of forecast are paired by using MET tools. An average of 680 points at a given pressure level are paired. Figure 4 to 6 present RMSE ratios of analysis time at 300, 500, and 850 mbar pressure levels, respectively. As you can see, experiment run 1 improves the analysis fields of temperature, relative humidity, and wind in all three pressure levels, whereas experiment run 2 does not impact those three fields.

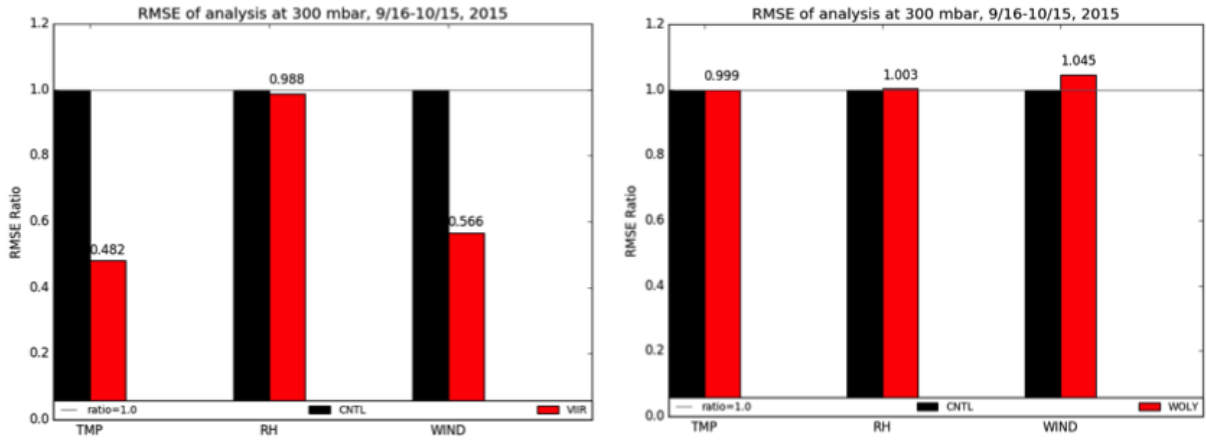


Fig. 4. RMSE ratios of two experiment runs at 300 mbar pressure level. The panel at left shows RMSE ratios of experiment run 1, and panel at right shows the ratios of experiment run 2. Black bars (ratios are always equal to 1) represent the ratios of control run. They are here as references, so that people can easily figure out how well the improvement is.

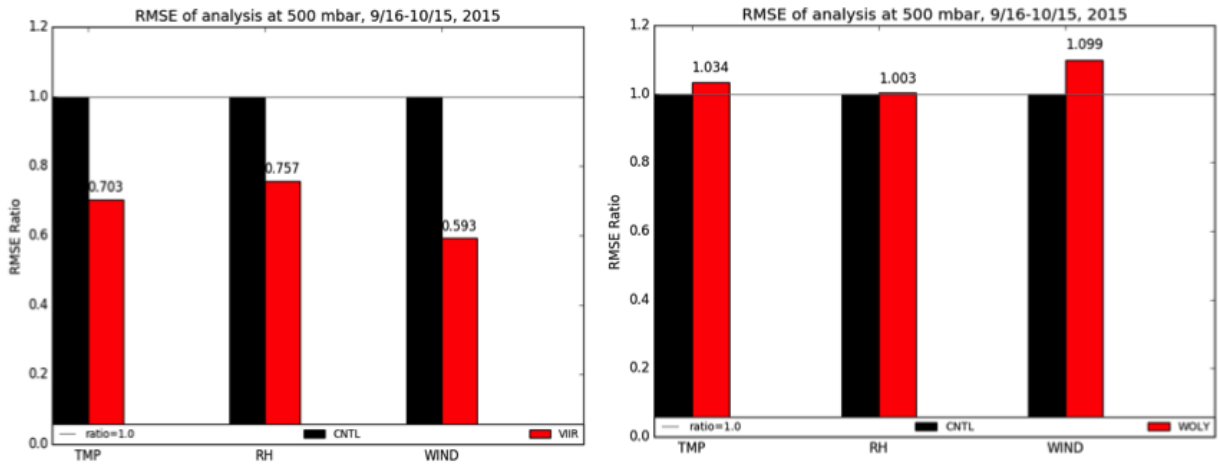


Fig. 5. RMSE ratios of two experiment runs at 500 mbar pressure level

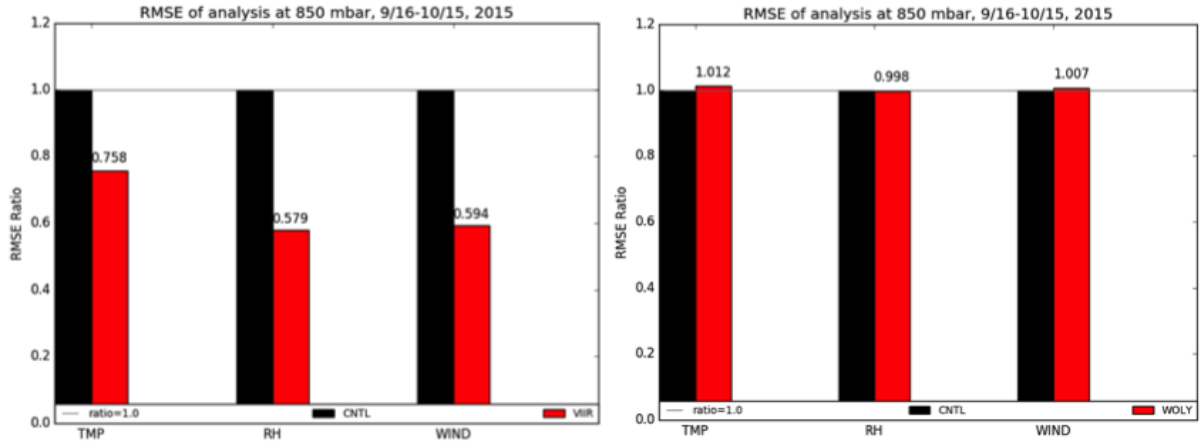


Fig.6. RMSE ratios of two experiment runs at 850 mbar pressure level

The RMSE ratios of two experiment runs at analysis time, 6, 12, 24-hour forecasts are included in Table 1.

Table 1. RMSE ratios for the two experiment runs

Forecast (hours)	P(mbar)	Experiment 1 (GDAS+VIIRS wind data)			Experiment 2 (VIIRS wind only)		
		T	RH	W	T	RH	W
0	300	0.482	0.988	0.566	0.999	1.003	1.045
	500	0.703	0.757	0.593	1.034	1.003	1.099
	850	0.758	0.579	0.594	1.012	0.998	1.007
6	300	0.902	0.985	0.993	0.987	1.007	1.042
	500	0.980	0.933	0.979	1.032	1.000	1.027
	850	0.988	0.934	0.998	1.026	1.030	1.019
12	300	0.899	1.012	0.957	0.997	1.001	1.022
	500	0.963	0.917	0.968	1.015	0.988	1.032
	850	0.960	0.933	0.986	1.015	1.106	0.998
24	300	0.952	1.011	0.985	0.995	0.998	1.017
	500	0.972	0.970	0.964	1.005	0.996	1.000
	850	0.996	0.980	0.976	0.998	1.018	0.982

RMSE ratios of experiment run 1 is analyzed first. RMSE ratios of analysis time are ~ 0.75 for all three variables at all three pressure levels. This

means experiment run 1 improves the model initial fields by ~25%. The ratios of RMSEs of wind are ~0.6 in three levels. It indicates wind field is improved most among three variables (~40%). The impact of data assimilation on forecast decrease with time. At 6-hour forecast, the ratios are ~0.95. In other words, the impact decrease to ~5%. The ratio is close to 1 in 12 and 24-hour forecasts. It tells us that the impact of data assimilation weakens in 12 and 24-hour forecasts.

RMSE ratios for experiment run 2 (VIIRS wind data only) show different results. Majority of ratios of all variables in three levels are greater than 1 in analysis, 6 to 24-hour forecasts. In other words, VIIRS wind only data assimilation does not impact the forecasts positively. Actually, assimilating very coarse data introduced “noise” to the initial fields, which make the modified initial fields even worse.

Conclusions

- Even if Alaska gets intensive polar-orbit satellites coverage, VIIRS wind data with “good” quality is too coarse in Alaska. Assimilation of the coarse data in the model initial field does not improve the model forecast.
- Conventional observational data has much stronger positive impact on model initial fields than VIIRS wind data only.

Reference

1. Jiang Zhu, Don Morton, E. Stevens, 2018: Assimilation of Polar Winds Data in a GINA-WRF model. Poster, 98th American meteorology society Annual Meeting, Jan 7-11, Austin,TX, USA
2. Jiang Zhu, 2018: Satellite Wind Data Assimilation in HRRRAK-like Model, presentation on 08/23/2018, at NCAR DTC, Boulder, CO

Acknowledgement

This research project was supported by DTC visiting program. Special thanks goes to Louisa Nance, Ming Hu, Trevor Alcott, Chunhua Zhou, and Jessica Johnson.