WAVELET SATELLITE THINNING FOR IMPROVING DATA ASSIMILATION AND FORECASTS USING TARGET OBSERVATION SCHEMES

164

Daniel L. Birkenheuer*, Yuanfu Xie, and Ning Wang[1]
NOAA/ESRL/GSD, Boulder Colorado
[1]Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Ft. Collins, CO

1. INTRODUCTION

The Forecast Applications Branch (FAB) under the Global Systems Division (GSD) at the Earth System Research Laboratory (ESRL) has been working on a project with THe Observing-system Research and Predictability EXperiment (THORPEX) to better incorporate satellite data through targeted data assimilation (TDA). Targeted data assimilation identifies a region upstream in space-time that is important to a future forecast state. Improving the data quality assimilated in this "targeted" area potentially improves the forecast. Targeted data assimilation has been proposed for a number of years (Reynolds et al. 2010, Aberson 2003). For satellite assimilation, the targeted area identifies the best region to apply full resolution satellite data. Ideally, higher satellite data density in this critical region will generate improved forecast quality in the desired area at the desired future time.

The work at ESRL/GSD/FAB focuses on two aspects of TDA. The gridded statistical interpolation (GSI) (Kleist et al. 2009) is the assimilation system of choice for the operational numerical modeling at the National Centers for Environmental Prediction (NCEP). This experiment ultimately will demonstrate whether GSI preserves high resolution satellite data in the region that is of interest upstream in space-time (target area), while maintaining nominal satellite data coverage in all other areas. The second goal of this TDA work is to put the satellite data in to a wavelet framework that is both easily stored and transmitted (compressed) that intrinsically contains high density data in the region of interest (target region) and low density information in all other locations. A GSI interface has been created to assimilate the wavelet data.

The major initiative in this paper is not different from what has been established in the fundamental idea of targeting observations; however what is unique here is the scheme of using the wavelet transform of spatial information (Yano et al. 2001, Kestin et al. 1998) to supply a targeted high resolution area imbedded inside a low resolution area. Wavelet schemes also provide another benefit. By their nature they can filter data, and in our application, remove high-frequencies in the regions of low density use, and transition in a smooth manner to high frequency in the region of high data density interest. The transform's ability to make a smooth transition between regions adds to its attractiveness in that numerical noise at the resolution boundaries is potentially reduced.

This targeted observation experiment has two components. The first of which was executed this past fiscal year. The goals of part one are to essentially:

- Run full-resolution Advanced Microwave Sounding Unit (AMSU) A and B satellite data through a comparable (30km) resolution grid to establish the "finest assimilation field" that GSI could produce
- Thin the same data set using GSI thinning to a coarser density, the density chosen was 120km (which is nominal thinning for GSI).
- Obtain and read the bufr AMSU satellite data files using GSI and perform the experiment on a Mercator grid set up over the north central Pacific Ocean.
- Compare the output of the two GSI runs to examine subjectively how and if they differ.

The above items were completed and further description continues below. Following these completed milestones, the next step in the process is to employ wavelet filtering to enable a new "revised" bufr file to be generated for GSI ingest. This new file would contain high resolution data in a selected area that would be the typical "targeted" area, for a targeted forecast test. The remainder of the satellite data field would be at coarser resolution. Wavelets would be the key in inserting the high resolution data inside the coarse area with a filtered transition for the bounding area surrounding the high density data. The subsequent step takes the wavelet filtered data and generates a new bufr data file that is both smaller in overall size, because it contains mostly coarse data, and it would offers high-resolution data in the region simulating the "targeted" area. The test would be to see what GSI does with this new bufr data. Will GSI retain the high density information where inserted? Will GSI output suffer from noise, potentially induced by edge affects in the boundary region around the full-resolution data? If a smooth transition can be achieved, then we have confidence that there is a way to both target an area for full-resolution data density and make the satellite data more efficient by assimilating thinned data elsewhere. Typically it is too expensive for GSI to assimilate full-resolution data routinely, and data thinning is the nominal approach in GSI satellite data assimilation to solve this problem.

2. MAJOR PROGRESS DISCUSSION

The current work culminated in the generation of full and thinned resolution GSI runs (Weather Research & Forecast model (WRF) network Common Data Form (netCDF) output fields) using AMSU A and B satellite data over the Pacific. The thinned resolution was done at 120km through the GSI thinning parameters that are part of the GSI system. The grid used was 30km Mercator and this matched in principle the 30km AMSU data at full

---

* *Corresponding author address:*
Daniel Birkenheuer, NOAA/ERL/GSD/FAB, GSD7,
325 Broadway, Boulder, CO 80305;
e-mail: daniel.l.birkenheuer@noaa.gov.

resolution. The full resolution (not thinned) AMSU data was run through the same Mercator Pacific grid (30km). Both runs generated output that graphed comparisons to the background field. The thinned GSI run demonstrated both weaker gradients and reduced amplitude in both maximum and minimum regions. Thus, the full resolution GSI fields had more detail and highlights after assimilation. The following figures detail the observations described above.

These results demonstrate that indeed the GSI thinning does reduce the robustness of the fields in some cases while possibly inducing spurious features in other areas (namely the U wind component, Fig. 2). In addition, results differ sufficiently that a wavelet-modified area at full resolution within a thinned field should become evident in GSI comparison to background as showing more detail in the embedded full-resolution region. Thus, we have sufficient confidence that this work can continue and is on a track that can show impact to assimilation. Testing can also commence to see whether the wavelet approach to embedding full-resolution data in the field causes GSI to generate and noise around the region in which the higher resolution data was injected.

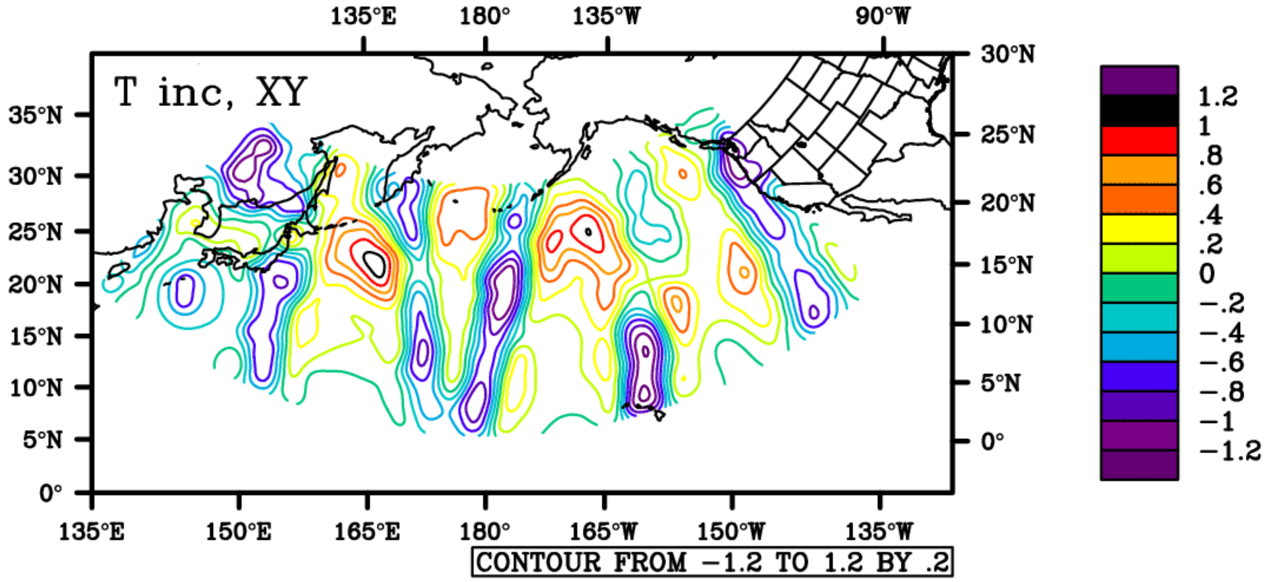
3. FIGURES AND OBSERVATIONS

The following figures show the results achieved in this test. It will be easier to show the thinned result alongside the full-resolution result. Four fields are shown as output by standard GSI tools. These fields are differences induced by assimilating satellite data at the different resolutions compared to the initial background state. Fields GSI typically generated are T (temperature), V (latitudinal wind component), U (longitudinal wind component), and Q (moisture in specific humidity). As you will see the satellite data do not affect moisture as much as the other fields, this is likely due to there being only two channels that GSI currently assimilates from AMSU B (the moisture instrument). The other moisture channels apparently have covariance error characteristics that eliminate them from assimilation consideration. The reader ***should be careful*** in examining all of the plots shown since the GSI plotting routines appear to select contour colors that are not related to consistent value. So in the two plots shown here the top plot has (for example) black assigned to +1K, while in b) this same color is assigned +0.5K. So direct comparison of colors is not possible and care should be used when reading these plots. This especially is evident in the region over China where colors in both plots are very dissimilar, but when examined, the corresponding values are roughly the same in magnitude in the plotted contours.

The GSI plotting routines do not currently support a Mercator projection. However, we chose this projection for our assimilation in the WRF model framework. As a result, we chose to have the data plotted here by GSI in a Lambert coordinate system. This is why the data appear to be curved in this projection. Therefore, it is important to understand that even though these data are plotted in a Lambert reference frame, the grid remains Mercator.

3.1 *Temperature (K)*

(A) 120km thinned



(B) 30km full-resolution

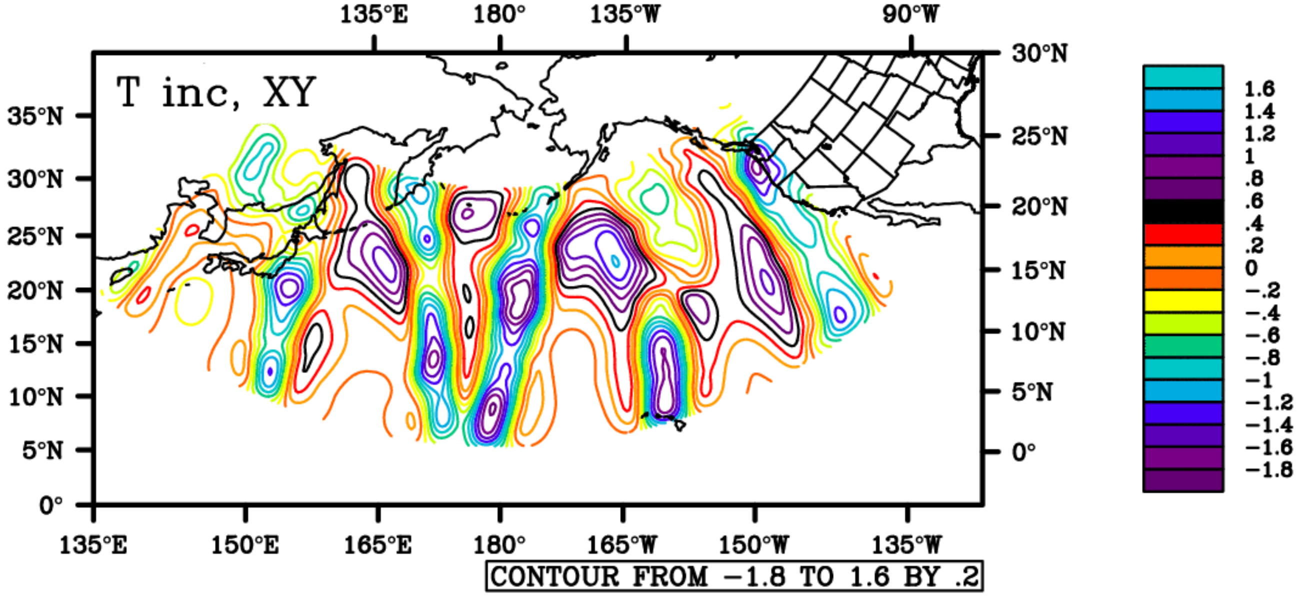
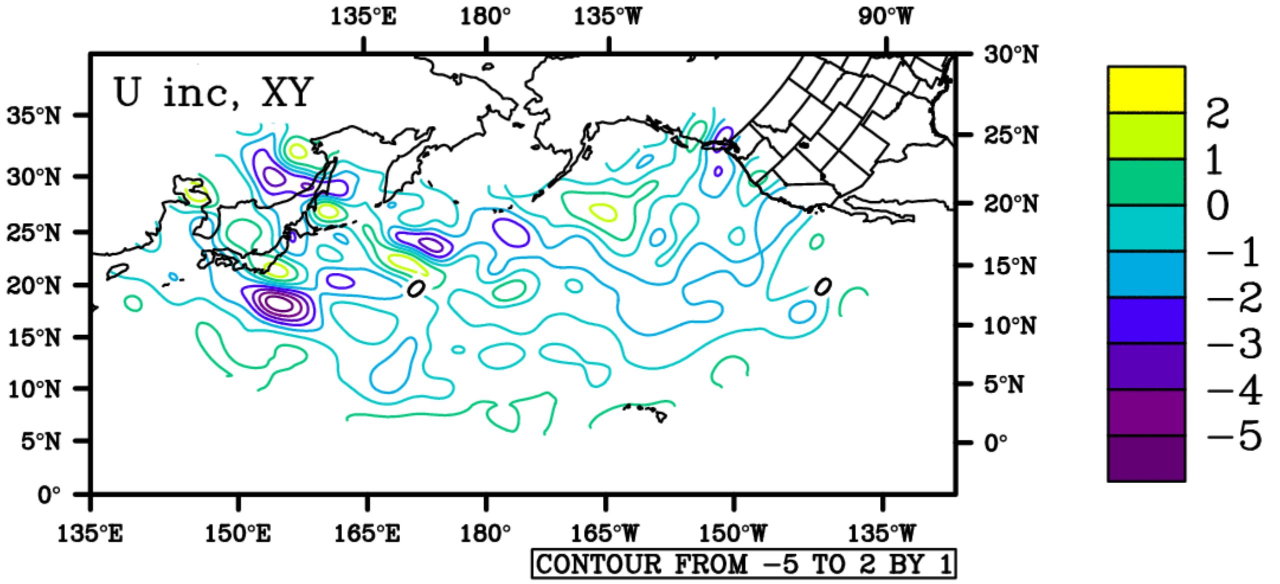


Fig 1. Temperature comparisons in a) and b) show that the thinned (a) field misses the full range of feature maxima and minima. The (b) field extends from -1.8 to +1.6 while the (a) field extends from -1.2 to +1.2 K. Also evident in the full-resolution assimilation are stronger gradients, one is particularly evident directly off of the CONUS west coast, another at about 135E.

It should be noted that AMSU A data are all thermal sensors and this satellite data assimilated case is by far the most directly impacted by thermal radiance data. Of all of the fields shown in this report, this is perhaps the most significant since it relates directly to the thermal atmospheric properties sensed. The following wind component impacts are really secondary analysis adjustments to the assimilated thermal field in Fig. 1. The final Q fields are based only on 2 satellite sensors (as explained above) and these are situated high in the atmosphere and as will be seen have minor impact on the result. But there are differences between thinned and full-resolution fields.

3.2 *U wind component (m/s)*

(A) 120km thinned assimilation



(B) 30km full-resolution

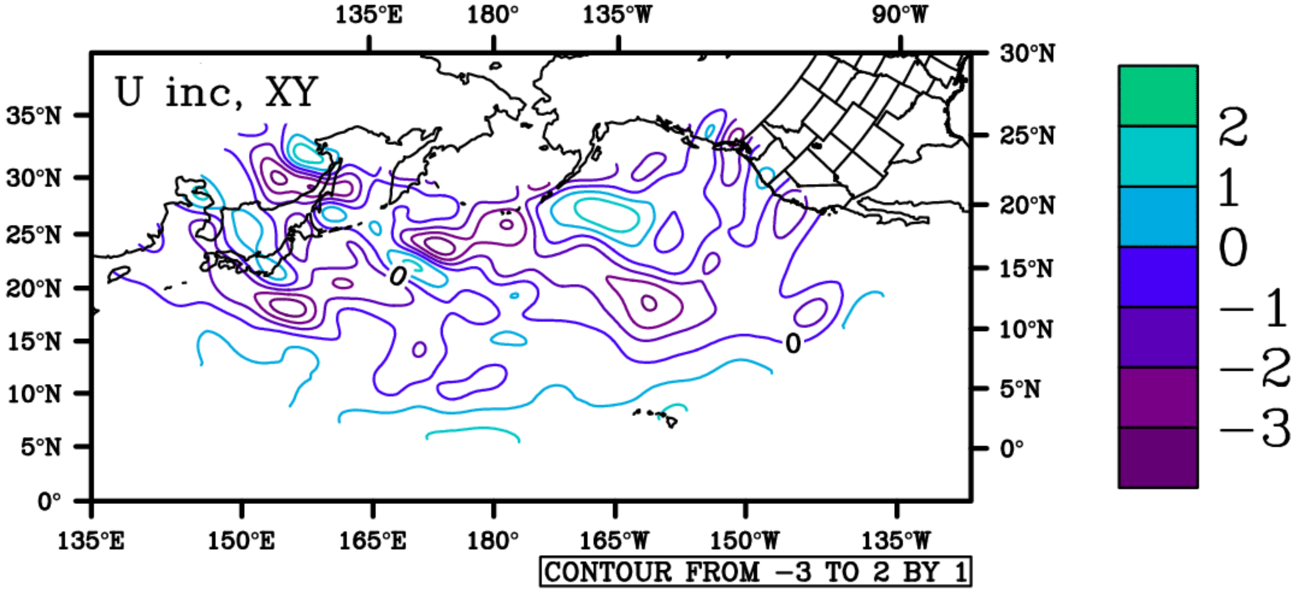
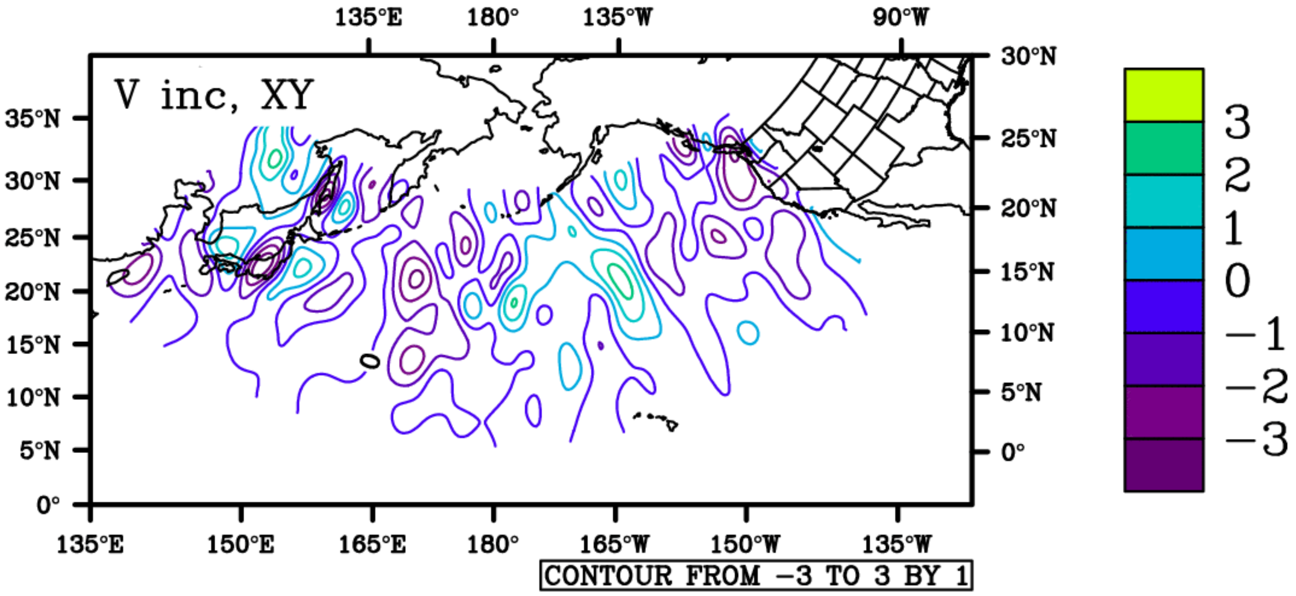


Fig. 2. (A) shows the thinned U component range: -5.0 to +2.0 m/s and (b) showing a smaller range: -3.0 to +2.0 m/s. Again, this is an inferred effect of the thermal assimilation and its effect primarily significant to an ensuing model run that would be initialized from this field. In this case, it appears quite possible that the thinned data set could be providing spurious high values. Not only that, there appear to be places where the change in the U component is of opposite sign over the full-resolution assimilation. If this is true, data thinning of satellite information could have negative effects in adding information to the solution. Indeed, this makes a strong case for higher resolution assimilation if not overall, certainly in target, and vital regions for good forecast generation.

3.3 *V wind component (m/s)*

(A) 120km thinned assimilation



(B) 30km full-resolution

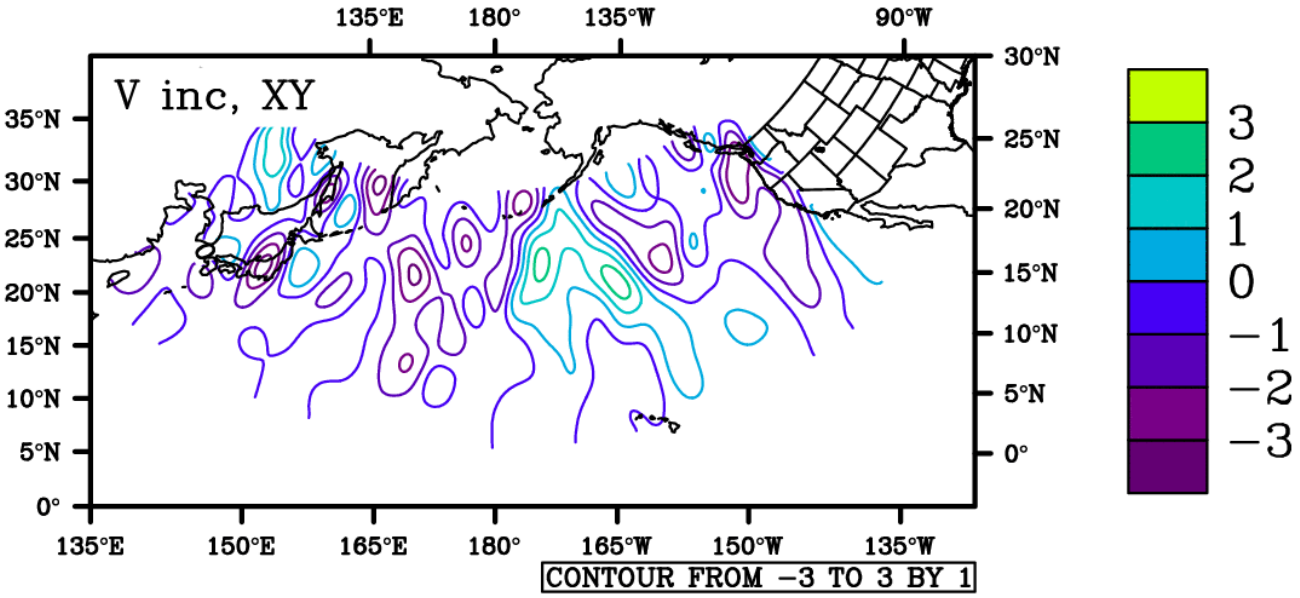
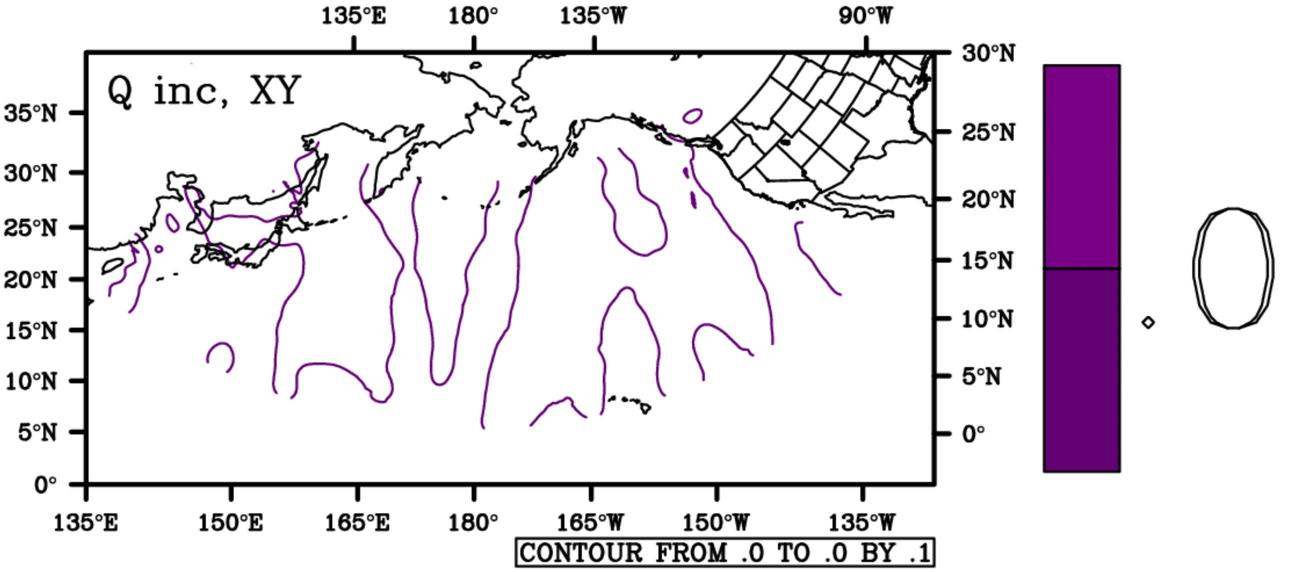


Fig. 3. The impact to the V wind field. Again this is a secondary effect, the ranges here are about equal (-3.0 to +3.0 m/s). In this case the contour colors are identical. We see that there are very pronounced differences in the central Pacific with a stronger wind gradient near the dateline in the full-resolution field. The differences over land should also be discounted since some of the AMSU satellite sensors are not effective over land due to microwave emissivity problems. GSI should take this into account, but we have no way to make sure this is really occurring. (The area specifically pointed out in regard to this potential problem is over central China and Japan. The apparent "added structure" in the thinned data set here is strongly speculated to be spurious. The increase in structural detail over the central Pacific is deemed genuine.

3.4 *Q (moisture, kg/m^3)*

(A) 120km thinned



(B) 30km full-resolution

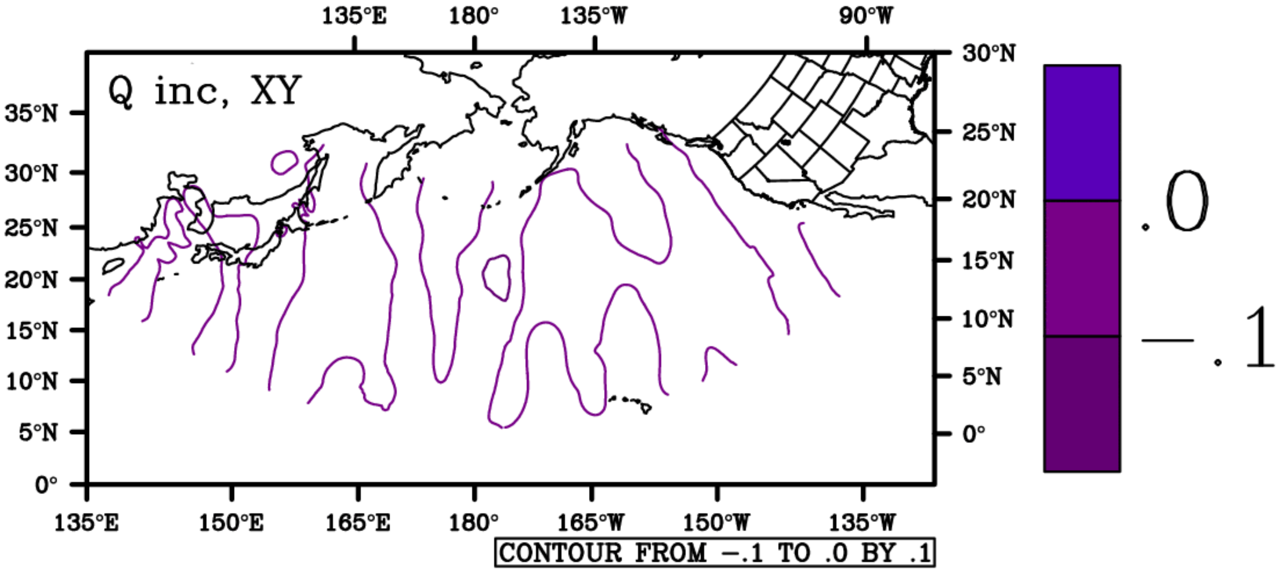


Fig. 4. The specific humidity field amplitude is small in both cases ranging from -0.1 to +0.1 in the full resolution solution with a high dynamic range than the thinned (0.0 to +0.1). We also see more structure in the contours especially in the central Pacific where there exists a closed contour in the full resolution plot that does not occur in the thinned data set.

4. SUMMARY AND CONCLUSIONS

This work has shown that the GSI satellite data thinning does result in significant differences in the background field versus using the data in full-resolution form. Which assimilation is better cannot be inferred from this work to date, but this work does allow us to move to future steps that can answer that question because we observe greater detail in resulting analyses. Ultimately, we will seek to see if the higher resolution features are preserved in a targeted region that is incorporated within thinned data. This full resolution experiment can also serve to evaluate whether the targeted regions by an NCEP targeted observation scheme are sufficient. Thus, the experiments in the past year serve the THORPEX project goals.

The basic framework to show that full-resolution data is measurably different than thinned one now exists and will allow advancing to the next step. With the framework to study wavelet filtering established, full-resolution regions within a coarser input satellite data file can be fairly compared in future assimilations to this experiment.

The next step builds on these completed milestones to take the same satellite AMSU data and run it through the wavelet to bufr system that will render a new bufr data file with coarse data everywhere except where full-resolution data are desired (over a targeted region). Comparison to the assimilation differences presented here will reveal the effectiveness of inserting full-resolution data into the GSI using the wavelet technique.

5. REFERENCES

Aberson, S. D., 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. Mon. Wea. Rev.,**131**, 1613–1628

Kestin, T. S., D. J. Karoly, J. I. Yano, and N. A. Rayner, 1998: Time-frequency variability of ENSO and stochastic simulations. *J. Climate,***11,** 2258–2272

Kleist, Daryl T., David F. Parrish, John C. Derber, Russ Treadon, Wan-Shu Wu, Stephen Lord, 2009: Introduction of the gsi into the ncep global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705.
doi: http://dx.doi.org/10.1175/2009WAF2222201.1

Reynolds, Carolyn A., James D. Doyle, Richard M. Hodur, Hao Jin, 2010: Naval research laboratory multiscale targeting guidance for t-parc and tcs-08. *Wea. Forecasting*, **25**, 526–544.
doi: http://dx.doi.org/10.1175/2009WAF2222292.1

Yano, Jun-Ichi, Mitchell W. Moncrieff, Xiaoqing Wu, Michio Yamada, 2001: Wavelet analysis of simulated tropical convective cloud systems. part 1: basic analysis. *J. Atmos. Sci.*, **58**, 850–867.
doi:http://dx.doi.org/10.1175/1520-0469(2001)058<0850:WAOSTC>2.0.CO;2