MODIS- and AVHRR-derived Polar Winds Experiments
using the NCEP GDAS/GFS

NA10NES4400011

Year 1 Second-half Progress Report
December 2010 to May 2011

28 June 2010

David Santek, PI

Proposed Work

Atmospheric Motion Vectors (AMV) are routinely generated from geostationary and polar orbiting satellites and they are incorporated into most global numerical weather prediction models throughout the world. However, advances to the AMV derivation process together with changes to assimilation systems and forecast models requires the strategies for use of the satellite-derived winds to be continually evaluated.

The focus of our proposal will be in three areas using AMVs generated from polar orbiting satellite data: (1) Quality control and thinning using the Expected Error; (2) Experiments assimilating polar winds derived from Advanced Very High Resolution Radiometer (AVHRR) images; and, (3) Experiments designed to simulate winds from the Visible/Infrared Imager/Radiometer Suite (VIIRS) instrument onboard the future NPOESS Preparatory Project (NPP) and National Polar-Orbiting Operational Environmental Satellite System (NPOESS) satellites [now restructured as the Joint Polar Satellite System (JPSS)] .

Year 1 plans from the proposal:
We propose to investigate using the EE for quality control and work with NCEP on its potential use as a quality control parameter. As time permits, we will investigate additional characteristics of the winds (such as clear sky vs. cloudy water vapor vectors and height assignment technique) to include in the determination of the EE thresholds. With guidance from NCEP personnel, all of these experiments will use the NCEP Global Data Assimilation System/Global Forecast System (GDAS/GFS). The resulting quality control procedure will be transitioned into NCEP operations at the end of the first year.

End of Year One Progress

During the second six months of this project, these four areas were addressed:

1. Gain an understanding of the current satellite AMV quality control
2. Modify GSI code to access the EE value and screen the satellite winds
3. Run control and experiments using the EE
4. Analyze observation and forecast impact
5. Investigate importing AVHRR polar winds

Satellite AMV quality control

The quality control system for satellite-derived winds is currently based on a set of three criteria; a wind observation is rejected if:

1. The pressure level of the observation is within 50 hPa of the tropopause.
2. The difference between the observed and background zonal *or* meridional flow is greater than a threshold value (qcU=qcV=7 ms^{-1})[1].
3. The observation is within 200 hPa of the surface *and* appears over land or ice.

While criteria (1) and (3) are practical considerations concerning the trustworthiness of satellite-derived wind observations near the tropopause or the Earth's surface, criterion (2) is meant only to reject observations when they disagree with the model background by more than a defined amount. Since the OMB is simultaneously a measure of observation quality *and* a measure of the potential innovation an observation can impose on the analysis, such a QC criteria suffers from being forced to reject what could be the most useful observations that remove large errors from the analysis, only because the (accurate) observations disagree too greatly with the model background.

The new QC criteria seeks to make use of the Expected Error, a measure of wind speed error that is derived from a regression using traditional Quality Indicators (QI) which compare characteristics of observations to each other and to the operational forecast, wind speed, and wind and temperature shear that constitute a measure of the synoptic-scale environment (Le Marshall et al. 2004). The Expected Error is largely decoupled from the wind speed OMB and serves as a measure of error that is more independent from the potential innovation than the OMB metric.

Modify GSI code

Changes were made to the routines read_prepbufr.f90 and setupw.f90 to extract the Expected Error value from the PREPBUFR file and store it in an unused element in the cdata_all array. The EE is stored in the word used for the StationID, cdata_all(8), which is normally not used for satellite winds. Appendix A provides details of the code changes, which includes debug and diagnostic code that we continue to use to evaluate and analyze the results.

Run experiments

We ran control and EE threshold experiments using the Moderate Resolution Imaging Spectroradiometer (MODIS) AMVs for the dates:
24 August – 01 October 2010
01 January – 15 February 2011
24 February – 30 April 2011

A first attempt is made to find a suitable threshold of Expected Error that can replace criterion (2), while leaving the other QC criteria in place. We used an Australian Bureau of Meteorology recommended value of eliminating winds where the EE > 5 ms^{-1} (LeMarshall et al. 2004). In order to retain higher speed winds, which are usually assigned a high EE value, we additionally required the EE to be larger than 0.1*speed before discarding. Therefore, criterion (2) becomes:

2. The expected error is greater than 5 ms^{-1} *and* the expected error is greater than 10% of the observed wind speed.

---

[1] qcU = qcV = (ObsSpd + 15)/3 (IR wind within 200 hPa of surface OR WV wind below 400 hPa) AND (GuessSpd +15)/3 < qcU

Observation impact:

Differences between the two model runs in terms of accepted observations were examined over a 10-day period: 10-19 September 2010. There were 2.5 million vectors; in the control 800,000 were accepted, while in the experiment only 200,000 passed the EE threshold criteria. However, the observation minus background (OMB) and observation minus analysis (OMA) were very similar for the control and experiment for the u- and v-components (Table 1)

Table 1a: U-component OMB and OMA statistics of the control and EE experiment for 10-19 September 2010. All quantities are in ms^{-1}.

	Control (Mean)	Control (StdDev)	EE experiment (Mean)	EE experiment (StdDev)
OMB	-0.1	2.5	-0.1	2.2
OMA	0.0	2.2	0.0	1.9

Table 1b: V-component OMB and OMA statistics of the control and EE experiment for 10-19 September 2010.

	Control (Mean)	Control (StdDev)	EE experiment (Mean)	EE experiment (StdDev)
OMB	-0.1	2.6	-0.1	2.3
OMA	0.0	2.2	0.0	1.9

These results are encouraging since using the EE provides a more quantitative screening of the data, whereas the previous quality control method discarded polar wind observations if either the u- or v- component deviated from the background by more than 7 ms^{-1}.

The relationship between wind speed OMB and the Expected Error is not particularly strong, but conforms to expectations with low OMB values correlated to low Expected Error (Figure 1).

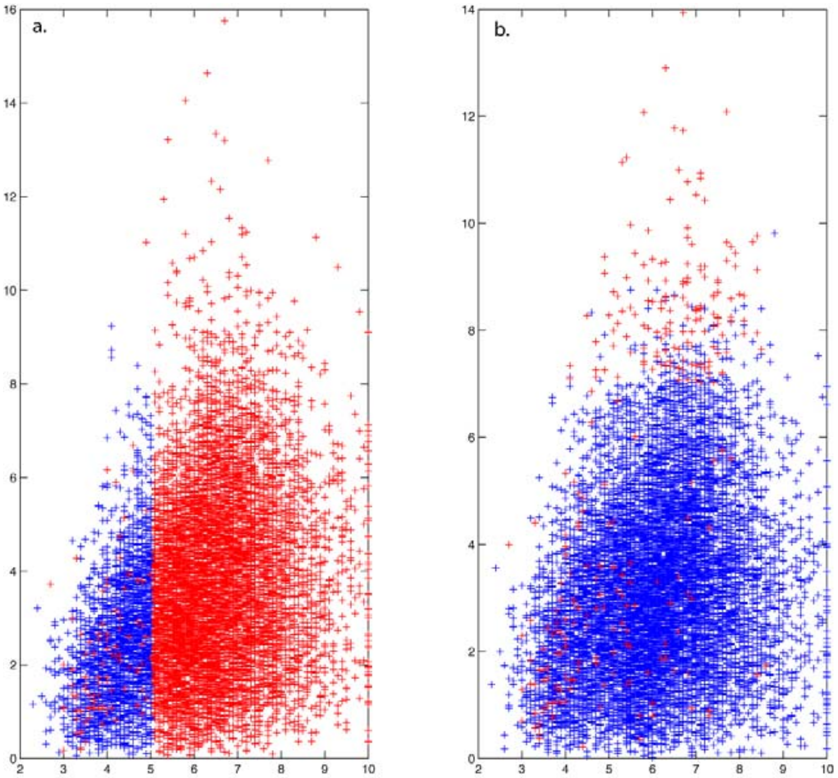

Figure 1. Scatter plot of OMB wind speed (ordinate) and expected error (abscissa) of MODIS satellite winds for 0000 UTC 21 September 2010. Red (blue) points are rejected (accepted) observations in the (a) experiment and (b) control.

While low expected error typically translates to low OMB wind speed, the reverse is not necessarily true; a high expected error can have an incredibly wide variance in OMB wind speed. Provided that the expected error can be thought of as a measure of wind speed error largely independent of OMB wind speed, Figure 1 illustrates that a low OMB wind speed does not necessarily mean that an observation should be trusted. While the old QC procedure (Fig. 1b) removed very few observations around the OMB > 7 ms^{-1} threshold, the new QC procedure (Fig. 1a) drastically reduces the number of accepted observations to only those with a low expected error.

Forecast impact

a) Anomaly correlation scores

Our expectation is that these modifications to polar winds will have a largely neutral impact on anomaly correlation (AC) scores computed globally or over a hemisphere, except for mid-range (5-7 day) forecasts in which polar or near-polar features play an important role. As such, we expect impact from these modifications to only be reflected in certain Day-5 to Day-7 forecasts, and remain neutral otherwise. It is also expected that these modifications will have a greater impact in the southern hemisphere, where satellite winds typically have a larger impact on the analysis (Zapotocny et al. 2007). A time-series of Day-5 AC scores reveals similar characteristics (Figure 2).

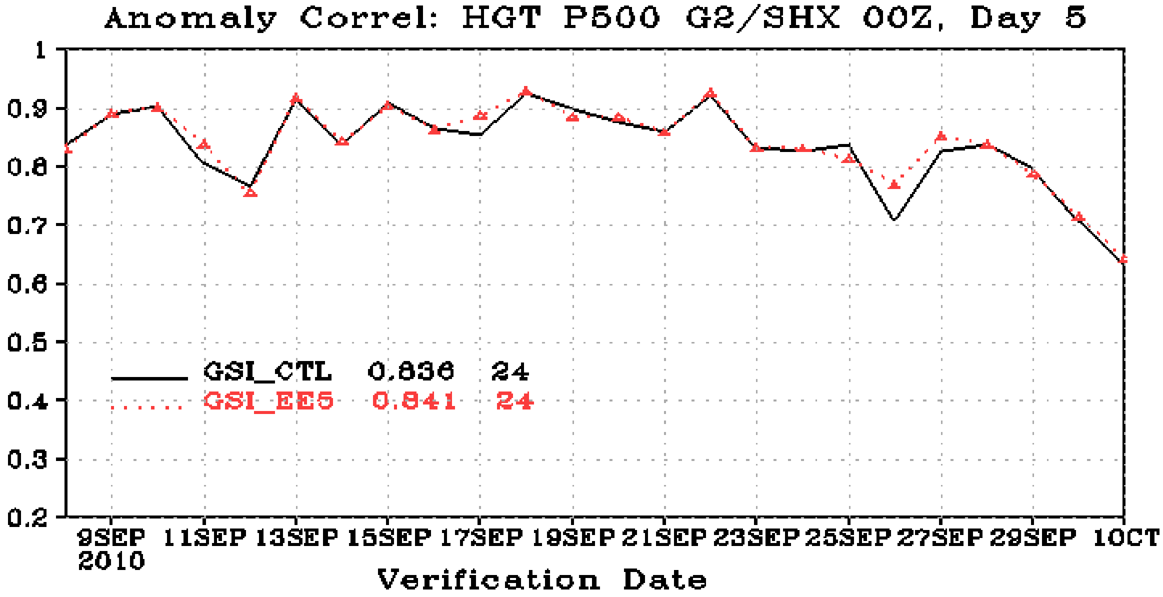

Figure 2. Day-5 anomaly correlation scores for 24 forecasts between 08 September 2010 and 01 October 2010. Scores are computed for 500 hPa geopotential heights over the southern hemisphere (20S-80S) for the control (black) and the experiment (red).

For the month of September 2010, the only significant impact occurs on the 26th, when the control simulation nearly busts with a Day-5 AC value of almost 0.7, while the experiment sees significant improvement. It is interesting to note that an even worse forecast is made for 01 October, though the experiment and control are practically identical. Both the 26 September and 01 October events are poorly forecast by the control run, while the EE experiment appears to improve only the 26 September event. One can assume that the discrepancy between these two events can be traced back to one or both of the following:

1. Differences in the synoptic evolution of both forecasts that allows for communication of improved initial conditions at polar latitudes to impact one event but not the other.
2. Differences in the amount of information provided by MODIS wind observations in the initial conditions of both forecasts.

i) Synoptic evolution: RMSE growth

Tracking the growth of 500 hPa geopotential height RMSE through 120 hours of the forecast reveals some fundamental differences between the two events (Figure 3).

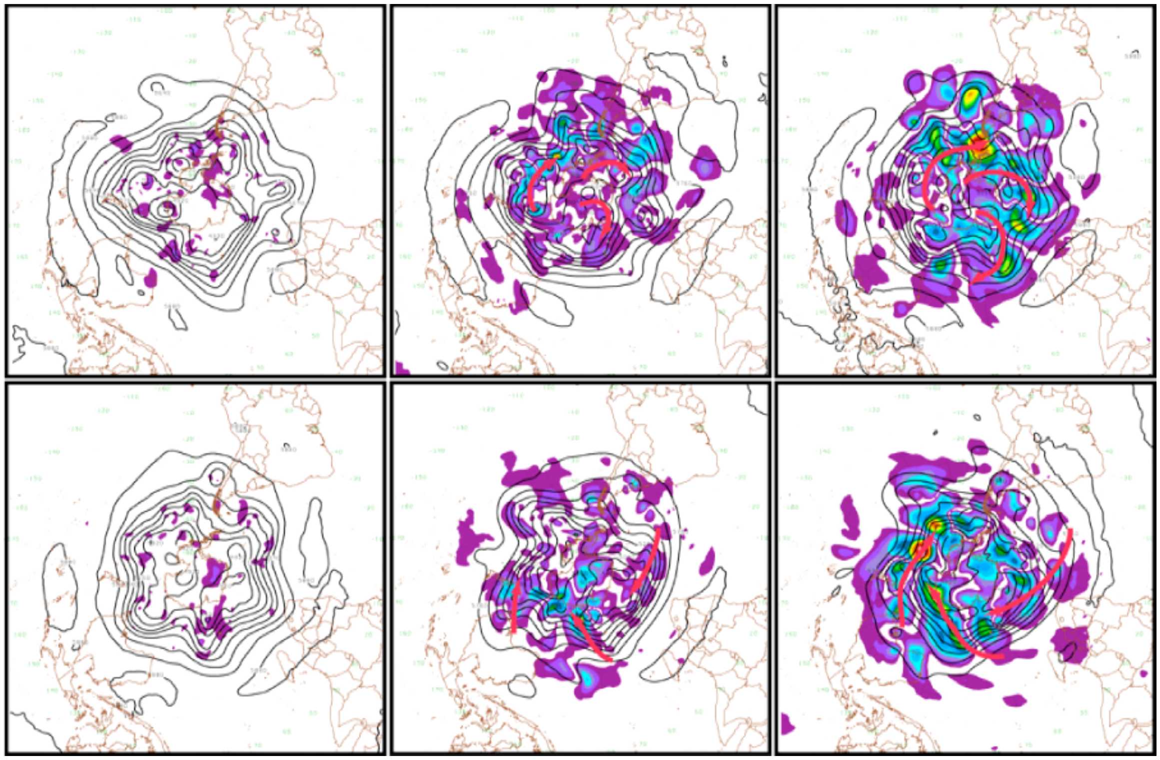

Figure 3. RMSE of 500 hPa geopotential heights at 24 hours (left), 72 hours (middle), and 120 hours (right) for 26 September 2010 (top) and 01 October 2010 (bottom) cases. Geopotential height contoured every 120 m and RMSE relative to (control) verifying analysis shaded every 20 m. Arrows indicate the direction of error propagation.

Errors in the 26 September event begin at high latitudes and propagate out into the mid-latitudes as they grow; improvement of initial conditions at polar latitudes is communicated out through mid-latitude wave activity and results in a significant improvement in Day-5 AC scores, and a reduction in RMSE of the meridional portion of the flow is in fact largely relegated to the equatorward branch of mid-latitude wave activity (not shown). In contrast, 500 hPa height RMSE in the 01 October case starts in mid-latitude troughs and works its way *toward* the pole while it grows; flow at polar latitudes plays a significantly reduced role in the evolution of synoptic-scale features, and likewise we should not assume that modifications to observations at polar latitudes should improve this case.

ii) Initial condition differences

A comparison of OMB/EE scatter plots for the analyses of both forecasts shows some limited differences, especially in the distribution of Expected Error (Figure 4).

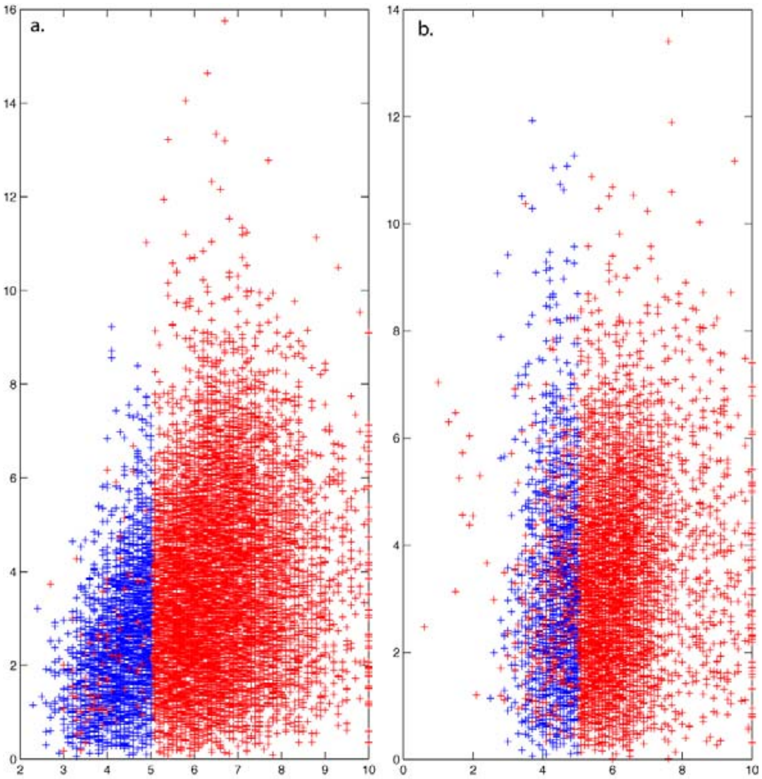

Figure 4. OMB wind speed (ordinate) versus Expected Error (abscissa) for MODIS wind observations assimilated at 0000 UTC on (a) 21 September 2010 and (b) 26 September 2010.

The initial conditions of the 26 September forecast include assimilation of 11,072 MODIS wind observations; the distribution of Expected Error includes a tapering toward low expected error at low OMB wind speeds, and a fairly large number of poor (EE>7.5 ms^{-1}) observations that were getting into the analysis using the original QC. On the other hand, MODIS winds only provide 6,976 observations to the initial conditions of the 01 October event, with a distribution of expected error clustered around a mean value of 5.7 ms^{-1}. There are many fewer poor observations that need to be scrubbed from the analysis by the improved QC.

The differences in the number of incoming observations speak to the possibility of a smaller impact of our modifications on the 01 October event, though the synoptic evolution of that event precludes much polar involvement at all. However, synoptic evolution notwithstanding, it appears that there is a wide variance in the number of MODIS wind observations that are brought into the analysis system on each cycle, as well as the distribution of Expected Error within those observations.

An analysis which contains many MODIS wind observations with a wide distribution of expected errors including a narrow tail of low-error observations and a wide tail of high-error observations (e.g. the distribution of the September 21 analysis, see Fig. 4a) stands to benefit from these modifications, because many poor observations are being removed that would have otherwise been admitted into the analysis, and there are still a large number of (good) observations admitted. However, an analysis which contains very few MODIS wind observations with a narrow distribution of expected error containing very few poor observations (see Fig. 4b) may suffer from the loss of too much data when the EE > 5 ms^{-1} threshold is applied. An already depressed number of observations is further drastically reduced, and very few of the newly rejected observations are exceptionally poor to begin with.

In fact, we can calculate the difference in Day-7 AC between the experiment and the control and regress it onto these very characteristics (Figure 5).

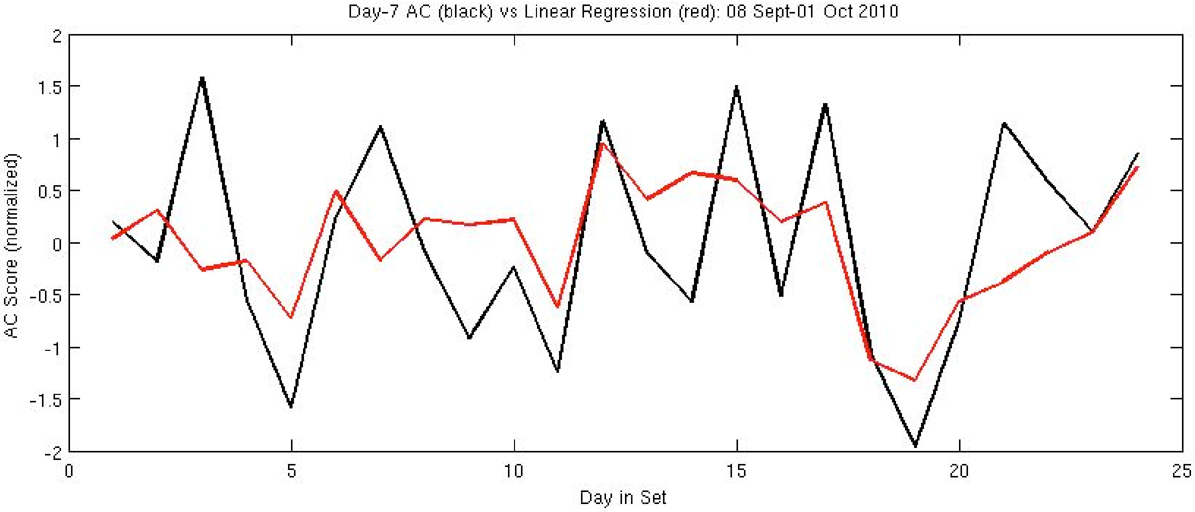

Figure 5. Timeseries of difference in Day-7 AC score (experiment – control) for 24 forecasts between 08 September 2010 and 01 October 2010. Calculated difference in AC (black) and linear regression (red) based on four characteristics of observations: total number of observations, mean expected error, standard deviation of expected error, and number of observations with EE > 7.5 ms^{-1} normalized by the total.

A linear regression of four observation characteristics onto the Day-7 AC score yields regression coefficients (normalized):

1. Total number of observations: 0.57
2. Mean expected error: -2.06
3. Standard deviation of expected error: -1.17
4. [Number of obs. EE > 7.5 ms^{-1} / total]: 2.98

The experiment tends to outperform the control when there are a large number of observations, the mean EE is low, the standard deviation is low, and there are many bad observations that need to be removed. Likewise, the experiment tends to underperform the control when there are few observations, the mean EE is higher, the distribution is wider, but there are few poor observations. We are exploring possible explanations for this unexpected relationship, including how the subsequent thinning of the winds may contribute.

Import AVHRR winds

In preparation for incorporating AVHRR satellite-derived winds, we began investigating different methods of reading in additional wind observations without relying on NCEP personnel to make the winds part of the usual dataset. We were able to use scripts written by Dennis Keyser several years ago to convert the original BUFR to PREPBUFR. Also, we are examining the possibility of reading in the observations directly into the GSI from text files. This would be done in the read_prepbufr.f90 routine. Both of these methods look promising at this time, but with changes being made XiuJuan Su, we will await her updated code before pursuing further.

References

Le Marshall, J., A. Rea, L. Leslie, R. Seecamp, and M. Dunn, 2004: Error characterization of atmospheric motion vectors. *Aust. Met. Mag.*, **53**, 123-131.

Zapotocny, T. H., J. A. Jung, J. F. Le Marshall, and R. E. Treadon, 2007: A two-season impact study of satellite and in-situ data in the NCEP Global Data Assimilation System. *Wea. and Forecasting*, **22**, 887-909.

Conferences and workshops

Santek and Hoover

> Attended and presented at the *JCSDA 9th Workshop on Satellite Data Assimilation* from 24-25 May 2011 at the University of Maryland, College Park, Maryland.

Appendix A: Code changes

read_prepbufr.f90:

```
$ diff gsi_code_BTH/read_prepbufr.f90 gsi_code/read_prepbufr.f90

230,233d229
< ! BTH - Add satwndee variable for EE information
<   real(r_kind) satwndee
< ! BTH - End
<
527,530d522
< ! BTH - Initialize satwndee as empty set
<       satwndee = -9999.0
< ! BTH - End
<
539,544d530
< ! BTH - Put EE information into satwndee
<         if( id ==257 .or. id == 258 .or. id ==259 ) then
<           call ufbint(lunin,satqc,4,1,iret,satqcstr)
<           satwndee = satqc(4)
<         endif
< ! BTH - End
1071,1085d1056
<
< ! BTH - modify and output data for MODIS winds
<           if(ictype(nc) == 257 .or. ictype(nc) == 258 .or. ictype(nc) == 259) then
< !      put EE information in station id slot of cdata_all
<            cdata_all(8,iout) = satwndee
< !      write out MODIS sat-wind data
<            write(6,*) 'BTH_C = ', cdata_all(1,iout), cdata_all(2,iout), cdata_all(3,iout), &
<                          cdata_all(4,iout), cdata_all(5,iout), cdata_all(6,iout), &
<                          cdata_all(7,iout), cdata_all(8,iout), cdata_all(9,iout), &
<                          cdata_all(10,iout), cdata_all(11,iout), cdata_all(12,iout), &
<                          cdata_all(13,iout), cdata_all(14,iout), cdata_all(15,iout), &
<                          cdata_all(16,iout), cdata_all(17,iout), cdata_all(18,iout), &
<                          cdata_all(19,iout), cdata_all(20,iout)
<           endif
< ! BTH - End
```

setupw.f90:

```
$ diff gsi_code_BTH/setupw.f90 gsi_code/setupw.f90
151,156d150
< ! BTH - Add variables for output to dayfile
<   real(r_kind) modisqc1, modisqc2, spdo, spdm, spdomg
<   real(r_kind) eeq, ee, bthpres
<   integer(i_kind) bthuse
< ! BTH - End
<
186d179
< ! BTH - NOTE: EEQF is being passed into array as id for satwinds
573,582d565
< ! BTH - Copy qcu and qcv to dedicated modis qc variables for output
<        modisqc1 = qcu
<        modisqc2 = qcv
< ! BTH - Copy EEQF data to eeq, pressure data to bthpres for output
<        eeq = data(id,i)
<        bthpres = data(ipres,i)
< ! BTH - Change EEQF to EE: EEQF = 100-(EE*10)
<        ee = (100-eeq)/10.
< ! BTH - End
<
584,588c567,569
< ! BTH - Remove replace  qcu/qcv check and base on ee rather than abs(dudiff), abs(dvdiff)
< ! BTH - ORIGINAL CODE:          abs(dudiff) > qcu .or. &                   !  u component check
< ! BTH - ORIGINAL CODE:          abs(dvdiff) > qcv .or. &                   !  v component check
<          ((ee .GT. 5.0) .AND. (ee .GT. 0.1*spdob)) .or. &     ! BTH - End
<          (presw > prsfc-r200 .and. isli /= izero))then ! near surface check
---
>           abs(dudiff) > qcu .or. &                   !  u component check
>           abs(dvdiff) > qcv .or. &                   !  v component check
>           (presw > prsfc-r200 .and. isli /= izero))then ! near surface check
803,805d783
< ! BTH - Save analysis usage flag for dayfile output
<        bthuse = rdiagbuf(12,ii)
< ! BTH - End
868,899d845
<
< ! BTH - Printout of MODIS data (using Li Bi's template from setupw.f90_EXP)
<
<     if ( ictype(ikx) == 257 .or. ictype(ikx) == 258 .or. ictype(ikx) == 259 ) then
< ! Define speed of obs (spdo) and speed of model guess (spdm)
<       spdo = sqrt(data(iuob,i)*data(iuob,i) + data(ivob,i)*data(ivob,i))
<       spdm = sqrt(ugesin*ugesin + vgesin*vgesin)
< ! Define speed of ob-minus-guess (spdomg)
<       spdomg = sqrt(dudiff*dudiff + dvdiff*dvdiff)
< ! Redefine bthuse as = 0 if not 1 or -1 (bthuse only a valid option for jiter = 1)
<       if (bthuse .ne. 1 .and. bthuse .ne. -1) bthuse = 0
< ! Output MODIS data from IR cloud drift (ictype=257)
<       if (ictype(ikx) == 257) write(6,87) 'bth257  ',spdo,spdm,spdomg, &
```

```
<       data(iuob,i),data(ivob,i),ugesin,vgesin, &
<       modisqc1,modisqc2,data(ilone,i),data(ilate,i),data(iskint,i), &
<       dtime,ictype(ikx),jiter,eeq,bthpres,bthuse
< ! Output MODIS data from WV cloud drift (ictype=258)
<       if (ictype(ikx) == 258) write(6,87) 'bth258  ',spdo,spdm,spdomg, &
<       data(iuob,i),data(ivob,i),ugesin,vgesin, &
<       modisqc1,modisqc2,data(ilone,i),data(ilate,i),data(iskint,i), &
<       dtime,ictype(ikx),jiter,eeq,bthpres,bthuse
< ! Output MODIS data from (WV clear air???) (ictype=259)
<       if (ictype(ikx) == 259) write(6,87) 'bth259  ',spdo,spdm,spdomg, &
<       data(iuob,i),data(ivob,i),ugesin,vgesin, &
<       modisqc1,modisqc2,data(ilone,i),data(ilate,i),data(iskint,i), &
<       dtime,ictype(ikx),jiter,eeq,bthpres,bthuse
<     endif
< 87  format(2x,a9,10f8.3,2f11.2,f8.2,x,i4,x,i2,f14.7,f14.7,i6)
<
< ! BTH - End
<
<
```