New Quality Indicators for GOES Derived Winds and their Potential Affect on NWP Data Impact Experiments

Howard Berger, C Velden and J LeMarshall

UW-CIMSS 1225 W Dayton Street, Madison WI, 53706 USA

Abstract

Data quality control is important to numerical model assimilation and subsequent forecast impacts. A promising new quality characterization method for Atmospheric Motion Vectors (AMVs) is the “expected error” (EE) index, as proposed by Dr. John LeMarshall and developed at the Bureau of Meteorology in Australia. This method extends and modifies the current AMV quality indicator (QI) scheme developed at EUMETSAT that is used operationally by numerical forecast centers to thin the AMV input. The EE algorithm linearly regresses several AMV parameters against co-located RAOBS. Coefficients from the regressions can then be used to come up with “expected errors” for each derived vector. It is shown that AMVs with lower EE have better quality. Therefore the EE indicators will be used for our proposed data impact experiments.

The expected error software can also be used to estimate AMV correlated error, which has been found to be a possible source of error in data impact experiments. It is proposed that an accurate estimate of the AMV correlated error can provide useful information to variational assimilation of the data. Model impacts experiments to test this are being designed.

INTRODUCTION

AMV quality control is important, particularly for NWP data assimilation applications. Poor data ingested into NWP models can lead to poor analyses and forecasts. This need for quality control has led AMV producers to create automatically generated indices as a measure of the quality of the wind. These indices provide a quality control check and can be related to the AMV’s observation error.

The two most prominent operational indices are the quality indicator (QI) and the recursive filter flag (RFF). The QI is a weighted average of vector tests in which each wind is assigned a score from 0 to 1. 0 Indicates a poor quality wind, and 1 indicates a high quality wind (Holmlund, 1998). The RFF is a 3-dimensional recursive filter object analysis in which the winds are fit to a high-density analysis field. The AMV’s height is also re-adjusted slightly based on a level of best fit. After the process, the wind is assigned a quality flag between 0 and 100, with 100 being the highest quality wind (Hayden and Purser 1995, Velden et al 1997 and 1998).

While both of these indicators have been successful in estimating the quality of AMVs, the normalized score of both quality indicators can be difficult to interpret. In order to convert these scores into something that better represents an observation error, which data assimilation centers can use directly, the indices must be converted to a velocity error. This conversion must be done through some sort of look-up table or through previously calculated statistical studies. An alternative approach is suggested in this paper: the expected error (EE).

EXPECTED ERROR

The EE indicator is an extension of the previously described QI (LeMarshall et al., 2004). Essentially, it linearly regresses the five QI tests along with other vector and model information against actual AMV/RAOB vector differences. The nine predictors are shown in table 1. The first five predictors are the normalized 0 to 1 scores as described in the previous section. The sixth and seventh predictors, the wind speed and pressure level, are derived from the wind vector itself, while the two gradient predictors are generated from the model guess. These predictors are generated for several weeks of data and then regressed against the AMV-RAOB vector difference in order to create regression coefficients. The coefficients can then be used to estimate the error for subsequent vectors.

1) QI Speed Test
2) QI Direction Test
3) QI Vector Difference
4) QI Local Consistency Test
5) Qi Forecast Test
6) Wind Speed
7) Assigned Pressure Level
8) Wind Shear (200 hPa Above – 200 hPa below)
9) Temperature Gradient (200 hPa Above – 200 hPa below)

Table 1: **Expected Error Predictors**

EXPECTED ERROR IMPLEMENTATION AT CIMSS

The EE software has been running in real-time at CIMSS since September 2005. Currently, we are generating regression coefficients and expected errors for the GOES-12 Infrared, Visible, Water Vapor, and Shortwave-IR winds in the Northern Hemisphere, and Infrared, Visible, and water vapor winds in the Southern Hemisphere. All of the included winds have been processed and quality-controlled by CIMSS real-time processing. For each real-time wind set, the AMVs and RAOBs are co-located, matched, and stored every six hours. Before the winds are stored, the EEs of the current winds are calculated using the previous coefficients. Once they are stored, new regression coefficients are generated from the current month's worth of data and the previous two months of data. These coefficients are used for the subsequent wind set. Table 2 shows the coefficients for the Northern-Hemisphere infrared (NHIR) dataset for May 10, 2006.

Predictor	**Value**
CONST	10.26
QI Speed Check	1.97
QI Direction Check	2.55
QI Visible	-2.82
QI Local Consistency	-2.61
QI Forecast Check	-3.48
Vector Speed	0.08
Vector Pressure Level	-4.8×10^{-3}
Wind Shear	0.038
Temperature Gradient	0.054

Table 2: **Expected error predictors and coefficients valid May 9th, 2006.**

EXPECTED ERROR PERFORMANCE

In order to evaluate the skill of the expected error, it is compared to the actual AMV – RAOB vector differences (matches 50 km or less apart) from February through April 2006. For each AMV, the RAOB wind speed and direction is interpolated linearly to the assigned AMV height level. Figure 1 shows scatter plots of the regressed EE (ms^{-1}) versus each of the nine predictors. The colors of the points represent the winds' speed. As the figures show, some of the predictors have a stronger relationship with their predictand than others. The QI-forecast check and wind speed have the clearest relationship (largest slope) while some of the other QI checks have a relatively weak relationship with EE. Higher wind-shear environments tend to have a higher minimum EE than lower sheared environments. This relationship might be related to height uncertainty in these sheared regions. The temperature gradient predictor's relationship is not clear. All of these relationships, however, are statistically significant to 95%.

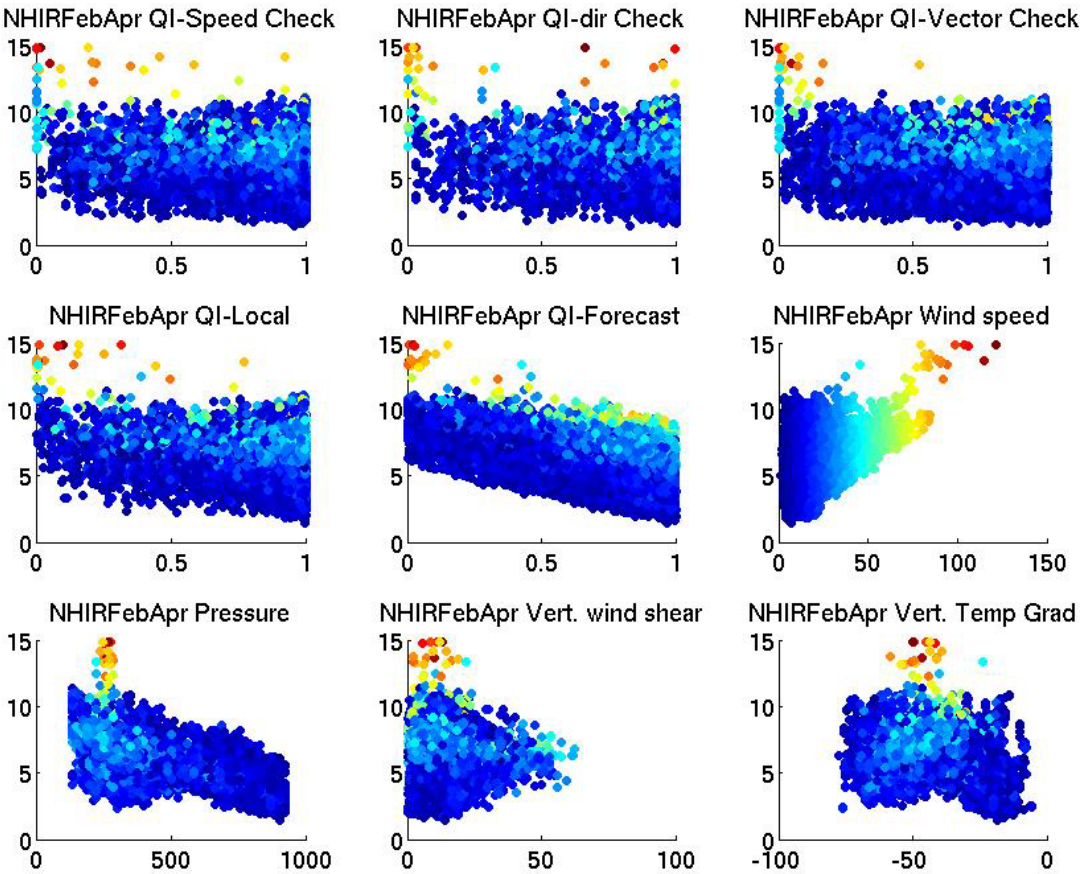

Figure 1: **Expected error versus each of the predictors of expected error (ms^{-1}). The darker colors (blues) represent slower winds while the lighter colors (reds) represent faster winds. The x-axis for the first five plots are unitless. The sixth plot has units of ms^{-1} for its x-axis. The 7^{th} is in hPa, the 8^{th} in ms^{-1} and the 9^{th} in K.**

It is also worth examining the overall statistical skill of the EE as compared to the actual AMV – RAOB difference. Table 3 summarizes the results of this three -month comparison for each of the Northern Hemisphere wind sets. As the table shows, the average EE is slightly higher than the actual AMV – RAOB difference for all the of the channels, although this bias is smaller for the visible and shortwave IR channels. The shortwave-IR and visible winds also have a lower RMS difference than the infrared and water-vapor channels. This improved relationship may be due to the relatively smaller errors (i.e. lower mean AMV – RAOB difference) for the shortwave IR and visible channels, and for the fact that the QI has a stronger relationship with AMV – RAOB difference than with the infrared and water vapor channels. Because the QI makes up five of the nine predictors for the EE, its skill in predicting AMV – RAOB difference will directly affect the skill of the EE. This relationship between QI and EE error will be further examined in the next section.

GOES DATASET	NHIR	NHVS	NHWV	SWIR
NUMBER OF MATCHES	14068	2100	13108	1629
MEAN EE (MS^{-1})	6.07	3.68	6.44	4.61
MEAN AMV-RAOB (MS^{-1})	5.34	3.61	5.51	4.47
BIAS (MS^{-1}) (EE VS ACTUAL)	0.726	0.068	0.93	0.142
RMS (MS^{-1}) (EE VS ACTUAL)	3.53	1.97	3.79	2.75

Table 3: **Overall regression statistics comparing the expected error to the actual AMV – RAOB differences for the period between February and April, 2006. The visible (NHVS) and shortwave-IR (SWIR) expected error have a lower bias and rms than the infrared (NHIR) and water-vapor (NHWV) when compared to the actual AMV – RAOB difference.**

QI VERSUS EXPECTED ERROR

The EE was created, in part, to improve the relationship between a given quality indicator and the AMV – RAOB difference. Thus, it is informative to compare the skill of the EE relative to the QI: the indicator the EE is enhancing. Figure 2 shows RMS statistics of AMV – RAOB differences of winds as binned by a central EE value (red circled curve) and a central QI (blue x curve). In other words, an EE bin of 5 ms^{-1} corresponds to winds that have EEs between 4.5 and 5.5. The rms of the AMV - RAOB differences for those winds are what is plotted in Figure 2. The numbers on each curve represent the number of winds in each bin. The four plots correspond to winds produced from different channels: NHIR (upper-left), NHVS (upper-right), NHWV (lower-left) and NHSWIR (lower-right). The dashed-line on each of the plots represents the 1-1 line. If the EE were a perfect predictor of the actual error (as defined as AMV – RAOB difference), each of the points would fall on this line. As the plots show, however, although the EE is often close to this line, it varies by about a meter per second from this line at various points. The actual error also shows sensitivity to the QI (corresponding to the top axis). As the QI increases, the actual error decreases slightly, but at a less rapid rate than for the EE. In general, the range of actual error values is larger for the EE ranges shown than for the QI ranges shown corresponding to a stronger sensitivity. These results imply that the EE is a better predictor of wind quality than the corresponding QIs.

It should also be noted that even if the EE fell on the 1:1 line, it can only measure the actual error as defined by the AMV - RAOB difference. The errors to which the EE is being compared also contain inherent raob instrument noise, as well as error introduced through the matching process (AMVs and raobs are not perfectly collocated in space or time). Therefore, the EE cannot be considered a true AMV “observation error” per se. The above errors must be accounted for before the EE can be considered a true observation error."

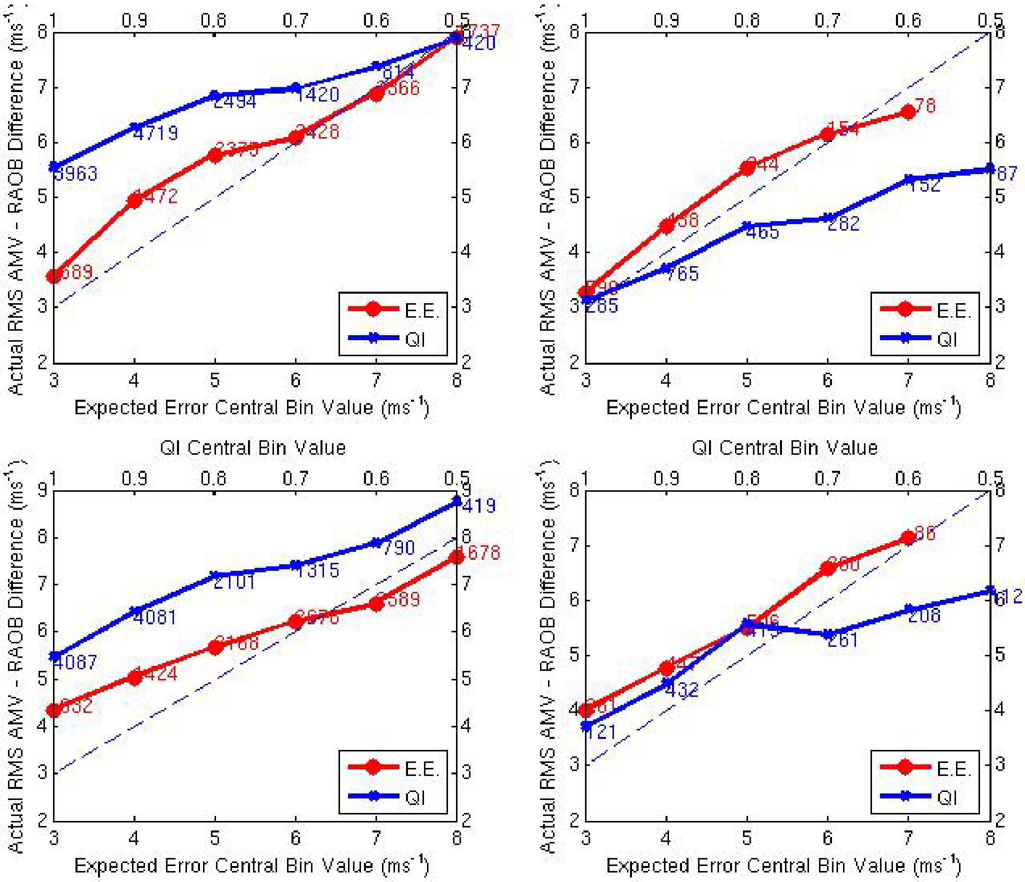

Figure 2: **Comparisons of expected error bins (red circles) and QI bins (blue x's) as a measure of the RMS of the actual AMV – RAOB difference for NHIR (upper-left), NHVS (upper-right), NHWV(lower left) and NHSWIR(lower – right). The lower axis corresponds to EE and the upper axis corresponds to the QI. The number of matches in each bin is also shown. In general, the EE shows a stronger relationship with the actual error as compared to QI for each channel.**

EXPECTED ERROR AND SPATIALLY CORRELATED ERROR

The EE also modifies the AMV spatial error correlations. These error correlations are particularly important for numerical weather prediction, as most current advanced data assimilation methods assume that observations are spatially uncorrelated. To examine the spatially correlated AMV error, a procedure similar to the one described in Bormann et al (2003) is used. An AMV is matched to the closest RAOB, although the matches were restricted so that a unique RAOB and AMV were chosen for each match. These matches were accrued from September 2005 – through April 2006. If one assumes that the RAOBs are spatially uncorrelated to each other, then correlations between AMV – RAOB pairs can be assumed to be entirely from the AMV. The correlations are then calculated as a function of AMV – RAOB pair separation using an EE maximum for the AMV data that was used. The correlations are then fitted to the same correlation function as in Bormann (2003) using a non-linear least squares method. The function was extrapolated to a 0-separation distance. These correlation functions are plotted for the various EE thresholds in figure 3.

As shown in the figure, the lower EE thresholds generally lower the spatial error correlations as a function of separation distance. Although the 0-separation correlation is lowered for the lower EE thresholds, the correlation distance (defined as the distance at which the correlation drops by a factor of $2e^{-1}$) is not necessarily reduced. Note also that Bormann's study involved a full year of data. It is possible these results have a seasonal bias included or have too few matches to be statistically significant. Regardless, these preliminary results show promise that the EE can be used to reduce some of the spatial error correlations inherent in the data.

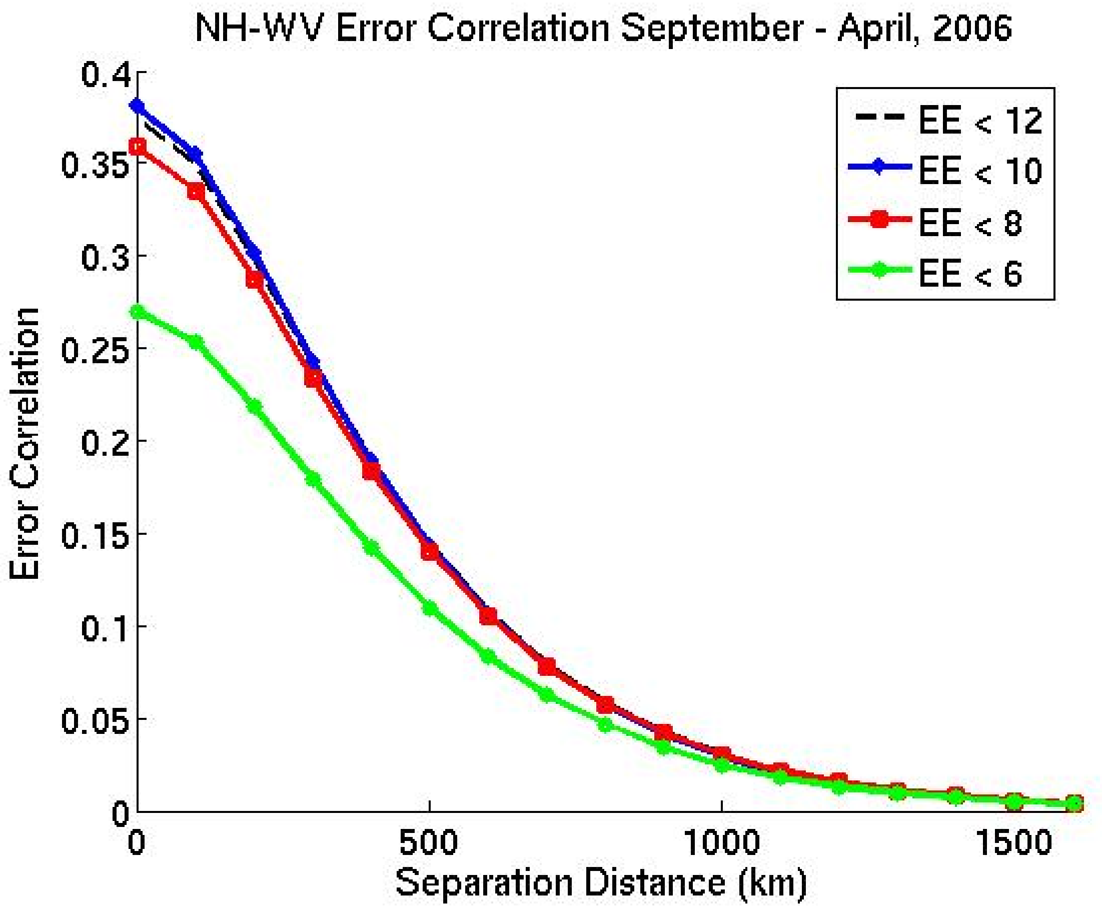

Figure 3: **Interpolated Correlation Function of NH-WV AMV spatial error correlations as a function of AMV separation distance. Each curve represents the correlation function using a different expected error maximum. In general, lowering the maximum expected error lowers the AMV error correlations as a function of distance.**

EXPECTED ERROR DISCUSSION

The EE has been shown to be a reasonable estimate of AMV quality. It is again emphasized that it is not a true measure of the observation error. The observation error, in a data assimilation context, is defined as the difference between an observation and the truth. EE also includes RAOB error and representative error between the AMV and the RAOB. Also keep in mind that the EE has been calculated after the quality control of CIMSS real-time processing. Most of the bad winds have already been rejected. Finally, the EE naturally increases with wind speed. Attempts to normalize EE by the AMV speed , however, did not add value to the indicator. Even with these contingencies, the EE shows skill over the existing QI in representing the quality of AMVs. This skill should prove useful in numerical weather prediction model impact studies.

DATA ASSIMILATION APPLICATIONS

As described in the previous sections, the EE appears to be useful as an AMV quality control descriptor. An obvious application of this new index is for assimilation in numerical weather prediction models. As part of a collaborative effort, CIMSS has partnered with the Office of Naval Research and the Naval Research Laboratory to examine the impacts of AMVs in numerical weather prediction models. No results are yet available at the time of this writing, but it is hoped the effects of employing EE information can be tested as part of this collaboration

Another avenue of AMV NWP impact research is through the NOAA THORPEX initiative. On-going work at CIMSS involves investigating the impact of the AMVs within the NCEP GFS model. We are investigating the impact of the operational GOES winds as well as specially-derived rapid-scan AMVs within the GDAS system. Experiments using the EE as a threshold quality control value for the AMV assimilation trials are envisioned.

FUTURE WORK

We will continue to investigate the performance of the expected error, both in a statistical sense versus RAOBS and in the context of the data assimilation experiments described previously. We also plan to investigate creating an expected error that does not rely on the QI-first guess check and the model shear information. When evaluation is complete, the expected error will be implemented into the CIMSS and NESDIS real-time AMV processing. Information about the spatially correlated error will also be included. It is expected that both of these error indices will be useful in future data impact studies.

REFERENCES

Bormann, N., Saarinen, S., Kelly, G., Thépaut, J. 2003: The Spatial Structure of Observation Errors in Atmospheric Motion Vectors from Geostationary Satellite Data Monthly Weather Review 2003 **131**: 706-718

Holmlund, K., 1998: The utilization of statistical properties of satellite-derived atmospheric motion vectors to derive quality indicators. Wea. Forecasting, **13**, 1093–1104.

Le Marshall, J, Rea A. 2004: Error characterization of atmospheric motion vectors. Aust Met. Mag. **53,**p. 123-131

Velden, C.S., C. M. Hayden, M. S. J. Nieman, W. P. Menzel, S. Wanzong, and J. S. Goers, 1997: Upper-Tropospheric winds derived from geostationary satellite water vapor observations. Bull. Amer. Meteor. Soc., **78**, 173–195.

Velden, C.S., T. L. Olander, and S. Wanzong, 1998: The impact of multispectral GOES-8 wind information on Atlantic tropical cyclone track forecasts in 1995. Part I: Dataset methodology, description, and case analysis. Mon. Wea. Rev., **126**, 1202–1218.