**Proceedings for the 14th International Winds Workshop,
23-27 April 2018, Jeju City, South Korea**

MIGRATION TO A FORECAST INDEPENDENT QUALITY INDICATOR AND FURTHER QC CHANGES AT THE MET OFFICE

James Cotton, Mary Forsythe

Met Office, FitzRoy Road, Exeter, UK

Abstract

AMVs are supplied with quality indicator (QI) values largely based on spatial and temporal consistency checks, but can also include a first-guess check against a numerical weather prediction (NWP) forecast. In this study, we investigate the migration from using the QI with the first-guess check, QI1, to the model-independent QI2 for data screening in NWP. We also relax the applied QI thresholds, since the QI2 is found to have little skill in discriminating "bad" data. The current set of thresholds (referred to as the strict thresholds) vary by satellite, channel and latitude band, and are very strict in the tropics. The revised set of QI thresholds (relaxed thresholds) simply vary by satellite and are less strict, such that no filtering is applied for GOES and Himawari-8. A large proportion of data have a very high QI2 value and so combining this with the relaxed thresholds allows a greater number of winds for assimilation.

The forecast impact of migrating to use QI2 is evaluated through a series of assimilation experiments in the Met Office global model. Using QI2 and the relaxed thresholds results in 20% more AMVs assimilated compared to the control case using QI1 and the strict thresholds. However, the background (T+6 forecast) U/V wind fit to the AMVs is degraded by 8%. To improve the quality of AMVs assimilated we tighten the background check against our own forecast model. This removes only a small number of winds but ensures the overall quality of the AMVs assimilated remains largely unchanged in the migration to QI2. Despite this, it is found that the background fit to the geostationary radiance data remains worse, especially for SEVIRI channels 9 and 10. To counter this increase, a minimum wind speed threshold is introduced for the AMVs to remove winds slower than 4 m/s.

The overall impact of these changes on forecast RMSE is neutral or slightly beneficial when verified versus observations and ECMWF analyses. In particular, beneficial impacts are seen for 250 hPa wind scores in the tropics.

MOTIVATION

Quality control (QC) is an essential step in data assimilation to remove data of "bad" quality, but also to remove observations that contain information that cannot be simulated by numerical weather prediction (NWP) models (Thepaut, 2003). AMVs are supplied with quality indicator (QI) values, which are largely based on spatial and temporal consistency checks on the derived wind components. They come in two flavours, QI1, which includes a first-guess check against the forecast from an NWP model, and QI2, which has no such check. For the QI1, the forecast consistency check is performed against the respective centres model of choice, e.g. for ECMWF in the case of EUMETSAT, GFS for NOAA/NESDIS.

From a data assimilation perspective, we want observations to be as model-independent as possible (at least of other NWP models, if not our own). The Met Office AMV QC screening has thus far used QI1, as this gave greater protection against large height assignment errors. With the extra QC methods introduced in recent years (Cotton, 2016) we now aim to migrate from QI1 to QI2 for data screening, removing the dependency on other NWP centres forecasts. At the same time, we also update the corresponding QI thresholds.

DOES QI2 HAVE SKILL?

A large proportion of winds have a very high QI2 value (Figure 1), more so than with QI1. We can also note from the minimum values of each distribution that the range of QI2 values disseminated varies by each AMV producer. For example GOES winds are only provided for QI2 > 50, whereas the Metop winds from EUMETSAT are provided with no pre-filtering. For MSG there is a trend for reduced root mean square vector difference (RMSVD) with increasing QI2. GOES data have a slight downward trend in RMSVD for QI2 > 60, but most are quite “flat”, particularly Himawari-8. The data from EUMETSAT have odd spikes in the O-B statistics values of QI2. Overall, we can conclude that QI2 is not very useful at discriminating “bad” data, but is of most use for MSG.

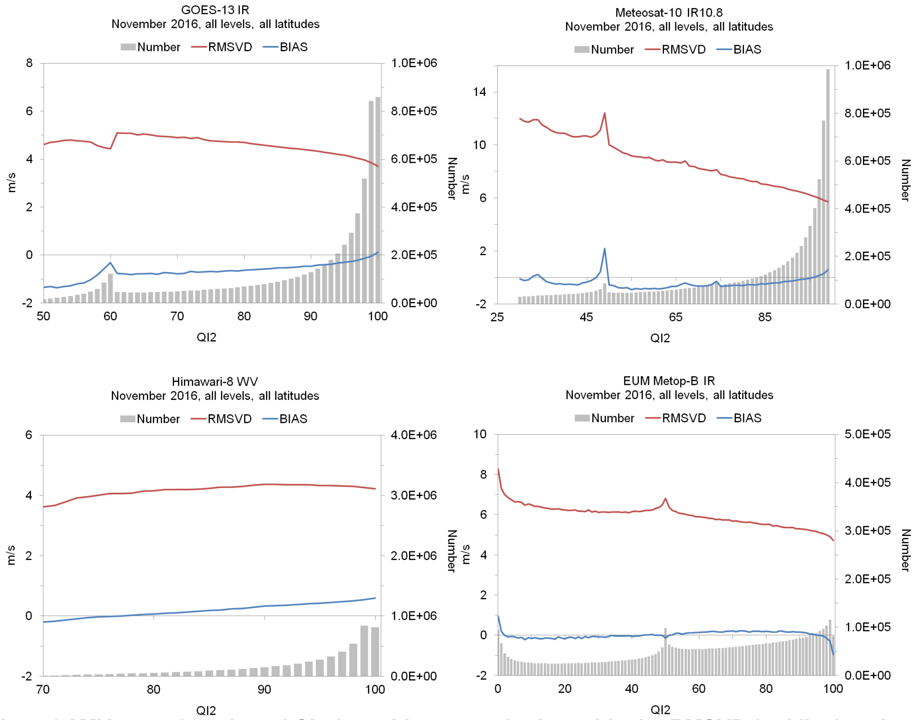

***Figure 1*: Distribution of AMVs as a function of QI2 (grey histograms), alongside the RMSVD (red line) and speed bias (blue line) for several different satellite channels: GOES-13 IR (top left), MSG IR10.8 (top right), Himawari-8 (bottom left), and Metop-B IR (bottom right).**

QI1 has formerly been used for screening observations, with the exception of VIIRS data that are only supplied with a QI2. Threshold values (Table 1) are set for different satellite, channel, latitude band (extra-tropics or tropics), and height band combinations. An important feature of this formulation is that much higher thresholds were used in the tropics. These strict thresholds, combined with the use of QI1, meant that we are essentially only assimilating winds that agreed very closely with the NWP forecast. Although this was done for a good reason in the past, this is clearly no longer optimal if we are to improve the model performance in the tropics.

A new set of thresholds are proposed for use with the migration to QI2 (Table 2). The new thresholds are simplified, as they only vary by satellite, and are more relaxed compared to the old values. For data types that show less skill with QI2, namely GOES and Himawari, we use thresholds that coincide with the minimum expected value. Hence we are not actively filtering, but just protecting against any lower values that may appear unexpectedly, e.g. in case of satellite navigation errors or change in the derivation scheme.

Satellite	QI	Channel	Extra-tropics (High/Mid/Low)	Tropics (All heights)
GOES-13/15	QI1	IR WV	85/80/80 80	90 90
Meteosat-8/10	QI1	IR VIS WV	85/80/80 65 80	90 90 90
Himawari-8	QI1	All	85	85
Metop (EUM)	QI1	IR	80	-
VIIRS	QI2	IR	60	-
LeoGeo	QI1	IR	70	-

***Table 1*: Old, "strict", QI thresholds previously used for screening. Thresholds can vary by satellite, channel, latitude, and height band. High level is above 400-hPa height, mid level is 400-700 hPa, and low level is below 700-hPa height. Tropics is defined as 20°N-20°S.**

Satellite	QI	Channel	Threshold	Active?
GOES-13/15	QI2	All	50	N
Meteosat-8/10	QI2	"	85	Y
Himawari-8	QI2	"	70	N
Metop (EUM)	QI2	"	60	Y
VIIRS	QI2	"	60	Y
LeoGeo	QI2	"	60	Y

***Table 2:* New, "relaxed", QI thresholds now used for screening. Thresholds simply vary by satellite all data sets use QI2. The active column denotes whether the threshold is actively screening data or just set to the minimum value expected.**

IMPACT EXPERIMENTS – PART I

To evaluate the impact of migrating from QI1 to QI2 for AMV quality control we perform a number of impact experiments. The QI2 trial experiments are compared to a reference experiment that uses QI1 and the "strict" thresholds (Table 1).

In the first instance we trial the use of QI2 with the updated set of "relaxed" thresholds (Table 2). This allows an increase of 20% in the number of AMVs assimilated and results in large changes to the mean wind field in the tropics, especially at 850 hPa (Figure 2). The largest changes are seen in the area covered by Meteosat-10 (0° service), e.g. slowing down the northeasterly winds to the south of West Africa. There is very little difference in the mean analysis over the area covered by Himwari-8 where we know the QI2 has no skill.

The change in the model background (short-range T+6hr forecast) fit to the observations is an important diagnostic for trial evaluation. This is because an improved analysis should result in a more accurate short-range forecast, which then agrees better with the next batch of observations in the subsequent cycle. In this experiment, we find that the background standard deviation (O-B) fit to the AMVs was degraded by around 8%. The background fit to some other observation types is also made worse, including a ~1% degradation for the SEVIRI clear sky radiances. This perhaps indicates that we are allowing data with too large innovations through. The impact on forecast root mean square error (RMSE), whilst largely neutral versus observations, has larger errors for 500 hPa-geopotential-height scores versus ECMWF analyses.

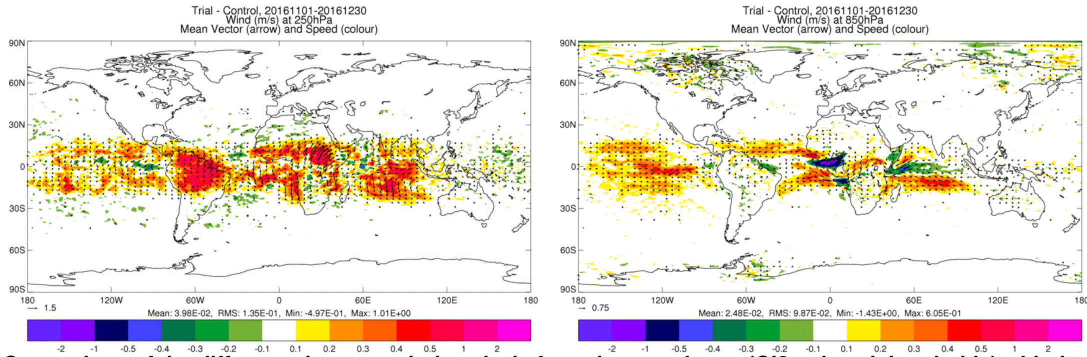

***Figure 2*: Comparison of the difference in mean wind analysis from the experiment (QI2, relaxed thresholds) with the reference (QI1, strict thresholds). Analysis at 850 hPa (left) and 250 hPa (right).**

IMPACT EXPERIMENTS – PART II

In the second round of testing, we try to ensure that the *overall quality* of AMVs used in the assimilation remains similar in the migration from QI1 to QI2. As already shown, the QI2 has limited skill so increasing the thresholds will not be of much use. Instead, we improve O-B for the AMVs by tightening the background check against our own model (Lorenc and Hammon (1988); Ingleby and Lorenc (1993)). Whilst not ideal, this approach at least means we are comparing to our own model, the sensitivity is tuneable, and the current setting is relaxed.

To tighten the background check we modify the parameter PdBadMult, a multiplier for the probability density of "bad" observations. The larger the value of PdBadMult, the stricter the check and the lower the RMS errors of the remaining data (Figure 3). The RMS error for QI2 matches the RMS error for the reference with QI1 (as shown by the dashed red/blue lines in Figure 3) when PdBadMult is increased to around 160. The key point is that we have only removed a very small number of winds to achieve this (the solid green line remains well above the dashed green line in Figure 3). Hence, it is an efficient way to improve O-B, and retains far more winds compared to using QI1.

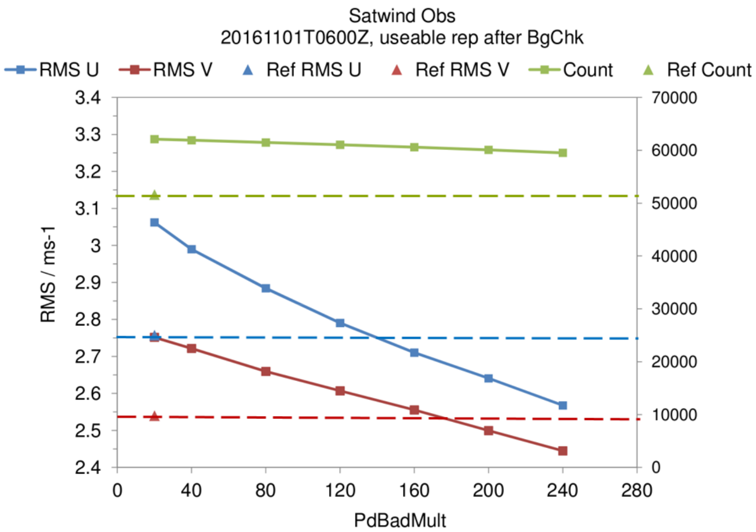

Figure 3: The sensitivity of wind RMS error and data volume to the strictness (PdBadMult) of the background check. U-wind RMS is the solid blue line, V-wind RMS is the solid red line, and solid green is the number of winds remaining. The dashed lines show the baseline levels from the reference case using QI1 and the strict thresholds.

A second impact experiment using QI2 with the relaxed thresholds, but applying the tighter background check still allows an increase of 18-19% in the number of AMVs assimilated (only ~1% fewer than in experiment I). We find that the O-B background fit to the AMVs is now more similar to the reference experiment; U-wind RMS is reduced by 2%; V-wind is increased by 0.5%. The background fit to other observation types is mostly neutral, with the exception of SEVIRI channels 9 and 10. These channels have weighting functions near the surface. Looking at map plots of the standard deviation O-B in these SEVIRI channels, the degradation in fit can be linked to an increase in the O-B for assimilated Meteosat-10 AMVs. The main driver appears to be an increase in wind direction (rather than wind speed) O-B near the equator west of Africa and in the Indian Ocean (Figure 4). These problem areas coincide with the areas of very low wind speeds, with average AMV speeds typically less than 5 m/s. Removing cases where either the AMV or model wind speed is less than 4 m/s significantly reduces the bias and directional variability (Figure 5).

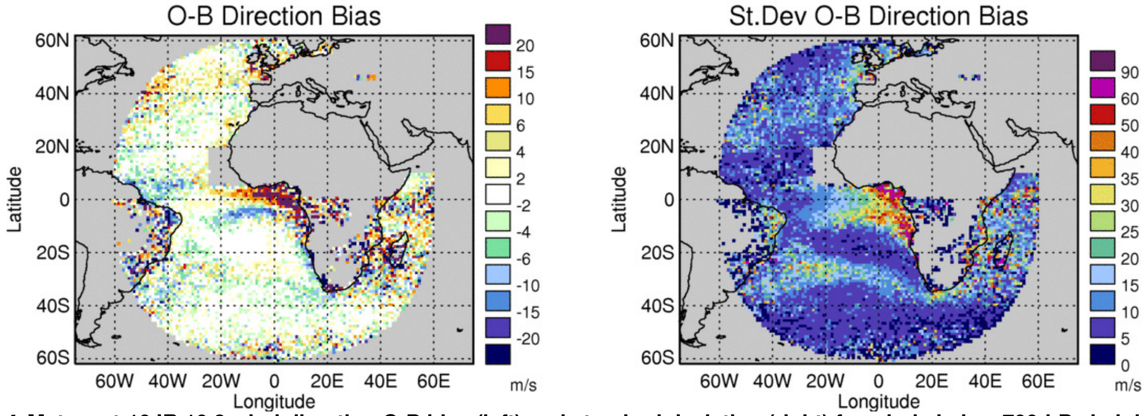

***Figure 4*: Meteosat-10 IR 10.8 wind direction O-B bias (left) and standard deviation (right) for winds below 700-hPa height. Data plotted are those used in the assimilation for a period of one month (December 2016). Note the large and variable direction bias around the equator to the south of West Africa.**

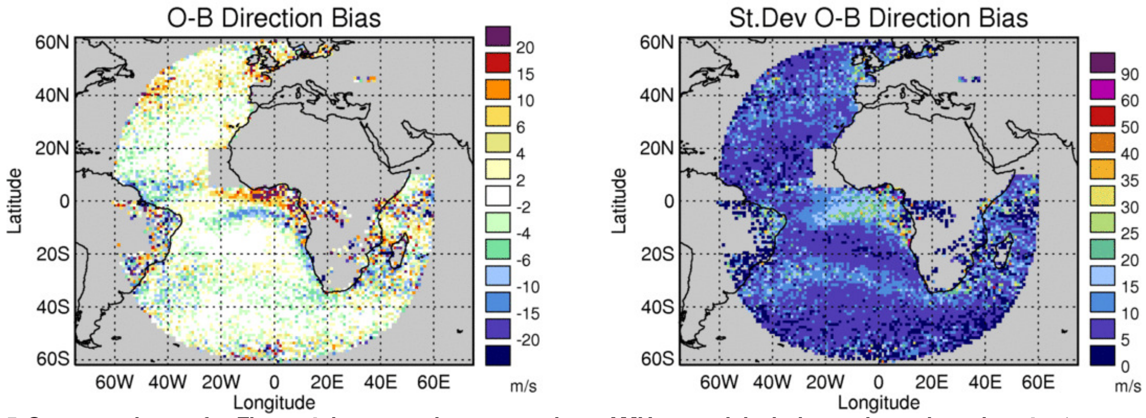

***Figure 5*: Same caption as for Figure 4, but removing cases where AMV or model wind speeds are less than 4 m/s.**

IMPACT EXPERIMENTS – PART III

In a third round of testing, we perform an impact experiment using the QI2 with the relaxed thresholds, applying the tighter background check, and removing the slow winds. This configuration still allows 11-13% more AMVs to be used compared to the reference case with QI1. The background fit to the AMVs remains similar to the second experiment: U-wind RMS reduced by 2%, V-wind RMS increased by 1.0%. The addition of the slow speed check improves the background fit to the geostationary radiances from GOES, SEVIRI and AHI. For SEVIRI channels-9/10 there is an improvement in the summer season but in the winter season, the O-B standard deviation remains 1% higher than in the reference. It would have been preferable to get this closer to neutral but the degradation is now limited to one season only.

The impact on forecast RMSE is largely neutral but there are some statistically significant impacts (Figure 6 and Figure 7). The number of significant positive/negative impacts is 60/13 in the winter season and 65/20 for the summer season and so the majority of significant impacts are beneficial. We can also observe consistent beneficial impacts for day 1-3 forecasts of winds at 250 hPa in the tropics, verified against both observations and ECMWF analyses. Near-surface (2 m) temperatures also show improvements against ECMWF at nearly all lead times.

Figure 6: **Winter season forecast scorecards. These show the change in RMS error for various forecast parameters on the y-axis, ordered from northern hemisphere (90°N-20°N) at the top third, tropics in middle third, and southern hemisphere (20°S-90°S) at the bottom third. On the x-axis is the forecast lead-time, from T+0 to T+144. On the left, we consider observations as "truth" (surface, sonde, and aircraft for W250); on the right, we use ECMWF analyses. Green triangles denote a positive impact (i.e. reduction in RMSE) and purple a negative impact. Impacts that are statistically significant at the 95% level are denoted by a shaded box.**

Figure 7: **Same caption as Figure 6 but for the summer season.**

CONCLUSIONS

The QI2 (without forecast information) shows little skill in discriminating “bad” data, with the exception of MSG, which does have a trend of reduced RMS at higher QI2 value. We have revised the QI thresholds used for data screening and these are generally more relaxed, such that no active filtering is applied for GOES-13/15 and Himawari-8. The relaxed thresholds allow a much larger volume of AMVs to be considered for assimilation and this leads to large changes in the low-level wind analyses in the tropics. The resulting changes in short-range forecast error are not all beneficial, with the background fit to AMVs significantly degraded.

In an attempt to ensure that the overall quality of AMVs assimilated remains about the same in going from QI1 to QI2, we tighten the AMV background check against Met Office forecast winds. This only removes a small number of observations compared to the use of QI1 and so is a far more efficient way to remove larger innovations. After tightening the background check, we still find a degraded T+6 forecast fit to the geostationary radiances, particularly SEVIRI channels-9/10. This was improved by implementing a slow-speed check to remove AMVs with observed or model speeds < 4 m/s. Impact experiments with the whole package of changes show a mostly beneficial impact on forecast RMSE. Statistically significant improvements are seen for winds at 250 hPa in the tropics at days 1-3. This set of changes were implemented in Met Office operations with OS40, 13 February 2018.

In this study, we only have the QI for data screening. With the new AMV BUFR format, we will have much more information from the derivation process carried through with the observations, such as the estimated cloud top pressure error, cloud optical thickness, optimal estimation cost, etc, which could be used for screening (or adjusting assumed observation errors). The new BUFR also has space for the so-called “Common QI” which as demonstrated in the recent AMV intercomparison study, results in a greater similarity between the different AMV data sets (Santek et al., 2018). This should further simplify the AMV screening process.

It should be noted that the GOES data considered in this study are from GOES-13/15. The next generation of US geostationary imagers GOES-16 is now operational, and features a new tracking and cloud height assignment scheme. The GOES-16 winds show a stonger dependence on data quality with QI2 than seen with the previous generation (Lean and Bormann, 2018).

REFERENCES

Cotton, J, (2016) Toward improved height assignment and quality control of AMVs in Met Office NWP. Proceedings of the 13th International Winds Workshop, 27 June – 1 July 2016, Monterey, California, USA.

Ingleby, N, Lorenc, A, (1993) Bayesian quality control using multivariate normal distributions. Q. J. R. Meteorol. Soc., **119**, 1195-1225.

Lean, K, Bormann, N, (2018) GOES-16 AMV data evaluation and algorithm assessment. Proceedings of the 14th International Winds Workshop, In Press., 23-27 April 2018, Jeju, South Korea.

Lorenc, A, Hammon, O, (1988) Objective quality control of observations using Bayesian methods. Theory and a practical implementation. Q. J. R. Meteorol. Soc., **114**, 515-543.

Santek, D, Dworak, R, Wanzong, S, Winiecki, K, Nebuda, S, Garcia-Pereda, J, Borde, R, Carranza, M, (2018) Third atmospheric motion vector intercomparison study, In Press., nwcsaf.org.

Thepaut, J N, (2003) Satellite data assimilation in numerical weather prediction: an overview. Proceedings of Seminar on recent developments in data assimilation for atmosphere and ocean, 8–12 September 2003, ECMWF, Reading, UK