10.6 CHARACTERISING THE SPATIAL STUCTURE OF OBSERVATION ERRORS IN SATELLITE-DERIVED ATMOSPHERIC MOTION VECTORS FOR DATA ASSIMILATION

Niels Bormann*, Sami Saarinen, Jean-Noël Thépaut, Graeme Kelly
European Centre for Medium-range Weather Forecasts (ECMWF), Reading, U.K.

1. Introduction

This study characterises statistically the spatial structure of observation errors in satellite-derived Atmospheric Motion Vectors (AMVs), AMVs are derived by tracking structures in image sequences from geostationary satellite data and provide excellent temporal and spatial coverage. They are an established input to data assimilation systems (e.g., Bouttier and Kelly 2001). Most wind producers now provide datasets at 160 km (at nadir) resolution or higher (e.g., Nieman et al. 1997, Schmetz et al. 1993).

A good specification of the random and systematic errors of any observation is essential to assimilate the information in a near optimal way. The errors assigned to the observations together with an estimate for the error in the First Guess determine the weighting of both in the analysis system (e.g., Daley 1993). For technical reasons, error specifications are frequently simplified. For instance, First Guess errors are assumed to be isotropic and observation errors are assumed to be uncorrelated (e.g., Derber and Bouttier 1999, Daley 1993).

Many authors have suggested that the assumption of uncorrelated errors is invalid for AMVs (e.g., Rohn et al. 2001, Butterworth and Ingleby 2000). Many aspects of the winds derivation are likely to introduce spatially correlated errors, for example, the use of forecast data in the height assignment. In an attempt to reduce the influence of correlated errors in the assimilation, AMVs are frequently thinned to a lower resolution (e.g., 1.25º≈140 km, Rohn et al. 2001), or observation errors are inflated, to avoid overfitting.

The characterisation of the error correlations for AMVs presented here provides important guidance for a refined use of AMVs in assimilation systems. A more detailed report is provided in Bormann et al. (2002).

2. Method and Data

2.1 *Overview*

The calculations presented in this study use a large number of pairs of collocations between an AMV and a radiosonde. The method to derive spatial error correlations from this database is similar to the one used for short-range forecast errors (e.g., Daley 1993, Hollingsworth and Lönnberg 1986). It is based on the assumption that observation errors from sondes are spatially uncorrelated. Therefore, any correlations between the AMV-sonde differences of two stations are attributed to spatially correlated AMV errors. Grouping the collocation pairs from a dense sonde network by station separation allows a characterisation of the average spatial structure of the AMV error correlations over this network. A more detailed discussion of the method can be found in Daley (1993).

The magnitude of the spatially correlated AMV error can be estimated by extrapolating the separation/correlation relationship to zero separation. A suitable correlation function is used to do this, as is further discussed below. The extrapolated correlation at zero separation partitions the variance of the AMV-sonde differences into spatially correlated and spatially uncorrelated parts. The former is the spatially correlated part of the AMV observation error, whereas the latter is made up of the following components: the spatially uncorrelated AMV error, the error for the sonde observations, and any errors arising from the mismatch of representativeness between sondes (point measurements) and AMVs (area averages).

2.2 *Collocations*

Our statistics are based on the AMV datasets summarised in Table 1, covering the period 1 January—31 December 2001, collocated with wind observations from all available radiosonde or pilot reports[1] (e.g., Fig. 1). Poor AMV or sonde observations have been eliminated by requiring a quality indicator of 60 % or more for the EUMETSAT data (e.g., Rohn et al. 2001; quality indicators were not yet available for the other wind datasets), and by rejecting any collocations with more than 18 m/s vector difference (about three times the standard deviation of the departures).

Table 1: AMV data used.

Satellite	*Producer*	*Data used*
GOES-8 & 10	NOAA/NESDIS	operational IR and WV winds
MET-5 & 7	EUMETSAT	operational IR and WV cloud track winds from 160 km (at nadir) segments
GMS-5	JMA	operational IR and WV winds

The collocation criteria were as follows: spatial: within 150 km; vertical: within 25 hPa; temporal: within 1.5 h (CGMS recommendations, e.g., Velden and Holmlund 1998). Further experimentation has revealed no significant sensitivity of our results to tightening the

* *Corresponding author address:* Niels Bormann, ECMWF, Shinfield Park, Reading RG2 9AX, U.K.

[1] For simplicity, radiosondes and pilot reports will be referred to as "sondes" in the remainder of the text.

spatial collocation criterion. Biases are removed for each station by subtracting the mean difference vector for the vertical layer considered.

The method can be used only over dense high-quality sonde networks and thus mainly over land (e.g., Fig. 1). Furthermore, we derived statistics for the entire year over larger regions (e.g., Northern Hemisphere extra-tropical part of the respective satellite disk[2]) and assume uniform errors and error correlations for these regions.

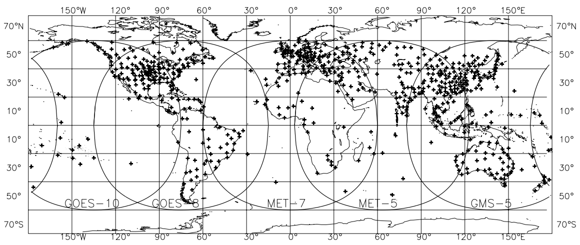

Figure 1: Map of radiosonde or pilot sites considered in this study. Also shown are the outlines of the disks viewed by each geostationary satellite.

2.3 Correlation function

For the isotropic component we derived a least squares fit of a correlation function to our empirical correlation data. We found the isotropic assumption a good approximation for the ⟨Δu,Δu⟩+⟨Δv,Δv⟩ correlations, and will therefore focus the discussion of the isotropic part on this quantity (Δu and Δv denote the difference between the collocated AMV and radiosonde in the u and v component, respectively). The correlation function is primarily used to extrapolate the correlation data in a statistically reasonable way to zero separation to estimate the AMV errors. We use the following function *R* of the station distance *r*:

$$R(r) = R_0 (1 + \frac{r}{L}) e^{-r/L} \qquad (1)$$

with the intercept $R_0 > 0$ and the length scale $L > 0$ as fitting parameters. This correlation function has been used by a number of other authors (e.g., Daley 1993, Thiébaux 1985). The fits were calculated from data grouped into 100 km bins with weights of each data point determined by 95 % confidence intervals for the correlations. The confidence intervals were estimated through a t-test, based on sub-sampled data.

3. Results

3.1 *Isotropic error correlations*

We will first present results for the isotropic part of the ½(⟨Δu,Δu⟩+⟨Δv,Δv⟩) departure correlations for the different satellites and geographical regions. Figure 2 shows the Northern Hemisphere correlations between the AMV-sonde differences as a function of station distance for WV winds from all levels.

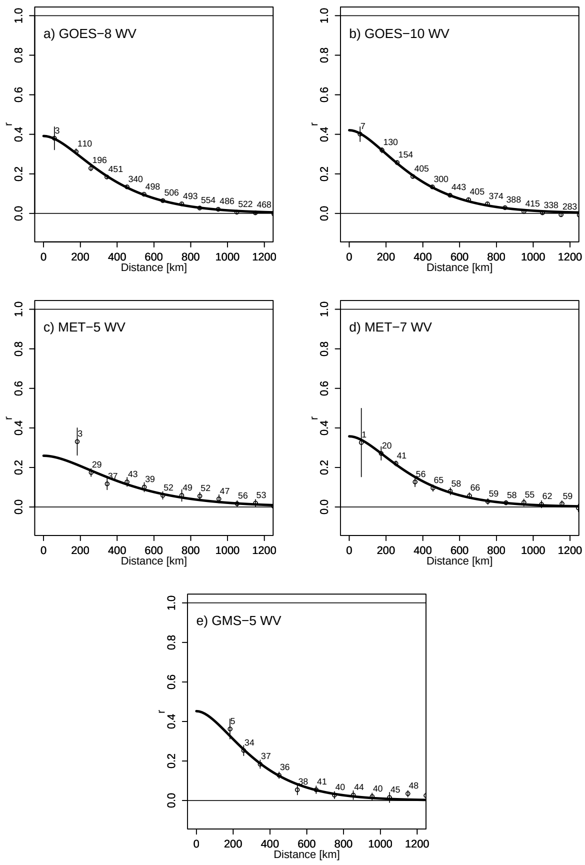

Figure 2: a) ½(⟨Δu,Δu⟩+⟨Δv,Δv⟩) AMV-sonde departure correlations for Northern Hemisphere GOES-8 WV winds from all levels as a function of station separation. Error bars indicate 95 % confidence intervals. The fits were calculated as described in the main text. b) As a), but for GOES-10. c) As a), but for MET-5. d) As a), but for MET-7. e) As a), but for GMS-5.

There are statistically significant correlations in the AMV-sonde differences for distances up to about 800 km. As expected, the correlations are close to zero for station distances larger than about 1000 km. Note that Fig. 2 shows the correlations between the AMV-sonde differences; to obtain an estimate of the AMV *error* correlations we would need to normalise the values to 1 at zero separation.

The departure correlations show no statistically significant differences for the five WV AMV datasets, except for Northern Hemisphere MET-5 winds which show significantly broader and flatter departure correlations. The reasons for the different behaviour of MET-5 WV winds are not well understood, but might be related to the East Asian sonde network and geography rather than the AMV data.

Winds from the IR channels of the five satellites show similar results (not shown). The results are less accurate as fewer collocations are available as IR winds cover a larger vertical range. Error correlation

[2] Geographical areas are referred to as follows: Northern Hemisphere (NH): North of 20N; Tropics: 20N—20S ; Southern Hemisphere (SH): South of 20S.

structures derived for different vertical layers for IR winds also do not differ significantly (not shown).

The departure correlations tend to be significantly flatter and broader over the Tropics for most winds (cf, Figures 2 and 3). Meaningful Southern Hemisphere results could only be obtained for GMS-5 WV winds, and these are not significantly different from their Northern Hemisphere counterparts (not shown).

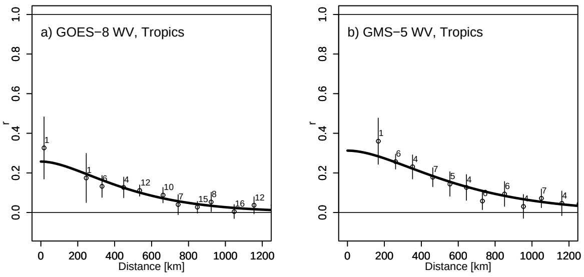

Figure 3: a) As Fig. 2a, but for tropical GOES-8 WV winds. b) As a), but for GMS-5 WV winds.

3.2 *Estimation of AMV errors*

We now estimate the spatially correlated part of the observation error in AMVs by extrapolating to zero distance the distance/correlation relationships using the fitted correlation functions. We use the ½(⟨Δu,Δu⟩ + ⟨Δv,Δv⟩) departure correlations and assume that the errors in each wind component are the same.

Our statistics give spatially correlated AMV errors for WV wind components in the Northern Hemisphere of 3.1—3.5 m/s, except for MET-5, where the problems noted earlier produce suspiciously low error estimates of 2.6 m/s (Table 2). The differences between the other four wind datasets are not statistically significant. Within the accuracy of the method used, errors for the Southern Hemisphere GMS-5 WV winds give similar values (not shown). The errors for the tropical winds tend to be significantly lower, with values of 1.6—2.5 m/s. IR winds give similar, but slightly smaller errors with overall Northern Hemisphere errors of around 2.7—3.3 m/s for all levels.

Table 2: Estimates of the spatially correlated AMV component error σ_{AMV} [m/s] and its uncertainty. The values are derived for AMVs from all levels. "n/a" indicates insufficient data for reliable statistics.

Data	WV σ_{AMV} [m/s]		IR σ_{AMV} [m/s]	
	NH	Tropics	NH	Tropics
GOES-8	3.3(1)	2.3(5)	3.3(1)	2.2(4)
GOES-10	3.5(1)	n/a	3.1(1)	n/a
MET-5	2.6(3)	n/a	2.7(6)	2.4(6)
MET-7	3.1(3)	n/a	2.8(3)	n/a
GMS-5	3.5(4)	2.5(5)	2.7(5)	n/a

For high and medium levels, the estimates for the correlated part of the AMV wind component error agree well with independent estimates for the total component error based on departures from the ECMWF First Guess together with estimates for the First Guess errors (not shown). This gives some indication that the spatially uncorrelated part of AMV errors is likely to be small, and the spatially correlated part dominates.

3.3 *Anisotropic error correlations*

Meaningful statistics for the anisotropic structure of the error correlations could only be produced for Northern Hemisphere WV winds, and Northern Hemisphere IR winds from GOES-8 and GOES-10. These reveal a considerably anisotropic structure of the error correlations. For example, the S—N correlation scales for the ⟨Δv,Δv⟩ departure correlations are much broader than the W—E correlation scales for most WV or IR winds considered (e.g., Fig. 4), except for GMS-5 WV winds (not shown). Apart from a slightly more diagonal orientation, the correlation structures shown in Fig. 4 show similarities with error correlations typical for short-term forecast errors (e.g., Daley 1993, Hollingsworth and Lönnberg 1986).

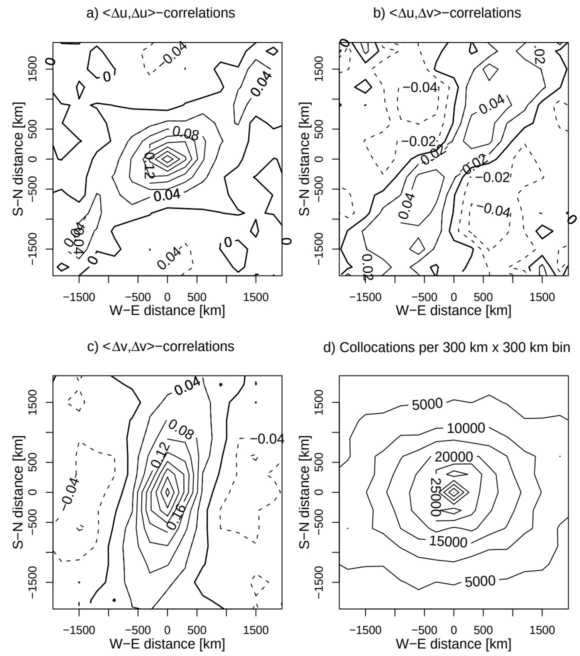

Figure 4: a) ⟨Δu,Δu⟩ AMV-sonde departure correlations for Northern Hemisphere GOES-8 WV winds as a function of W—E and S—N distance from an AMV located at (0 km, 0 km). The binning box size is 300 km x 300 km. b) As a), but for the ⟨Δu,Δv⟩ departure correlations. c) As a), but for the ⟨Δv,Δv⟩ departure correlations. d) The number of collocations per 300 km x 300 km bin.

4. Discussion and Conclusions

We have characterised the spatial structure of errors in satellite-derived AMVs by analysing 12 months of pairs of collocations of AMVs and sondes and assuming spatially uncorrelated errors in the sonde observations. The analysis provides estimates of the spatially correlated part of the AMV component error. The main findings are:

- AMVs show statistically significant spatial error correlations on scales up to about 800 km. The correlations are similar for winds from different satellites, spectral channels, or vertical levels. Tropical error correlations tend to be broader than midlatitude ones.
- The AMV-sonde departure correlations show considerable anisotropy. For instance, $\langle \Delta v, \Delta v \rangle$ departure correlations are broader in S—N direction than in W—E direction for most winds.
- The spatially correlated part of the annual mean AMV component error is about 3.1—3.5 m/s for extra-tropical WV winds above 400 hPa and 2.7—3.3 m/s for IR winds.

The similarities found for the spatial error correlations for different wind datasets are striking, particularly given the differences in the processing and quality control between different winds producers. This confirms that spatial error correlations are inherent in the AMV approach. The use of temperature forecasts for the height assignment is likely to be the largest contribution to these correlated errors. This is supported by the fact that the correlations found in this study share some of the characteristics of correlations in short term forecast errors. For instance, the length scales for the AMV error correlations compare favourably with length scales in temperature forecast errors (e.g., Derber and Bouttier 1999). Considering the entire AMV dataset, the correlated observation errors imply that the larger-scale spatial structures represented in the AMV dataset have larger errors than the small-scale structures (e.g., Bormann et al. 2002).

The findings highlight that error characteristics assigned to AMVs at many data assimilation centres are considerably suboptimal. The satellite winds indeed invalidate the assumption of uncorrelated observation errors. Furthermore, the correlation scales found in this study are much larger than the thinning scales typically applied to AMVs in an attempt to suppress the impact of spatially correlated errors (1.25—2.5º ≈140—275 km, e.g., Rohn et al. 2001). Nevertheless, even a suboptimal specification of the error covariances can be used to extract some information from observations, even though the reduction in the analysis error will not be optimal. This has been shown in theoretical studies and by practical experience with positive forecast impact from the assimilation of AMVs with such suboptimal settings (e.g., Bouttier and Kelly 2001, Daley 1993).

The discrepancies between our results and current assimilation practise suggest that there is some scope for improvement in the assimilation of AMVs through a revision of the assumed error characteristics, based on correlated errors, revised thinning or error inflation. More realistic spatially correlated observation errors in a data assimilation system will alter the filtering properties of the system compared to uncorrelated observation errors. Further investigations are necessary to characterise the influence and relevance of the error correlations in AMVs and possibly other observations in today's data assimilation systems.

Acknowledgements

This study was funded by the ECMWF/ EUMETSAT Fellowship agreement. Constructive comments and suggestions from Erik Andersson, Mike Fisher and Anthony Hollingsworth are gratefully acknowledged.

References

Bormann, N., S. Saarinen, G. Kelly, and J.-N. Thèpaut, 2002: *The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data.* EUMETSAT/ECMWF Fellowship Programme Research Report 12, ECMWF, Reading, U.K. [submitted to *Mon. Wea. Rev.*].

Bouttier, F., and G. Kelly, 2001: Observing-system experiments in the ECMWF 4-DVAR assimilation system. *Quart. J. Roy. Meteor. Soc.*, **127**, 1496—1488.

Butterworth, P., and N. B. Ingleby, 2000: Recent developments in the use of satellite winds at the UK Met. Office. In *Proceedings of the Fifth International Winds Workshop*, Lorne, Australia, EUMETSAT, 151—159.

Daley, R., 1993: *Atmospheric data analysis.* Cambridge University Press, Cambridge, UK, 460 pp.

Derber, J., and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195—221.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111—136.

Nieman, S.J., W.P. Menzel, C.M. Hayden, D. Gray, S.T. Wanzong, C.S. Velden, and J. Daniels, 1997: Fully automated cloud-drift winds in NESDIS operations. *Bull. Amer. Meteorol. Soc.*, **78(6)**, 1121-1133.

Rohn, M., G. Kelly, and R. W. Saunders, 2001: Impact of a new cloud motion wind product from METEOSAT on NWP analyses. *Mon. Wea. Rev.*, 129, 2392—2403.

Turner, J. and D.E. Warren, 1989: Cloud track winds in the polar regions from sequences of AVHRR images. *Int. J. Remote Sensing*, **10(4)**, 695-703.

Schmetz, J., K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, and L. van de Berg, 1993: Operational cloud motion winds from METEOSAT infrared images. *J. Appl. Meteorol.*, **32**, 1206-1225.

Thièbaux, H.J., 1985: On approximations to geopotential and wind-field correlation structures. *Tellus*, **37A**, 126—131.

Velden, C. S., and K. Holmlund, 1998: Report from the working group on verification and quality indices (WG III). In *Proceedings of the Fourth International Winds Workshop*, Saanenmöser, Switzerland, EUMETSAT, 19—20.