Operational Assimilation of QuikSCAT data at ECMWF

H Hersbach, L. Isaksen and P.A.E.M. Janssen
European Centre for Medium-Range Weather Forecasts
Reading, England

Since 22 January 2002, near-surface wind observations from QuikSCAT has been assimilated in the operational 4D-Var system at ECMWF. A concise description of implementation of QuikSCAT data, its performance and ongoing research will be presented.

1 Operational setup

At ECMWF, the 4D-Var assimilation system is using an incremental approach (Courtier et al. 1994, Rabier et al. 2000) where innovations are calculated at T511 (40 km) resolution and the minimization step of the analysis is performed at T159 (120 km) resolution. The 25 km resolution at which the SeaWinds Real-Time BUFR Data product is provided (for a description, see Leidner et al. 2000), is, therefore, too high to be assimilated directly. Instead of thinning the data, which was for instance the strategy followed for ERS-2 scatterometer data and other satellite data, it was chosen to create a 50 km resolution product that makes use of all information contained in the 25 km product[1]. The 50 km wind product is based on all backscatter measurements available within that cell, i.e., four 25 km sub-cells. For this a Maximum Likelihood Estimator (MLE) is optimized expressing the misfit between observed σ_0 and modelled σ_m backscatter values:

$$\mathrm{MLE} = -\sum_i \frac{(\sigma_{0i} - \sigma_{mi}(\mathbf{u}))^2}{(\sum_j \sigma_{mj}(\mathbf{u})^2)} \tag{1}$$

The sum contains up to 16 backscatter measurements; the modelled backscatter values are based on the QSCAT-1 geophysical model function. An example of the MLE surface as function of the 50 km wind vector is displayed in the left panel of Figure 1. There are typically four local maxima, corresponding to four possible wind solutions (see middle panel of Figure 1). Out of these ambiguities, the most likely solution and the solution that is most anti-parallel to this solution are presented to 4D-Var. They are assimilated by the use of a simplified double-well cost function (see right panel of Figure 1):

$$J(\mathbf{u}) = \frac{J_1 J_2}{(J_1^4 + J_2^4)^{0.25}}, \quad , J_{1,2}(\mathbf{u}) = \frac{||\mathbf{u} - \mathbf{u_{1,2}}||^\mathbf{2}}{\sigma_o^2} \tag{2}$$

In the vicinity of each solution, this cost function looks quadratic. In this way, the de-aliasing of the proper wind solution is performed dynamically during the assimilation. The observation error σ_o is $2\,\mathrm{ms}^{-1}$ in each wind component.

[1] developed and implemented by Mark Leidner in 2000, when he was a visiting scientist at ECMWF

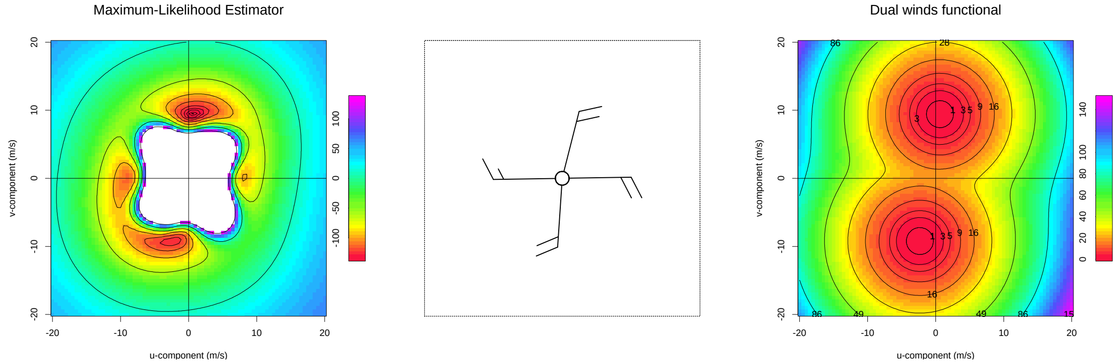

Figure 1: Typical distribution of the MLE in wind-vector space (left panel) and the four solutions corresponding to the local maxima (middle panel) for NSCAT. For QuikSCAT observations similar results apply. Right panel: the double-well cost function as used in the 4D-Var optimization. Courtesy of M. Leidner and R. Hoffman (AER).

QuikSCAT data suffers from rain contamination. Since March 2000, a JPL rain flag is provided in the SeaWinds Data product (JPL rain flag). Backscatter measurements from such flagged cells are excluded from inversion algorithm (1). In addition, 50 km cells in which more than one sub-cell is rain-contaminated are rejected. The beneficial effect on the data quality is illustrated in Figure 2. The large cloud of erroneously high observed winds in the left panel is removed after the JPL rain flag quality control has been applied.

For strong winds, the 50 km winds are bias corrected (see right panel of Figure 2 as well).

In the outer 200 km at both sides of the swath, there are no inner-beam measurements. Therefore, due to the lack of sufficient independent observations, the quality of the wind product is poor, as can be seen from Figure 3. Besides, there is no JPL-rain flag information available for these cells. Therefore, these outer swathes

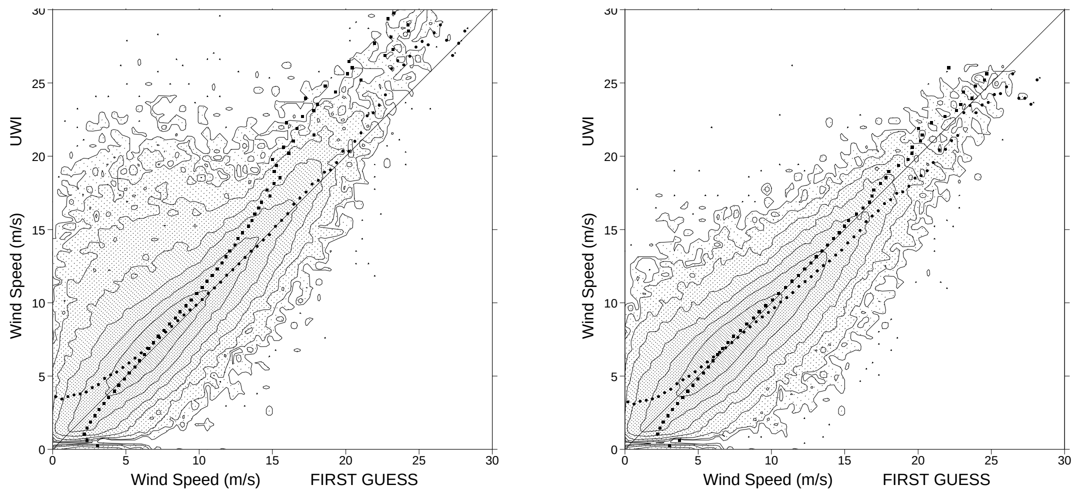

Figure 2: Scatter plot between QuikSCAT 50 km wind speeds and collocated ECMWF first guess fields, for all data (left panel), and for data that were not rejected on the basis of rain contamination (right panel). In addition, for the right panel 50 km winds have been bias-corrected.

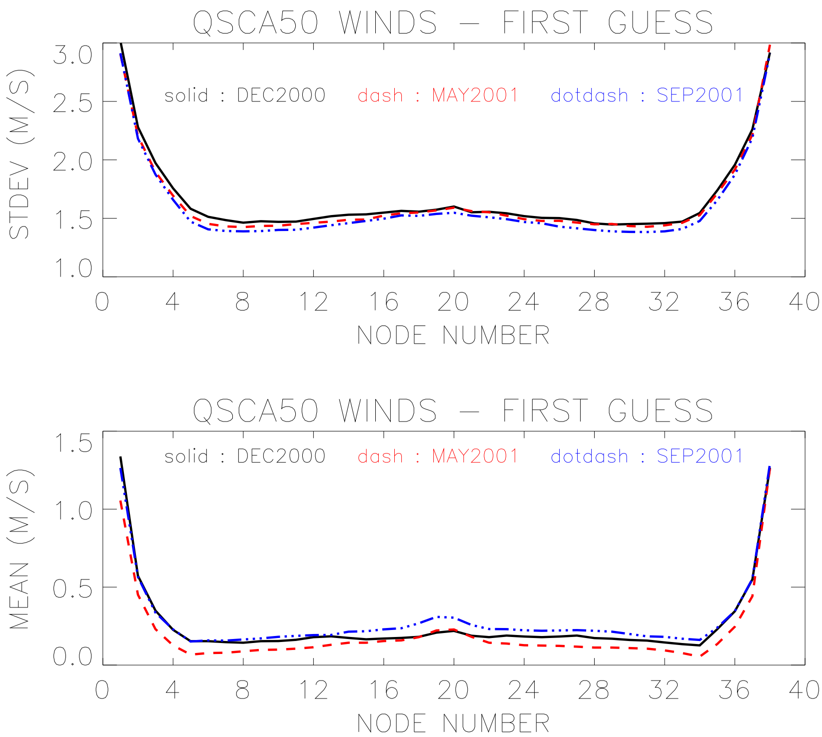

Figure 3: Average bias and standard deviation of the statistics between 50 km QuikSCAT wind speeds and collocated ECMWF winds as a function of 50 km across node number. Nodes 1-4 and 35-38 do not contain only inner-beam measurements and are not assimilated.

are excluded from assimilation, leaving a swath width of 1,400 km.

In the nadir part of the swath, the azimuth diversity between the two inner and the two outer measurements is limited. This leads to shallow maxima for the MLE in wind direction. It often occurs that the four inverted wind solutions are pointing in similar directions. In order to exclude such cells, it is demanded that the angle between the two wind solutions used in cost function (2) should be at least 135 degrees. This quality control is activated for around 50% of the cells in the nadir part of the swath (node 15-24). The quality of the remaining winds in this part of the swath is fair as can be seen from Figure 3.

Prior to the assimilation, a global check on instrument performance is performed on 6-hourly data batches. This quality control is based on normalized average values of the optimal MLE values in (1). Too high average values indicate a potential instrument anomaly. Since its introduction on 17 April 2002, it has not been activated. However, it would have been activated for the instrument anomaly that occurred at 19 March 2002 (see Figure 4). This global check is purely self-consistent and does not rely on any ECMWF model fields.

Finally, like all data assimilated in 4D-Var, QuikSCAT data is subject to variational quality control (Anderssen and Järvinen 1999). This means that if the deviation of the analysis wind is too large from the observed wind, a gross error is suspected, in which case the weight on the observation is significantly reduced.

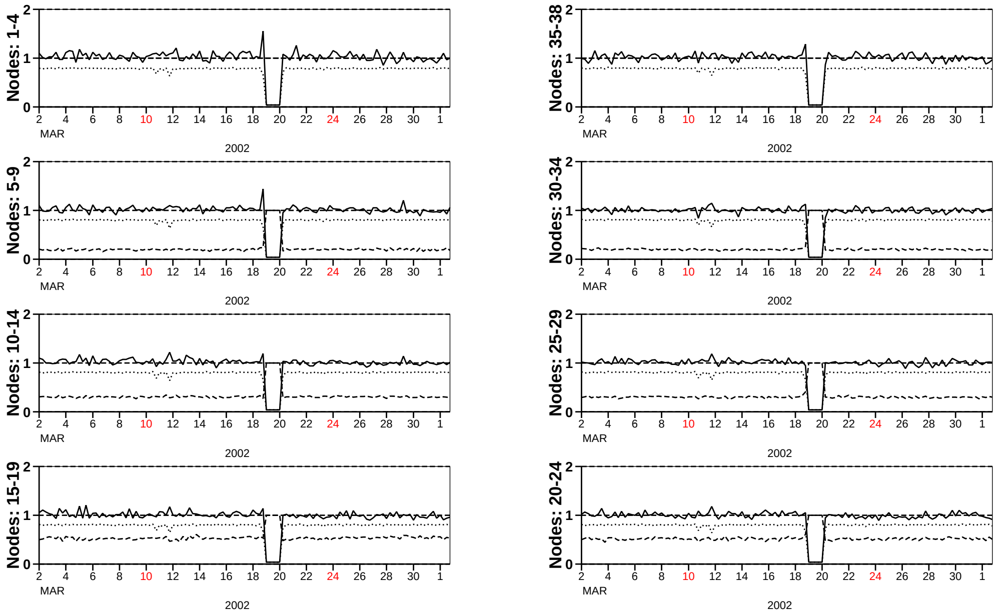

Figure 4: Time series for the normalized MLE averaged over 6-hourly data batches (solid lines), for the 38 across-node 50 km cells. The peaks for 19 March 2002 09-15 UTC indicate an instrument anomaly.

In Figure 5, a typical example of QuikSCAT data, as presented to the assimilation system, is displayed for the North Atlantic. Each dot represents a 50 km wind vector cell; the color coding indicates its status (see caption). Only green cells are assimilated, which is about 50% of the total amount of data.

2 Performance

The assimilation of QuikSCAT data has a positive impact on forecast performance, especially on the Southern Hemisphere. This is not surprising, because of the larger (ocean) data volume and lower amount of conventional data. The average impact for a full-resolution assimilation experiment for the period of May 2001 is presented in Figure 6.

On average, the analysis of tropical cyclones has been improved. However, this is not true for all cases. Too much good quality data around tropical cyclones is rain flagged, leaving only moderate wind speeds far from the center. An example is presented in Figure 7, which shows hurricane Gustav. For this case the analysis has improved; surface winds are enhanced, and the hurricane is moved towards its observed position. The JPL rain flag, however, is set for all 25 km cells near the center of Gustav, and, therefore, no 50 km product could be determined for this area.

3 Ongoing research

Although the JPL rain-flagged data usually corresponds to rainy areas (an example is given by the collocated TRMM data in the lower left panel of Figure 7), a visual inspection of 25 km QuikSCAT winds suggests that such data is not seldom of acceptable quality. This especially applies in regions of extreme conditions, like tropical cyclones. An example is given in the left panel of Figure 8, which shows Hurricane Lily for 20021003 00 UTC, less than 24 hours before landfall. Most winds are rain-flagged. However, the flow and the wind intensity (up to 80 knots) look quite realistic. None of these winds were assimilated. Currently experiments using a different rain-contamination quality-control scheme are conducted. This alternative quality control (developed by Portabella and Stoffelen, 2002), is based on the deviation of the normalized MLE of the 25-km analogue of (1). Much more potentially high quality data is retained, as can for instance be seen from the right panel of Figure 7. The successful assimilation of the anticipated strong winds (up to 80 knots for Lili) requires an adaptation of the variational quality control for QuickSCAT.

In the next version of the ECMWF assimilation system, a new minimization algorithm will be implemented. Its method will be of the conjugate gradient type; its convergency critically requires a strictly quadratic cost function. This disables the

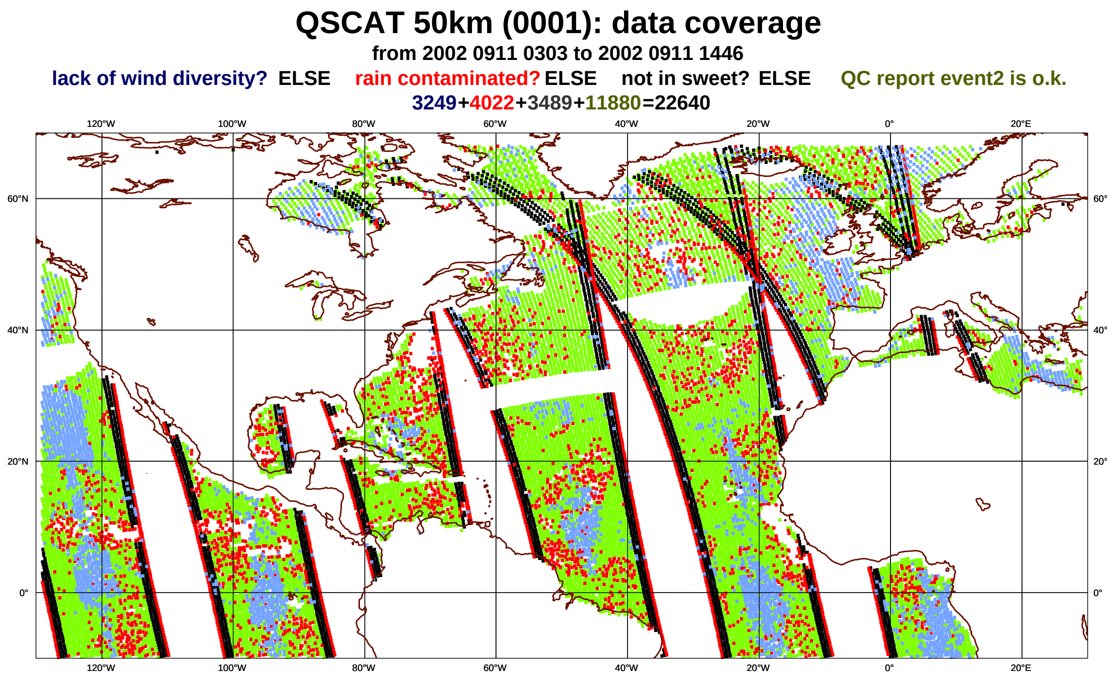

Figure 5: Example for assimilated QuikSCAT data over the Northern Atlantic. Each dot represents a 50 by 50 km wind-vector cell. Green cells were assimilated. Black (outer swath), blue (lack of azimuth diversity) and red (rain contamination) cells were rejected.

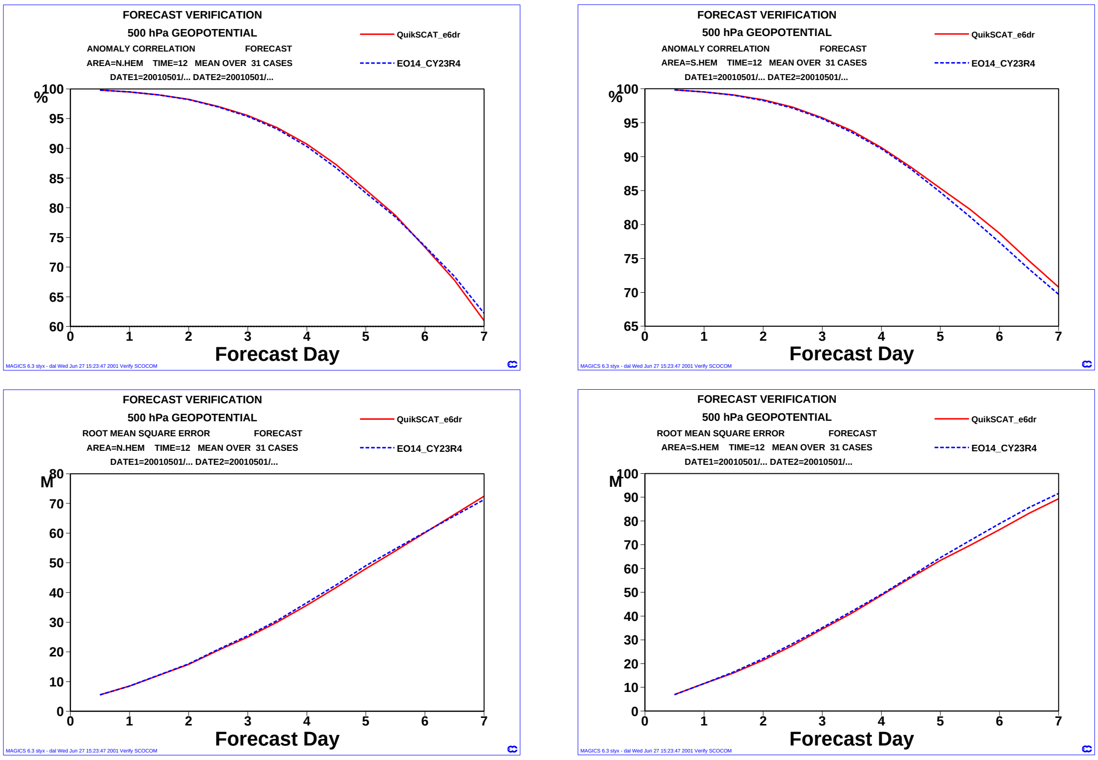

Figure 6: Anomaly correlation coefficient and root mean square error for two experiments in May 2001, averaged over the Northern hemisphere (left panels) and Southern Hemisphere (right panels). Red solid curves are for the pre-operational setup of the ECMWF assimilation and forecasting system that became operational on 22 January 2002, including QuikSCAT data. Blue dashed curves are for the same setup, without the assimilation of QuikSCAT data

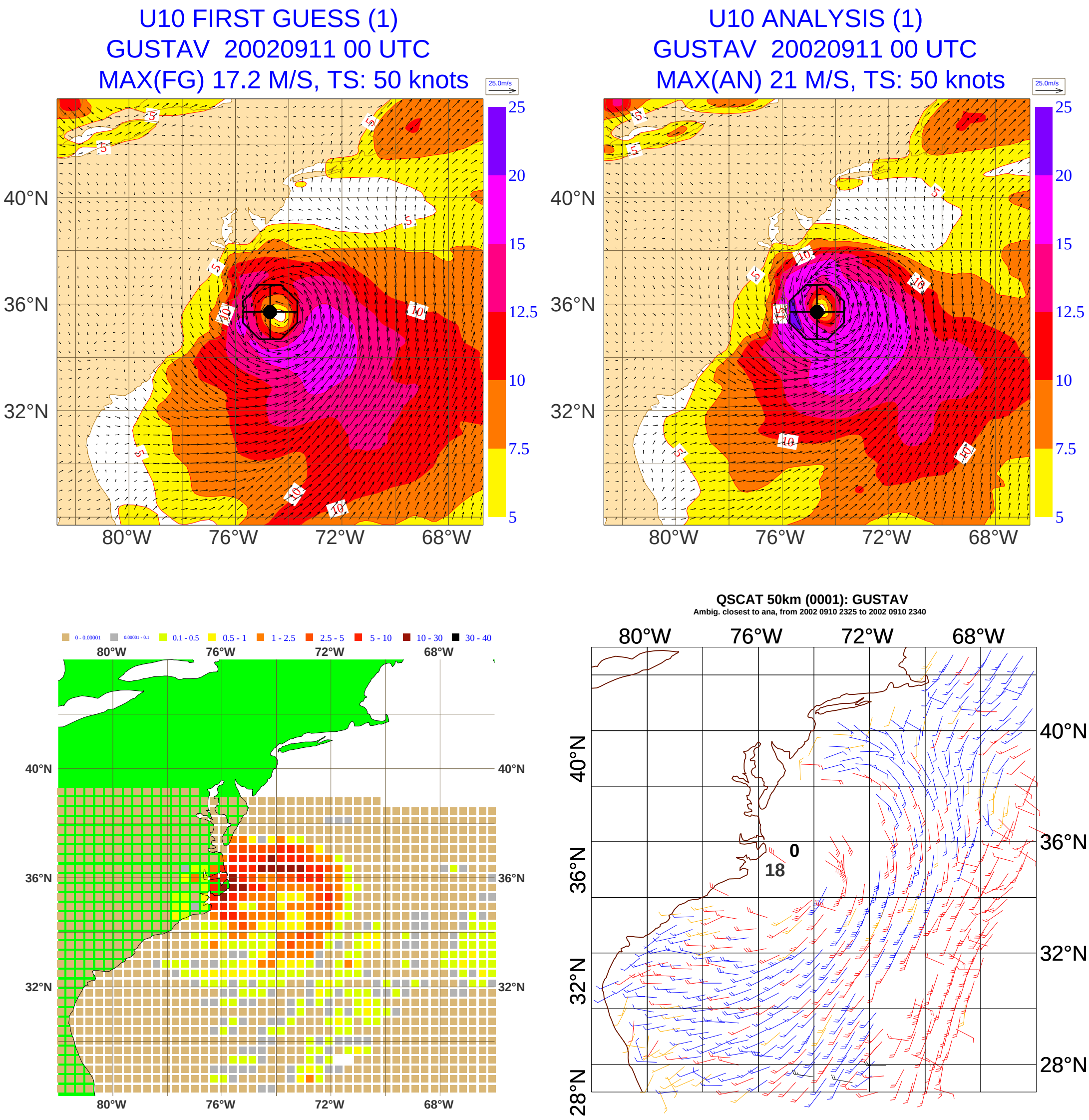

Figure 7: Assimilation of Hurricane Gustav for 20020911 00 UTC. First-guess winds and analysis winds are presented in the top left, top right panels respectively. The observed center location is indicated by the black cross; the maximal observed wind was 50 knots. Assimilated QuikSCAT winds are displayed in the lower right panel. Only blue vectors were active. Red winds are either rain contaminated or in the outer swath. Yellow vectors represent winds that are rejected on the basis of a lack in azimuth diversity. Collocated rain estimates based on TRMM measurements are given in the lower left panel.

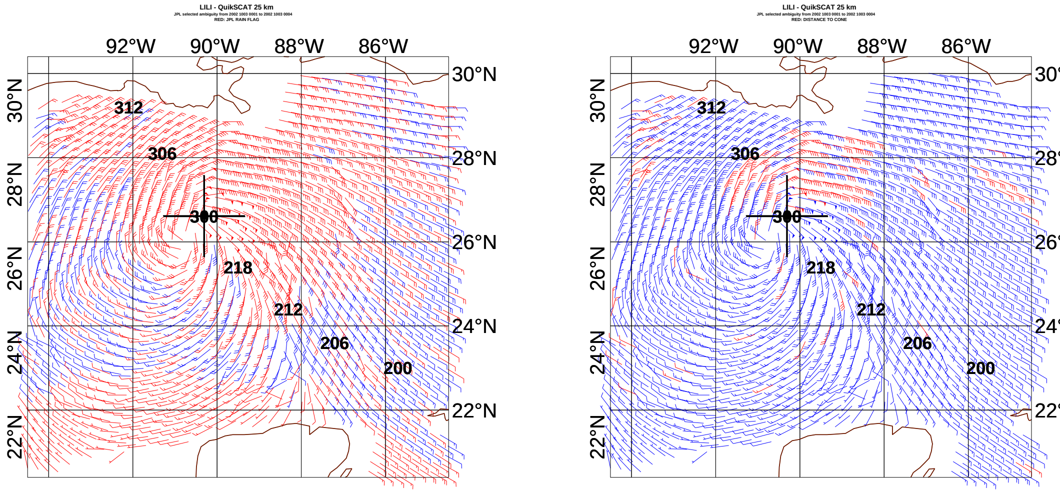

Figure 8: JPL selected wind vectors for 25 km QuikSCAT data for Hurricane LILI at 20021003 00 UTC. Red barbs are based on the JPL rain flag and KNMI cone-distance quality control for the left, respectively right panel. The observed center is indicated by the cross-wire; maximum observed winds were 125 knots (category 4).

use of a double-well cost function (2) and therefore the dynamical de-aliasing within the 4D-Var minimization steps. However, the de-aliasing can be performed in the trajectory calculations instead. The effect of this new strategy will be discussed.

Acknowledgements

Part of the work presented in this paper was funded by the European Space Agency (Project Ref. 15988/02/I-LG).

References

Andersson, E. and H. Järvinen, 1999: Variational quality control. *Q. J. R. Meteorol. Soc.* **125**, 697-722.

Courtier, P., J.N. Thépaut and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367-1388.

Leidner, S. M., R.N. Hoffman and J. Augenbaum, 2000: SeaWinds Scatterometer Real-Time BUFR Geophysical Data Product, User's Guide Version 2.3.0, NOAA/NESDIS.

Portabella, M., and A. Stoffelen, 2002: Quality control and wind retrieval for SeaWinds, *final report of the EUMETSAT QuikSCAT fellowship*, KNMI, de Bilt, the Netherlands.

Rabier, F., H. Järvinen, E. Klinker, J.F. Mahfouf and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**, 1143-1170.