

Document NWPSAF-KN-VS-017

Version 1.0

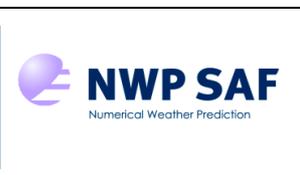
13-12-2016

Optimization of ASCAT data assimilation in global NWP

Wenming Lin¹, Marcos Portabella¹, Ad Stoffelen², Jur Vogelzang², Giovanna De Chiara³

1. *Institut de Ciències del Mar (ICM-CSIC), Barcelona, Spain*
2. *KNMI, De Bilt, The Netherlands*
3. *ECMWF, Reading, UK*



		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

Optimization of ASCAT data assimilation in global NWP

Wenming Lin¹, Marcos Portabella¹, Ad Stoffelen², Jur Vogelzang², Giovanna De Chiara³

1. Institut de Ciències del Mar (ICM-CSIC), Barcelona, Spain.

2. KNMI, De Bilt, The Netherlands

3. ECMWF, Reading, UK

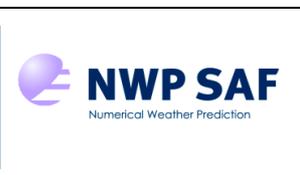
This documentation was developed within the context of the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF), under the Cooperation Agreement dated 29 June 2011, between EUMETSAT and the Met Office, UK, by one or more partners within the NWP SAF. The partners in the NWP SAF are the Met Office, ECMWF, KNMI and Météo France.

Copyright 2016, EUMETSAT, All Rights Reserved.

Change record			
Version	Date	Author / changed by	Remarks
0.1	11-10-2016	Wenming Lin	Initial draft
0.2	09-12-2016	Marcos Portabella	Review
0.3	11-12-2016	Ad Stoffelen	Review
0.4	12-12-2016	Marcos Portabella	Review
1.0	13-12-2016	Wenming Lin	Final version

Contents

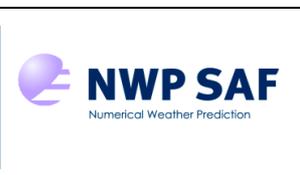
1 INTRODUCTION	4
2 DATA.....	7
3 QUALITY CONTROL	8
3.1 <i>QC INDICATORS</i>	8
3.2 <i>SENSITIVITY ANALYSIS</i>	9
3.3 <i>QUALITY CATEGORIZATION AND O/B ERRORS</i>	12
4 NEW ASCAT PRODUCTS	17
4.1 <i>AGGREGATION PROCESS</i>	17
4.2 <i>AN EXAMPLE</i>	18
4.3 <i>ERROR ANALYSIS</i>	20
4.4 <i>QUALITY CONTROL</i>	21
5 IFS EXPERIMENTS.....	23
5.1 <i>BIAS CORRECTION</i>	23
5.2 <i>EVALUATION PROCEDURE</i>	25
6 CONCLUSIONS.....	27
REFERENCES	29

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

1 Introduction

The first Advanced Scatterometer (ASCAT-A) onboard Metop-A satellite was launched on 19 October 2006, and the second ASCAT (ASCAT-B) onboard Metop-B satellite was launched on 17 September 2012. Over the last few years, considerable development of the ASCAT Wind Data Processor (AWDP), and in particular of the Two Dimensional Variational ambiguity removal (AR) scheme (2DVAR), has been carried out [Vogelzang *et al.*, 2007; Vogelzang *et al.*, 2009]. This has led to robust and consistent ASCAT Level 2 wind data products at various resolutions (25 km and 12.5 km) and backscatter (σ^0) processing types, i.e., nominal (Hamming window averaging) and coastal (box-car averaging) processing [Verhoef *et al.*, 2012]. More recently, the AWDP 2DVAR has been further improved in the context of a Numerical Weather Prediction Satellite Application Facility (NWP SAF) Associated Scientist study [Lin *et al.*, 2016a]. The background wind field provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) does not well resolve the mesoscale sea surface wind flow under increased wind variability conditions, such as in the vicinity of low-pressure centres, frontal lines, and moist convection. By using the default 2DVAR scheme in AWDP, the ECMWF background winds affect the resulting 2DVAR analysis, which in turn leads to ASCAT ambiguity removal errors under increased wind variability conditions. In [Lin *et al.*, 2016a], two important components of the variational approach have been further examined and improved for ASCAT AR purposes: first, the narrower Gaussian model error structure functions (GSF) have been replaced by a wider numerical structure function (NSF) [Vogelzang and Stoffelen, 2012]; second, the fixed observation (O) and background (B) error variance values have been replaced by situation-dependent O/B errors. A method to automatically estimate O/B errors from the ASCAT data itself, which takes into account the spatial representativeness of both error variances, has been proposed and thoroughly validated by Lin *et al.* (2015a).

It is shown that the influence of the misspecification of the ECMWF background errors in 2DVAR can be reduced by using NSF and flexible (i.e., situation dependent) O/B errors and that the ASCAT AR (selected) wind quality improves when compared against continuous buoy winds. Furthermore, in comparison with the “old” 2DVAR analysis (which uses GSF and fixed O/B errors), the new 2DVAR analysis winds are not only much closer to the selected ASCAT winds, but also of substantially higher quality and resolution, i.e., they show smaller-scale convergence/divergence features not present in the old 2DVAR analysis fields. This may be further exploited to improve higher dimensional variational (e.g., 4D-Var) data assimilation of scatterometer winds, such as that

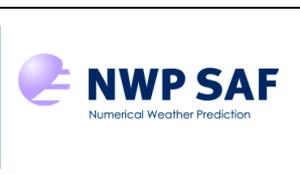
		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

of ECMWF. However, as reported by *Stoffelen et al.* (2014), besides the importance of suitable model error structure functions and situation-dependent O/B error estimation, several other important issues also need to be addressed in order to improve the impact of scatterometer data assimilation into global NWP models, including:

- The removal of unrepresentative winds in an improved wind Quality Control (QC). A more aggressive QC may bring higher forecast impact, but not necessarily a closer analysis fit to ASCAT. Since the inversion residual or Maximum Likelihood Estimator (MLE) is a good proxy for local wind variability, it is strongly correlated with high observation minus background (O-B) departures. Usually, unrepresentative observations are handled by a first guess check, but this depends on the highly variable quality of the background in such highly variable cases, which correspond to squall lines, gust fronts, etc. A more independent and observation-based QC, based on the MLE and Singularity Exponent (SE) indices needs to be tested following *Lin et al.* [2015a].
- A new ASCAT wind product more representative of the resolved NWP model scales. In global scatterometer wind data assimilation, particularly in the tropics, the spatial representativeness error is quite variable and spatially correlated naturally. At the moment scatterometer observations are thinned and assimilated with low weight, and as a consequence the small-scale information in scatterometer winds is not assimilated. Correlation in the spatial representativeness error is probably better dealt with by averaging the relatively-high resolution ASCAT wind vector cells (WVCs) to lower resolution in an aggregation process. Aggregated **O-B** error SDs are smaller than **O-B** SDs from thinned observations due to the lower representativeness error. Consequently, the aggregated winds have low representativeness error and thus low error correlation and can be assimilated by the global NWP without thinning.
- A suitable wind-speed bias correction. Biases are quite detrimental for mesoscale initialization, but exist due to dynamical closure and parameterisations in NWP models. Moreover, wind bias corrections to these artefacts brought by observations only last very short in the forecast cycle (i.e., a few hours). To provide more representative wind observations to the NWP data assimilation cycle, systematic model biases due to drag parameterization, mixing in stable conditions, moist convection, diurnal cycle or ocean currents could be geographically accounted for.

<p>The EUMETSAT Network of Satellite Application Facilities</p>		<p>Optimization of ASCAT data assimilation in global NWP</p>	<p>Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016</p>
---	---	--	--

In addition, initialisation of the short-lasting and generally modest-amplitude dynamic 100-km size (after thinning) atmospheric structures in NWP models is very challenging: it requires wind observations with not only high accuracy, but also abundant spatial coverage and temporal frequency. An international scatterometer constellation may provide the ocean surface component covering such challenging requirement, which however is beyond the scope of this work. This study mainly focuses on the above mentioned tasks, since they can potentially improve the ASCAT wind data assimilation in ECMWF 4D-Var. Section 2 describes the different types of wind sources used in this study. Section 3 reviews two essential quality indicators and their application in wind quality control and quality classification. In section 4, an aggregation process is described and applied to generate new ASCAT wind products which are more representative of the ECMWF resolved (wind) scales than the nominal product. Then a triple collocation analysis is adopted to quantify the spatial representativeness errors and the standard deviation errors of wind zonal (u) and meridional (v) components. In section 5, the experiments associated with IFS data assimilation are presented. First, the wind speed biases between the various ASCAT wind products and the ECMWF background are estimated, and a set of bias correction look-up tables is derived. Second, the impact experiments to be carried out soon at ECMWF are summarized. Finally, the conclusions can be found in Section 6.

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

2 Data

The data set used to derive the Observation (O) and Background (B) error Standard Deviations (SDs) consists of five years (January 2009- December 2013) of ASCAT-A Level 2 (L2) data collocated with ECMWF forecasts and moored buoy winds. Both ASCAT 25-km and 12.5-km coastal products are evaluated. The ASCAT data in Binary Universal Format Representation (BUFR) are provided by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF). The ECMWF forecast winds are acquired by interpolating in space and time three ECMWF 3-hourly forecast winds on a 0.5625° lat/lon grid to the ASCAT data acquisition location and time. ASCAT L2 BUFR files already include ECMWF winds. The buoys used in this study include the National Data Buoy Center (NDBC) moored buoys off the coasts of U.S.A., the Ocean Data Acquisition System (ODAS) buoys in the north-east Atlantic and British Isles inshore waters, the National Oceanic Atmospheric Administration (NOAA) Tropical Ocean Atmosphere (TAO) buoy arrays in the tropical Pacific, the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) Triangle Trans-Ocean Buoy Network (TRITON) buoys in the western Pacific, the Prediction and Research Moored Array in the Atlantic (PIRATA), and the Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA) at the tropical Indian Ocean.

The buoy winds are reported hourly by averaging the wind measurements over 10 minutes, and distributed through the Global Telecommunication System (GTS) stream, and subsequently quality controlled and archived at the ECMWF Meteorological Archival and Retrieval System (MARS). Note that the individual buoy observations are segregated into 1 m/s speed bins and 10° direction bins. The collocation criteria for this buoy data set are 30 minutes distance in time and 25-km distance in space from the ASCAT measurements. Only the ASCAT measurement closest to the buoy acquisition is used in case more than one WVC meets the collocation criteria. The total amount of collocations with MARS buoy winds is about 148,000 for the mentioned period.

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

3 Quality control

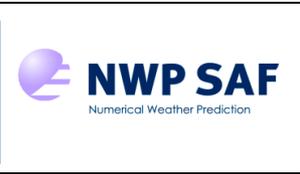
3.1 QC indicators

The wind QC method is developed by analyzing the characteristics of certain quality-sensitive parameters derived from the scatterometer data itself, and by optimizing the thresholds of these parameters in order to (filter) preserve as many (un-)representative quality winds as possible. A commonly used QC indicator is the inversion residual or maximum likelihood estimator (MLE), which depicts the minimum distance between the backscatter measurements and the scatterometer Geophysical Model Function (GMF) as follows

$$\text{MLE} = \frac{1}{3} \sum_{i=1}^3 (z_{mi} - z_{si})^2 \quad (1)$$

where $z_{mi} = (\sigma_{mi}^0)^{0.625}$ is the backscatter measurement of the i th beam in z-space, and $z_{si} = (\sigma_{si}^0)^{0.625}$ is the backscatter simulated through the geophysical model function (GMF), i.e., CMOD5n [Verhoef *et al.*, 2008; Verspeek *et al.*, 2010], using the solution wind vector as input. In general, the measurement sets $\{z_{mi}\}$ are located close to the GMF surface, indicating low MLE values and good quality wind retrievals. Occasionally, a large inconsistency between the triplets and the GMF is induced by other geophysical conditions than a WVC-mean wind, resulting in large MLE values. A MLE sign is defined by Portabella *et al.* (2012) for ASCAT, in order to better segregate different sea surface geophysical conditions. Triplets located inside the GMF cone surface are assigned with a positive MLE value, while those located outside the cone surface are assigned with a negative MLE value. Lin *et al.* (2015a) show that large positive MLE values correspond to high wind variability conditions, and in turn, to retrieved winds of reduced quality.

An image processing technique, known as singularity analysis (SA), has proven to be more effective than MLE in terms of detecting rainy wind vector cells (WVCs) [Lin *et al.*, 2014]. The singularity exponents (SEs) derived from SA provide quantitative information about the sub- and/or inter-WVC variability, and is a complementary QC indicator to MLE. The combination of SE and MLE has been proposed to improve the ASCAT wind QC in [Lin *et al.*, 2015b], and to classify and quantify the buoy, ASCAT and NWP wind uncertainty as a function of the wind variability, based on a triple collocation (TC) analysis of ASCAT, buoy, and the European Centre for Medium-range Weather Forecasting (ECMWF) model output [Lin *et al.*, 2015a]. In particular, a new SA implementation scheme, which combines the gradient measurements of multiple parameters in the analysis, is used to reveal the underlying geophysical phenomena. In this new implementation, SEs

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

are derived using several scatterometer-derived parameters (i.e., wind u , v components and MLE) simultaneously as input to SA [Lin *et al.*, 2016b], as shown below,

$$h(\mathbf{x}) = \frac{\log \left[\frac{T_\psi \|\nabla s\|(\mathbf{x}, r)}{\langle T_\psi \|\nabla s\|(\cdot, r) \rangle} \right]}{\log r_0} + o \left(\frac{1}{\log r_0} \right) \quad (2)$$

where $T_\psi \|\nabla s\|(\mathbf{x}, r)$ is the wavelet projection of the analyzed signal s at the point location \mathbf{x} and the scale factor r , $\langle T_\psi \|\nabla s\|(\cdot, r) \rangle$ is the mean value of the wavelet projection over the whole signal. The scale r_0 is defined as the smallest accessible scale. In the combined singularity analysis, $T_\psi \|\nabla s\|(\mathbf{x}, r)$ is given by

$$T_\psi \|\nabla s\| = \sqrt{\frac{(T_\psi \|\nabla u\|)^2}{[\sigma(T_\psi \|\nabla u\|)]^2} + \frac{(T_\psi \|\nabla v\|)^2}{[\sigma(T_\psi \|\nabla v\|)]^2} + \frac{(T_\psi \|\nabla \text{MLE}\|)^2}{[\sigma(T_\psi \|\nabla \text{MLE}\|)]^2}} \quad (3)$$

The estimated standard deviations $\sigma(T_\psi \|\nabla s\|)$ for u , v and MLE fields from one month (January 2013) of ASCAT 25-km data (in parenthesis the scores for ASCAT 12.5-km product) are:

$$\begin{cases} \sigma(T_\psi \|\nabla u\|) & = 0.7(0.6) \\ \sigma(T_\psi \|\nabla v\|) & = 1.4(1.2) \\ \sigma(T_\psi \|\nabla \text{MLE}\|) & = 3.7(2.8) \end{cases} \quad (4)$$

These fixed values are then used in our SA software for the entire process. As expected, the SD factors of the ASCAT 25-km product over 25 km are larger than those of the 12.5-km coastal product over 12.5 km, since the gradients of the former product are estimated over a larger distance than the latter. However, given that the distance is doubled, the values are rather close.

3.2 Sensitivity analysis

The sensitivity of MLE and SE to wind quality is evaluated by the Vector Root-Mean-Square (VRMS) difference between ASCAT and buoy winds,

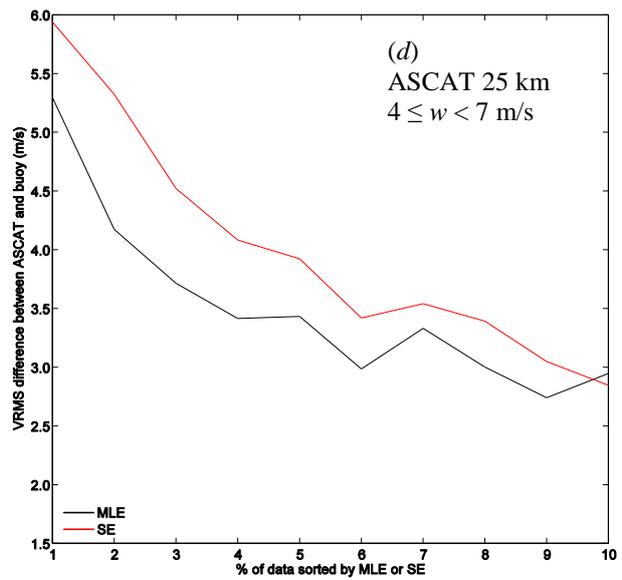
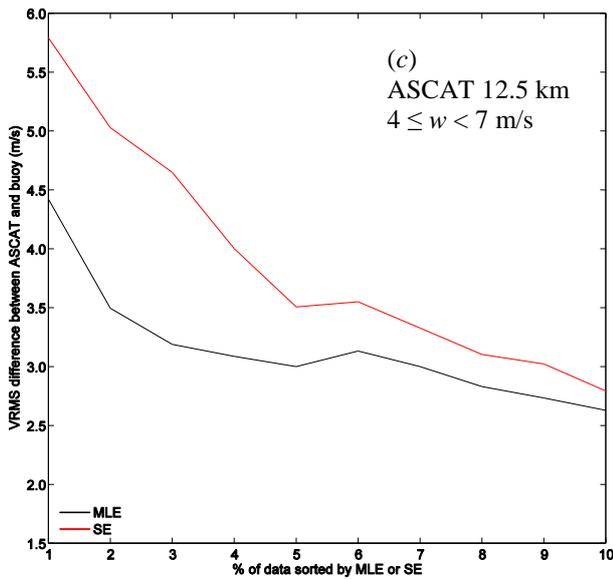
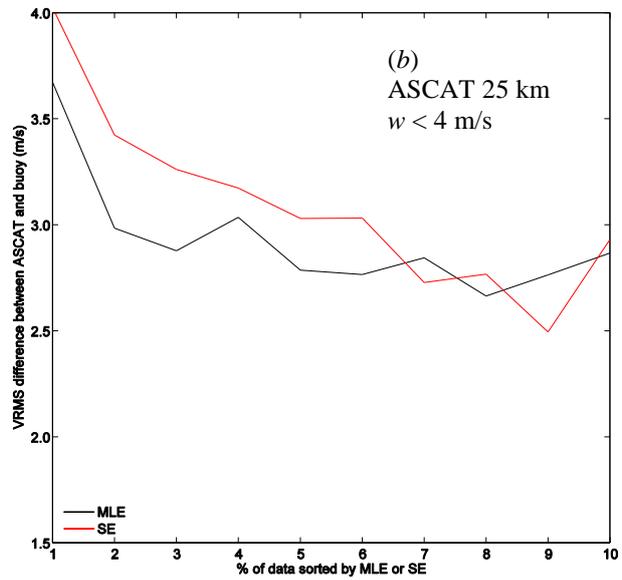
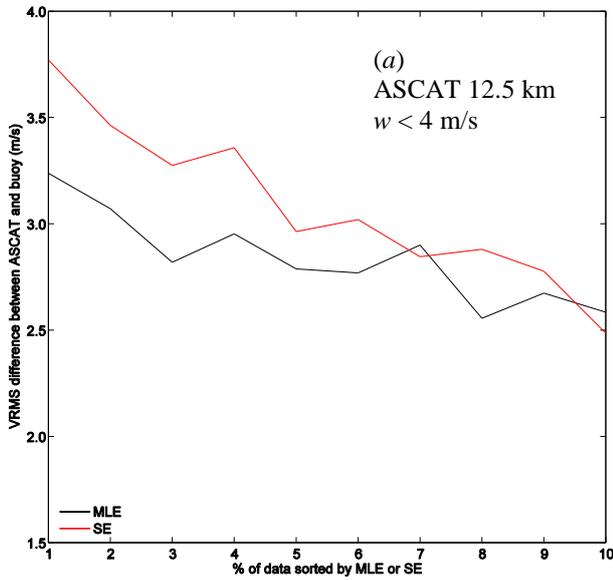
$$\text{VRMS} = \sqrt{\frac{1}{N} \sum_i^N \left[(u_i^{\text{scat}} - u_i^{\text{ref}})^2 + (v_i^{\text{scat}} - v_i^{\text{ref}})^2 \right]} \quad (5)$$

where $(u_i^{\text{scat}}, v_i^{\text{scat}})$ and $(u_i^{\text{ref}}, v_i^{\text{ref}})$ are the i^{th} scatterometer and reference wind vector (i.e., buoy in this case) respectively; and N is the total number of collocations. Overall, the ASCAT wind quality is a monotonic function of SE, i.e., the wind quality decreases as SE decreases. Regarding the signed MLE, in line with [Portabella *et al.*, 2012], the ASCAT wind quality decreases rapidly as

<p>The EUMETSAT Network of Satellite Application Facilities</p>		<p>Optimization of ASCAT data assimilation in global NWP</p>	<p>Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016</p>
---	---	--	--

MLE increases for $MLE \in (0, +\infty)$, while the wind quality degradation is generally small for $MLE \in (-\infty, 0)$. WVCs with negative MLE values are never rejected according to the current AWDP QC. We focus on the positive MLE domain, where the wind quality is indeed a monotonic function of MLE. Consequently, it is straightforward to inter-compare the different parameter sensitivities by binning the sorted data by the same percentage interval (of the total amount of data) as the MLE threshold decreases and the SE threshold increases. The threshold values corresponding to the different parameter bins are not relevant in the comparison anymore. On the other hand, it is easy to set the threshold values for filtering/sorting a given percentage of most-variable winds. In addition, since the two indicators are sensitive to wind speeds as well, the collocations are roughly separated into four categories: $w < 4$ m/s; $4 \leq w < 7$ m/s; $7 \leq w < 10$ m/s; and $w \geq 10$ m/s, where w is the ASCAT retrieved wind speed. The number of collocations is about 25, 50, 44 and 29 thousand for the four categories respectively.

Fig. 1 shows the VRMS difference between ASCAT and buoy winds as a function of the sorted bins by MLE (black curves) and SE (red curves) under different wind speed conditions. The binning is set to every 1% of the studied collocations. The left side shows the sensitivity analysis for the ASCAT 12.5-km coastal product, and the right side for the ASCAT 25-km product. Given a certain percentile, the VRMS scores of the 12.5-km product are generally smaller than those of the 25-km product. This does not necessarily mean that the QC indicators derived from the former are less sensitive to wind quality than the latter. Actually, this indicates that the former product resolves surface winds on a scale closer to buoy measurements than the latter, as evidenced by the lower global VRMS score of the former (2.33 m/s) as compared to that of the latter (2.38 m/s). Opposite results are obtained when using ECMWF winds as reference: the VRMS scores of the 12.5-km product are generally higher than those of the 25-km product (not shown), indicating that the former product resolves surface winds on smaller scales than the latter where the latter is closer to ECMWF. The scale analysis is further presented in Section 3.3. Nonetheless, it demonstrates that SE is more capable than MLE in terms of filtering the most unrepresentative (1% - 10%; depending on the speed category and WVC resolution) ASCAT winds w.r.t. buoys. Moreover, due to its inherent spatial integration, SE proves to be a more effective quality indicator and thus classifier for the finer (12.5-km) WVC resolution product. Since the gradients are estimated over a shorter distance (due to smaller WVC spacing), the regularity of the ASCAT-derived fields at a certain position is estimated more locally in case of higher WVC resolution.



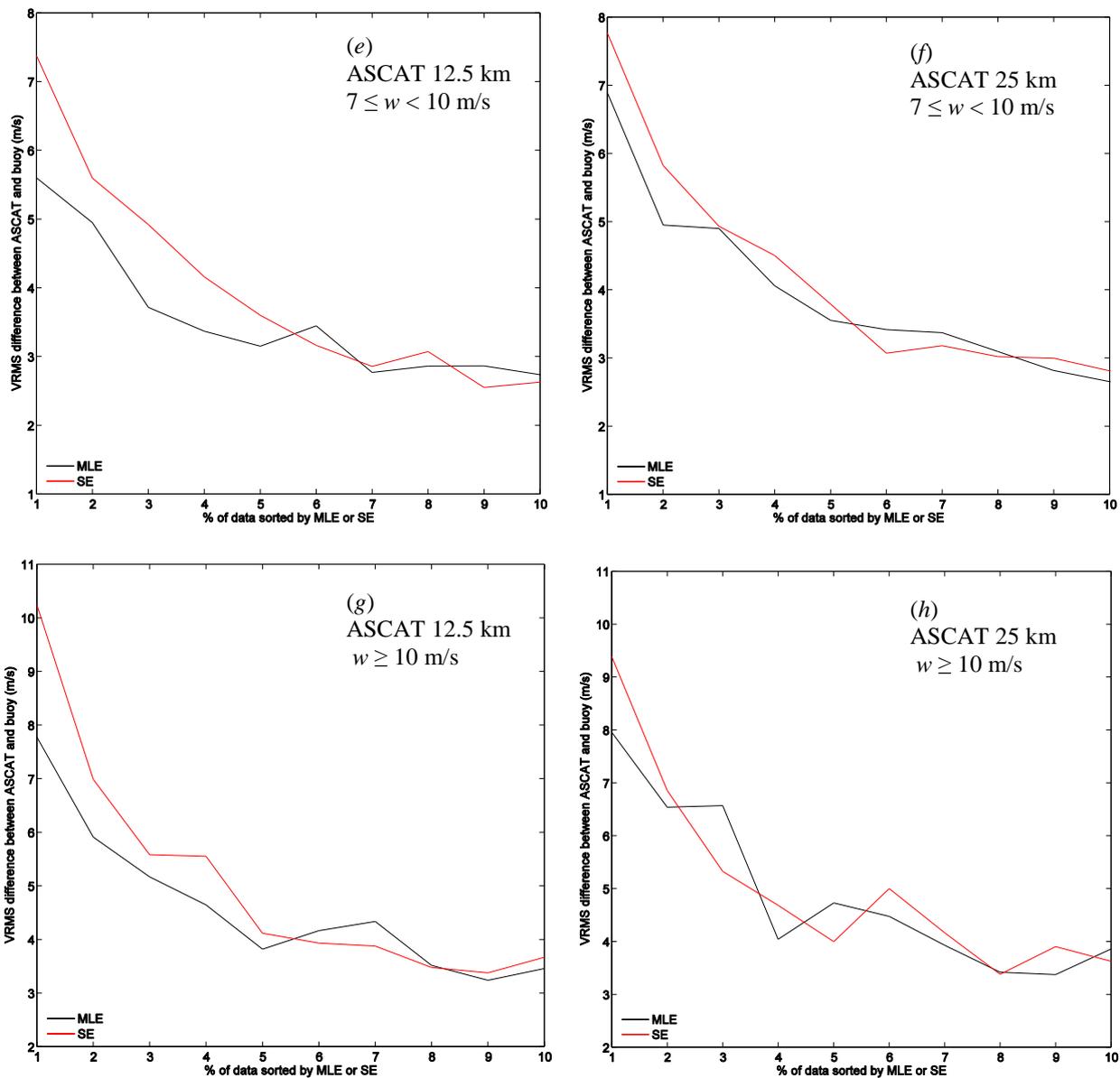
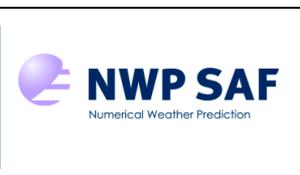


Fig. 1 Vector Root-Mean-Square (VRMS) difference between ASCAT and buoy winds as a function of the sorted bins by MLE (black curves) and SE (red curves), for the ASCAT 12.5-km (left panels) and the ASCAT 25-km (right panels) products, and different wind speed intervals, i.e., from top to bottom: $w < 4$ m/s; 4 m/s $< w < 7$ m/s; 7 m/s $< w < 10$ m/s; $w > 10$ m/s.

3.3 Quality categorization and O/B errors

In [Lin *et al.*, 2016a], MLE and SE intervals were used to classify the wind quality of the ASCAT 12.5-km product in different wind quality categories. However, the ASCAT data set was only separated into 6 categories due the limited amount of collocations. In addition, the categorization

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

was not optimized. For instance, categories 3 and 4 actually have similar wind quality while category 6 has too many data (~87%) [Lin *et al.*, 2016a]. The quality categorization is further refined in this section using the ASCAT-derived MLE and SE (Section 3.2) with the current extended collocation data set (see Section 2). To separate the data set in a more meaningful way than [Lin *et al.*, 2016a], the following criteria are used to determine the categorization:

- the winds within each category should have similar error statistics; in each category, the SE and MLE thresholds are chosen such that the SD of the ASCAT wind differences w.r.t. MARS buoys of WVCs with $MLE \in [MLE_1, MLE_2]$ are similar to those of WVCs with $SE \in (SE_1, SE_2]$. By doing this, the quality of the wind data in each MLE interval is similar to that of each SE interval;
- the ratio of the VRMS scores (i.e., VRMS difference between ASCAT and buoy) between neighbouring categories should not be too small; in practice, the difference of neighbouring VRMS scores is about 20%;
- the number of data in each category should be as large as possible, in order to reduce the discretization errors associated with the final O/B error fields;
- the minimum bin size is set to $> 1\%$.

Note that the above criteria only provide intuitive guidance to define the categories, but do not determine a unique setting of MLE and SE thresholds. For instance, it is proposed to separate the collocated ASCAT 25-km data set into 9 categories (see the corresponding MLE/SE intervals in the left side of Fig. 2). The percentiles from the most variable category (C1) to the most stable category (C9) are 1.1%, 1.2%, 1.4%, 2.5%, 4.0%, 12.5%, 18.4%, 24.1%, and 34.8% respectively, while those corresponding to the 6 categories defined in [Lin *et al.*, 2016a] are 1%, 1.3, 1.3%, 3%, 6%, and 87.4%, respectively. Note that the new categorization better discriminates the wind quality of the relatively more stable wind conditions as compared to the old categorization. A similar separation is applied to the collocated ASCAT 12.5-km data set, in which MLE/SE intervals are somewhat different than those shown in Fig. 2, but the category percentiles are similar to those of the 25-km product.

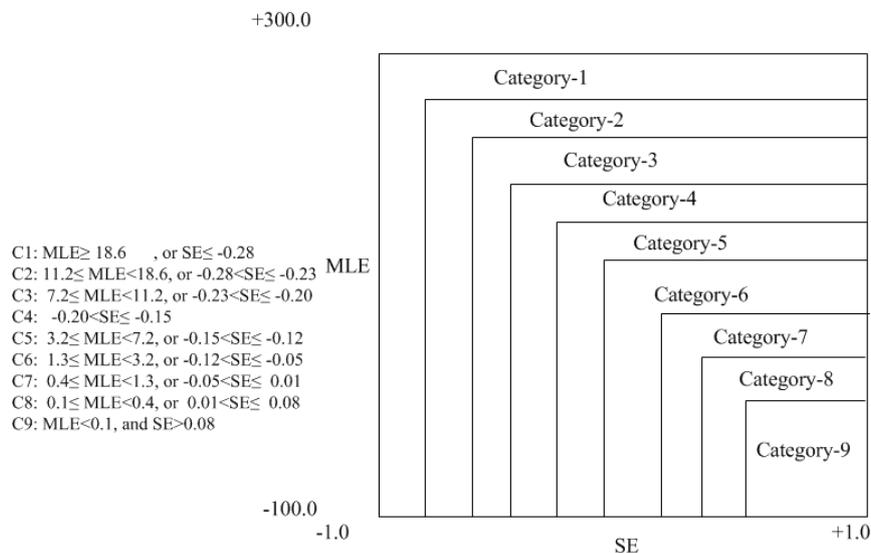


Fig. 2 Schematic representation of the defined categories used to develop situation-dependent O/B errors for the collocated ASCAT 25-km data set. The numerical intervals of SE and MLE are presented in the x - and y -axis respectively.

The random errors of wind data from three collocated sources can be quantitatively evaluated from the triple collocation (TC) method [Stoffelen, 1998]. The key parameter in TC analysis is the representativeness error, r^2 , which denotes the common small-scale variance present in higher spatial resolution systems, such as in-situ and scatterometer wind measurements, but absent in relatively lower resolution systems like the global NWP wind output (e.g., ECMWF). Conventionally, r^2 is estimated by integrating the difference between ASCAT and ECMWF wind spectra from the minimum scatterometer scale of 25 km to a scale of 800 km where the ECMWF deficit becomes negligible [Vogelzang et al., 2011]. In [Vogelzang et al., 2015], cumulative variance is calculated as a function of spatial scale, and the representativeness error is found to be given by the difference in cumulative variance of scatterometer and ECMWF wind components at a scale of 200 km. However, high variability wind regions, i.e., categories 1 to 3, are generally very localized, thus an alternative method to compute r^2 , introduced by Lin et al. (2015a), is used instead. The TC analysis is repeated for different r^2 values until a close intercalibration of the different wind sources is achieved. The r^2 value that determines a bias close to zero for both the calibrated buoy and ASCAT winds (w.r.t. ECMWF winds) is considered as the best estimated spatial representativeness error. To simplify the analysis, the r^2 values of u and v wind components are assumed to be identical for a certain category. Fig. 3 shows the estimated r^2 values for the above categories. As expected, the estimated r^2 values of the ASCAT 12.5-km product are larger than

those of the ASCAT 25-km product. Fig. 4 shows the estimated error SDs (i.e., $\sqrt{\frac{SD_u^2 + SD_v^2}{2}}$; on ECMWF scale) for ASCAT and ECMWF winds under different wind categories. Note that the ECMWF errors are independent of the ASCAT grid resolution.

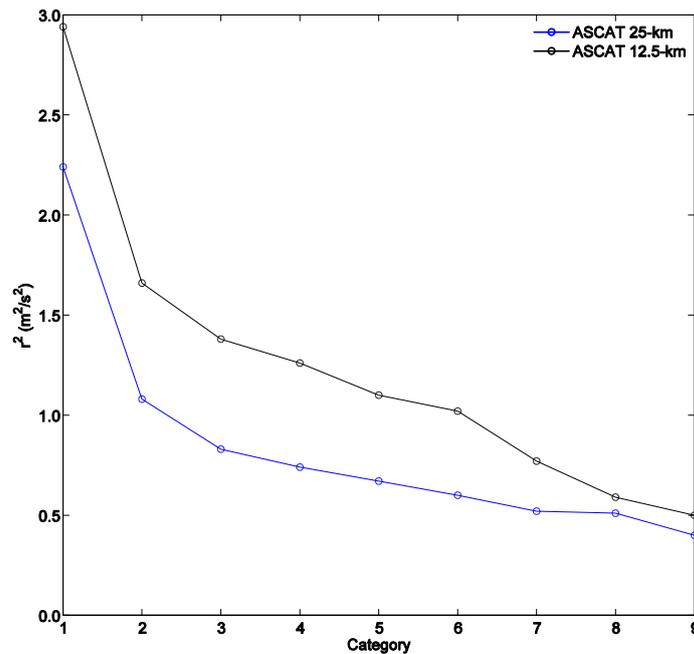


Fig. 3 The estimated r^2 values for the ASCAT 25-km (blue curve) and 12.5-km (black curve) products.

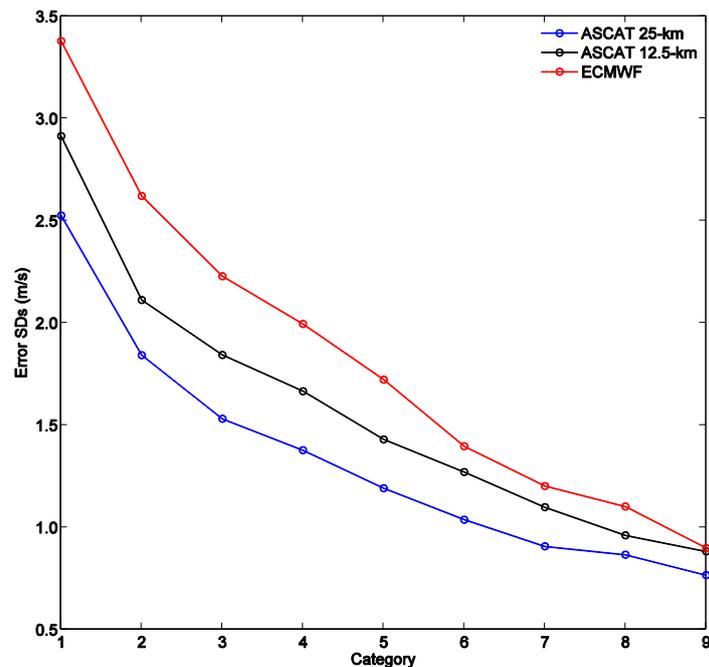
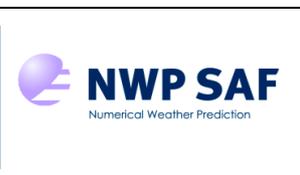


Fig. 4 The estimated error SDs of ASCAT 25-km (blue), ASCAT 12.5-km (black) and ECMWF (red) winds on the scale resolved by ECMWF for different wind categories.

The EUMETSAT Network of Satellite Application Facilities		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
---	---	--	---

The situation-dependent O/B errors can be estimated in near real time, since MLE and SE are all derived from the ASCAT data itself. However, the use of such **O/B** errors in 4D-Var data assimilation, such as the Integrated Forecasting System (IFS) at ECMWF, is not trivial. To date, one cannot set **B** errors freely in IFS, but can easily adjust the **O** (ASCAT) errors during the data assimilation. Consequently, it is proposed to modify the O errors according to the ratio of **O** and **B** errors (i.e., $R_{error} = \frac{SD_O^{TC}}{SD_B^{TC}}$; see Fig. 4) when testing the impact of situation-dependent O/B errors on data assimilation. That is, the observation error is set to $R_{error} \cdot SD_{IFS}$, where SD_{IFS} is the background error used by IFS.

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

4 New ASCAT products

4.1 Aggregation process

Currently, the ASCAT 25-km wind product with an inflated error variance value of 1.5 m/s and thinned by a factor of 16 to 100 km (i.e., only one WVC out of four WVCs along and across track is used, as shown in Fig. 5a) is assimilated into ECMWF IFS. This strategy does not assimilate the small-scale information in scatterometer winds, and does not reduce the inherent representativeness error of ASCAT observations. The latter is a particularly relevant issue that can potentially degrade the impact of scatterometer data assimilation into global NWP, since the representativeness error is spatially correlated. There are two ways to address the representativeness error in global data assimilation:

- The current data thinning technique. However, it may be more effective to deal with the correlation in the spatial representativeness error by assimilating denser ASCAT observations, but with reduced observation weight (*De Chiara et al., 2012; De Chiara et al., 2016*);
- To avoid thinning and provide aggregated wind observations on scales more representative of those resolved by the NWP model. This is called super-obbing in NWP data assimilation.

In this study, a lower ASCAT grid resolution is generated for further testing (without thinning) in the ECMWF IFS. Such new product is derived by averaging the winds within an $N \times N$ box (see an example with $N=3$ in Fig. 5b) over the ASCAT 25-km grid product. Since the ASCAT wind product on a 25-km grid is actually of about 50 km spatial resolution [*EUMETSAT, 2014*], one high-resolution row/column space is left out between neighboring low-resolution grid points in order to avoid spatial correlation due to the instrument. In this way, a 3×3 25-km WVC wind averaging is actually presented as a 100-km gridded product.

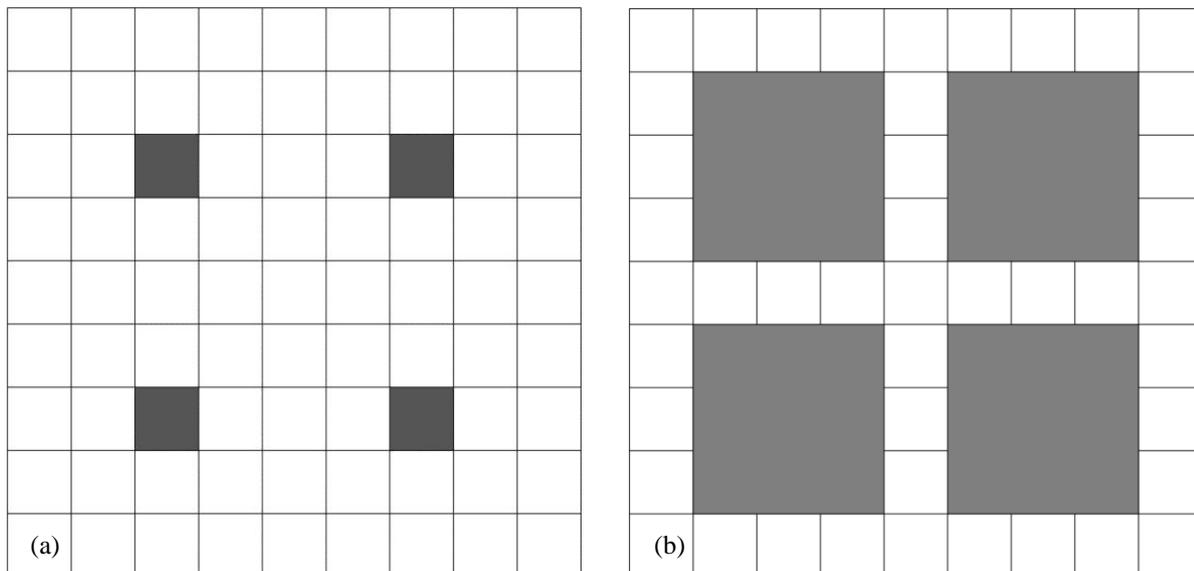


Fig. 5 Schematic representation of the grid points (each small box represents a 25 km× 25 km WVC) used in the global NWP data assimilation: (a) when a thinning of 4 (i.e., one WVC every 4 along and across track WVCs) is applied; (b) when the new low resolution product is used. Note that in the right panel 3×3 high-resolution WVC winds are averaged to produce the low resolution product.

Several new low-resolution ASCAT data products are generated in this Section for testing purposes in IFS, including:

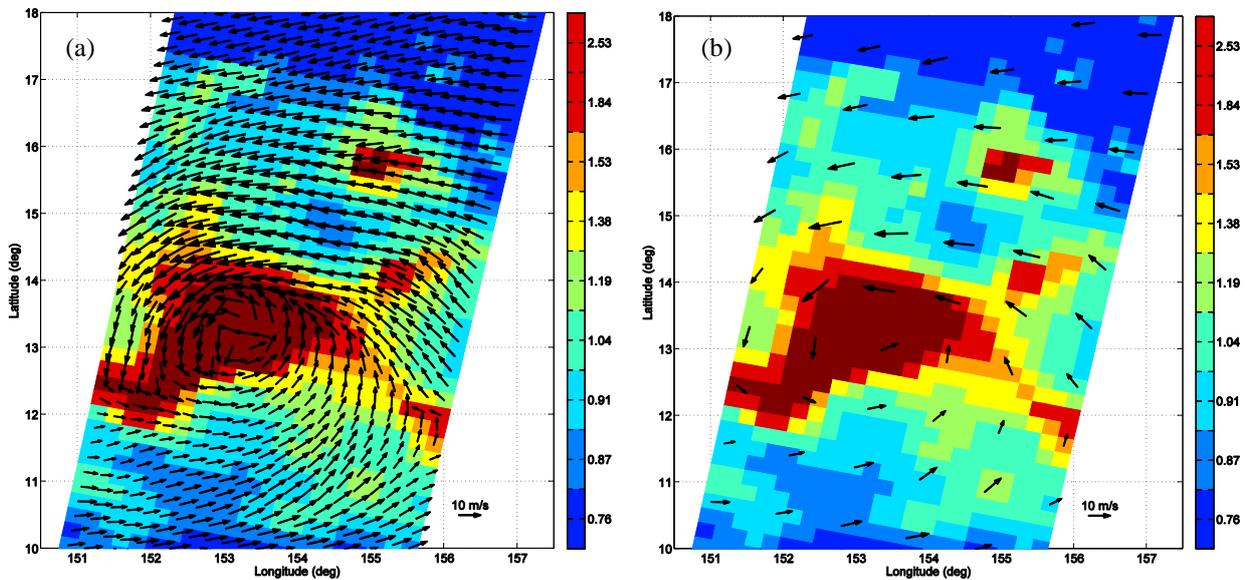
- 1) A 100-km low resolution derived from ASCAT 25-km product, for which the three out of four WVC averaging strategy shown in Fig. 5b is used (2×3 is used in case of edge WVCs); the number of WVCs in each swath side is **6**.
- 2) A 50-km low resolution derived from ASCAT 12.5-km coastal product, using the same averaging strategy as in 1); the number of WVCs in each swath is **10**.
- 3) A 62.5-km low resolution derived from ASCAT 12.5-km coastal product, for which the four out of five averaging strategy is used; the number of WVCs in each swath is **8**.

A thorough analysis of the representativeness errors and the **O/B** errors for the above low resolution products is carried out in Section 4.3.

4.2 An example

Fig. 6(a) shows an ASCAT 25-km wind field with its corresponding SD errors (see background colours), as estimated in Section 3. Note that several ASCAT selected wind vectors near the cyclone centre appear erratic. As discussed in [Lin *et al.*, 2016a], AR errors are mainly due to the

mislocation of background low pressure centre. In some cases, these errors can be mitigated by assigning more weight to the observation term in 2DVAR. The operational IFS only assimilates the ASCAT dual ambiguities at the WVCs indicated by the arrows in Fig. 6(b). Apparently, the thinned wind field does not well represent the wind frontal areas and the low pressure centre location, since it misses small-scale information. The relevant wind error statistics (such as O/B errors and representativeness errors) in Fig. 6(b) are though very similar to those in Fig. 6(a). Fig. 6(c) and (d) illustrate the ASCAT 100-km and 50-km low resolution winds derived from the aforementioned aggregation process. In both plots, the observation errors are reduced as compared to those of the operational ASCAT 25-km (Fig. 6a and 6b) and 12.5-km (not shown) products. The smallest observation errors correspond to the ASCAT 100-km grid product (Fig. 6c). Note that the positions are averaged in the same way as the observations, thus the two outer WVCs in Fig. 6(c) should be 12.5% closer to their neighbours than the other ones.



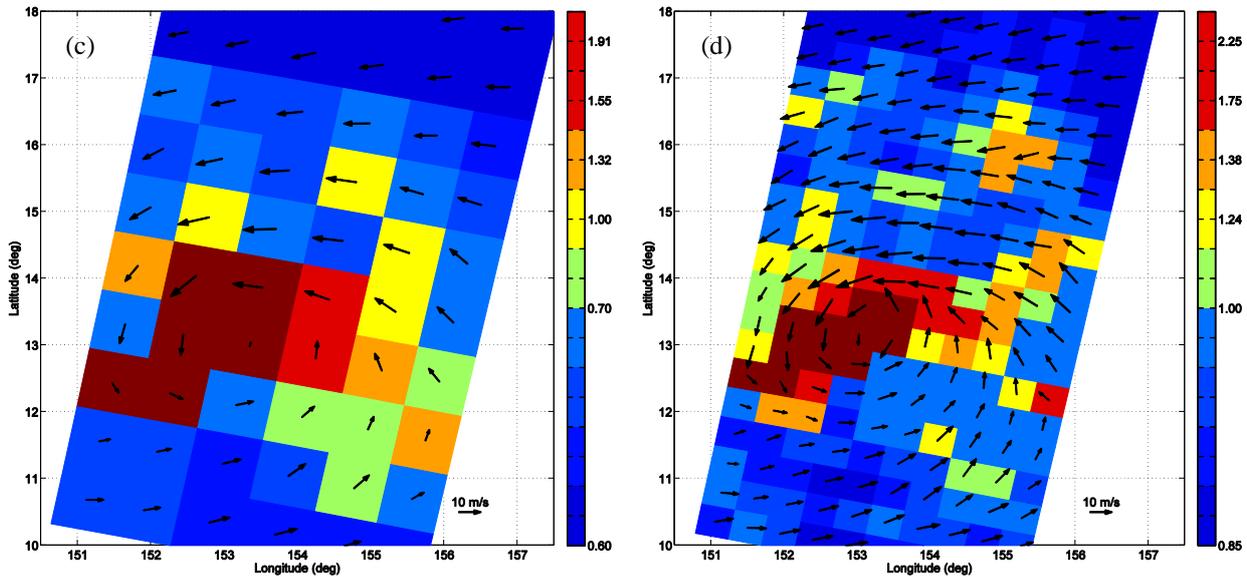


Fig. 6 ASCAT wind fields superimposed on the observation errors (SD): (a) ASCAT 25-km wind field acquired on July 31 2015, around 22:58 UTC; (b) ASCAT 25-km wind field thinned one out of four; (c) ASCAT 100-km and (d) 50-km low resolution winds derived from the aggregation process.

4.3 Error analysis

The aggregation process reduces the ECMWF-scale observation errors and the representativeness errors. Both types of errors are quantified by the aforementioned TC analysis in this section. To simplify the error reduction analysis, the triple collocated data sets are all separated into three quality categories (i.e., 4%, 6% and 90% of each data set) rather than 9, according to the criteria in Section 3.3. Fig. 7(a) shows the representativeness error as a function of wind quality category (x -axis) for the different ASCAT data sources (see the colours). As expected, the r^2 value decreases with decreasing ASCAT grid resolution and increasing category number. For high wind variability conditions (C1), the reduction of r^2 values with decreasing resolution is more remarkable than for stable wind conditions (C3). More interestingly, the r^2 values for ASCAT 100-km data seem to be independent of wind variability conditions, i.e., $r^2 \cong 0.2 \text{ m}^2/\text{s}^2$ for all three categories, indicating that this product is not only the most representative of ECMWF resolved scales but also equally representative at all wind variability conditions. Fig. 7 (b) shows the estimated observation error SDs and their corresponding error bars (indicating the uncertainty in the error estimation). In general, the distribution of observation errors is similar to that of the representativeness errors (Fig. 7a). Moreover, the error reduction due to the aggregation process is more remarkable for high wind

variability conditions than for stable conditions. In conclusion, the 100-km product is the most representative of ECMWF scales. In turn, it is the product with the smallest error (at ECMWF scales) and therefore the product with the highest weight in data assimilation.

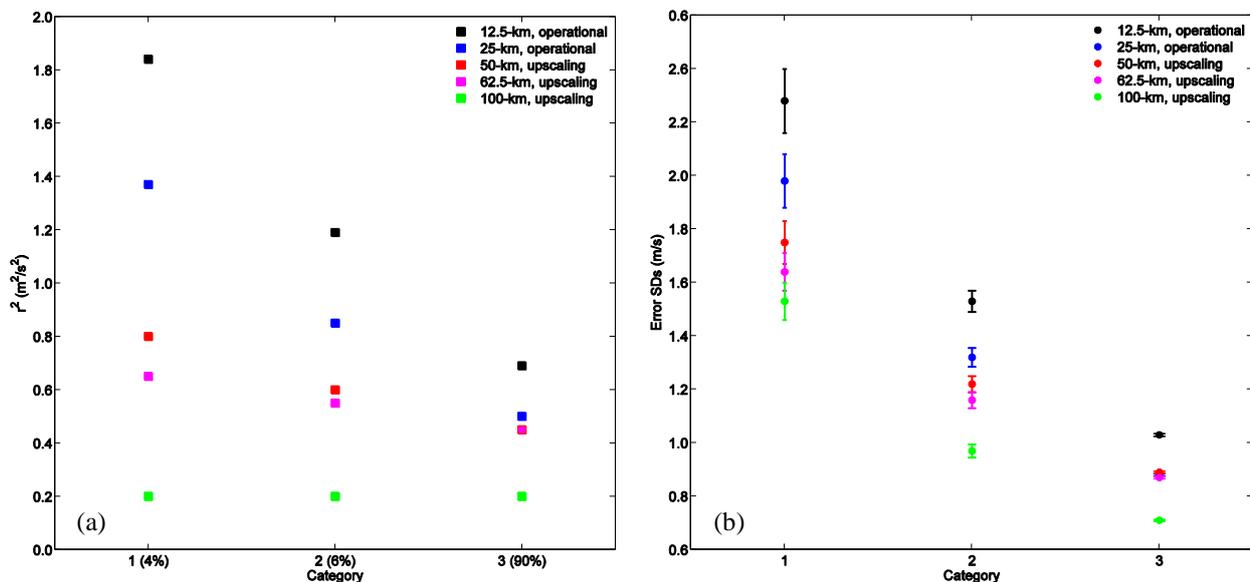


Fig. 7 (a) The representativeness error and (b) the estimated error SDs as a function of wind quality category for different ASCAT wind data sources. The error bars denote the uncertainty in the estimated observation errors.

4.4 Quality control

As shown in Figures 4 and 7, the ASCAT wind quality is generally higher than that of the ECMWF background at the scales resolved by the latter. Note that although IFS does perform its own quality control for all assimilated observation types, including ASCAT winds, in this section we propose an ASCAT QC for each of the developed low-resolution products for completeness. A similar QC procedure to that developed for the ASCAT 25-km and 12.5-km products [Lin *et al.*, 2016b] is applied to the low resolution products. A set of MLE/SE QC thresholds is proposed for the low resolution products, such that the QC rejection ratios of the low resolution data are similar to those presented in [Lin *et al.*, 2016b], i.e., 1.3%, 1.2% and 0.7% of data for winds in the following ranges: $4 \text{ m/s} < w < 7 \text{ m/s}$, $7 \text{ m/s} < w < 10 \text{ m/s}$ and $w > 10 \text{ m/s}$, respectively. Note that potentially less data may be rejected at low resolution such as 100 km, as ASCAT is more representative of ECMWF at these scales. Here we keep rejecting the same amount of data for different products to make the comparison be straightforward. Table 1 shows the QC thresholds (i.e., WVCs with $\text{MLE} > T_{\text{MLE}}$ or

$SE < T_{SE}$ are rejected) for each ASCAT product. Table 2 shows the overall VRMS statistics for the QC-accepted and rejected data separately.

Table 1: The proposed QC thresholds for different ASCAT data sets. For the ASCAT low resolution products, the MLE and SE values of the high-resolution product are averaged, similar to the wind averaging described in section 4.1.

Wind speed (m/s)		12.5-km	25-km	50-km	62.5-km	100-km
MLE	4 - 7	18.6	18.6	11.0	10.5	12.5
	7 - 10			15.0	12.0	13.0
	≥ 10			13.0	13.0	19.0
SE	4 - 7	-0.26	-0.26	-0.25	-0.24	-0.25
	7 - 10	-0.28	-0.30	-0.25	-0.24	-0.28
	≥ 10	-0.34	-0.44	-0.31	-0.29	-0.33

Table 2: The VRMS difference [m/s] between ASCAT and buoy winds (or ASCAT and ECMWF winds, in parenthesis) for the different data categories. Note that only ASCAT winds above 4 m/s are used.

	12.5-km	25-km	50-km	62.5-km	100-km
QC-accepted	2.25 (2.35)	2.31 (2.22)	2.35 (2.23)	2.41 (2.19)	2.62 (2.04)
QC-rejected	7.3 (6.6)	7.1 (6.0)	7.1 (5.6)	7.5 (5.6)	7.5 (5.2)

In line with the error analysis in section 4.3, the ASCAT winds are generally closer (more discrepant) to ECMWF (buoy) winds as the resolution decreases. Note though that for the ASCAT 12.5-km QC-rejected category, the VRMS score w.r.t. buoy winds is unexpectedly higher than that of the ASCAT 25-km and 50-km products. This is an indication that the rejected ASCAT 12.5-km winds are more variable than the rejected ASCAT 25-km or 50-km winds at buoy scale. This may be due to the fact that non-wind geophysical phenomena (such as rain, wind bursts) become more significant in the wind retrieval at higher resolution, which however needs further investigation.

5 IFS experiments

5.1 Bias correction

For mesoscale initialization, biases may be quite detrimental. A suitable wind-speed bias correction is thus required to fit with the model dependent parameterizations. In practice, the scatterometer wind biases are analysed off-line as a function of wind speed and WVC number, and then a look-up table is created and delivered to the DA system. That is, the wind speed bias is calculated as follows:

$$\Delta_{i,j}^A = \frac{1}{N} \sum_{k=1}^N (\Delta w_{i,j}^A) \quad (6a)$$

$$\Delta_{i,j}^E = \frac{1}{N} \sum_{k=1}^N (\Delta w_{i,j}^E) \quad (6b)$$

where Δw means the wind speed difference between ASCAT and ECMWF, i and j stand for the speed bin and WVC number respectively. The superscript A (E) indicates that the bias is estimated as a function of ASCAT (ECMWF) speed bin. Consequently, the average of 6(a) and 6(b) is taken as the bias correction factor, and the corrected ASCAT wind speed is written,

$$w' = w + \frac{\Delta_{i,j}^A + \Delta_{i,j}^E}{2} \quad (7)$$

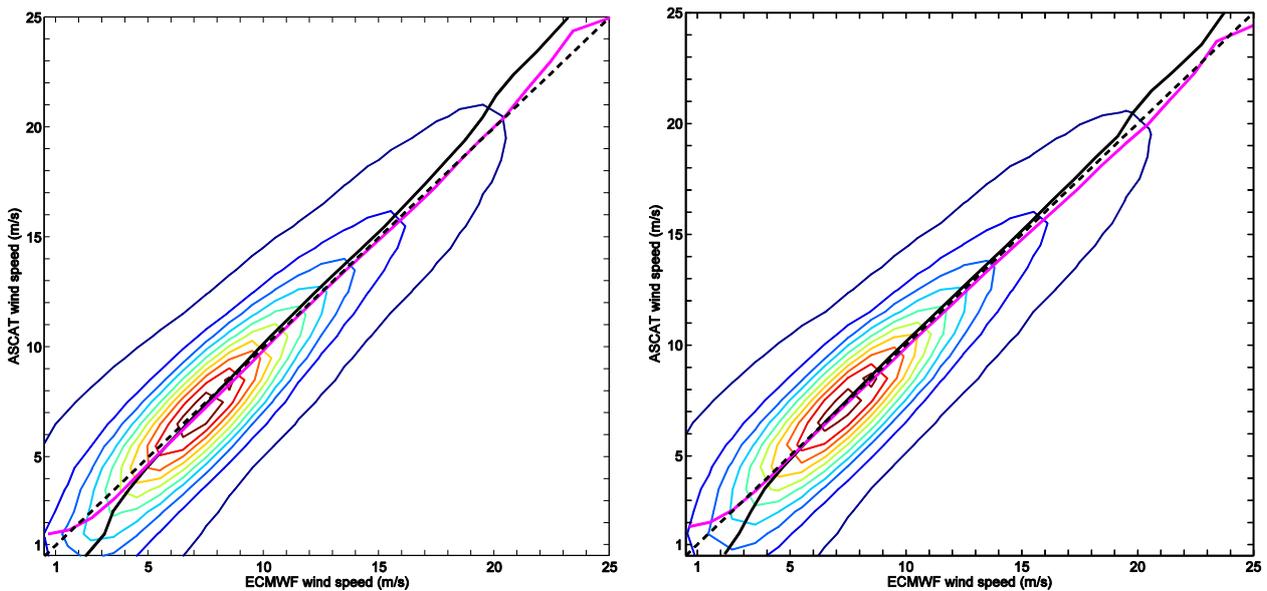


Fig. 8 2-D histogram of ASCAT 25-km wind speed versus ECMWF wind speed for WVC #1: (a) without bias correction; (b) after applying bias correction. The black (magenta) curve illustrates the mean ECMWF (ASCAT)

<p>The EUMETSAT Network of Satellite Application Facilities</p>		<p>Optimization of ASCAT data assimilation in global NWP</p>	<p>Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016</p>
---	---	--	--

wind speed at a set of ASCAT (ECMWF) speed bins, in which the binning is set to 1 m/s. Note that the ASCAT 25-km data in July-August 2015 are used in these plots.

Fig. 8(a) shows the 2-D histogram of ASCAT 25-km wind speed versus ECMWF wind speed for WVC #1 (i.e., the outermost left-swath WVC). The black curve shows the averaged ECMWF wind speed as a function of ASCAT wind speed bin, and the magenta curve shows the averaged ASCAT speed as a function of ECMWF wind speed bin. In both cases, the wind speed binning is set to 1 m/s. The wind speed bias (i.e., the asymmetry of these two curves along the diagonal) typically happens at low ($w < 4$ m/s) and/or high wind speed ($w > 15$ m/s) conditions: the operational ASCAT wind speed is lower (or higher) than ECMWF background at low (or high) winds. Fig. 8(b) shows the same 2-D histogram but for the ASCAT data after bias correction. It is clear that the asymmetry of black and magenta curves along the diagonal has been significantly reduced as compared to that of Fig. 8(a). Eq. (7) is thus effective in correcting the ASCAT wind-speed bias.

Fig. 9(a) shows the wind-speed bias correction values for different ASCAT left-swath WVC numbers. At low winds ($w < 4$ m/s), the correction values are positive for WVCs with relatively high incidence angle (WVC number ≤ 13), but negative for WVCs with relatively low incidence angle (WVC number > 13). In contrast, at medium and high winds, the correction values have the opposite trend. Fig. 9(b) presents the bias correction values for WVC #1 and for different ASCAT data products. The biases of ASCAT 25-km and 12.5-km product are actually identical. However, the biases of the low grid resolution products are slightly different from those of the higher resolution products. Since the aggregation process is implemented in u and v wind components, the low resolution wind speed (after averaging) is generally lower than that of the operational products, such that a larger positive (smaller negative) correction value is required for the aggregated low resolution data for low (medium and high) winds as compared to the high-resolution data. However, the WVC-dependent differences tend to be much larger than the resolution-dependent differences.

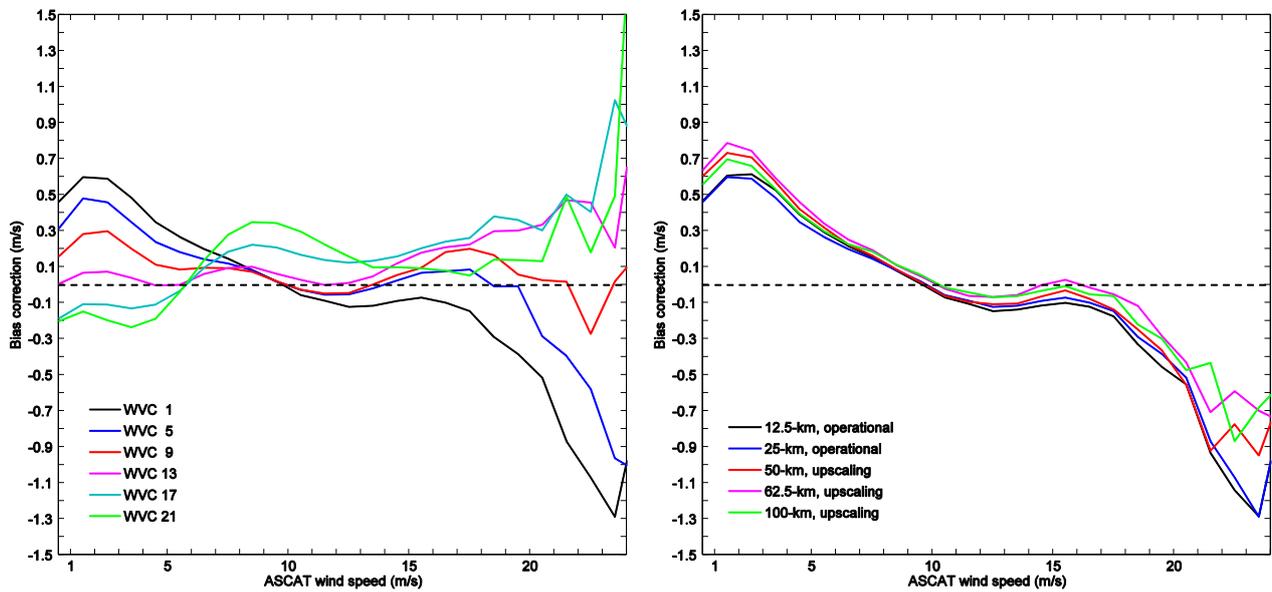


Fig. 9 The bias correction values as a function of ASCAT wind speed for (a) ASCAT 25-km data and different WVC numbers, and for (b) WVC #1 and different ASCAT data products.

The wind speed bias correction scheme ignores biases that are related to dynamical closure and parameterisation that are known to cause geographical biases between scatterometer and ECMWF wind vectors.

5.2 Evaluation procedure

Two months of ASCAT low resolution data (July-August 2015) are used to evaluate the impact of the refined QC and the super-obbing technique on the IFS data assimilation. Besides, several interesting cases, such as that of Fig. 6, will be analysed in particular to better understand the impact over Tropical areas. To evaluate the impact of each new feature added in this study (i.e., O/B situation-dependent errors, aggregated products), it is proposed to carry out the experiments in an incremental way (i.e., adding one feature at a time). Consequently, the following impact experiments will be carried out:

- 1) Operational IFS data assimilation approach, using ASCAT-25 km during the mentioned period;
- 2) Same operational data assimilation, but with the observation errors set according to the strategy proposed in Section 3.3;
- 3) New data assimilation with the aggregated 100-km product. In this case, the same amount of grid points and same grid spacing as those of 1) and 2) (except for the two outer WVCs at each

<p>The EUMETSAT Network of Satellite Application Facilities</p>		<p>Optimization of ASCAT data assimilation in global NWP</p>	<p>Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016</p>
---	---	--	--

swath) are assimilated. Note that a second 180° ambiguity solution is added for each low resolution grid point, since the current ASCAT assimilation scheme uses two ambiguities (180° apart) rather than the selected solution.

- 4) Data assimilation with the 12.5-km coastal product. This high resolution data is thinned by a factor of 16 to 50 km (i.e., only one WVC out of four WVCs along and across track is used);
- 5) New data assimilation with the aggregated 50-km and 62.5-km products respectively.

<p>The EUMETSAT Network of Satellite Application Facilities</p>		<p>Optimization of ASCAT data assimilation in global NWP</p>	<p>Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016</p>
---	---	--	--

6 Conclusions

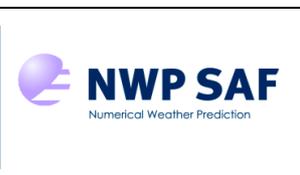
To improve the impact of ASCAT wind data assimilation in 4D-Var is more challenging than improving ASCAT AR in the AWDP 2DVAR, since the objectives of AR and NWP model initialisation prove slightly different. In 4D-Var, one cannot increase the weight of ASCAT observations as proposed for 2DVAR, because it may aggravate the overfitting problem above and near the areas with ASCAT observations, while this is of no detriment to AR. As a result, scatterometer winds are normally thinned and assimilated with low weight to avoid such overfitting effects. However, this leaves a large and quite variable spatial representativeness error which is only statistically accounted for by an enhanced observation error.

Several issues are addressed in this study in order to achieve a more effective assimilation of ASCAT winds at ECMWF. First, we increase the amount of triple collocations (buoy-ASCAT-ECMWF) in order to reduce the uncertainties of error estimation. The ASCAT/ECMWF wind data sets are further separated into 9 categories according to the wind quality (or wind variability) characteristics, and then a thorough analysis is carried out to evaluate the **O/B** errors of the operational ASCAT 25-km and 12.5-km products. In this way, the error structure of the ASCAT wind fields can be more precisely estimated than in *Lin et al.* (2015a). The analysis confirms that the observation error of the ASCAT 12.5-km product is higher than that of 25-km product at the scale resolved by ECMWF, which is mainly due to the fact that the former has a higher representativeness error than the latter.

Second, several ASCAT lower resolution (or aggregated) products are defined and generated to be further tested in the ECMWF data assimilation scheme. The relevant steps of the ASCAT processing, including quality classification, **O/B** errors estimation, and bias correction, are thoroughly analysed before the new products are delivered to ECMWF. The analysis shows that these low resolution products are more representative of the winds at ECMWF scale, and that the observation errors (including the representativeness errors) are also reduced as compared to the operational ASCAT (25-km and 12.5-km) products. In particular, the ASCAT 100-km product has not only the smallest observation errors but also the smallest representativeness errors, which latter are nearly constant regardless of the wind variability conditions. Finally, the bias correction is evaluated and a set of bias correction look-up tables are created and delivered to ECMWF together with the low resolution test data sets. In the near future, impact experiments of the low resolution

The EUMETSAT Network of Satellite Application Facilities	 NWP SAF Numerical Weather Prediction	Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
---	--	--	---

ASCAT wind products will be carried out at ECMWF. The outcome of these experiments will be later included as an appendix to this report.

		Optimization of ASCAT data assimilation in global NWP	Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016
--	--	--	---

References

De Chiara, G., P. Janssen, H. Hersbach, and N. Bormann, “Assimilation of scatterometer winds at ECMWF,” *Proceedings of 11th International Winds Workshop*, 2012.

De Chiara, L. Isaksen, and S. English, “Assimilation of satellite surface winds at ECMWF,” *ECMWF/ESA Workshop: Tropical modelling, observations and assimilation*, Reading, UK, Nov. 2016.

EUMETSAT, “Estimation of ASCAT normalised radar cross section: ATBD,” EUM/TSS/SPE/14/762689, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany, 2014. [Available at www.eumetsat.int.]

Lin, W., M. Portabella, A. Stoffelen, A. Turiel, and A. Verhoef, “Rain identification in ASCAT winds using singularity analysis,” *IEEE Trans. Geosci. Remote. Sens. Lett.*, 11(9), 1519–1523, 2014.

Lin, W., M. Portabella, A. Stoffelen, J. Vogelzang, and A. Verhoef, “ASCAT wind quality under high sub-cell wind variability conditions,” *J. Geophys. Res. Oceans*, 120, pp.5804-5819, Aug. 2015a.

Lin, W., M. Portabella, A. Stoffelen, A. Verhoef, and A. Turiel, “ASCAT wind quality control near rain,” *IEEE Trans. on Geoscience and Rem. Sens.*, 53(8), pp.4165-4177, Aug. 2015b.

Lin W., M. Portabella, A. Stoffelen, J. Vogelzang, and A. Verhoef, “On mesoscale analysis and ASCAT ambiguity removal,” *Q. J. R. Meteorol. Soc.*, 142, pp. 1745-1756, Apr. 2016a.

Lin, W., M. Portabella, A. Turiel, A. Stoffelen, and A. Verhoef, “An improved singularity analysis for quality control: application to low winds,” *IEEE Trans. on Geoscience and Rem. Sens.*, 54(7), pp. 3890-3898, Jul. 2016b.

Portabella, M., A. Stoffelen, A. Verhoef, and J. Verspeek, “A new method for improving scatterometer wind quality control,” *IEEE Trans. Geosci. Remote. Sens. Lett.*, 9(4), 579-583, 2012.

Stoffelen, A., “Toward the true near-surface wind speed: Error modeling and calibration using triple collocation,” *J. Geophys. Res.*, 103(C4), 7755–7766, doi:10.1029/97JC03180, 1998.

Stoffelen, A., J. Figa-Saldaña, A. Cress, J. Cotton, G. De Chiara, T. Valkonen, C. Payan, and G.J. Marseille, “Scatterometer mesoscale NWP data assimilation,” *12th International Winds Workshop*, Splinter discussion session summary, Copenhagen, Denmark, 15-20 June, 2014.

Verhoef, A., M. Portabella, A. Stoffelen, “High-resolution ASCAT scatterometer winds near the coast,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2481-2487, Jul. 2012.

Verhoef, A., M. Portabella, A. Stoffelen, and H. Hersbach, “CMOD5.n- the CMOD5 GMF for neutral winds,” KNMI, De Bilt, The Netherlands, Ocean and Sea Ice SAF Technical Note, *SAF/OSI/CDOP/KNMI/TEC/TN/3* 165 v.1. Available from <http://www.knmi.nl/scatterometer/publications/>.

<p>The EUMETSAT Network of Satellite Application Facilities</p>		<p>Optimization of ASCAT data assimilation in global NWP</p>	<p>Doc ID : NWPSAF-KN-VS-017 Version : 1.0 Date : 13-12-2016</p>
---	---	--	--

Verspeek, J., A. Stoffelen, M. Portabella, H. Bonekamp, C. Anderson, and J. Figa-Saldaña, “Validation and calibration of ASCAT using CMOD5.n,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 386–395, Jan. 2010.

Vogelzang, J., “Two dimensional variational ambiguity removal (2DVAR),” *Report NWPSAF-KN-TR-004 Version 1.0*, UKMO, UK, 2007.

Vogelzang, J., A. Stoffelen, A. de Verhoef, and H. Bonekamp, “Validation of two-dimensional variational Ambiguity removal on SeaWinds scatterometer data,” *J. Atmos. Ocean. Technol.*, vol. 26, no. 7, pp. 1229–1245, Jul. 2009.

Vogelzang, J., and A. Stoffelen, “NWP model error structure functions obtained from scatterometer winds,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2525–2533, Jul. 2012.

Vogelzang, J., A. Stoffelen, A. Verhoef, and J. Figa-Saldaña, “On the quality of high-resolution scatterometer winds,” *J. Geophys. Res.*, 116, C10033, doi:10.1029/2010JC006640, 2011.

Vogelzang, J., G.P. King, and A. Stoffelen, “Spatial variances of wind fields and their relation to second-order structure functions and spectra,” *J. Geophys. Res.Oceans*, 120, doi:10.1002/2014JC010239, 2015.