Invitation to Tender (ITT)
ESA: AO 1-8365/15/NL/CT

SCIENTIFIC ASSESSMENT OF TDS-1 GNSS-R SCATTEROMETRIC MEASUREMENTS (TGSCATT)

Technical Note 5: Evaluation of SGR-ReSI Level 2 wind quality: triple collocation analysis

Wenming Lin, Marcos Portabella

Institute of Marine Sciences (ICM-CSIC)

25 May 2018

DOCUMENTATION CHANGE RECORD

Issue / Revision :	Date :	Change :	Description :
Version 0.1	2018-05-25	Wenming Lin	Draft version

Table of Contents

1. Introduction

Work Package (WP) 5000 uses the knowledge acquired in previous TGSCATT WPs to consolidate, validate and document the Level-1 to Level-2 inversion algorithms for TDS-1. Validation entails the development of extended TDS-1 matchup datasets with new independent measurements to validate the performance of the inversion algorithms over a wide range of conditions. The task should result in documentation of the TDS-1 Level 2 inversion algorithms in the form of Algorithm Theoretical Basis Documents (ATBD).

Over the course of the TGSCATT project, a new Level 2 SGR-ReSI wind dataset, using the so-called Calibrated Bistatic Radar Equation (CBRE) approach [1], has been made available to the team. Such wind dataset addresses several systematic and random errors present in earlier SGR-ReSI data versions and is therefore of high interest to the project. As such, WP5000 has been somewhat redefined and is now focused on the objective validation of the new wind dataset.

2. Data and methodology

In this study, the triple collocation (TC) analysis [2] is used to evaluate the quality of SGR-ReSI wind speed retrieved with the CBRE algorithm. Two months of collocated SGR-ReSI wind data (May – June 2015), ASCAT L2 winds and ECMWF background winds are analyzed. The ASCAT data is provided by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice (OSI) Satellite Application Facility (SAF). It already includes collocated ECMWF model wind output, which are estimated by interpolating three ECMWF 3-hourly forecast winds on a 62.5-km grid both spatially and temporally to the scatterometer data acquisition location and time, respectively. The collocation criteria for this data set are 30 minutes distance in time and 25 km distance in space from the SGR-ReSI acquisitions. The total amount of collocations is about 51 thousand, with 15 thousand in Eclipse mode and 36 thousand in sunlight mode.

The TC method has been widely used as a tool for intercalibration and individual error assessment of three different collocated geophysical variables, provided that errors are additive, error distributions are close to Gaussian, and that the collocated data sources are independent. Given three measurement systems W_i, i = 1, 2, 3, which represent ASCAT, SGR-ReSI and ECMWF wind respectively, the measurements and measurement errors are approximated by the following linear expression,

$$W_i = a_i w + b_i + \delta_i \tag{1}$$

where w is the common quantity in this study, i.e., the true wind speed certain spatial scale, a_i and b_i stand for the scaling and bias calibration coefficients respectively, and δ_i for the random measurement error. δ_i is assumed to be unbiased, and its variance does not change with w. The random observation errors are assumed to be uncorrelated with w, $\langle \delta_i w \rangle = 0$. ASCAT and SGR-ReSI winds resolve smaller turbulent scales than ECMWF, and the variance common to these smaller scales, $r^2 = \langle \delta_1 \delta_2 \rangle$, is part of the observation errors δ_1 and δ_2. By definition, r^2 is the correlated part of the representativeness errors of W_1 and W_2. Furthermore, since W_3 does not include these smaller scales, its observation error δ_3 is independent of δ_1 and δ_2, so $\langle \delta_i \delta_3 \rangle = 0$ (i=1, 2). Note that in the TC definition, w contains only scales resolved by all three systems. In summary, wind errors of different systems are all assumed to be uncorrelated, except for the spatial representative error [2], due to the turbulent scales only resolved by systems 1 and 2.

Since ASCAT winds have been proven to be of high quality, the wind system 1 is chosen as calibration reference, then the other two calibration factors are given by respectively

$$\begin{cases} a_2 = \dfrac{M_{23}}{M_{31}} \\ a_3 = \dfrac{M_{23}}{M_{12} - r^2} \end{cases} \tag{2}$$

where $M_{ij} = \langle W_i W_j \rangle$ (i, j =1, 2, 3) stands for the mixed second-order moment of system i and j. The bias correction factors b_i are given by,

$$b_i = M_i - a_i M_2 \quad (i\text{=2,3}) \tag{3}$$

where M_i stands for the first order moment of the i^{th} system. After calibration, the error variances estimated on the scale of w (ECMWF) for each wind system are given by

$$\langle \delta_i^2 \rangle = M_{ii} - \varepsilon^2 \tag{4}$$

The quantity ε^2 has several different expressions, e.g., $\varepsilon^2 = M_{12} - r^2 = M_{23} = M_{13}$ and denotes the common true variance in the three measurement systems. Note that to obtain the error variances on the scale of wind system 2, r^2 has to be subtracted from the above buoy and scatterometer error variances, and added to the ECMWF error variance. In practice, the TC analysis is implemented using an iterative approach. A 4-sigma quality control (QC) is used to filter out the outliers before applying Eqs. (2)-(4).

3. Representativeness error estimation

Ideally, the three data sources are well inter-calibrated when the TC analysis converges. This can only be achieved with consistent r^2 value. Setting a wrong r^2 value leads to miscalibrated results among the three data sources. Note that the r^2 value only affects the calibration coefficients of system 3, but not system 1 or 2. Therefore, an effective way of estimating r^2 is to repeat the TC analysis for different r^2 values until an "optimal" intercalibration of system 3 with respect to systems 1 and 2 is achieved. Practically, the regression slope of system 2 versus system 1 (denoted as s_{21}) does not equal to 1.0 after TC calibration, therefore one cannot find an r^2 value which leads to s_{13} = 1.0 (regression slope of system 1 versus system 3) and s_{23} = 1.0 (regression slope of system 2 versus system 3) simultaneously, because one has $s_{13} = s_{23}/s_{21}$ in the symmetric regression implementation. In this study, we first find the r^2 values which lead to s_{23}=1 (denoted as $r^2_{s_{23}}$) and s_{13}=1 ($r^2_{s_{13}}$), and then define the overall r^2 value as $(r^2_{s_{23}} + r^2_{s_{13}})/2$. Actually, this is equivalent to find the r^2 value which leads to $s_{23}= s_{31}$ or $s_{32}= s_{13}$ after TC calibration.

To verify the above proposed method for determining the r^2 value, the slope values as a function the representativeness error for the triplet buoy-ASCAT-ECMWF (year 2009 - 2014) are examined in Fig. 1. Here the ASCAT winds are selected as calibration reference. The three panels stand for the results for wind zonal (u), meridional (v) components and wind speed respectively. The determined r^2 values for the u, the v, and the wind speed components are 0.47, 0.44, and 0.52, respectively. The representativeness error values for u and v (notably the former) are close to those estimated by [3], [4] in a different period of time. Therefore, the above method is effective in searching the r^2 value for a triple collocated wind data set. Note though that the TC analysis of the wind speed component requires many more iterations (i.e., 30) than that of the u/v components before it converges. This is probably due to the fact that, in contrast with the u and v components, the wind speed distribution is non-Gaussian.

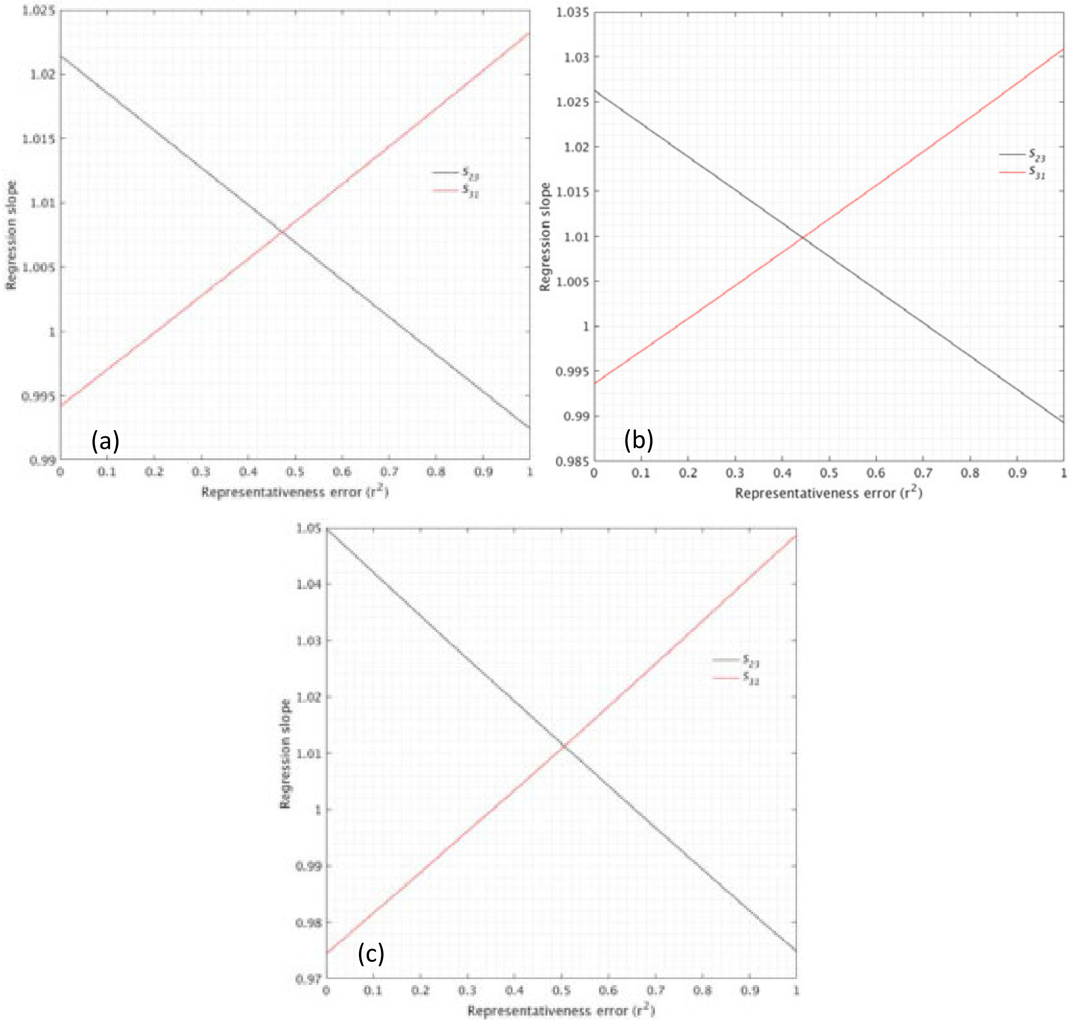

Fig. 1 Slope values as a function of the representativeness error (r^2) for the buoy-ASCAT-ECMWF triplet, for the (a) *u* component, (b) *v* component, and (c) wind speed.

Table 1 summarizes the TC results of the buoy – ASCAT – ECMWF wind speed triplet, with an r^2 value of 0.52. It shows that the estimated SD error of ASCAT wind speed is generally lower than that of ECMWF wind speed.

Table 1: TC calibration coefficients and the estimated SD errors of wind speed on the scales resolved by ASCAT (fourth column) and ECMWF (fifth column)

	Regression slope	**Scaling factor**	**Bias correction**	**ε [m/s] SCAT scale**	**ε [m/s] Model scale**
ASCAT	$S_{23} = 1.017$	1.00	0.00	0.34 ± 0.001	0.80 ± 0.003
ECMWF	$S_{13} = 0.984$	1.05	0.18	1.25 ± 0.007	1.02 ± 0.005

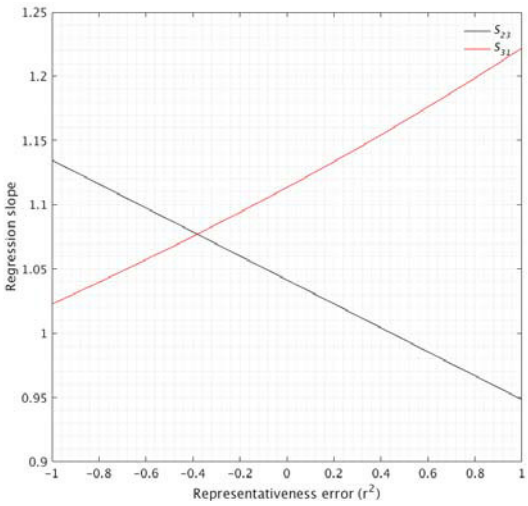

Fig. 2 Slope values as a function of the representativeness error (r^2) for the ReSI-ASCAT-ECMWF wind speed triplet. Note that only SGR-ReSI data acquired in sunlight mode are used in here.

Using the above r^2 value of 0.52 in the TC analysis of the ASCAT – SGR-ReSI –ECMWF wind speed triplet could be a straightforward way. However, one should note that SGR-ReSI actually resolves winds at a larger scale than those of the buoy and/or ASCAT, which implies that the r^2 value for the triplet ASCAT – SGR-ReSI –ECMWF should be smaller than 0.52 [5]. Alternatively, one could estimate the r^2 value using the same approach as in Fig. 1. Fig. 2(a) shows the slope values as a function of the representativeness error for the ASCAT – SGR-ReSI –ECMWF triplet. A negative r^2 value (-0.38) is found, which is indeed not physical, and thus implies that the SGR-ReSI winds are too noisy, or the spatially representative scale of SGR-ReSI winds may be the largest of all three data sources. This needs further investigation.

Alternatively, we propose to search the r^2 value which leads to the regression slope (ASCAT vs ECMWF) of the ASCAT –SGR-ReSI –ECMWF being the same as that of Table 1, with the assumption that the regression slopes of ASCAT and ECMWF winds should be identical in both the buoy-ASCAT-ECMWF and the ASCAT-SGR-ReSI-ECMWF triplets. [Note that the sampling in both triplets is different because there are simply not enough quadruplets (buoy – ASCAT - SGR-ReSI - ECMWF) to perform TC (see section 4). Although the different sampling may lead to differences in the TC analysis, these are assumed to be small.] Using the above mentioned assumption, the determined r^2 values for the triplets in sunlight mode and Eclipse mode are 0.26 and 0.09, respectively.

4. Validation results

Figure 3 shows the contour plots for the triple collocated data set in sunlight mode before and after TC. The red lines represent the linear regression of the plots. The top panels show the two-dimensional (2-D) histograms of SGR-ReSI wind speed versus ASCAT winds, the middle panels show the 2-D histograms of SGR-ReSI versus ECMWF wind speed, and the bottom panels show the 2-D histograms of ECMWF versus ASCAT wind speed, before TC (left panels) and after TC (right panels). Similar results are achieved for the triplet obtained in Eclipse mode (not shown). Several conclusions can be drawn:

a) The SGR-ReSI winds are closer to ASCAT winds than to ECMWF winds;
b) The TC calibration improves the correlation coefficients among the three data sets, and the statistical scores of ECMWF winds w.r.t. ASCAT;
c) The TC calibration does not improve the statistical scores of SGR-ReSI winds w.r.t. ASCAT or ECMWF.

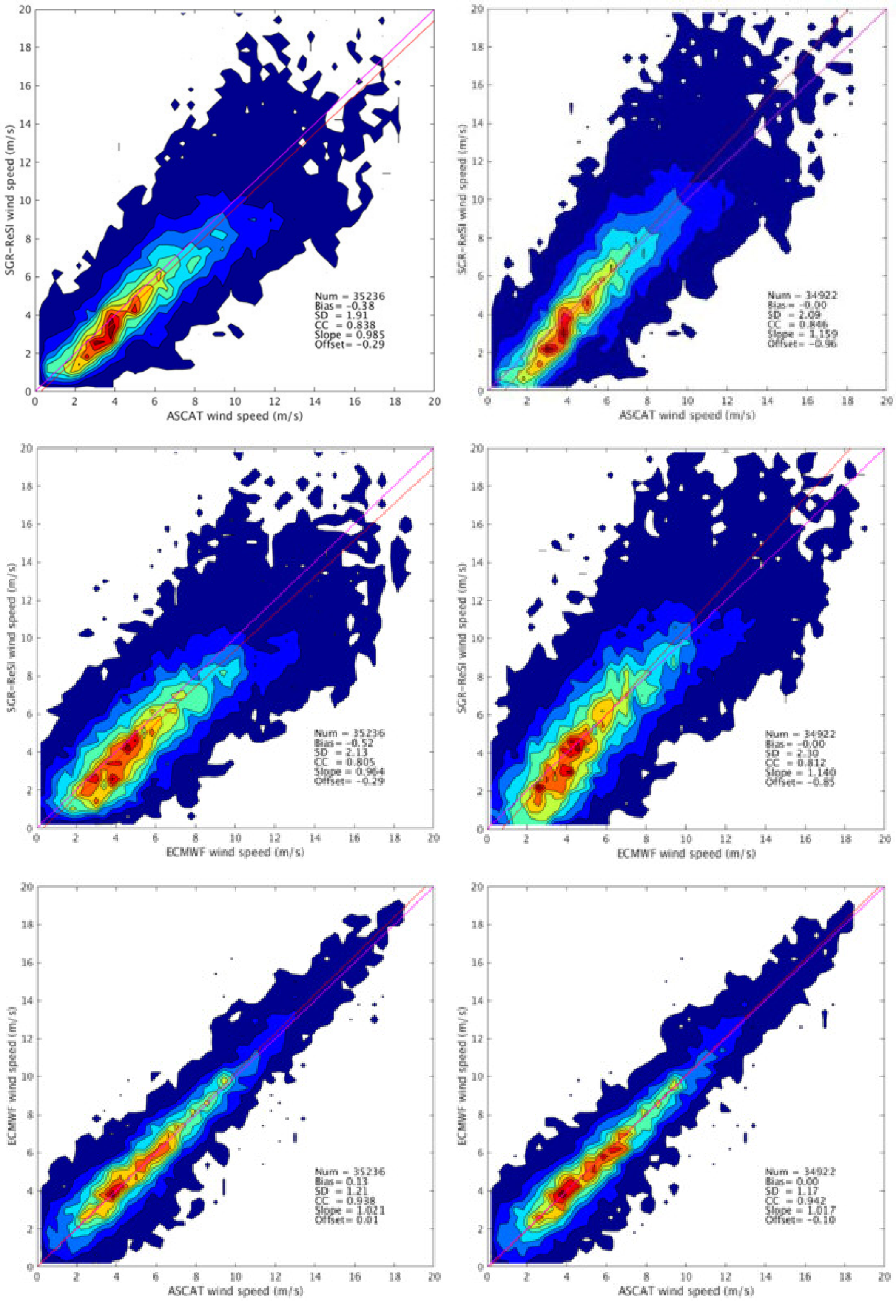

Fig. 3 Two-dimensional histograms of the wind speeds from SGR-ReSI (system 2) versus ASCAT (system 1, top panels), SGR-ReSI versus ECMWF (system 3, middle panels), and ECMWF versus ASCAT (bottom panels), before (left) and after (right) TC, with a representativeness error r^2 of 0.26. The red lines denote the linear regression between the analysed data sets. The legend shows the number of data points (Num), the bias, the standard deviation (SD) and the correlation coefficient (CC) of the compared data sets, and the regression coefficients (slope and offset).

Table 2 summarizes the results of TC analysis for the collocations in sunlight mode and in Eclipse mode respectively. It can be concluded that:

a) The representativeness errors (i.e., 0.26 for sunlight and 0.09 for eclipse) are significantly lower than that of the triplet buoy-ASCAT-ECMWF (i.e., 0.52), indicating that SGR-ReSI resolves larger scales than ASCAT.
b) As expected, the ASCAT wind errors are substantially smaller than those of SGR-ReSI and ECMWF wind speeds at the scales resolved by both SGR-ReSI and ECMWF.
c) The SGR-ReSI wind error is the largest of the three data sources.
d) The SGR-ReSI winds acquired in Eclipse mode are closer to ASCAT and ECMWF winds than those acquired in sunlight mode (not shown). Moreover, the former have larger errors than the latter;
e) The ASCAT and ECMWF wind speed SD errors at ECMWF scale in Table 2 are smaller than those in Table 1. This is probably due to the fact that most of the collocated buoy-ASCAT-ECMWF data are in the Tropics and coastal regions, while the collocated ASCAT-SGR-ReSI-ECMWF winds are sampled over a larger latitudinal range.

Table 2: TC calibration coefficients and the estimated SD errors on the scales resolved by ECMWF (fourth column) and SGR-ReSI (fifth column) for the collocated ASCAT - SGR-ReSI - ECMWF data.

In Sunlight	r^2	**Scaling factor**	**Bias correction**	ε **[m/s] SGR-ReSI scale**	ε **[m/s] ECMWF scale**
ASCAT	0.26	1.00	0.00	0.47 ± 0.003	0.70 ± 0.005
SGR-ReSI		1.22	0.66	2.04 ± 0.032	2.10 ± 0.034
ECMWF		1.05	-0.11	1.09 ± 0.015	0.94 ± 0.012
In Eclipse	r^2	**Scaling factor**	**Bias correction**	ε **[m/s] SGR-ReSI scale**	ε **[m/s] ECMWF scale**
ASCAT	0.09	1.00	0.00	0.47 ± 0.004	0.56 ± 0.005
SGR-ReSI		1.09	0.50	1.90 ± 0.043	1.93 ± 0.044
ECMWF		1.00	0.066	0.94 ± 0.018	0.89 ± 0.017

A similar approach is applied to the collocated data set only for high wind variability conditions. Following Lin et al. (2015), we use the singularity exponent (SE) derived from the ASCAT data to discriminate different wind variability conditions. Due to the lack of data, we only analyze a single high wind variability category, which includes all the collocations with SE value lower than -0.05 (i.e., the most variable 16% of the wind dataset). The results of TC analysis for the collocations with high wind variability are summarized in Table 3 and below:

a) For reference, note that for the collocated buoy-ASCAT-ECMWF dataset, the estimated SD errors of ASCAT and ECMWF for the same high wind variability category at ECMWF model scales are 1.13 and 1.41 m/s, respectively. These are significantly larger than those in Table 3, probably due (as already mentioned) to the different latitudinal sampling.

b) It is also important to note that the lack of data prevents a robust estimation of the representativeness errors, and as such, the r^2 values in Table 2 are also used to generate the results in Table 3. Note that under high wind variability conditions, larger r^2 values are expected though.

c) Comparing Tables 2 and 3, the quality degradation in terms of added error variance (i.e., SD^2) is similar for all three sources, although slightly larger for ASCAT and ECMWF. This discrepancy though is within the uncertainty of the method, which is larger for this specific analysis for the above reasons (see points 1 and 2).

Table 3: The same as Table 2, but for the collocations with ASCAT SE value lower than -0.05.

In Sunlight	r^2	**Scaling factor**	**Bias correction**	ε **[m/s] SGR-ReSI scale**	ε **[m/s] ECMWF scale**
ASCAT	0.26	1.00	0.00	0.76 ± 0.016	0.92 ± 0.021
SGR-ReSI		1.17	0.52	2.08 ± 0.083	2.14 ± 0.089
ECMWF		0.96	-0.41	1.23 ± 0.044	1.12 ± 0.039
In Eclipse	r^2	**Scaling factor**	**Bias correction**	ε **[m/s] SGR-ReSI scale**	ε **[m/s] ECMWF scale**
ASCAT	0.09	1.00	0.00	0.86 ± 0.030	0.91 ± 0.030
SGR-ReSI		1.16	0.95	1.98 ± 0.120	2.00 ± 0.120
ECMWF		0.97	-0.27	1.15 ± 0.061	1.11 ± 0.061

5. Additional buoy analysis

As already mentioned, the amount of collocated SGR-ReSI and buoy data is too limited to perform a thorough TC analysis. There are only 263 collocations in the period May - June 2016. Figure 4 shows the scatter plot of SGR-ReSI versus buoy wind speeds. However, note that the statistical scores (see legend) are comparable to those shown in Table 2.

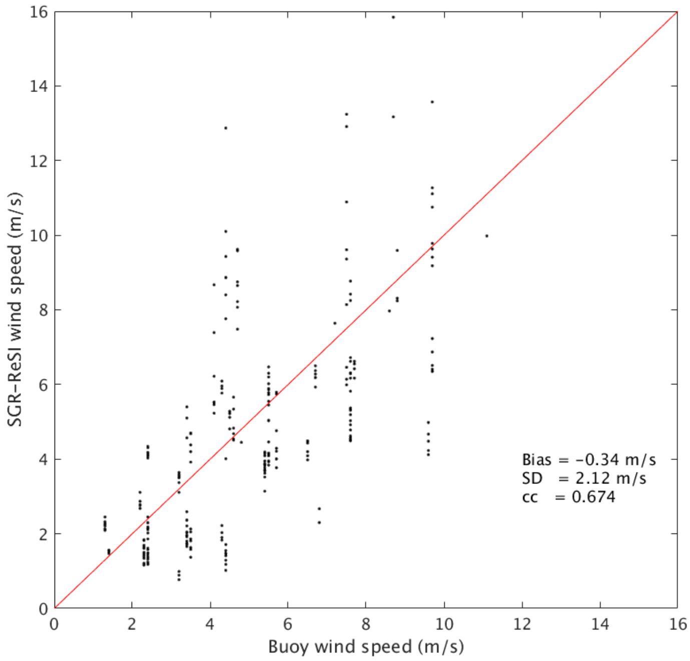

Fig. 4 Scatter plot of SGR-ReSI wind speed versus buoy wind speed. The legend shows the wind speed bias, SD error, and correlation coefficient.

6. Conclusions

A triple collocation analysis has been carried out using ASCAT-SGR-ReSI-ECMWF triplets in the period May-June 2016. For reference, a parallel TC of buoy-ASCAT-ECMWF (although different sampling) has been carried out. The results show that the overall SGR-ReSI SD errors are around 2 m/s. The SGR-ReSI errors are slightly larger in sunlight than in eclipse conditions and are certainly wind speed dependent. In addition, the estimated representativeness errors show that although SGR-ReSI footprint is comparable to that of ASCAT, the retrieved winds are of significantly lower spatial resolution. This may be due to the large noise in the backscatter measurements, which masks the small-scale signal.

The development of a more sophisticated TC model (accounting for wind-speed dependent errors rather than the current additive error assumption) may lead to more precise estimations of SGR-ReSI wind speed errors. For such development, a substantially larger dataset is needed.

References

[1] Foti et al., 2017.

[2] A. Stoffelen, “Toward the true near-surface wind speed: Error modeling and calibration using triple collocation,” *J. Geophys. Res.: Oceans*, vol. 103, pp. 7755–7766, Apr. 1998.

[3] J. Vogelzang, A. Stoffelen, A. Verhoef, and J. Figa-Saldaña, “On the quality of high-resolution scatterometer winds,” *J. Geophys. Res.: Oceans*, vol. 116, p. C10033, Oct. 2011.

[4] W. Lin, M. Portabella, G. De Chiara, A. Stoffelen, J. Vogelzang, “Towards an improved assimilation of scatterometer wind data in global NWP：Development of upscaled ASCAT data products,” to be submitted to *Quarterly Journal of the Royal Meteorological Society*, 2018.

[5] W. Lin, M. Portabella, A. Stoffelen, J. Vogelzang, and A. Verhoef, “ASCAT wind quality under high subcell wind variability conditions,” *J. Geophys. Res.: Oceans*, vol. 120, pp. 5804-5819, Aug. 2015.