

THE COMPLEXITY OF VALIDATION; A COMPARISON OF *IN SITU*, SCATTEROMETER AND ECMWF MODEL WINDS

Ad Stoffelen

Royal Netherlands Meteorological Institute, Postbus 201, 3730 AE de Bilt, the Netherlands,
E-mail: stoffelen@knmi.nl

ABSTRACT

We analysed the systematic and random errors in *in situ*, scatterometer and ECMWF model winds, with the main goal to obtain a refined calibration of the ERS scatterometer winds. Given the small dynamic range relative to the typical measurement uncertainty, the application of standard validation or calibration methods, such as regression or bin-average analyses, will often result in *pseudo* biases. Also, non-linear transformation, for instance between wind components and speed and direction, will generally give rise to *pseudo* biases. In fact, validation or calibration can only be done properly when the full error characteristics of the data are known. The problem is that in practise prior knowledge of the error characteristics is seldom available. Only by using triple collocations, random error modelling and calibration of two of the systems with respect to the third may be achieved. The *in situ* winds are shown to have the largest error variance, followed by the scatterometer, and the ECMWF model winds proved the most accurate. When using the *in situ* winds as a reference, surprisingly only the across-track scatterometer wind component was biased low by 5 %, while the other component proved unbiased. The ECMWF model winds are biased high by 6 % on both components for the period studied here. Further analysis, using a more extended triple collocation data set is recommended to confirm our conclusions. It is recommended to use the methodology of triple collocations also for the calibration of other noisy systems.

1. INTRODUCTION

The current operational ERS scatterometer processing uses the transfer function CMOD4 to derive winds from the backscatter measurements. CMOD4 was derived with a Maximum Likelihood Estimation (MLE) procedure using ERS measurements and ECMWF analysis winds (operational winds in november 1991) as input (Stoffelen and Anderson, 1995). The transfer function was verified against winds from the ESA-led Haltenbanken field campaign, and winds from the global forecast model of the British Meteorological Office, UKMO, (Offiler, 1994), and selected as the preferable function amongst some other proposals.

Winds from Numerical Weather Prediction (NWP) models are only a good reference, when they in turn are monitored against *in situ* winds from conventional platforms. Further, in a calibration exercise it is important that a representative sample of the day-to-day weather events is present. With hindsight, the Haltenbanken campaign was perhaps too limited in extend to guarantee this. In this study we will use a one-year data set of triple collocations of buoys operationally available at Météo-France, scatterometer winds and ECMWF model winds. We will re-address the wind calibration of CMOD4.

In Stoffelen (1996) it is demonstrated from statistical theory that substantial *pseudo* biases ($> 10\%$) may occur in regression or bin-average analyses, when:

- the random error characteristics of the observation systems, and
- the deformation of symmetric error distributions by non-linear transformation

are not taken into account. Also the theoretical interpretation of the commonly used scatter density plot is discussed here. Here we will repeat some of the results of Stoffelen (1996) and address the problem of *pseudo* biases by studying the full error characteristics of *in-situ*, scatterometer and ECMWF model winds through intercomparison.

In section 2 we will discuss the selection of a measurement domain where the errors are simple to describe. *Pseudo* biases after non-linear transformation will also be discussed. The wind components rather than speed and direction are shown to be the most convenient to provide an accurate description of observation errors. Without prior knowledge it is not possible to resolve both random observation error characteristics and calibration in the case of intercomparison of two noisy systems. In Stoffelen (1996) it is shown that with three noisy systems, it is possible to calibrate two of the systems with respect to the third, and at the same time provide an error characterisation for all three systems. We have used the *in-situ* winds as a reference and scaled the scatterometer and ECMWF model winds to have the same average strength.

Using a climatological wind component spectrum the representativeness error of the scatterometer and *in-situ* data with respect to the ECMWF model was estimated (Stoffelen, 1996). This important part of the observation error accounts for the spatial scales resolved by the one measurement system, but not by the other. The variance of the representativeness error of the scatterometer with respect to the ECMWF model as used in the computation is $0.75 \text{ m}^2\text{s}^{-2}$. Section 3 provides the subsequent obtained error model parameters and calibration scaling factors. Section 4 discusses the implications of this study for scatterometer data processing and wind data interpretation.

2. ERROR DOMAIN

When trying to characterise measurement errors, it is practical to select a parameter domain where the "cloud of doubt" is simple to describe. When it is symmetric then first and second order statistical moments may be sufficient to describe the errors and simulate the measurements. Although we need not to limit ourselves to these, for wind the two physical choices are either wind components (u, v) , or wind speed and direction (f, ϕ) . These are non-linearly related. We discussed in the introduction that random errors in the one domain may generate a serious *pseudo* bias in the other domain.

One way to approach error characterisation is to look in detail at the error sources. The anemometer characteristics for *in situ* winds will vary, but will generally not be the dominant error source. Interpretation errors, including height correction and platform motion correction errors, may be more substantial for the conventional winds, but some components of it are well characterised in the (f, d) domain, whilst other components are better characterised in the (u, v) domain. A major contribution to the observation error for conventional winds when comparing to scatterometer data or ECMWF model winds will be the representativeness error and this part of the total observation error is well

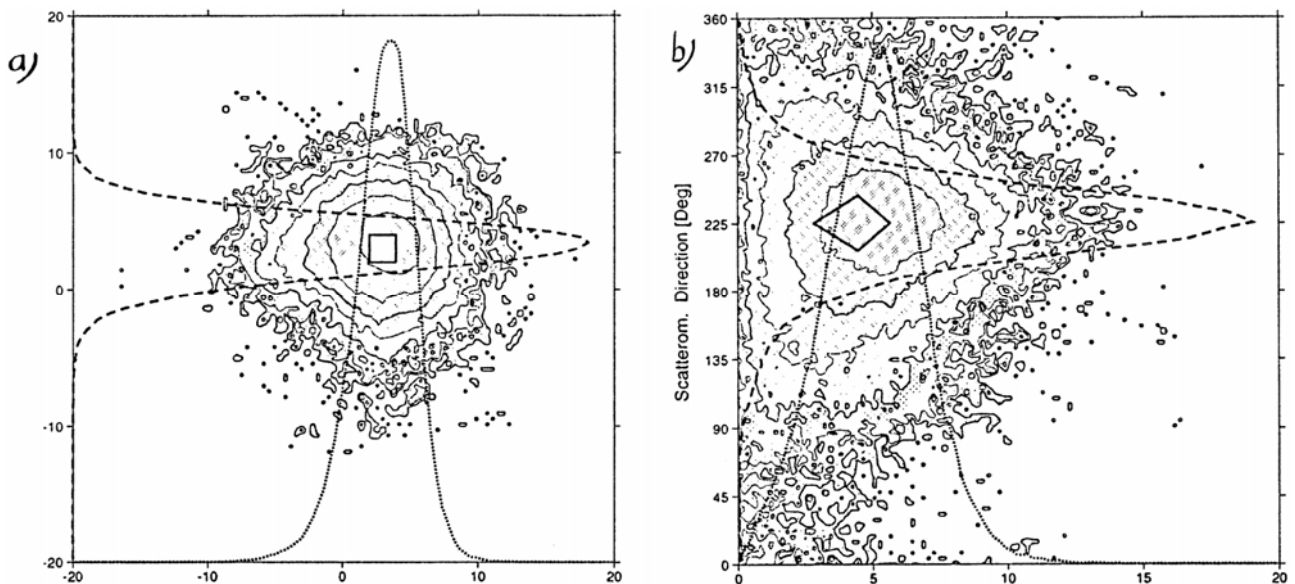


Figure 1: Scatter density plot of scatterometer winds for ECMWF winds with component values in between 2 and 4 ms^{-1} . The distribution as a function of the components (a) and the distribution as a function of speed and direction is shown (b). Density contours are logarithmic. The distributions along the vertical and horizontal axes are given by the dashed and dotted lines respectively. Component errors are simpler to describe than speed and direction errors.

characterised in the wind component domain (Stoffelen, 1996). Scatterometer winds are empirical and it is very difficult to assess which geophysical elements (e.g. waves, stability, or sea surface temperature) determine the interpretation error. The error sources in the ECMWF model that project onto the surface wind are even more difficult to elaborate on. It may be clear that a characterisation of the total observation error from a quantification of all the error sources contributing to it will be undoable. Therefore, an empirical approach was adopted here.

In figure 1 the distribution of scatterometer winds for a fixed ECMWF model wind subdomain is shown, in both physical spaces. Since, the ECMWF model winds are not perfect, the subdomain of "true" winds will be larger than the subdomain of the ECMWF model winds, i.e. it is clear that the distribution shown is effected by errors in both the ECMWF model and the scatterometer. We can see that the component errors are well-captured by a symmetric (normal) distribution. On the other hand, the wind direction random errors clearly depend on wind speed, and the wind speed error is not symmetrically distributed for light winds, i.e. the mean error for a given true light wind speed will always be positive (Hinton and Wylie, 1985). Latter is related to the fact that measured negative wind speeds can not occur. Hinton and Wylie used a truncated Gaussian function for the error distribution that did not allow negative speeds, to correct for the bias. This procedure is rather unsatisfactory, since it is not likely that the true error distribution contains discontinuities. Moreover, the "cloud of doubt" in the (f, ϕ) space is quite complicated and can not be described by second order statistics, whereas in (u, v) space the "cloud of doubt" seems much simpler to describe. Therefore, as is common practise in meteorological data assimilation, we will define an error model in the wind components.

In practise it is found that the error on both the u and v components is similar, as one may expect (see e.g. figure 1a). Also, by verifying the error distributions at higher speeds, we found little evidence of speed dependent component errors in the observation systems studied (see e.g. figure 2). Therefore, an error model with normal distributed component errors is well suited. It implies for speed and direction that the expected RMS wind speed difference $\langle (f_x - f_y)^2 \rangle$ increases monotonically with windspeed, and the wind direction RMS $\langle (\phi_x - \phi_y)^2 \rangle$ increases monotonically to a value of 104 degrees for decreasing wind speed (random direction). A good way to verify our approach is to simulate the wind speed and direction difference statistics with the error model we have obtained for

the wind components. Figure 2 shows such a comparison. We can see that the average wind speed difference indeed varies as a function of wind speed, and that it can be as large as 1 ms^{-1} . The standard deviation of the wind speed difference and the vector RMS difference go to a small value for low wind speed, as is observed for the real data as well. As expected, the wind direction standard deviation increases for decreasing wind speed and the wind direction bias is very small. Thus, our error model is as well able to simulate the observed difference statistics in the wind components as the difference statistics in wind speed and direction.

3. ERROR MODELLING AND CALIBRATION WITH THREE SYSTEMS

In this section we show the results of the calibration of scatterometer and ECMWF winds relative to the anemometer winds. The real-time available *in situ* winds were obtained from Météo-France, but further screened by the ECMWF monthly updated blacklists. Just over 50 % of the reports arrived from the WMO buoy identifiers 62111, 62112, 62118, and 62112. The scatterometer data were processed at ECMWF with PRESCAT (Stoffelen and Anderson, 1995). The ECMWF winds were from the First Guess at Appropriate Time (FGAT), which means that they are valid for the time of observation of the scatterometer. The spatial interpolation of the forecasts to the scatterometer node is bi-linear in the wind components. The *in situ* data is within 3 hours and 100 km from the scatterometer measurement time and node respectively, which presents a rather lax collocation constraint.

The average wind components of the *in situ* winds, scatterometer and FGAT are very close (within a few tenths of a ms^{-1}) and the systems thus have no absolute bias. A quality control procedure is applied to exclude gross errors with a rejection rate of $\sim 1\%$ of points. A wind direction bias correction was performed where the resulting corrections are 5.1 degrees for the scatterometer and 1.1 degrees for the ECMWF model (see also Stoffelen, 1996).

The resulting calibration scaling factors are shown in table 1. Remarkably, the scatterometer along-track component is not biased, whereas the across-track component is biased by 5 % (too low). This result is very striking, since it implies a wind direction dependent speed scaling; in directions upwind and downwind to the mid beam speeds need to be upscaled and at crosswind they need to remain the same. The wind direction change implied by this calibration is quite small and at maximum 1.4 degrees at angles under 45 degrees with upwind, downwind and crosswind.

The representativeness error estimate only influences the calibration coefficient of the ECMWF model as we would expect from equation (2), where its effect is only modest. ECMWF FGAT winds seem to be biased high with respect to the buoys (by 6 %).

	<i>u</i> component scaling	<i>v</i> component scaling
Scatterometer	1.00	0.95
ECMWF FGAT	1.06	1.06

Table 1: Calibration scaling factors for the along-track (*u*) and across-track (*v*) wind components. The scatterometer calibration is different for the two components and the ECMWF model is biased high.

Figure 4 shows the joint distributions of the wind components of *in situ* and scatterometer, scatterometer and FGAT, and FGAT and *in situ* data. It is evident that the scatter in the scatterometer and FGAT plot is smallest. This means that the *in situ* winds have the largest error. The *in situ* and scatterometer plot shows the largest scatter, which indicates that the FGAT winds are the most accurate. Table 2 shows the results of our estimates, which confirm the subjective analysis.

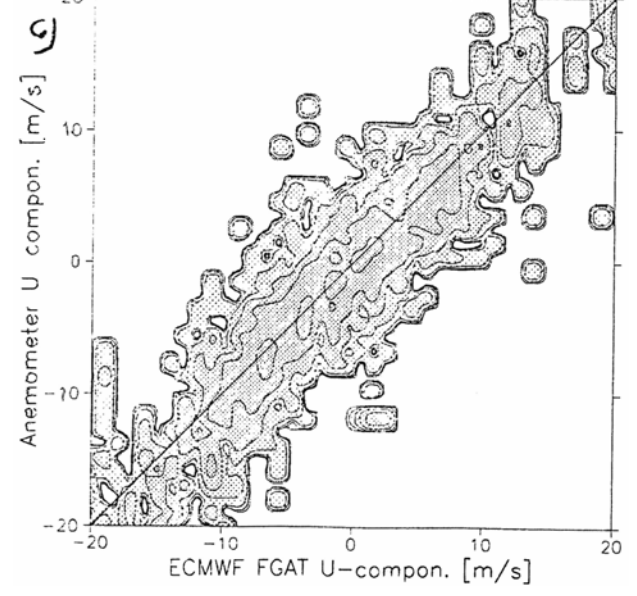
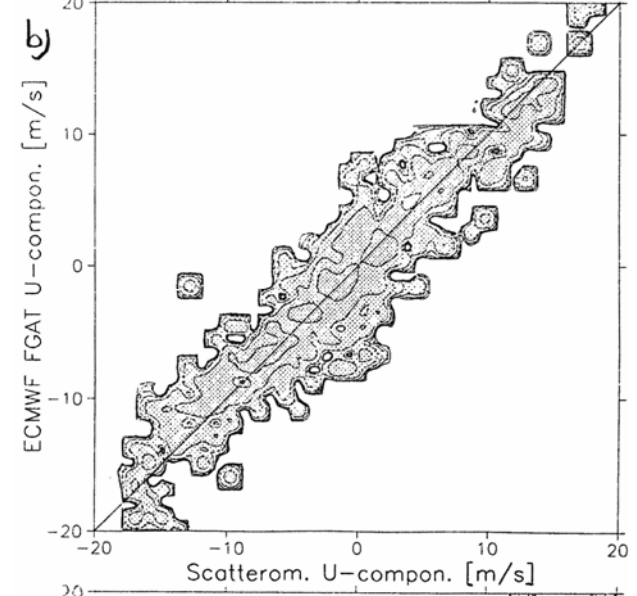
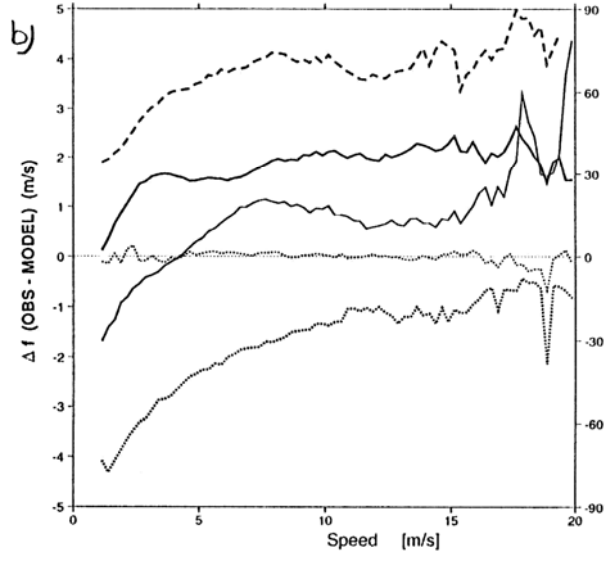
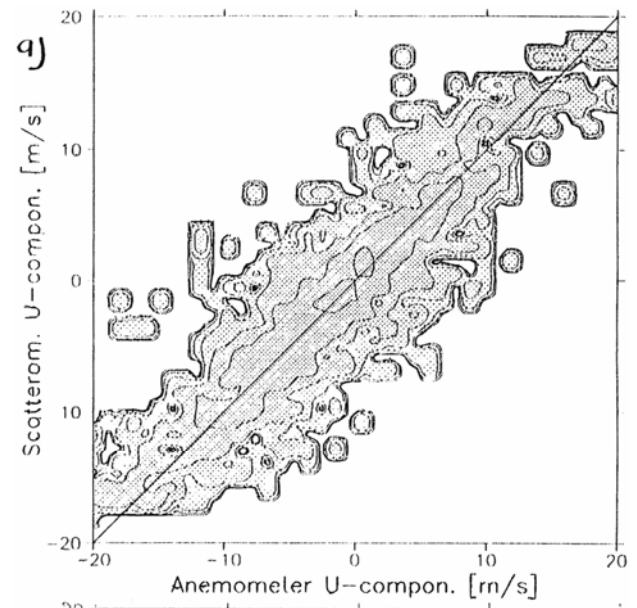
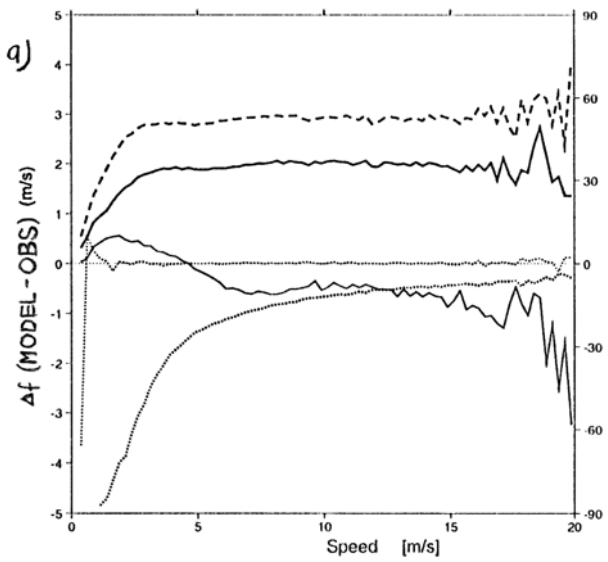


Figure 2: Above: Simulated (a) and true (b) wind speed and direction difference statistics for scatterometer and FGAT as a function of average wind speed. Speed bias (thin solid), standard deviation (thick solid), direction bias (thin dotted), standard deviation (thick dotted), and vector RMS (dashed) of differences are shown. The simulation in (a) is done with unbiased and normal distributed random wind component errors obtained from our error analysis, where the scatterometer winds were taken as truth. The speed and direction error characteristics are well simulated, including the pseudo biases. However, the errors in b) are slightly larger.

Figure 3: Right: Scatter density plot for the along-track (u) wind components of *in situ* anemometer and scatterometer (a), scatterometer and FGAT (b), and of FGAT and anemometer winds (c). The plots for the across-track (v) wind component look similar (not shown). Density contours are logarithmic.

The error estimates for the u and v component are quite similar for all systems, but compare best for the *in situ* anemometer winds. We further note that in general the errors on the v component are slightly smaller than the errors on the u component for this data set. The scatterometer and ECMWF model random error estimates compare well to a spectral analysis of these data (Stoffelen, 1996).

The anemometer measurements are local, but the variance measured on scales smaller than those represented by the ECMWF model will not be verified here and therefore treated as error. Also, the lax collocation constraints used for the buoys will contribute to the random anemometer error. When we subtract these errors we estimate the conventional data local error estimate is reduced to 1.80 and 1.71 ms^{-1} for the u and v components respectively. However, the local wind is not as relevant for operational (synoptic) meteorology as an area-averaged quantity like from the scatterometer.

	u component	v component
True variance	6.70 (6.74)	6.53 (6.59)
<i>In situ</i> error	2.64 (2.50)	2.58 (2.43)
Scatterometer error	1.87 (1.65)	1.65 (1.41)
FGAT error	0.95	0.86

Table 2: Estimates of RMS true variance and *in situ*, scatterometer and FGAT errors for the along-track (u) and across-track (v) wind components. The representativeness error contribution is computed with respect to the scales resolved by FGAT, but the numbers between brackets indicate the values at the spatial representativeness of the scatterometer winds.

4. CONCLUSIONS

In order to calibrate one observing system with respect to the other, one may use, either explicitly or implicitly, a simplifying assumption on the random errors of the two systems. For instance, it is common practise to assume that the errors of two systems that are compared are equal, or to assume that one system is much more accurate (i.e. is "truth") than the other. Given our results in table 2 and figure 2, it is obvious that both of these choices would have been crude. In the introduction we have shown that such assumptions may lead to substantial *pseudo* bias effects. It is impossible to calibrate one noisy system against another without such an assumption or other prior knowledge on the error characteristics of one or both systems (Stoffelen, 1996). As shown here, a proper calibration of an observing system can be done by using a reference system and at least one other observation system, that together can provide triple collocations.

It was found that the selection of a simple measurement domain where second order statistics are sufficient to describe the uncertainty of the measurements is preferred. More specifically, we have shown that wind error modelling using wind components is preferable to error modelling using speed and direction. Errors in speed are asymmetric and direction errors are strongly speed dependent for light winds. By assuming normal distributed wind component errors these features are well modelled, and it would on the other hand be quite complex to describe them in terms of speed and direction. Thus, wind component error statistics represent a simple method to describe complex errors in speed and direction.

We have shown that substantial *pseudo* wind speed biases can occur through the non-linear transformation of unbiased wind component errors to the wind speed and direction domain. In a direct wind speed calibration, where usually unjustly symmetric error distributions are assumed, the *pseudo* biases would be taken out, leading to biased wind components (see also Hinton and Wylie, 1985). Wind component error modelling as proposed here elegantly solves this problem.

A method to calibrate noisy systems has been developed using triple collocations. Furthermore, in a pair-wise comparison of the observation systems, the second order moments were used to estimate the true variance resolved by both systems and error variance of the observations.

Anemometer winds turned out to be the least accurate amongst the scatterometer and ECMWF model winds. After accounting for the lax space and time collocation constraints, and the variability on scales smaller than the resolution of the ECMWF model, the local accuracy of the anemometer winds was close to the scatterometer accuracy over its 50 km footprint. The extension of the triple collocation data set could be improved by a station to station height correction scheme and quality monitoring scheme. Alternatively, the procedure in this report could be repeated with the off-line NOAA buoy data set (see e.g. Wilkerson and Earle, 1990). Although not of high accuracy, the conventional wind observations provide currently the only means of NWP model and scatterometer system calibration.

We found that the CMOD4-derived scatterometer winds are biased low by 5% on the component along the mid beam direction, but are not biased on the other component. This means essentially that the wind speed bias is wind direction dependent. The effect of the component scaling on wind direction is fairly small.

The ECMWF model appears to be very accurate, but probably biased high by 6 % for the period we examined ('94). The error in the ECMWF model is determined by an extrapolation error and a dynamical error. Given the fact that the scatterometer minus FGAT statistics are very similar in both hemispheres (not shown), and that the dynamical errors are known to be larger in the SH, we conclude that the largest random error contribution is from the extrapolation.

The method of error characterisation by triple collocation is not only useful for the scatterometer, but, may also be applied to determine the prognostic skill of forecast models with increased sampling.

Acknowledgements

This work has been greatly stimulated by discussions in the ESA ASCAT SAG and has been partly sponsored by ESA. Herve Roquet provided me with the collocation data set.

REFERENCES

Hinton, Barry B., and Donald P. Wylie, (1985) "A correction for the errors in ship reports of light wind", J. Atm. Ocean. Techn. 2, pp.353-356.

Offiler, D., (1994) "The calibration of ERS-1 satellite scatterometer winds", J. of Atm. and Ocean Techn.

Stoffelen, Ad and David Anderson, (1995) "The ECMWF contribution to the characterisation, interpretation, calibration and validation of ERS-1 scatterometer backscatter measurements and winds, and their use in numerical weather prediction models", ESA contract report and executive summary for contract 9097/90/NL/BI, published by ECMWF, Shinfield Park, Reading, Berkshire, UK.

Stoffelen, Ad, (1996) "Error modelling of scatterometer, *in situ*, and ECMWF model winds; A calibration refinement", draft ESA contract report by KNMI, available from the author.

Wilkerson, John C., and Marshall D. Earle, (1990) "A study of differences between environmental reports by ships in the voluntary observing program and measurements from buoys", J. Geoph. Res. 95 C3, pp.3373-3385.