Quikscat progress report

A.C. Voorrips

December 20, 1999

1 Introduction

The Quikscat satellite was launched on June 19, 1999. It carries the Seawinds instrument, which is a conically scanning pencil-beam scatterometer, designed for measuring ocean surface winds over a wide swath. Although it is still in the Cal/Val phase, it seems that the instrument is working according to expectations.

From June until December 1999 I joined the KNMI Quikscat team, consisting of Marcos Portabella, Ad Stoffelen, and me. I took over from Julia Figa, who went to EUMETSAT. The project includes a wide range of activities which should lead to the production of high-quality wind information from Quikscat data: visualization and validation of σ^0 backscatter data, quality control to remove corrupt data, inversion from σ^0 data to (ambiguous) wind vector data and validation of these winds against independent data, ambiguity removal, preprocessing for assimilation (construction of an observation operator), and the set-up of a system for routine processing in quasi-real-time.

This report describes my activities in the last half year, in so far as they are relevant to the continuation of the work by the group. Since I was new in scatterometry, part of the effort simply has been getting to grips with the subject. Since Marcos is the "stable factor" in the group, I have relied heavily on software he (and Ad, and Julia) already made. Another major part of my time has gone into getting to know this software, and making small corrections/improvements here and there that are not worth mentioning in this report.

In describing the various subjects I've been working on, I will try to make as many references as possible to software I have developed to carry out the analyses. All this software is available either on my workstation (bgwd58) or on the WM machine (bgowm1).

2 Visualization in measurement space

It is useful to visualize the “raw” σ^0 observations somehow: for checking the validity of the geophysical model function (GMF) that is used in the inversion, for a visual check of the data (noise, outliers), and perhaps for getting ideas about the geometry of the inversion problem. The most fruitful visualization that was used for ERS is to make a selection $\sigma^0_{\text{fore}} + \sigma^0_{\text{aft}} \approx$ constant, and to plot $\sigma^0_{\text{fore}} - \sigma^0_{\text{aft}}$ against σ^0_{mid} (Stoffelen, 1998). Why this is fruitful can be understood if we first give a rough approximation of the GMF:

$$\sigma^0(V, \phi; \theta, \alpha, p) \approx A_0 + A_1 \cos(\phi - \alpha) + A_2 \cos(2\phi - 2\alpha) \tag{1}$$

where V is the wind speed, ϕ wind direction w.r.t. the cross-flight direction, p the polarization, θ the incidence angle, and α the azimuth look direction with respect to the cross-flight direction; $A_i = A_i(V, \theta, p)$, and, especially for V-polarization, $A_1 \ll A_2$. If we realize that $\alpha_{\text{fore}} = -\alpha_{\text{aft}}$, we find

$$\sigma^0_{\text{fore}} + \sigma^0_{\text{aft}} = 2A_0 + 2A_1 \cos\alpha \cos\phi + 2A_2 \cos 2\alpha \cos 2\phi \tag{2}$$

$$\sigma^0_{\text{fore}} - \sigma^0_{\text{aft}} = \qquad + 2A_1 \sin\alpha \sin\phi + 2A_2 \sin 2\alpha \sin 2\phi \tag{3}$$

where α is really α_{fore}. For ERS, $\alpha_{\text{fore}} = 45^\circ$, and this has the nice property that for the first equation, the third term vanishes; because $A_1 \ll A_0$, this means that taking a cross-section $\sigma^0_{\text{fore}} + \sigma^0_{\text{aft}} =$ constant comes down to keeping the wind speed (almost) constant. Therefore, in such a figure one can easily draw the theoretical (GMF) curve (varying wind direction with given, constant wind speed) together with the measurement points, and hence check the validity of the GMF.

For Quikscat the situation is more complicated, because α is varying with the distance from the satellite ground track and hence with the WVC number. In fact, it is not clear what would be a sensible cut through the (2- or 4-dimensional) cone. Alternatively, we can simply plot σ^0_{fore} against σ^0_{aft} for *all* quadruplets, both for the outer (VV) and for the inner (HH) beam. Examples are shown in figures 1–3. Figure 1 shows scatter plots for simulated data, without noise, for various parts of the swath. The Lissajous-type curves represent NSCAT-2 curves for constant wind speed and varying wind direction. It is clear that the form of these curves depends very much on the azimuth angle α (note that for $\alpha = 45^\circ$ (lower left panel), the modulation of $\sigma^0_{\text{fore}} + \sigma^0_{\text{aft}}$ for given wind speed is indeed small). Already it is clear that the outer edge of these curves should coincide with the outer edge of the cloud of measurements, if the NSCAT-2 GMF is perfect and the measurements contain no noise. Also clear is that, due to the geometry of the GMF, the density of points increases near the "inner edge" of the Lissajous-type curves. Figures 2 and 3 show the same scatter plots, but now for, respectively, simulated data with noise added according to the K_p values in the measurement files, and for actual observed data. It is clear that the picture is blurred compared to the no noise-results: the measurement points extend to over the outer edge of the theoretical curves, and also the density jump near the inner edges is not so clear. But already we can conclude that the Quikscat measurements follow the NSCAT-2 GMF at least to a fair degree over the whole swath , and with a noise level which does not deviate much from the noise derived from the specified K_p's (no big differences between figures 2 and 3).

We have started work to refine the analysis above by making bidimensional histograms of the scatter plots and sections through them: see figure 4. In the contour plot (left panel) we have added the theoretical "outer edge" of the curves in the previous plots, allowing easier comparison with the data points (the contours themselves do not show the edge so nicely, probably because of the binning and because of the contouring algorithm). The section through the contour plot (constant $\sigma^0_{\mathrm{fore}} + \sigma^0_{\mathrm{aft}}$) shows nicely the positions of the outer edge, *and* the increase in density of points near the inner edge of the curves; this type of plots allows a more detailed analysis of the correctness of the NSCAT-2 GMF. To see how good this works in practice requires a bit more work; the plot shown is for simulated data without noise, and even there the edges do not show up very sharply, probably because of the binning and because the number of points used is limited (even with 5 orbits of data!). Other approaches include the plotting of the "inner edge" of the Lisajous curves concatenated, and plotting only a 180° range of wind velocities (from one "outer edge" to one "inner edge") at a time.

NB The data used in this section were Quikscat orbits 00980–00984. The observations were read and processed on the bgowm1 with ~/marc/hdf/qretrieve25/retrievals_hdf25_mlesim.f (i.e, the "simulation" version of the inversion/processing software). The data were transferred to the bgwd58:/nobackup/users/voorrips/Qscat/simulation/Nonoise, -/Noise, and -Obs directories. The plotting was done with PVWAVE routines specified in the captions of the plots.

3 Validation of retrieved wind velocities

3.1 Scatter plots

I wrote software to produce simple statistics and contour plots of bidimensional histograms of inverted (KNMI or JPL) and NWP wind velocities; see for example figures 5 and 6. The procedure works as follows:

- With the inversion software on the bgowm1 (directory ~/marc/hdf/qretrieve25; script QR_hdf25_mle calling the retrievals_hdf25.f program, or QR_hdf25_mlesim calling retrievals_hdf25_mlesim.f for simulated σ^0's), data files are produced;
- The bidimensional histograms and the statistics are carried out at the bgowm1, directory ~/marc/biascor, scripts scbias or scbias_two. The first script is used when comparing either JPL or KNMI inversions with NWP (needing only 1 data file), the second when comparing JPL with KNMI inversions (needing two data files in the present configuration). The scripts create data files for various parts of the swath.
- The data files are transferred to the bgwd58 and plotted with PVWAVE-script qhist_mle2.pro, using the (bgwd58) script ~/Qscat/Wave/Hist/Two-step/qrun. Alternatively, the data can be transferred with the script qget and qhist_mle2.pro can be run interactively from a PVWAVE session.

The analysis procedure was split up in two steps because direct PVWAVE processing of the individual measurements is too slow when using more than a few orbits of data.

So far, we have no good reference data to check the quality of the inversions; we are still waiting for collocations with ECMWF NWP winds. Results of the comparison with sample data and the low-quality NCEP winds are discussed in some detail by Portabella and Stoffelen (1999).

3.2 Normalized RMS as a validation tool

One can validate the inverted wind directions against a reference wind direction (usually NWP) by taking the solution which is closest to the reference, and computing the RMS difference ("error"). However, it is clear that the more ambiguous solutions are provided by the inversion, the smaller the RMS error will be, because the chance that one of the solutions will be close to the NWP estimate will increase; in the limit of an infinite amount of solutions, the RMS will even go to 0! At the same time, the information content of the set of solutions in reality *decreases* with an increasing number of solutions, because there is no a priori way to say which of the solutions is the correct one. Therefore, I have attempted to define a "normalized root mean square error" (NRMS), which would give a more honest indication of the quality of the "closest solution" than the usual RMS error does. This NRMS should contain a normalization factor which is equal to the expected value in the case that there is no skill in the system:

$$\mathrm{NRMS} = \mathrm{RMS}\left(\frac{\phi^c - \phi^t}{\langle(\phi^c - \phi^t)^2\rangle^{1/2}}\right) \tag{4}$$

where ϕ^t is the "true" (reference) wind direction, ϕ^c is the solution closest to the true solution, and $\langle\rangle$ gives the expected value in case there is no skill, i.e., in the case ϕ^t is randomly distributed. Of course one must make sure that always $|\phi^c - \phi^t| \leq \pi$.

To define the NRMS, we have to specify the denominator of (4). Suppose we have n solutions, sorted such that $0 \leq \phi_1 < \phi_2 < \cdots < \phi_n < 2\pi$. For convenience, we define $\phi_0 \equiv \phi_n - 2\pi$, $\phi_{n+1} \equiv \phi_1 + 2\pi$. The boundaries between the solutions, which define their angular sector, are defined by

$$\phi_i^{(b)} = \frac{\phi_{i-1} + \phi_i}{2}, \quad i = \{1, \ldots, n+1\} \tag{5}$$

My definition of "no skill in the solutions" is to take a constant probability distribution for the true wind direction: $p(\phi^t) = 1/2\pi$. In that case we get

$$\langle(\phi^c - \phi^t)^2\rangle = \int_0^{2\pi} (\phi^c - \phi^t)^2 p(\phi^t) d\phi^t = \frac{1}{2\pi} \sum_{i=1}^{n} \int_{\phi_i^{(b)}}^{\phi_{i+1}^{(b)}} (\phi_i - \phi^t)^2 d\phi^t = \frac{1}{24\pi} \sum_{i=1}^{n} (\phi_{i+1} - \phi_i)^3 \tag{6}$$

where, in the last equality, we have used (5).

This variance depends both on the number and on the distribution of the solutions. Some appealing features of this definition can be easily found from (6). The first one is that the variance decreases with an increasing number of solutions, hence giving a higher "penalty" in the NRMS. For instance, if the solutions are regularly distributed over the circle ($\phi_{i+1} - \phi_i = 2\pi/n$), we find $\langle(\phi^c - \phi^t)^2\rangle = \pi^2/3n^2$. Another nice property is that, in the case that artificial "double minima" solutions are produced by the inversion procedure (so, for certain j, $\phi_{j+1} - \phi_j \approx 0$), this will have no effect on the normalization.

Ad has proposed a slightly different definition of "no skill in the solutions": instead of assuming ϕ^t to be equally distributed around the whole circle, he uses the fact that we already know which solution is the one closest to the reference. If we order the solutions such that always $\phi_1 = \phi^c$, instead of $p(\phi^t) = 1/2\pi$ we get

$$p(\phi^t) = \begin{cases} \frac{1}{\phi_2^{(b)} - \phi_1^{(b)}} & \text{if } \phi_1^{(b)} \leq \phi^t \leq \phi_2^{(b)} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Then we find, instead of (6),

$$\langle(\phi^c - \phi^t)^2\rangle = \frac{1}{12(\phi_2 - \phi_0)} \left[(\phi_2 - \phi_1)^3 + (\phi_1 - \phi_0)^3\right] \tag{8}$$

It can easily be checked that for $n = 1$ and $n = 2$, (6) and (8) yield the same result. Also in the case that the n solutions are regularly spaced, we find the same result. In other cases, however, these variances are different. The most important difference is that, whereas formulation (6) depends only on the distribution of the solutions, (8) also depends on which of the solutions turns out to be the "closest" solution. The simplest case is when we have $n = 3$, and solutions $\{\phi_1 = -\phi, \phi_2 = 0, \phi_3 = \phi\}, 0 \leq \phi \leq \pi$. We will not give the expressions for the computed variances (take care to order the solutions such that $\phi_1 = \phi^c$ when using (8)!), but the numerical result is plotted as a function of ϕ in figure 7. It can be seen that as long as the angular sectors are not too small, the results of (6) and (8) are comparable (and note that they are exactly equal when $\phi = \pi/3$). However, if the sector of the closest solution approaches 0 (which is the case for the central solution ϕ_2 if ϕ goes to 0), the normalizing variance in (8) also approaches 0. This type of behavior is to be expected always when the inversion procedure finds 3 or more solutions with wind directions close to each other. It will not lead to "disasters" in the computation of the NRMS (since also the maximum deviation $\phi_2 - \phi^t$ will go to 0 in this sector), but the expectation is that the average NRMS will be somewhat larger when using (8) than when using (6). It is probably worthwhile to try them both and see which one appears to be most useful.

4 Towards the 4D-VAR observation operator

4.1 MLE-dependence of the observation operator

In order to assimilate scatterometer data in a NWP model with a 4D-VAR data assimilation scheme, an "observation operator" has to be constructed. This operator has the form of a cost function term (Lorenc, 1988)

$$J_{\mathrm{o}}^{\mathrm{SCAT}}(\mathbf{V}) = -2\ln\{p(\sigma_o^{\mathbf{0}}|\mathbf{V})\} \tag{9}$$

Suppose that for a certain measurement $\sigma_o^{\mathbf{0}}$, the inversion procedure results in a set of n solutions $\{\mathbf{V}_1, \mathbf{V}_2 \ldots, \mathbf{V}_n\}$. Following Stoffelen (1998), to good approximation we can write

$$p(\sigma_o^{\mathbf{0}}|\mathbf{V}) = \sum_{i=1}^{n} w_i N(\mathbf{V}_i, \varepsilon_i) \tag{10}$$

The weight w_i and the width ε_i should depend respectively on the probability and on the accuracy of the solutions. In the case of ERS, where usually two solutions are found with almost the same MLE, w_i was set to 1/2 and ε_i to a constant. In the case of Quikscat, however, the number of solutions ranges from 1 to 4 (or more, but we do not consider those cases), with the MLE strongly varying from one solution to the next. Since the MLE is a measure of the distance of a measurement σ_o^0 from the wind cone, it seems reasonable to assume that with increasing MLE, the probability decreases that a certain solution is the "selected" solution (i.e., the one closest to the true wind velocity). To get a first impression of this, see figure 8. Here, the distribution of (JPL) inversion solutions is plotted as function of normalized MLE R, together with the number of solutions closest to the NWP estimate[1]. It can clearly be seen that the fraction of "selected" (i.e., closest to NWP) solutions decreases rapidly with increasing R, indicating that the MLE is a good measure for the probability of a solution.

The result of figure 8 prompted us to look for a way to determine the weights w_i on the basis of the normalized MLE of the solutions. To do this, we formulate the following

Assumption: *There exists a function $p_s(x)$, such that, if we have a set of inversion solutions $\{\mathbf{V}_1, \mathbf{V}_2 \ldots, \mathbf{V}_n\}$ with normalized MLE $\{R_1, R_2, \ldots, R_n\}$, the probability that rank j is the one closest to the true wind velocity is given by*

$$P(s = j | \{R_1, R_2, \ldots, R_n\}) = \frac{p_s(R_j)}{\sum_{i=1}^{n} p_s(R_i)} \tag{11}$$

(with $s = j$ I mean "rank j is the selected rank (i.e., closest to the truth)")

To determine $p_s(x)$, we concentrate first on only those cases which have exactly two solutions. Figure 9 shows a scatter plot of those cases. It can be seen again that, at least for not too high MLE's, rank 1 is the selected one most of the times, and rank 2 is only chosen if R_2 is not much larger than R_1. For larger MLE's, the difference between R_1 and R_2 seems to be less important.

If we take many more points than shown in figure 9, we can construct a bidimensional histogram showing the relative probability of selecting either the first or the second rank, as a function of R_1 and R_2. But according to our assumption, by applying (11) with $n = 2$, we find that the probability of selecting rank 1 is given by

$$P(s = 1 | \{R_1, R_2\}) = \frac{p_s(R_1)}{p_s(R_1) + p_s(R_2)} = [1 + p_s(R_2)/p_s(R_1)]^{-1} \tag{12}$$

Hence, the bidimensional histogram gives us an estimate of $p_s(R_2)/p_s(R_1)$ for every combination $\{R_1, R_2\}$. Figure 10, left panel, shows such experimentally determined ratios as a function of $R_2 - R_1$. As can be seen, for not too high values of R_1 the ratio seems to be a fairly constant function of $R_2 - R_1$, suggesting that $p_s(x)$ is a decaying exponential function of x. However, there is also a tendency that the ratio becomes a slower decaying function of $R_2 - R_1$ when R_1

---

[1] The normalized MLE R is defined as in Portabella and Stoffelen (1999), where we take the NWP wind velocity for the normalization instead of the velocity from the inversion.

increases. Therefore, we have attempted to fit the function

$$p_s(x) = \exp\left[-\left(\frac{x}{a_1 + a_2 x}\right)\right] \tag{13}$$

to the data. This function has the nice asymptotical property that for large R_1 and R_2, the probabilities for selecting rank 1 and rank 2 become equal, which is the behavior suggested in figure 9. We found a reasonable fit (figure 10, right panel) for the values $a_1 = 0.30$ and

$$a_2 = \begin{cases} 0.03 & \text{if } x \leq 2.5 \\ 0.03 + 0.015(x - 2.5) & \text{if } 2.5 < x \leq 4.5 \\ 0.06 & \text{if } x > 4.5 \end{cases}$$

Having determined $p_s(x)$, we can use this formulation to predict how often a certain solution rank corresponds to the "true" solution, for various numbers of solutions, and for various regimes of MLE distributions. Table 1 compares the predicted distribution over the ranks with the one actually observed in a sample of five orbits, for all cases and for cases stratified using the conditions $R_2 - R_1 \leq 1$ and $R_2 - R_1 > 1$. The correspondence is striking. Therefore we can conclude that our assumption in the beginning of this section is a useful one, and that equation (13) can be used to determine the weights w_i in the observation operator.

4.2 Dependence on the directional distribution of the solutions

As was discussed by Figa and Stoffelen (1999), the directional distribution of the solutions determines the "a priori" probability that a certain rank will be the closest to the true wind velocity: this probability is directly proportional to the angular sector that this solution represents (see figure 6.1) in the reference mentioned above). Suppose that we have again n solutions with wind directions ϕ_i and corresponding angular sectors $(\Delta\phi)_i$, then the "directional" probability for rank j is given by

$$p_d(j) \equiv P(s = j | \{\phi_1, \phi_2, \ldots, \phi_n\}) = \frac{(\Delta\phi)_j}{2\pi} \tag{14}$$

Hence, the MLE distribution and the directional distribution are two different (and presumably independent) pieces of information which can help to determine the probability of each solution. A straightforward (but not necessarily the best) way to combine them is given by

$$P(s = j | \{R_1, R_2, \ldots, R_n\}, \{\phi_1, \phi_2, \ldots, \phi_n\}) = \frac{p_s(R_j) p_d(j)}{\sum_{i=1}^{n} p_s(R_i) p_d(i)} \tag{15}$$

It turns out that the sizes of the angular sectors are evenly distributed over the various ranks; hence, a table like table 1 calculated on the basis of (14) instead of (11) gives equal probabilities for all ranks, and a table based on (15) gives almost exactly the same results as a table based on (11). Nevertheless, (15) is most likely the better choice as the basis for the observation operator, it only does not show up in this type of tables (actually, I used (15) to compute table 1, but I checked that it makes no difference with (11)).

NB The data used in this section were Quikscat orbits 00980–00984. The observations were read and processed on the bgowm1 with ~/marc/hdf/qretrieve25/retrievals_hdf25_mlesim.f (i.e, the "simulation" version of the inversion/processing software; but since I only used JPL MLE's, JPL inversions and NWP winds this could have been done just as well with retrievals_hdf25_sim.f). The data were transferred to bgwd58:/nobackup/users/voorrips/Qscat/simulation/Noise, and read in with the PVWAVE routine ~/Qscat/Wave/NMLE/mle_read.pro.

5 Conclusions

We have described progress in the visualization of the measurements, validation of inverted solutions, and the construction of the 4D-VAR observation operator. The visualization tools can already be used to qualitatively check the data and the correspondence with the NSCAT2 GMF, and we point out how they can be refined to have a more quantitative look. The tools for standard statistical validation and plotting of the inverted winds are ready; here the main bottleneck is to finally get collocations with high-quality reference (ECMWF) wind data. A normalized RMS error for the wind directions has been defined, which is a more meaningful quality estimator than the usual RMS error in the case of ambiguous wind solutions. Finally, a method is proposed and checked to compute the relative probability of ambiguous wind solutions on the basis of their respective MLE values. This method has been combined with the method to compute a priori probabilities on the basis of the distributions of the inverted wind directions, and can now be applied to construct the 4D-VAR observation operator.

We are approaching a full Quikscat scatterometer processing system, with visualisation, inversion, quality control and validation tools already available in some form, although fine-tuning and further development of all these tools still needs to be done. Also assimilation experiments can start soon, since the "ground work" for the observation operator has been done. One of the things which have not been attempted so far is to formulate a Quikscat ambiguity removal scheme; no doubt the availibility of PRESCAT and 2D-VAR will allow a rapid development of such a scheme.

Acknowledgements

Thanks to Julia, Ad and especially Marcos, without whom I would not have come far. Good luck to John!

References

Figa, J. and Stoffelen, A. (1999). Towards an improved Ku-band scatterometer wind product, *Final Rep. EUMETSAT Fellowship*, Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands.

Lorenc, A. C. (1988). Optimal nonlinear objective analysis, *Q. J. R. Meteorol. Soc.* **114**: 205–240.

Portabella, M. and Stoffelen, A. (1999). *EUMETSAT Quikscat Fellowship: progress report*, Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands.

Stoffelen, A. (1998). *Scatterometry*, PhD thesis, University Utrecht, Netherlands.

List of Figures

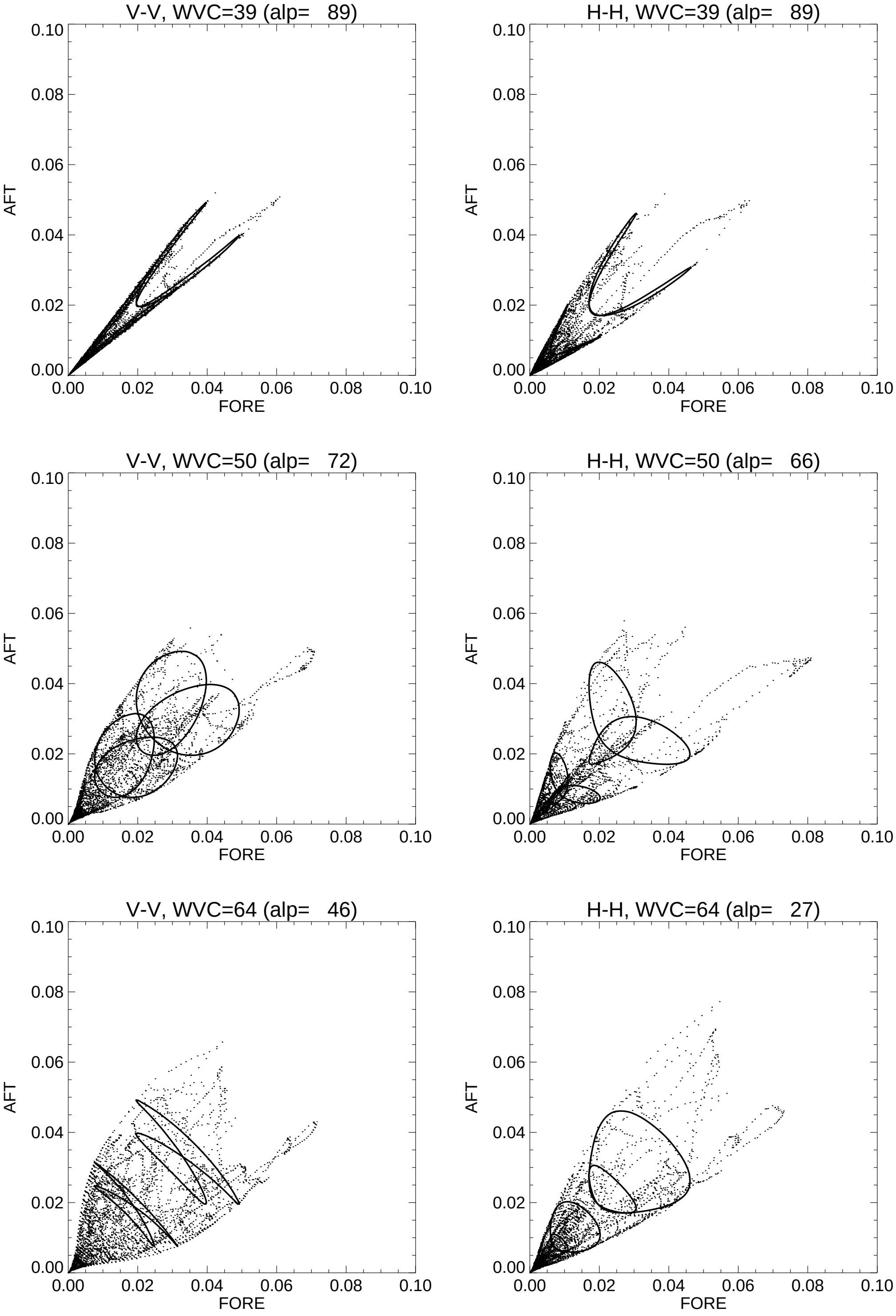


Figure 1: Scatter plot of σ_0 values for the outer beam (left panels) and the inner beam (right panels), for simulated measurements without noise. Results are shown for WVC 39 (upper panels), 50 (middle panels), and 64 (lower panels). In the plot titles, the azimuth angle α of the antenna is given. The Lisajous-type figures correspond to the NSCAT-2 GMF for 10 and 15 m/s (and 5 m/s, but this one is lost in the cloud of points near the origin).

Plot made with bgwd58:~/Qscat/Wave/Nscat2/sigplot.pro.

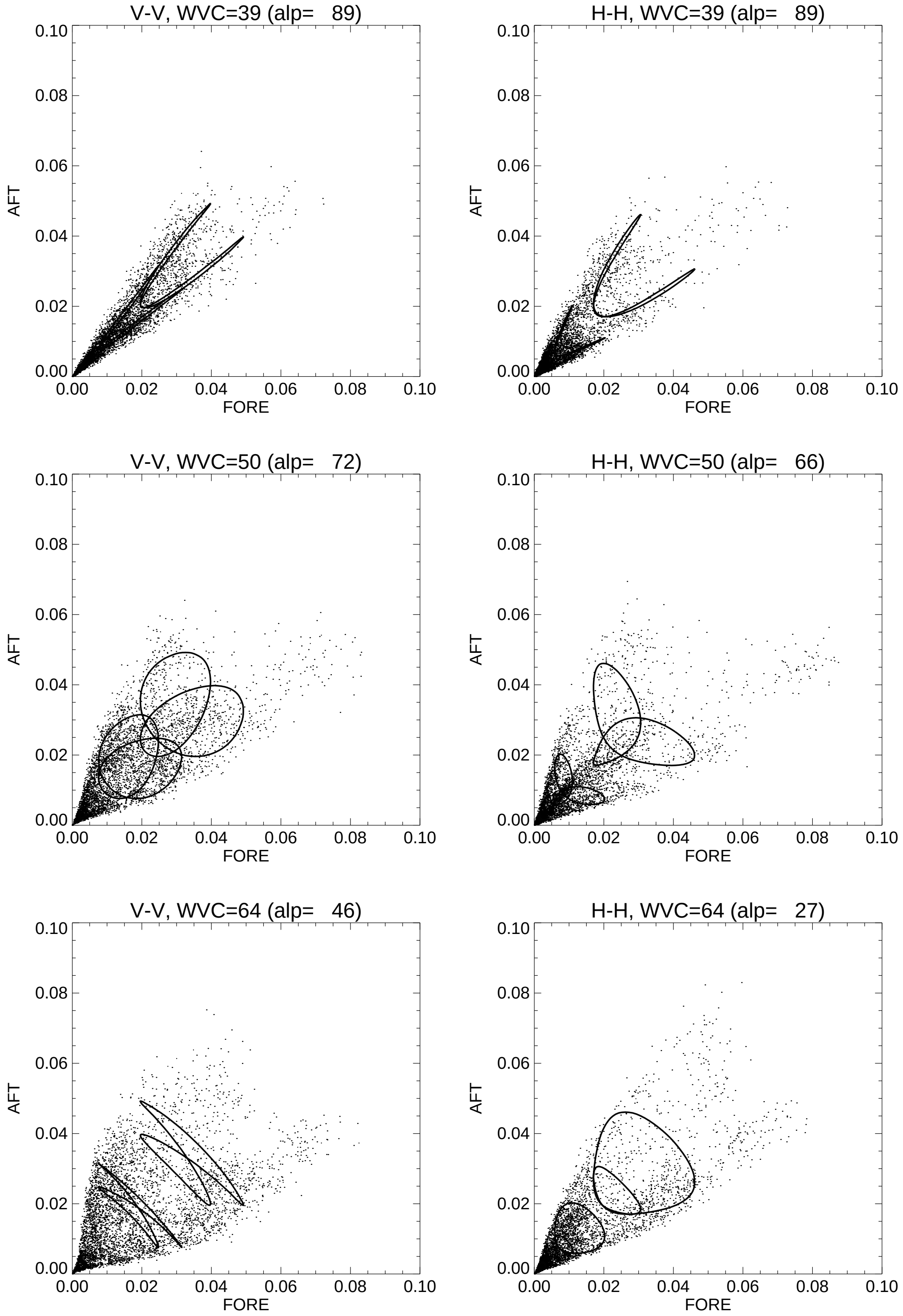


Figure 2: As figure 1, but now for simulated measurements with noise added.

Plot made with bgwd58:∼/Qscat/Wave/Nscat2/sigplot.pro.

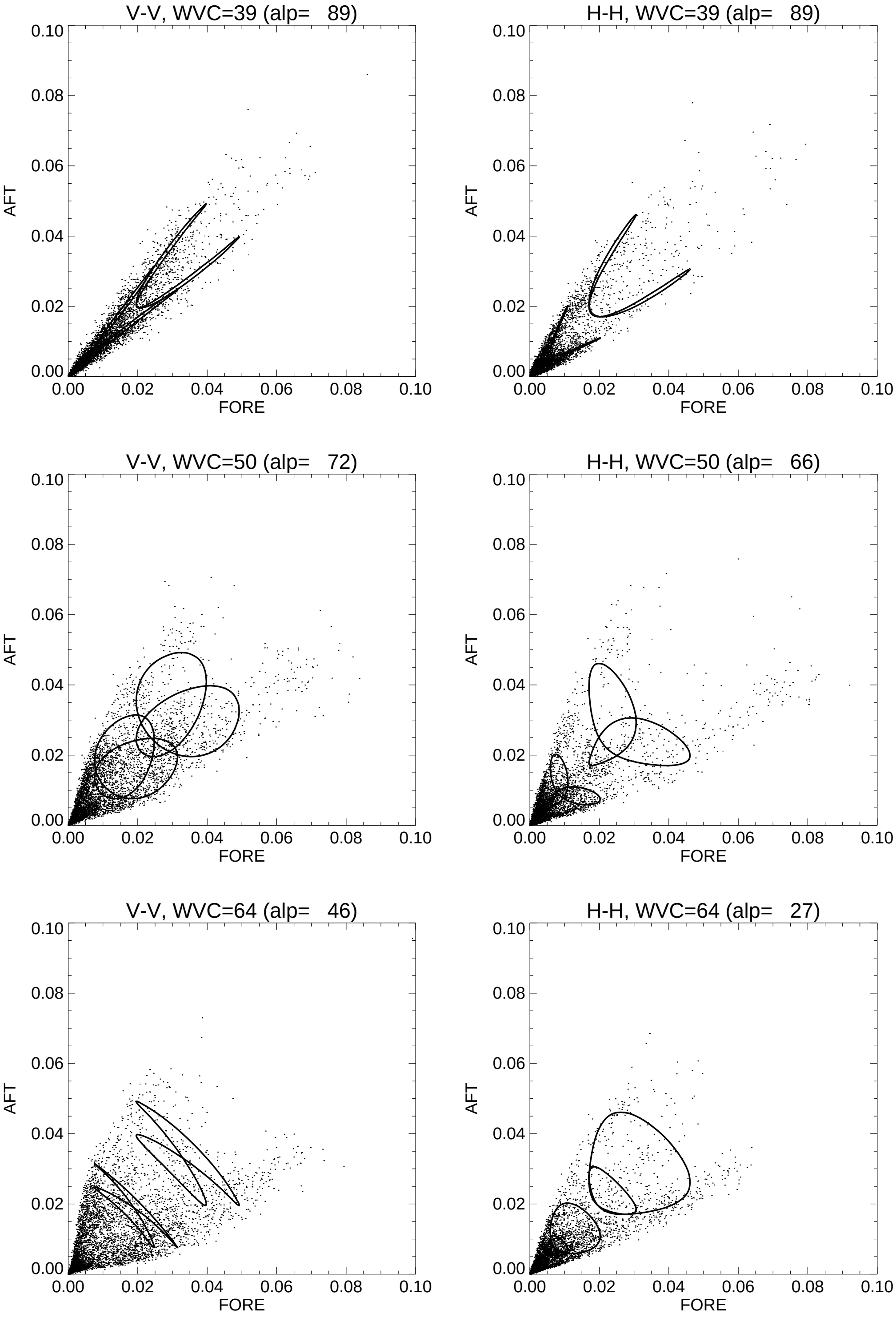


Figure 3: As figure 1, but now for real observations.

Plot made with bgwd58:~/Qscat/Wave/Nscat2/sigplot.pro.

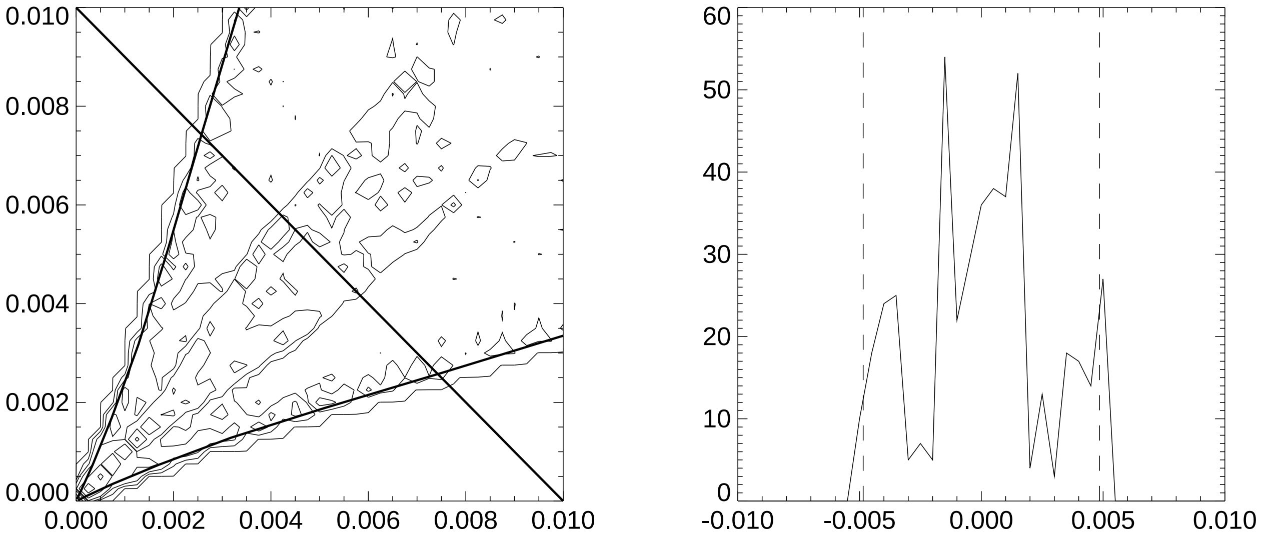
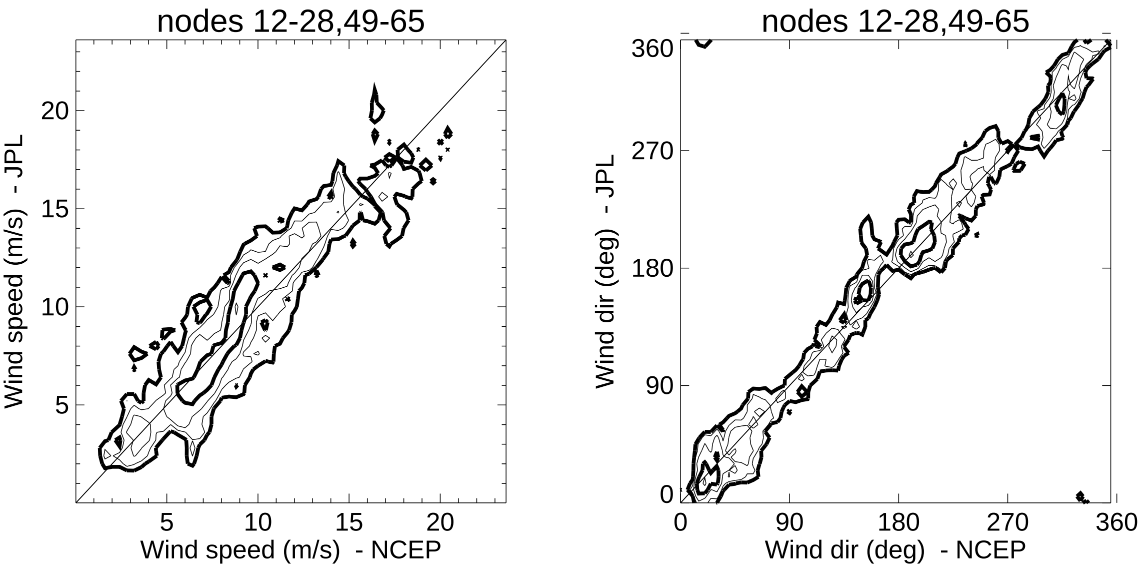


Figure 4: Left panel: contour plot (bidimensional histogram) based on scatter plots as in the previous figures (for simulated data without noise, inner beam, WVC 50). The fat lines extending from the origin are the boundaries of the NSCAT-2 cone. The right panel shows a section through the bidimensional histogram (represented by the third fat line in the left panel); the dashed lines are the boundaries of the NSCAT-2 cone.

Plot made with bgwd58:∼/Qscat/Wave/Nscat2/section.pro.

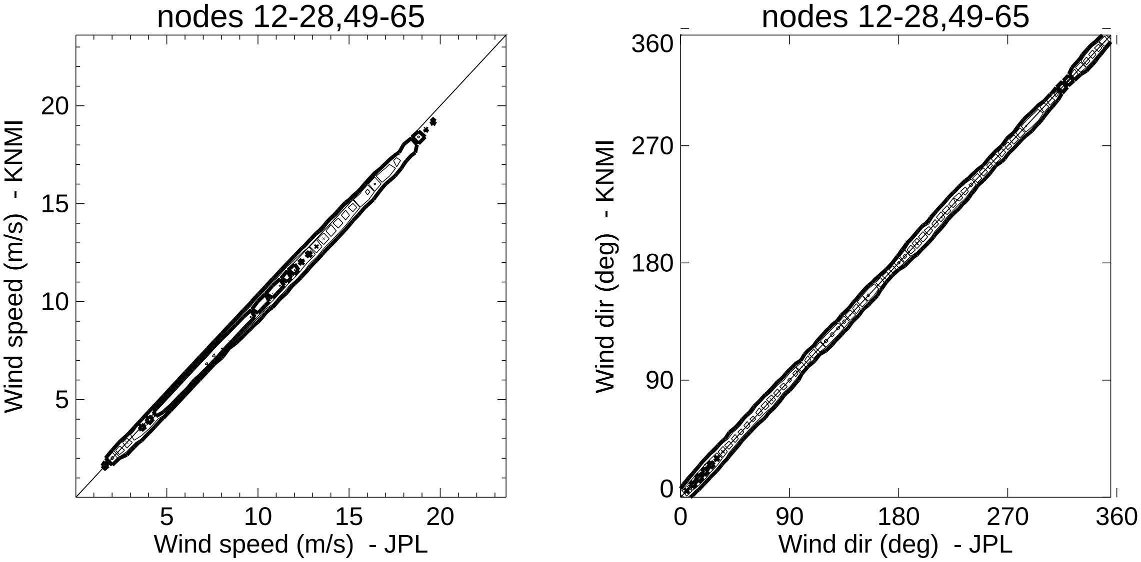


N= 19159
mx= 9.63 my= 9.82
m(y-x)= 0.19 s(y-x)= 2.08
cor_xy= 0.87

N= 17166
mx= 171.25 my= 173.04
m(y-x)= 1.78 s(y-x)= 18.67
cor_xy= 0.98

Figure 5: Joint distributions of the NCEP and JPL inverted (solution closest to NWP) wind speeds (left panel) and directions (w.r.t. cross-track direction, right panel) (sample data).

Plot made with bgwd58:∼/Qscat/Wave/Hist/Two-step/marcos.pro. This is a simplified version of qhist_mle2.pro.



N= 19182
mx= 9.77 my= 9.47
m(y-x)= -0.30 s(y-x)= 0.33
cor_xy= 1.00

N= 17721
mx= 155.70 my= 155.76
m(y-x)= 0.06 s(y-x)= 5.54
cor_xy= 1.00

Figure 6: Joint distributions of the JPL and KNMI inverted (KNMI solution closest to first-rank of JPL) wind speeds (left panel) and directions (w.r.t. cross-track direction, right panel) (sample data).

Plot made with bgwd58:~/Qscat/Wave/Hist/Two-step/marcos.pro.

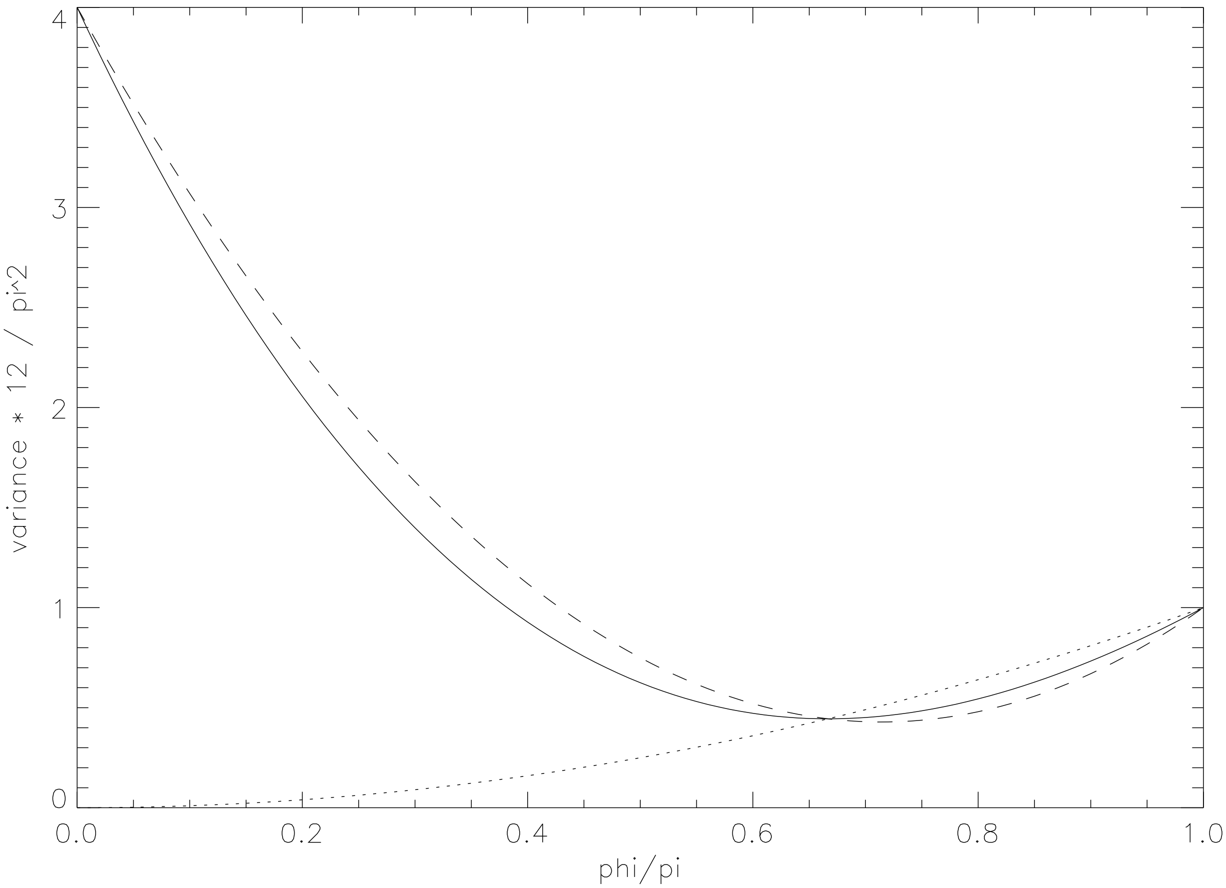


Figure 7: Variance $\langle(\phi^c - \phi^t)^2\rangle$ for the case $n = 3$, with solutions $\{\phi_1 = -\phi, \phi_2 = 0, \phi_3 = \phi\}$, as a function of ϕ. The solid line gives the result obtained with equation (6). The dotted line gives the results obtained with (8) for the central solution (ϕ_2), the dashed line the result obtained with (8) for the other two solutions.

Plot made with bgwd58:~/Qscat/Wave/NRMS/dirvar.pro.

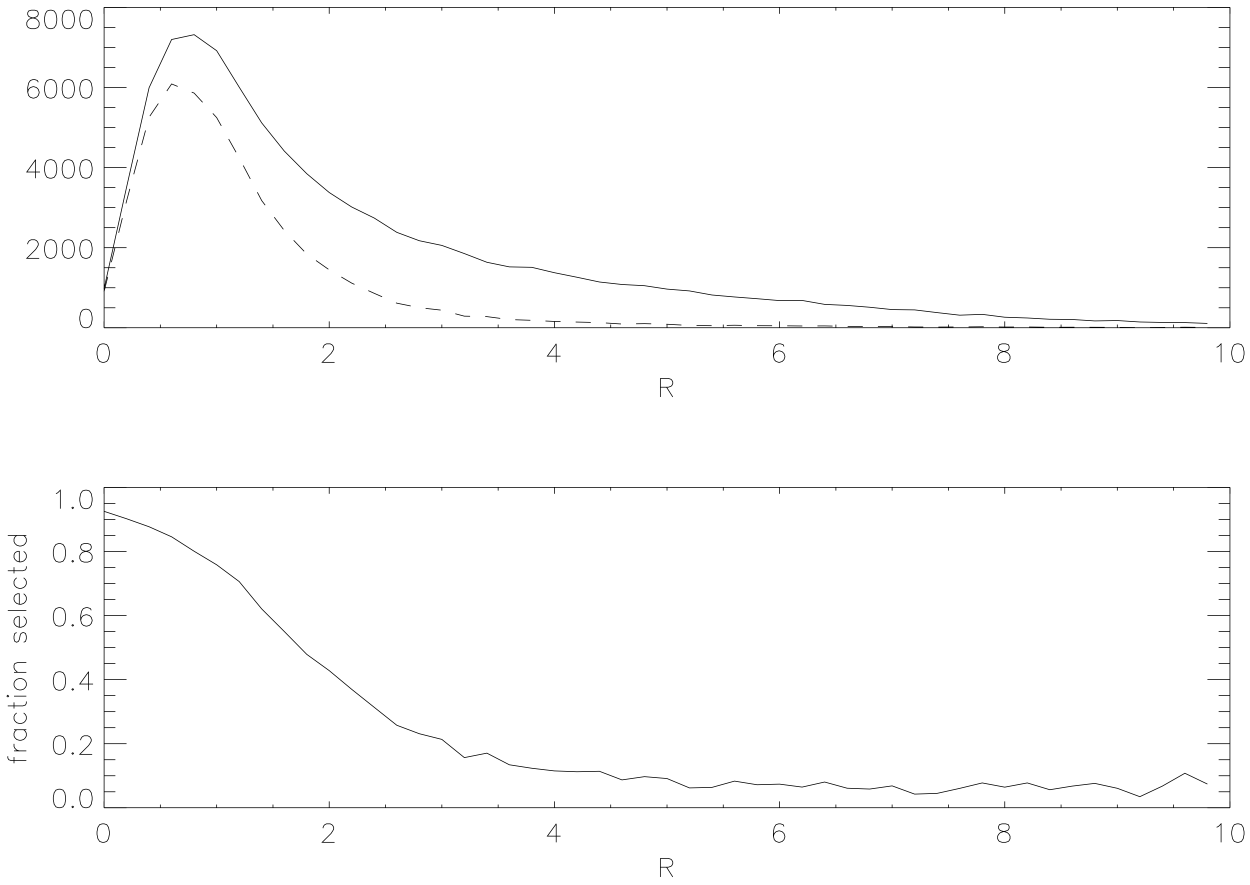


Figure 8: Upper plot: number of inversion solutions occurring (solid line) and number of solutions closest to the NWP wind velocity (dashed line), as a function of normalized MLE R, for five orbits of data, and for the NSCAT-like part of the swath. Lower plot: ratio of these two quantities.

Plot made with bgwd58:~/Qscat/Wave/NMLE/mle_frac.pro.

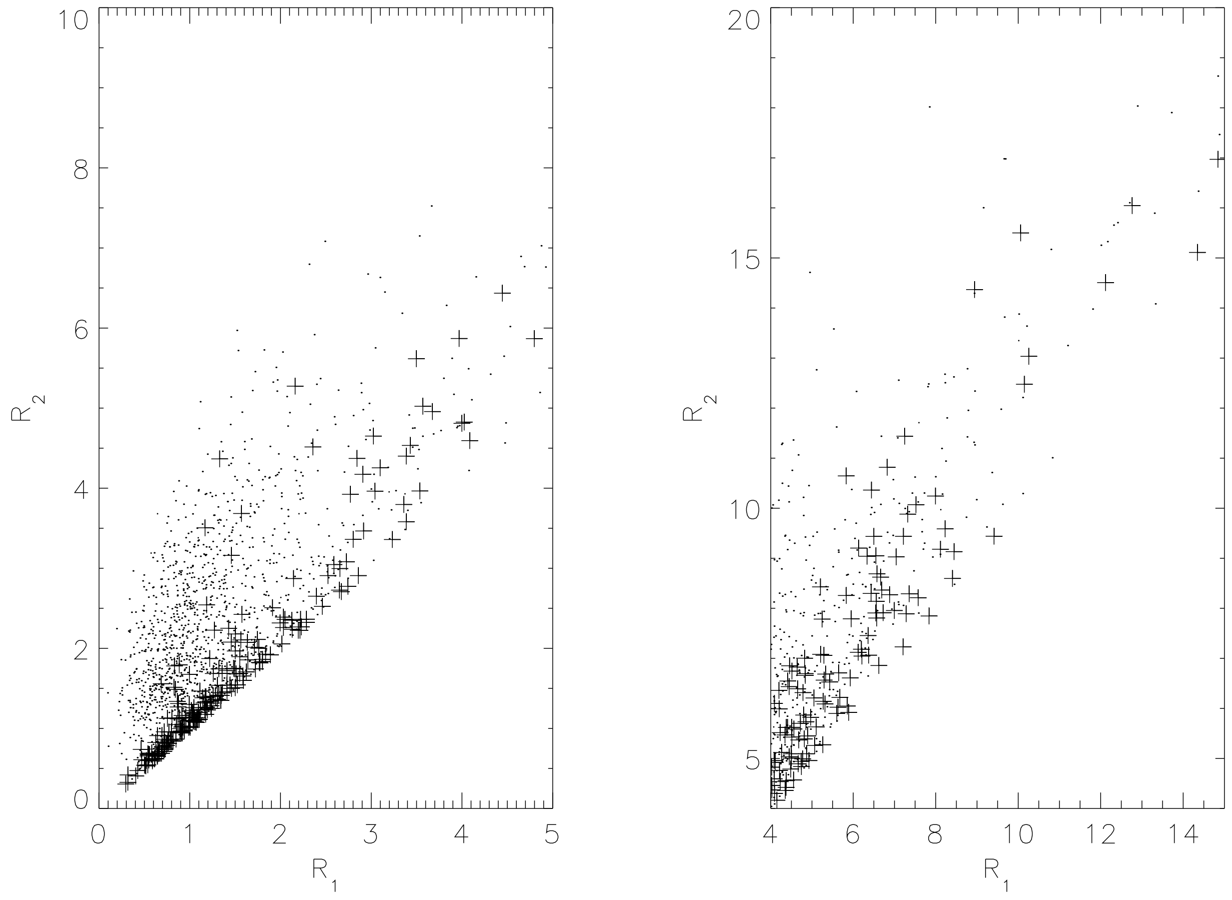


Figure 9: Scatter plot of first rank normalized MLE R_1 against second rank normalized MLE R_2, for cases with exactly two (JPL) solutions. Dots indicate pairs where the first rank is the one closest to the NWP wind velocity, pluses are pairs where the second rank is selected. Left panel highlights the lower range R_i's (1500 pairs used), right panel the higher range (15000 pairs used).

Plot made with bgwd58:~/Qscat/Wave/NMLE/scatfig.pro.

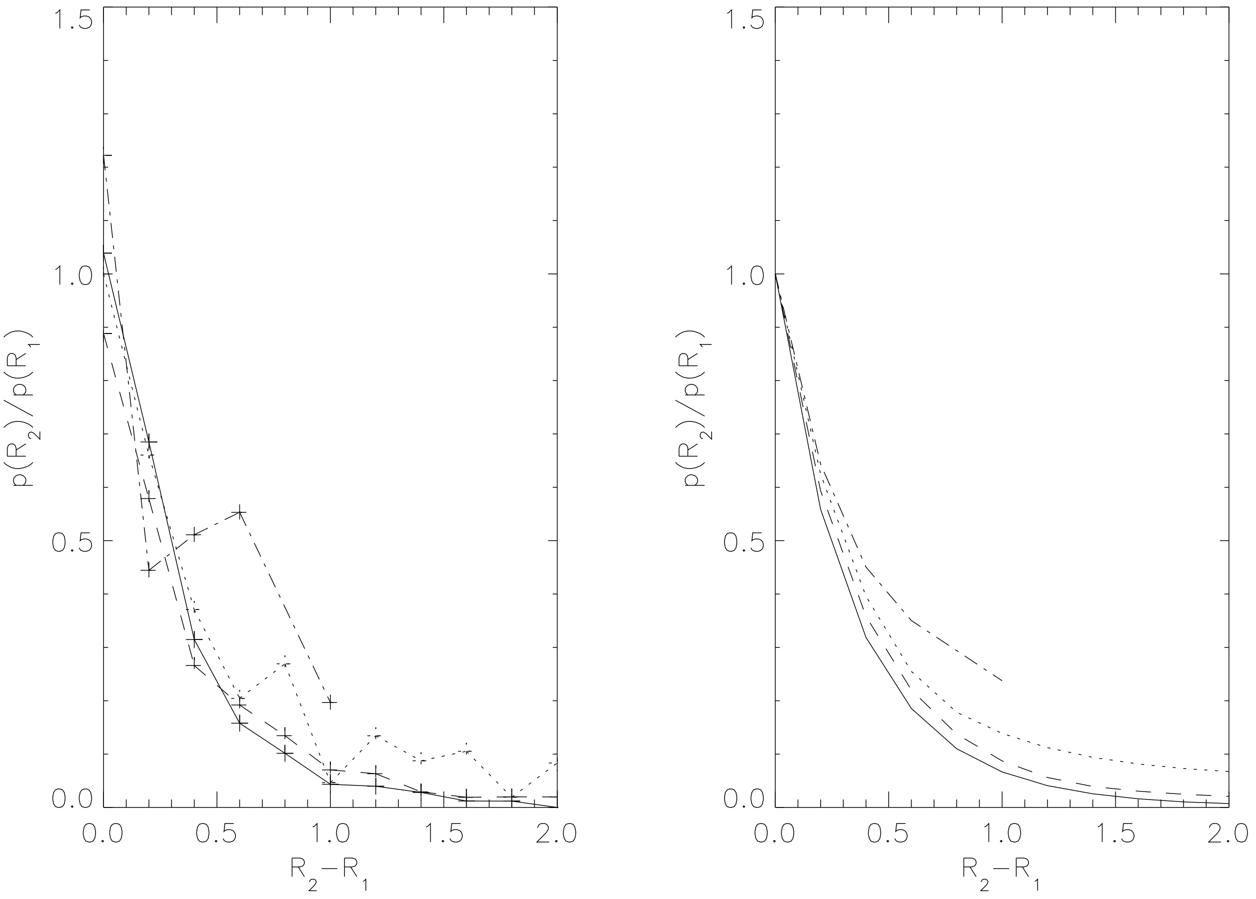


Figure 10: Ratio $p_s(R_2)/p_s(R_1)$ as a function of $R_2 - R_1$, for $R_1 = 0.5$ (solid), 1.1 (dashed), 1.7 (dotted), and 2.1 (dash-dot). Left panel: ratio determined experimentally using 5 orbits of data and applying equation (12). Right panel: ratio determined using the fit (13).

Plot made with bgwd58:~/Qscat/Wave/NMLE/plot_prob_rep.pro.

List of Tables

All data				
	Nr of solutions			
	2	3	4	all
N	45407	32881	47872	126160
rank 1	88 (89)	77 (74)	71 (70)	79 (78)
rank 2	12 (11)	18 (19)	19 (20)	16 (17)
rank 3	-	5 (8)	7 (6)	4 (4)
rank 4	-	-	3 (3)	1 (1)

$R_2 - R_1 \leq 1$				
	Nr of solutions			
	2	3	4	all
N	14742	18980	33905	67627
rank 1	70 (70)	65 (61)	62 (62)	65 (64)
rank 2	30 (30)	27 (28)	26 (26)	27 (27)
rank 3	-	8 (11)	8 (8)	6 (7)
rank 4	-	-	3 (4)	2 (2)

$R_2 - R_1 > 1$				
	Nr of solutions			
	2	3	4	all
N	30665	13901	13967	58533
rank 1	97 (98)	94 (90)	93 (90)	95 (94)
rank 2	3 (2)	5 (6)	4 (6)	4 (4)
rank 3	-	1 (4)	2 (3)	1 (1)
rank 4	-	-	1 (2)	0 (0)

Table 1: Distribution (in %) of the "selected" rank (i.e., the solution which is closest to NWP) over all ranks (in increasing order of JPL MLE value). Data are stratified w.r.t. total number of solutions (columns: 2, 3, or 4 solutions, or all cases together). N is the number of cases, summed up over all ranks. Percentages between brackets are the actual distributions observed in 5 orbits of data; percentages without brackets are the distributions predicted using eqs (11) and (13). The upper panel shows statistics over all data, the middle panel over data with $R_2 - R_1 \leq 1$, and the lower panel over data with $R_2 - R_1 > 1$.

Table generated using bgwd58:~/Qscat/Wave/NMLE/estimate.pro.