Royal Netherlands Meteorological Institute
Ministry of Infrastructure and Water Management

The EUMETSAT Network of Satellite Application Facilities

OSI SAF
Ocean and Sea Ice

EUMETSAT

Ocean and Sea Ice SAF

ASCAT land correction

SAF/OSI/CDOP3/KNMI/TEC/TN/384

03 June 2022

Jur Vogelzang and Ad Stoffelen

Royal Netherlands Meteorological Institute (KNMI)
De Bilt, the Netherlands

DOCUMENTATION CHANGE RECORD

Reference: SAF/OSI/KNMI/TEC/TN/384

Issue / Revision :	Date :	Change :	Description :
Version 0.9	Mar 2022	draft	Jur Vogelzang
Version 1.0	June 2022	first public version	Jur Vogelzang and Ad Stoffelen

Summary

EUMETSAT plans to add a land fraction to the full resolution radar cross section L1B product which is the base for coastal processing in the operational ASCAT Wind Data Processor (AWDP). This opens the possibility to correct full resolution measurements that fall partly over land and retrieve wind much closer to the coast than possible with the current operational products.

Simply replacing the current, rather conservative land fraction used in operational processing by the new one does not lead to any significant improvement. Therefore a land correction algorithm is presented, based on the assumption that within the area contributing to a wind vector cell land and sea have constant radar cross section. To reduce the frequency of K_p flag settings near the coast a weighted averaging is introduced to calculate the radar cross section in a wind vector cell, with Gaussian weights depending on the regression error.

The land-corrected ASCAT winds are compared with ECMWF forecast winds and buoy winds. The wind speed distributions close to the coast (50 km offshore at most) of the land-corrected product behave more realistic than those of the ECMWF model, which is known to possess flaws near the coast. A crucial parameter is the maximum land fraction that is accepted in the regression analysis. Its optimal value lies between 0.20 and 0.50. Buoy comparisons show that the standard deviations of the difference in the zonal and meridional wind components between the scatterometer and buoy measurements increases when approaching the coast. Part of the increase is caused by unrealistic high ASCAT wind speeds in several buoys; one in a sheltered position in Alaska, one off the coast of Haiti and the others in the Great Lakes. Several parameters are tested for their usability as quality control indicator, and the maximum regression bias error gave the best results. It reduces the VRMS difference with buoys up to 5 km offshore to 3.8 m/s for the land-corrected product with maximum land fraction 0.20, and to 4.2 m/s for the product with maximum land fraction 0.50. However, imposing a threshold on the maximum regression bias error also filters out a lot of WVCs that compare well with buoys. Further comparison with high-resolution wind data is recommended in order to further assess the accuracy.

The land-corrected product contains useful information close to the coast, but its real value has to be assessed by users.

Contents

1 Introduction

1.1 Problem formulation

Local wind fields, such as land-sea breezes and katabatic wind flows, convective cells, and coastal jets, strongly affect the microclimate in coastal regions. They determine to a large extent the advection and dispersion of pollutants in the atmosphere and coastal waters (by generation of local wind driven currents). Since most of the world's population lives in coastal areas and most pollutants are released into the environment near coasts, the generation of local winds is of great relevance for environmental purposes. As such, near-shore sea-surface wind field information can be important in a number of nowcasting, climate and ocean applications, such as those in semi-enclosed seas, straits, along marginal ice zones and in coastal regions. Wind measurements in the coastal zone are very valuable for all kinds of purposes. Scatterometer measurements near the coast may be contaminated by echoes from the land, and therefore only measurements far offshore are currently being used.

A first improvement in coastal processing came with the ASCAT coastal product at 12.5 km grid size. Here, the radar cross section per beam of a wind vector cell (WVC) is not calculated from the full resolution measurements as a Hamming-windowed average over a square area of 50 km × 50 km, as in the EUMETSAT L2A product, but as a boxcar average over a circle with 15 km radius [*Verhoef et al.,* 2012]. This decreases the minimum distance between the center of a WVC and the coast from about 35 km to about 20 km or slightly less.

The ASCAT coastal product does not take the shape of the Spatial Response Function (SRF) into account. The ASCAT SRF is to a good approximation an ellipse with a major axis of about 20 km and a minor axis of about 5 km [*Lindsley and Long*, 2015]. Due to the combined action of the Doppler effect and the radar pulse chirp which varies per beam, the ellipses for each beam have about the same orientation. Depending on the coast line, full resolution measurements much closer to the coast can be used.

This property is used in the new ASCAT land fraction. It is based on the ASCAT SRF [*Lindsley et al.,* 2016] combined with the Global Self-consistent Hierarchical High-resolution Geography (GSHHG) database [*Wessel and Smith*, 1996]. The GSHHG database is used to construct a very high resolution land-sea map on which the ASCAT SRF's are projected and the land fraction for each SRF is calculated. The new ASCAT land fraction is evaluated in [*Vogelzang and Stoffelen*, 2020], and it was recommended to apply a relatively conservative land fraction of 0.02.

A further improvement in coastal processing is to estimate the radar cross section of land and that of sea for coastal measurements, and to correct for the land contamination. This is the subject of this report.

1.2 Aims and scope

In this report we describe a simple algorithm for correcting land contamination of coastal scatterometer measurements. The algorithm is based on linear regression, and the optimal parameter settings are determined for ASCAT on a 12.5 km grid using buoy measurements and ECMWF forecasts as a reference.

1.3 Introductory remarks

This report closely follows the line of research during the project. In chapter 2 the algorithm for land correction is presented.

Chapter 3 contains a first analysis of the wind-corrected product for a limited data set. It is found that the optimal maximum land fraction is between 0.20 and 0.50, and that the regression parameters are not well suited for quality control. It is also found that many coastal WVCs have the KNMI QC flag set because of high K_p flag setting frequency.

In chapter 4 a weighting procedure with width dependent on the regression error is introduced. This reduces K_p and hence KNMI QC flag setting frequency. The best results are obtained for Gaussian weights.

Chapter 5 discusses the comparison of three selected ASCAT wind products (operational and land-corrected with maximum land fraction 0.20 and 0.50) with the ECMWF model winds as a function of the distance to the coast. It is shown that the ASCAT land-corrected winds behave much more realistic close to the coast than the ECMWF winds, that are known to 'feel' the coast too far offshore.

Chapter 6 discusses the comparison of the three aforementioned ASCAT wind products with buoy measurements. The buoy measurements are obtained from three different sources: the ECMWF MARS archive, the NDBC data set, and the Copernicus Marine Service in-situ data set. The metadata that are needed to convert the buoy winds to 10 m neutral winds (anemometer height, temperature sensor height, and humidity sensor height) are often not available, limiting the number of usable buoys to 300. The standard deviation of the difference between buoy and ASCAT wind components (u and v) increases with decreasing distance to the coast. Part of this increase is caused by unrealistic high scatterometer winds at a few buoy locations: one in a mountainous area with an irregular coast line off the coast of Alaska, one off the coast of Haiti, and the others all in the Great Lakes. Several parameters are tested for their usability as quality control, and the maximum regression bias error gives the best result: it reduces the VRMS difference with buoys up to 5 km offshore to 3.8 m/s for the land-corrected product with maximum land fraction 0.20, and to 4.2 m/s the product with maximum land fraction 0.50. However, imposing a threshold on the maximum regression bias error also flags a large number of WVCs that compare well with buoys.

In chapter 7 the winds of a high-resolution land-corrected product (6.25 km grid size) is compared with those retrieved from a SAR image. This comparison supports the conclusions from the previous chapters.

The conclusions and recommendations are given in chapter 8.

2 Land correction algorithm

Scatterometer wind data are given on a regular grid of wind vector cells (WVCs). ASCAT wind processing requires a triplet of radar cross section values from the fore, mid, and aft beam, respectively. Each cross section in the triplet is an unweighted average of all full resolution measurements that fall within a specified distance from the center of the WVC [*Verhoef et al.*, 2012]. For a grid size of 12.5 km the distance is 15 km at most, for a grid size of 6.25 km it is 7.5 km. A full resolution measurement is rejected if its land fraction is more than 0.02.

Suppose now that for a given WVC in a coastal region the radar cross section of the contributing full resolution measurements depends linearly on the land fraction f as

$$\sigma^0 = af + b \tag{2.1}$$

For zero land fraction this yields the radar cross section of the sea, $\sigma^0_{sea} = b$; for unit land fraction $\sigma^0_{land} = a + b$. This assumes that σ^0_{sea} and σ^0_{land} are essentially constant within the area represented by the WVC, which in some cases may be too strong an assumption, but, anyway, the wind retrieval uses only one mean ocean backscatter value for each view, which is estimated by b. Figure 2.1 shows a real example (row 106 and node 38 for the first ASCAT-B file of 2017)

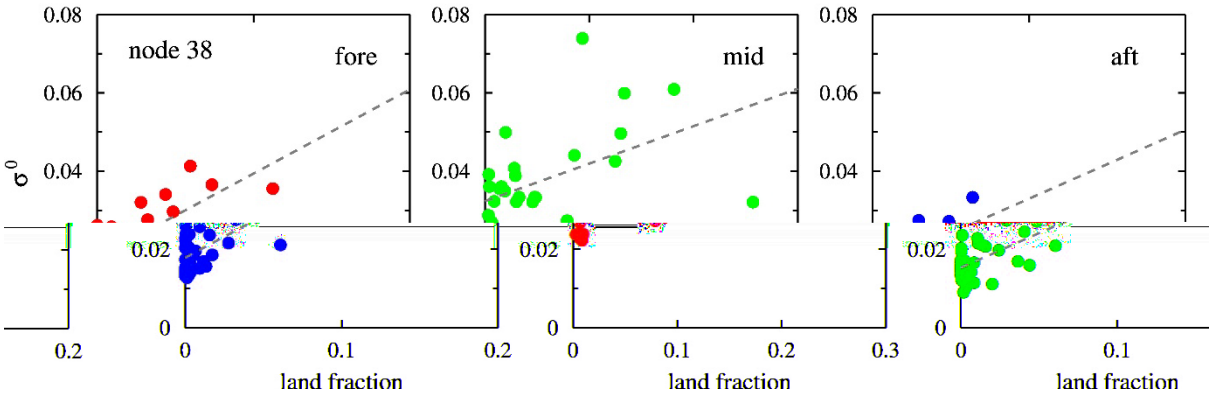

Figure 2.1 Example of land correction by regression.

The coefficients a and b can be obtained by a linear regression as

$$a = \frac{C_{f\sigma}}{C_{ff}} \tag{2.2}$$

$$b = M_\sigma - aM_f \tag{2.3}$$

where M_x is the first moment of quantity x and M_{xy} the second moment, with $x, y = \sigma, f$. Further, $C_{xy} = M_{xy} - M_x M_y$.

Once the regression coefficients are known, each full resolution radar cross section σ^0 can be corrected as

$$\sigma^0_{sea} = \sigma^0 - af \tag{2.4}$$

The σ^0_{sea} values are averaged per beam for each WVC to obtain the triplet of mean radar cross sections used for wind retrieval.

The mean square error (MSE) of the regression, σ_e, is given by

$$\sigma_e^2 = \frac{n}{n-2}\left(C_{\sigma\sigma} - 2aC_{f\sigma} + a^2 C_{ff}\right) \tag{2.5}$$

with n the number of (f, σ) pairs in the regression. Assuming Gaussian errors (a rather optimistic assumption in practice), the standard deviations in the regression coefficients a and b, denoted as σ_a and σ_b, respectively, read

$$\sigma_a^2 = \frac{\sigma_e^2}{nC_{ff}} \tag{2.6a}$$

$$\sigma_b^2 = \sigma_a^2 M_{ff} \tag{2.6b}$$

The errors σ_e, σ_a, and σ_b may be used for quality control.

Within a WVC and for a given beam, the land fractions will lie in a range $[f_{min}, f_{max}]$, and a threshold on f_{max} can also be applied for quality control. One also may consider a weighted average with weight $(1 - f)$ to account for the fact that σ^0_{sea} in (2.4) represents only part of the area covered by the full resolution measurement. This will returned to later in this report. However, such weighted averages are standard applied to calculate the geographical latitude and longitude of a WVC as well as the incidence angle and azimuth angle of each of its beams.

These considerations lead to the following algorithm for aggregation and land correction:

1. Standard averaging of all cross sections (and auxiliary data as geographical position and observation geometry), rejecting full resolution footprints with fraction between f_{max} and $1 - f_{max}$, with $f_{max} = 0.02$. If no footprint within the WVC is rejected, the WVC is over the open ocean or over land and no further processing is needed in the latter case;
2. If any footprint in the WVC has a land fraction exceeding f_{max}, then land correction is applied using only those footprints with land fraction below a threshold land fraction f_T, with f_T to be determined. A minimum number of three footprints is required.
3. If the land fraction fails the quality control in step 2 while step 1 led to a useful triplet of σ^0 values, the result of step 1 is retained.

In this way land correction is an additional processing step for coastal WVCs only.

As a first encouraging example, figure 2.1 shows the wind field in the Gulf of Tarente according to the current operational product (left hand panel) and new the land-corrected product without any additional quality control applied (right hand panel). Note that a significant number of WVCs near land is now unflagged and look well in agreement with the winds measured further offshore.

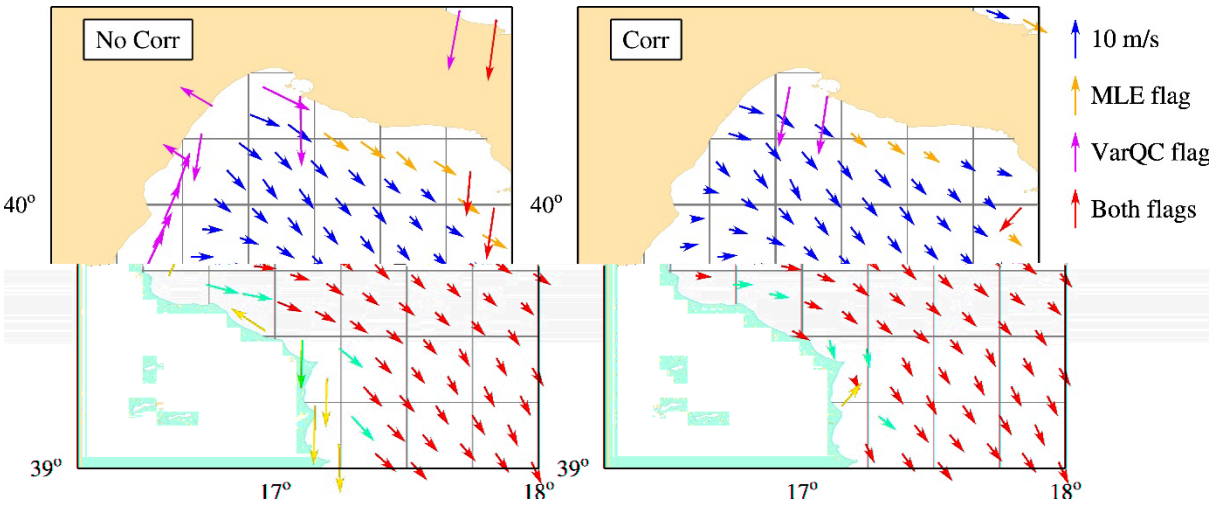

Figure 2.1 Wind field in the Gulf of Tarente on a 12.5 km grid. Left: current operational product. Right: product with land correction ($\boldsymbol{f_{max} = 0.02}$ and $\boldsymbol{f_T = 0.50}$).

3 Analysis

3.1 Classification

In order to investigate the various land correction settings in a systematic way, the operational OSISAF ASCAT wind product on a 12.5 km grid will be used as a reference (also referred to as the old product). The new product, of course, refers to the land-corrected wind product. Since the new land fraction was shown to have some problems around Antarctica [*Vogelzang and Stoffelen,* 2020], only WVCs between 60 °S and 60 °N were considered. These were divided into five classes according to the maximum value of the old land fraction in the reference product, f_L^{max}, over all three beams:

- class 0: $f_L^{max} \leq 10^{-6}$, same selection index and cell quality flag (identical open ocean cells);
- class 1: $f_L^{max} \leq 10^{-6}$, but different selection index and/or cell quality flag;
- class 2: $f_L^{max} > 10^{-6}$, both products have a valid selection index, but different selection indices and/or cell quality flags;
- class 3: $f_L^{max} > 10^{-6}$, only new product has a valid selection index;
- class 4: $f_L^{max} > 10^{-6}$, only old product has a valid selection index.

Class 0 is in the open ocean; the other classes in the coastal zone. Class 0 is not interesting, so this class is not considered in the further analyses. For class 4 one expects very few cases. This is indeed the case, see section 3.2, and most cases are caused by differences in the land map. Figure 3.1 shows as an example the Mississippi delta, an area with rapid morphological changes. The old product shows wind vectors that are very close to the coast and in some cases even unflagged winds above land. The new product without land correction, differing only from the old product in the land/sea map, follows the coast line much better. Application of land correction adds WVCs closer to the coast, but never above land. Another example (not shown) is Lake Rukwa in Africa, which apparently has a completely different shape in the old land map than in the new one. It is known that the water level of Lake Rukwa changed considerably over the years, so this is very likely the cause of the differences. These examples show that class 4 may be safely neglected in the following analyses.

3.2 Regression threshold land fraction

To speed up the calculations, the analyses presented in this section have been performed for one day of data, January 1, 2017. Only land corrected WVCs have been considered.

As stated in chapter 2, one of the parameters in the land correction processing is the threshold land fraction for regression, f_T. Figure 3.2 shows the number of WVCs not flagged by the KNMI Quality Control (KNMI QC) or the Variational Quality Control (VarQC) as a function of f_T. The curve type indicates the class, the color indicates the data used: red with land correction and blue without. The

number of class 4 WVCs decreases with f_T, stabilizing to about 30 WVCs for $f_T > 0.3$. The number of class 3 WVCs rises with f_T over the whole range of values.

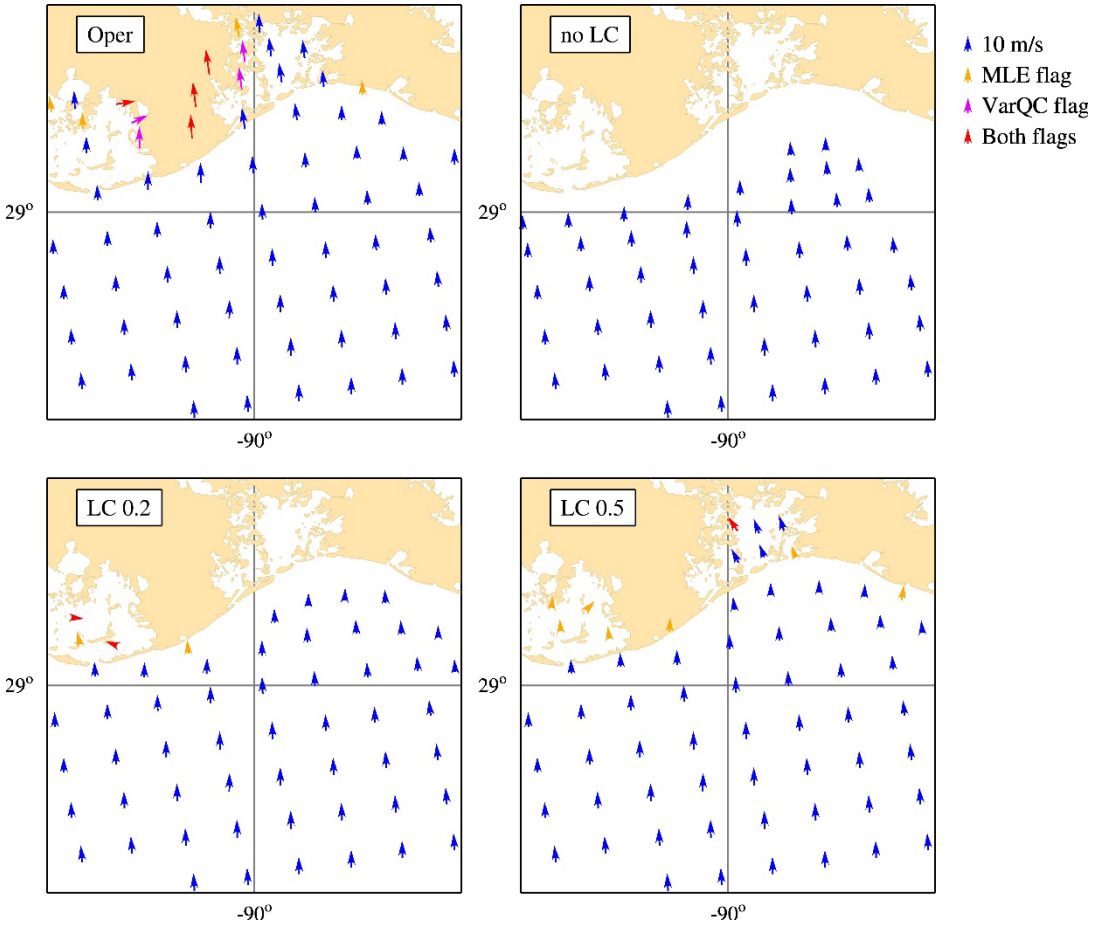

Figure 3.1 Mississippi delta. Upper left: old product; upper right: new product without land correction; lower left: land corrected product with regression threshold land fraction 0.20; lower right: land corrected product with regression threshold land fraction 0.50.

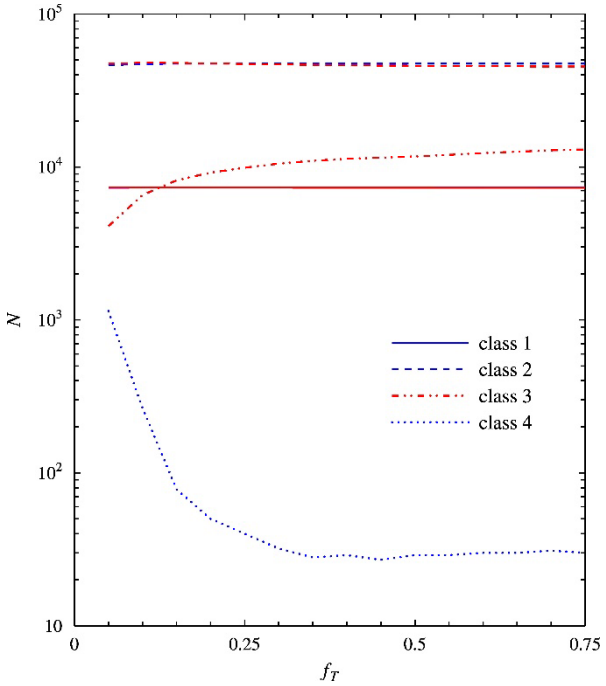

Figure 3.2 Number of unflagged WVCs per class as a function of threshold land fraction.

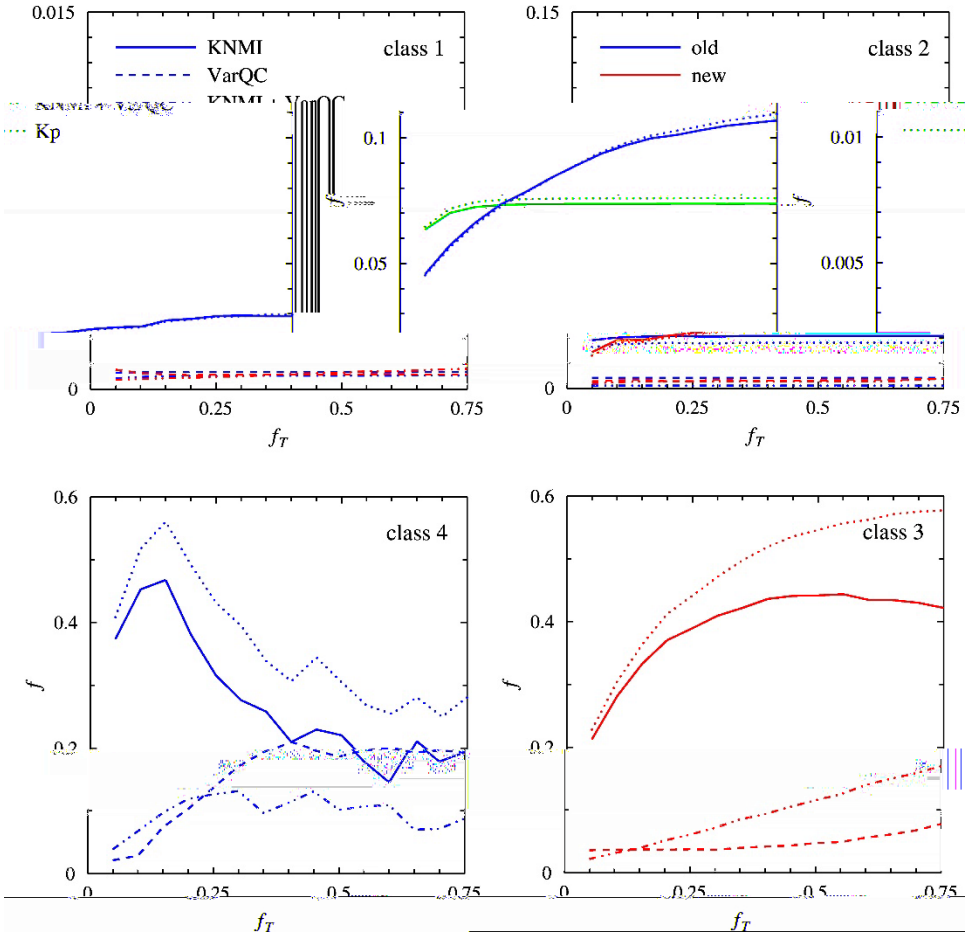

Figure 3.3 Flag setting frequencies as a function of threshold land fraction.

Note that there are relatively many WVCs in class 1. Closer inspection of the data showed that a large fraction of these have the same ambiguities, but in a different order. This is most likely caused by small numerical changes in the MLE caused by small land corrections in the radar cross sections. As such, most of the class 1 solutions remain practically the same.

Figure 3.3 shows the flag setting frequencies of various flags as a function of f_T. The red curves labeled "new" are with land correction, the blue curves labeled "old" without. Figure 3.3 shows that also the flag setting frequencies increase with f_T, in particular for classes 2 and 3. Note the frequent setting of the K_p flag (dotted curves). This will be returned to in chapter 4.

Figures 3.2 and 3.3 show that the number of unflagged WVCs or the flag setting frequencies give no clear criterion on the optimal value of the threshold land fraction. Such a criterion has to be derived from other properties of the data.

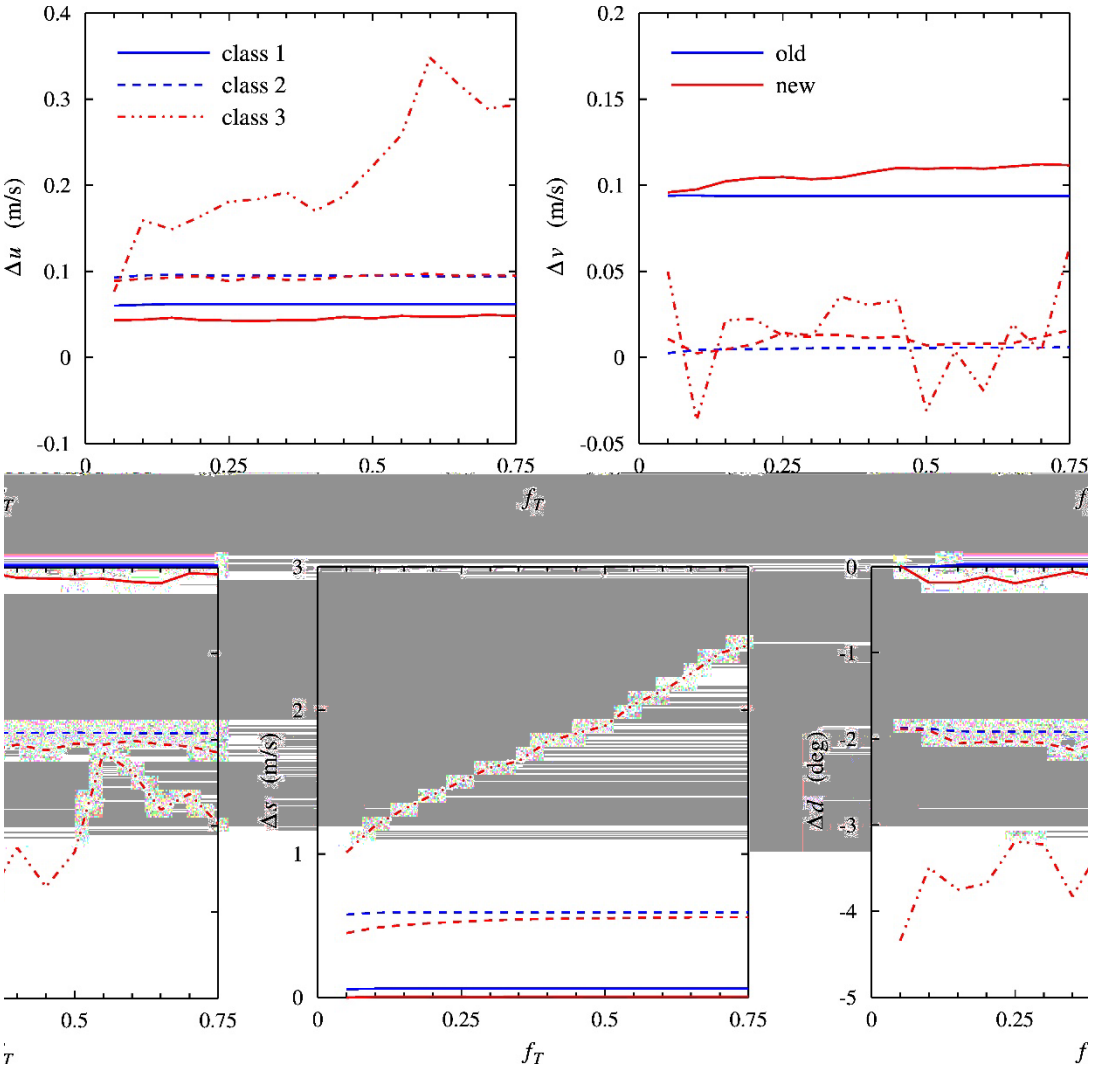

Figure 3.4 Average difference between scatterometer wind and ECMWF forecast as a function of $\boldsymbol{f_T}$ for the wind components $\boldsymbol{u}$ (upper left) and $\boldsymbol{v}$ (upper right), wind speed (lower left), and wind direction (lower right).

Figure 3.4 shows the average difference between the scatterometer wind and the collocated ECMWF forecast for the wind components, the wind speed, and the wind direction as a function of f_T. For classes 1 and 2 the results are close together. Land correction (red curves) decreases the difference for u and for wind speed, but increases it for v and wind direction. For class 3 the difference in u and in wind speed increase with f_T. It is fairly constant for v and slightly decrease for wind direction. The class 3 differences are clearly larger than those for classes 1 and 2. This does not necessarily mean that the class 3 winds have much poorer quality, because these winds are very close to the coast, where also model winds may have substantial errors.

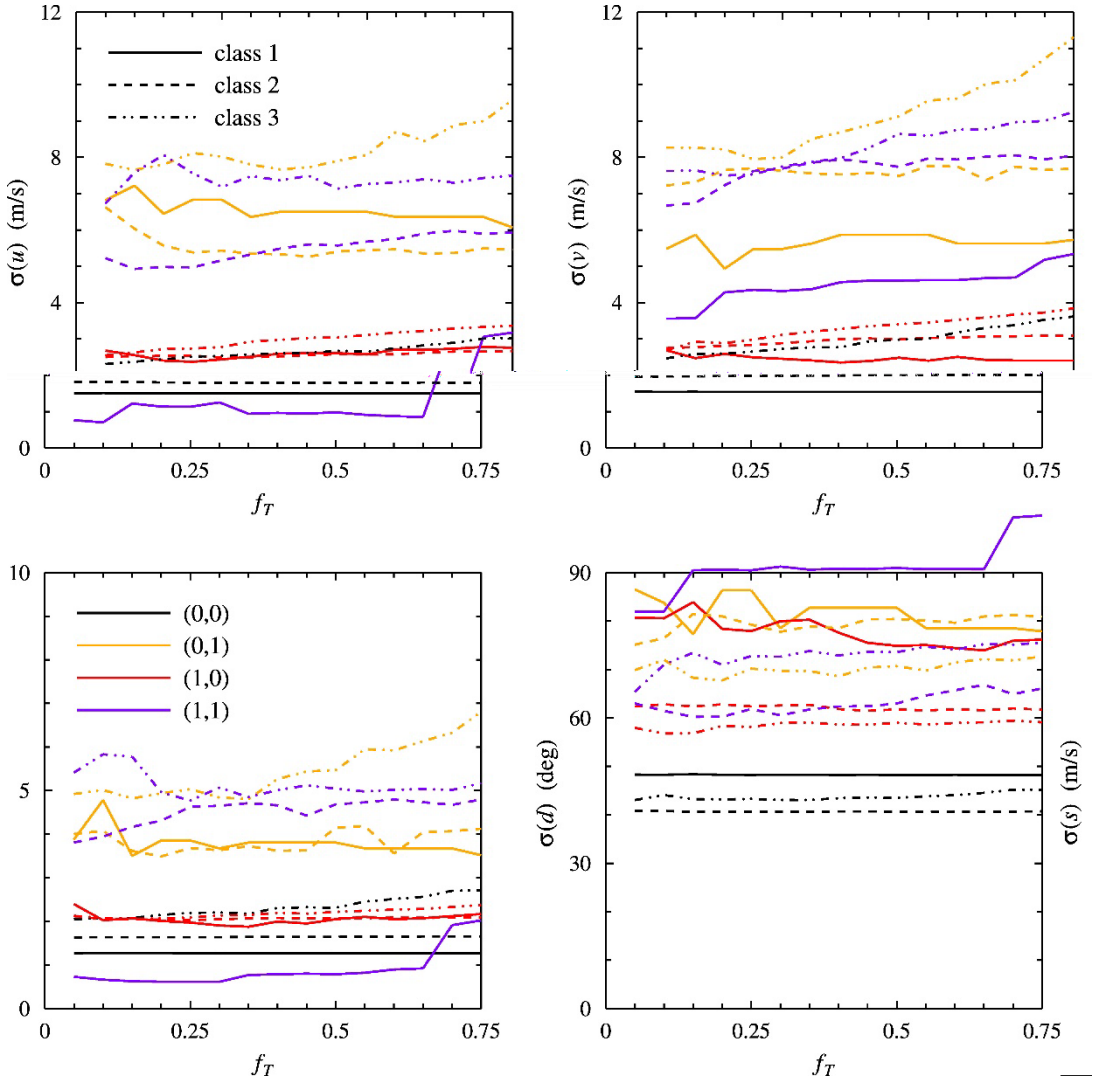

Figure 3.5 Standard deviation of the difference between land corrected scatterometer wind and ECMWF forecast as a function of f_T for the various classes and QC flag settings.

Figure 3.5 shows the standard deviations of the difference between the land corrected scatterometer winds and the collocated ECMWF forecasts as a function of f_T. The curve type indicates the class as given in the legend; the curve color gives the (KNMI QC , VarQC) flag values, the value 1 meaning

that the flag is set. All standard deviations show little variation with f_T, so comparison with the ECMWF forecast also gives no clue on the optimal regression threshold land fraction.

A possible clue may be given by figure 3.6 which shows the wind field over the Azores Isles for the same wind products as in figure 3.1. The southernmost island, Pico Island, has a volcano of 2300 m height that should disturb the wind around the isle. This is indeed the case for the land corrected products with $f_T = 0.20$ (lower left) and $f_T = 0.50$ (lower right), but the product with $f_T = 0.50$ appears too smooth. From visual inspection of this example and others a regression threshold land fraction between 0.20 and 0.50 seems optimal.

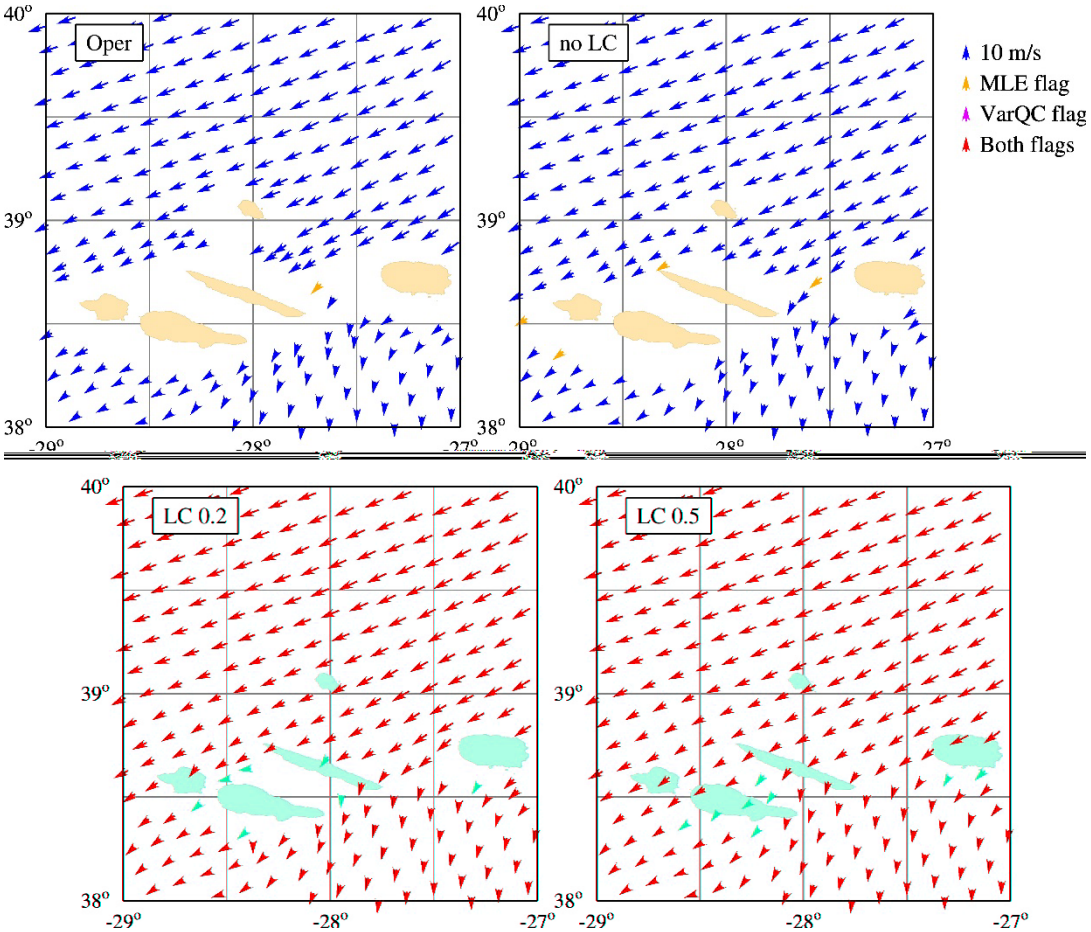

Figure 3.6 Wind field over the Azores. Same wind products as in Figure 3.1.

3.3 Regression parameters

We now take a closer look at the distributions of the values of the regression parameters to see if they may be used as quality indicators, choosing a regression threshold land fraction $f_T = 0.50$. Figure 3.7 shows the histograms of the regression scaling for the fore beam, for various classes as indicated by the

colors of the curves and for various QC flag settings, indicated upper right of each panel as (knmiQC,varQC).

Figure 3.7 shows a clear peak at $a = 0$ for the class 1 WVCs. In many cases most of the full resolution measurements will have only small land fractions, and the regression will be over a small range, resulting in a scaling close to zero. The other classes show broad peaks with center at $a \approx 0.05$. For all classes and all QC settings also large values of a are found. Similar results are found for the mid and aft beams (no results shown).

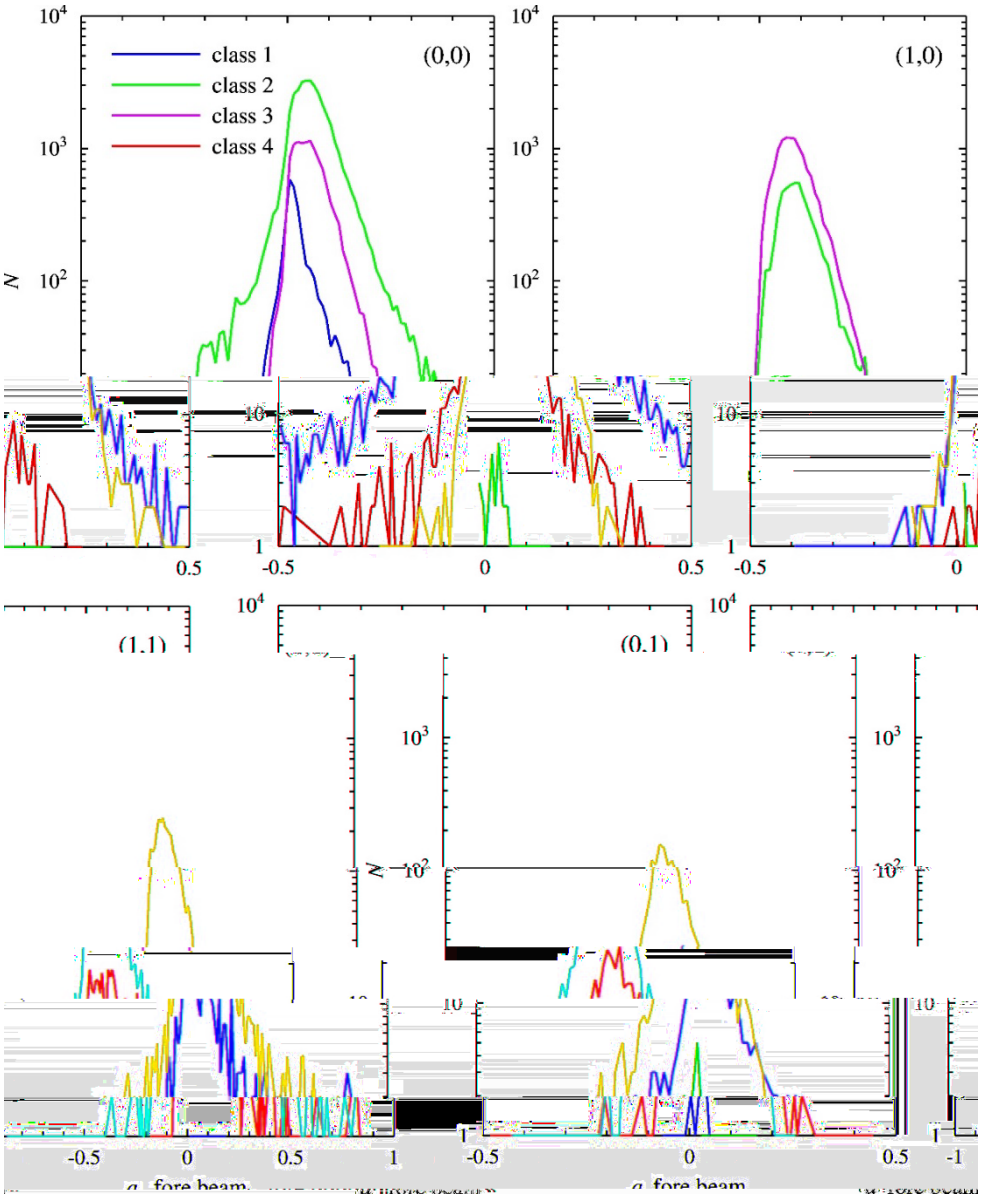

Figure 3.7 Fore beam regression scaling for various classes and QC settings

Figure 3.8 shows the histograms of the mid-beam regression bias for the various classes (indicated by the curve colors) and QC settings (indicated upper right in each panel). The regression bias *b* equals the radar cross section of the sea according to the regression formula in section 2. Note that there is a substantial number of negative biases, in particular for class 2 and 3 when the KNMI QC flag has been set. The mid beam contains more negative biases than the fore and aft beams (no results shown), indicating that the mid beam regression bias may be used as quality indicator.

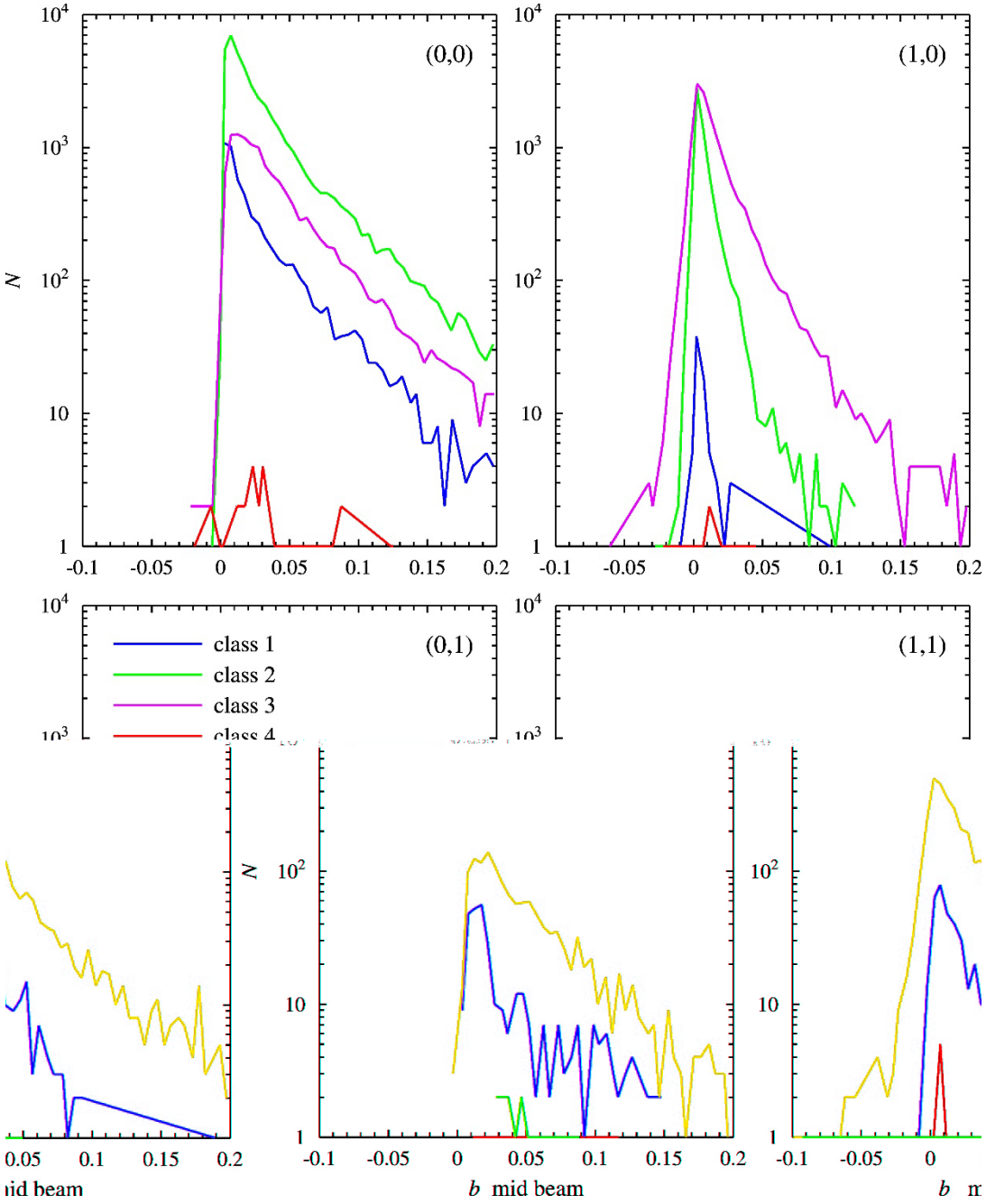

Figure 3.8 Mid beam regression bias for various classes and QC settings.

3.4 Regression errors

Figure 3.9 shows the regression error σ_e^2 (see equation 2.5) for the fore beam, again for the various classes and QC flag settings. The regression errors are smallest for the unflagged WVCs, and a maximum regression error between 0.002 and 0.004 will filter out some WVCs with the KNMI QC or VarQC flag set. The aft beam shows the same result, while the mid beam has larger regression errors (no results shown). Here, a maximum regression error of 0.006 to 0.012 seems appropriate.

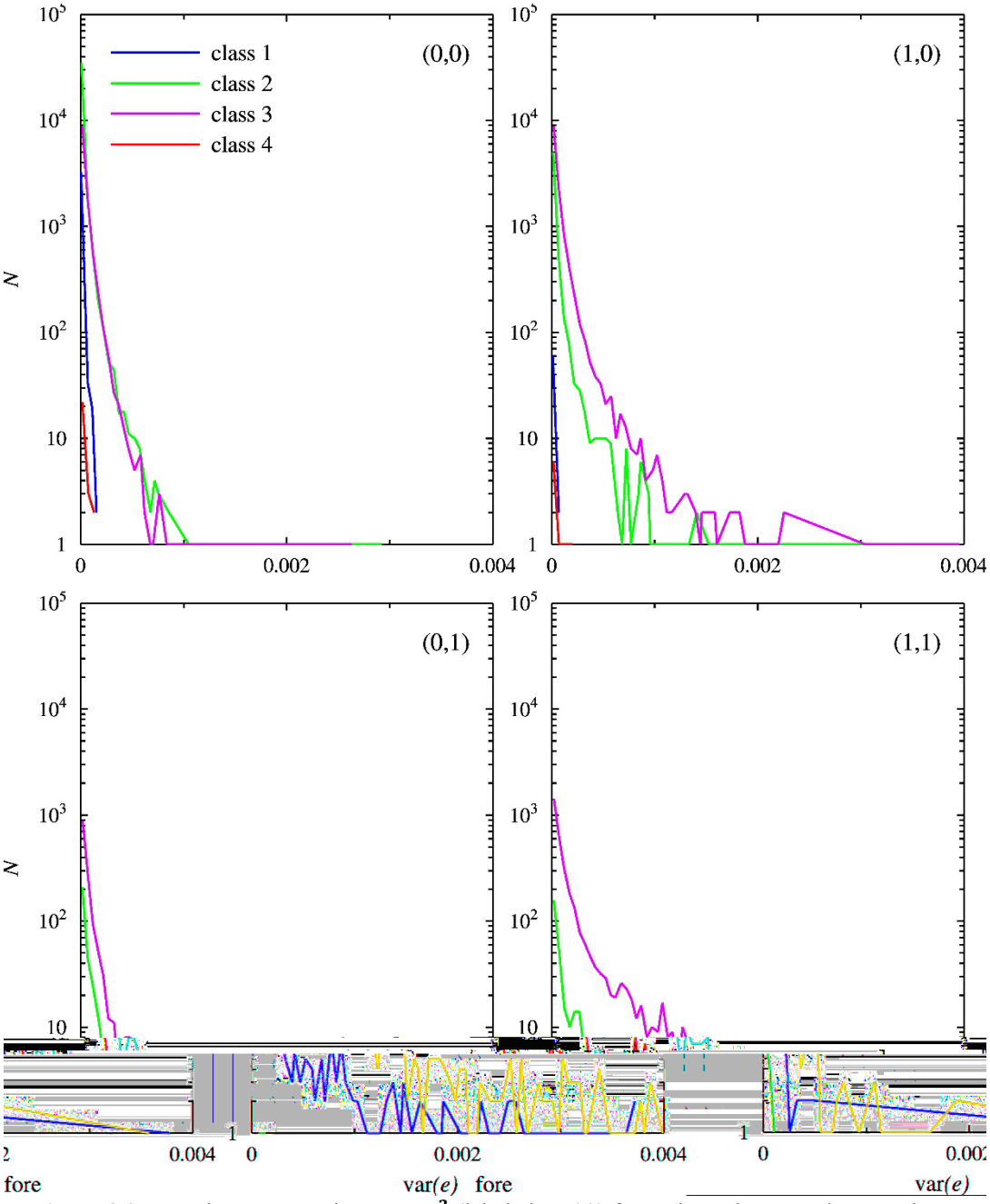

Figure 3.9 Fore beam regression error σ_e^2 (labeled var(e)) for various classes and QC settings.

Figure 3.10 shows histograms of the bias error σ_b^2 (equation 2.6b) of the fore beam. Clearly, the unflagged winds have the smallest bias error, and a maximum value of 0.00005 will filter out some WVCs with the KNMI QC or VarQC flag set. The aft beam shows similar results as the fore beam (no results shown). The mid beam has higher bias errors, and no clear threshold is visible (no results shown).

Figure 3.10 Fore beam bias error $\boldsymbol{\sigma_b}$

Figure 3.11 shows histograms of the regression scaling error σ_a^2 (equation 2.6a) of the fore beam for various classes and QC flag settings. The largest scaling errors are found for the unflagged winds. The mid and aft beams show similar patterns (not shown), so the regression scaling error has no potential to be used as QC indicator.

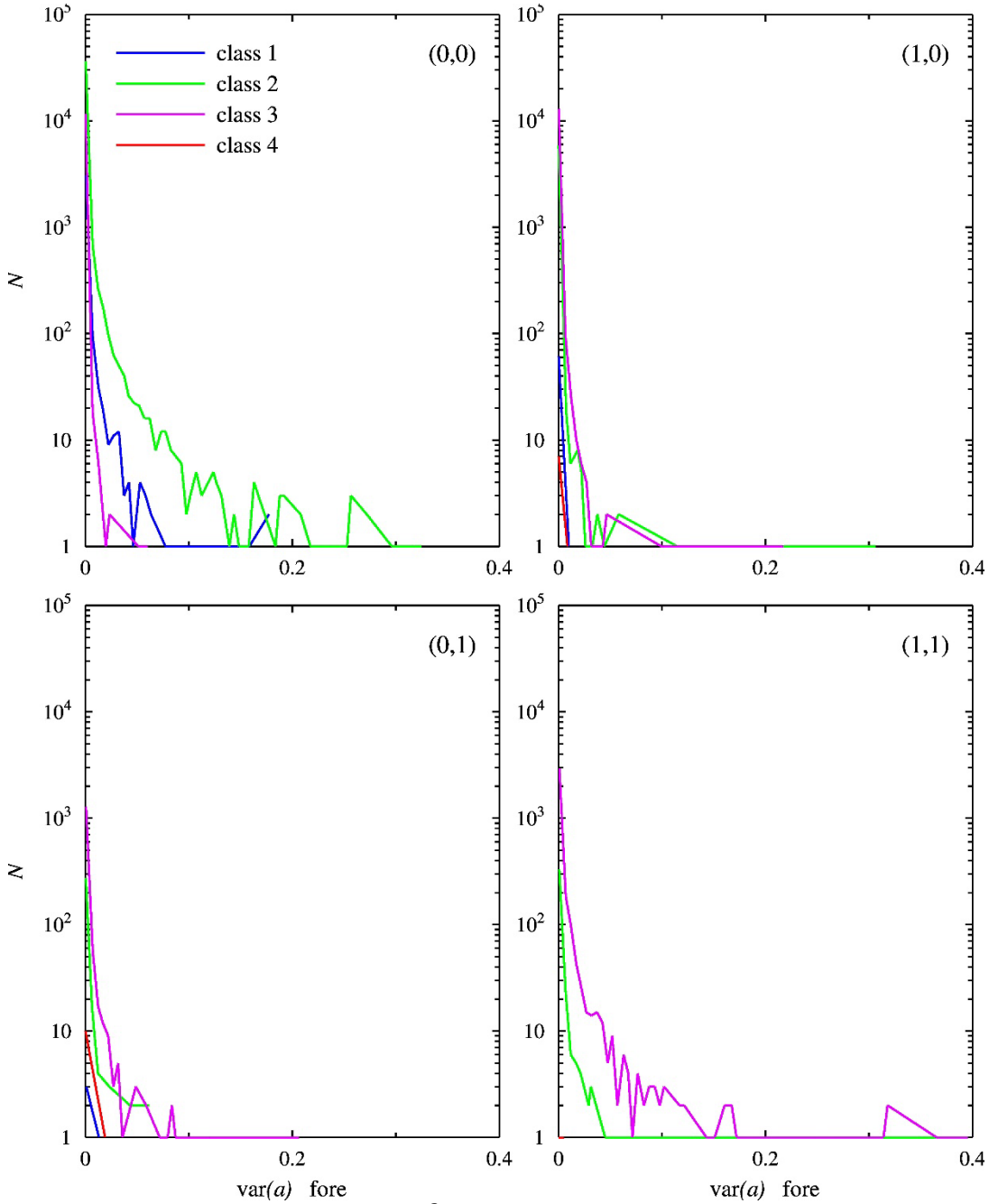

Figure 3.11 Fore beam scaling error $\boldsymbol{\sigma_a^2}$ (labeled var(a)) for various classes and QC settings.

3.5 *Correction to* σ^0

Figure 3.12 shows histograms of the relative correction to σ^0 defined as $\Delta\sigma^0/\sigma^0$, with $\Delta\sigma^0$ the difference between the radar cross section from the land corrected product and that from old product. The histograms are shown for class 1 (solid curves) and class 2 (dashed curves) and for the fore, mid, and aft beams (blue, red, and green curves, respectively), for the four QC classes.

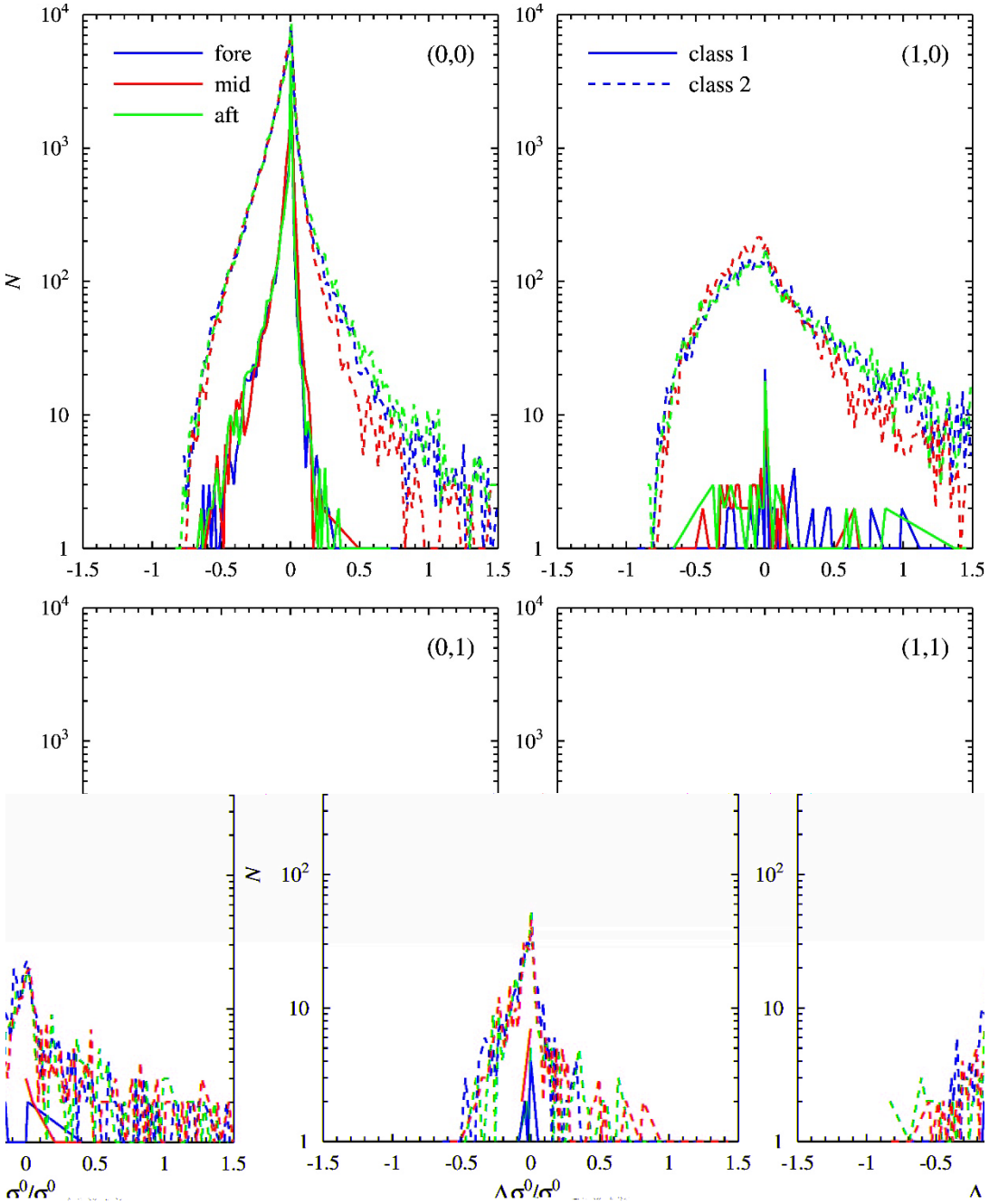

Figure 3.12 Histograms of the relative correction in σ^0.

Figure 3.12 clearly shows that for the unflagged winds (upper left panel) the majority of the corrections is negative, i.e., that the land corrected σ^0 is lower than the uncorrected σ^0. This clearly shows the ability of the regression algorithm to mitigate the effect of strong land reflections, in particular for the two top panels. The negative corrections are restricted to a minimum of about -0.6 for class 1 and -0.8 for class 2. Positive land corrections also occur, as the radar cross section of the land may be lower than that of the sea, notably at high winds. Here the flagged winds tend to have more and higher positive corrections than the unflagged ones, and no clear threshold value is visible.

3.6 Resume

The results of this section can be summarized as follows:

- The regression threshold land fraction, f_L^{max}, is hard to estimate. The number of unflagged WVCs or the flag setting frequencies give no clue, and neither does comparison with the ECMWF forecasts. Visual inspection of some cases suggests a threshold between 0.20 and 0.50.
- The regression error and the bias error may be used as QC indicators. The preliminary thresholds are given in table 3.1. However, these thresholds are based on the existing quality control mechanisms and will have only a cosmetic role by reducing the number of flagged WVCs. Moreover, they are estimated from one day of data only, so the statistics are unreliable. These points will be returned to in chapter 6.

Beam	σ_e^2	σ_b^2
Fore	$\lesssim 0.004$	$\lesssim 0.00005$
Mid	$\lesssim 0.012$	--
Aft	$\lesssim 0.004$	$\lesssim 0.00005$

Table 3.1 Preliminary QC thresholds

It should further be noted here that the additional land correction step adds only a few seconds to the total processing time of about 1 minute for a full orbit of ASCAT data.

4 Radar cross section weights

4.1 Introduction

After land correction by equation (2.4), the land corrected σ^0 values are averaged. This averaging can be done unweighted or weighted. With the latter choice uncorrected σ^0 values that lie far from the regression line can be given a lower weight, so they contribute less to the average σ^0 value used for wind inversion.

The distance of a given full resolution σ^0 value to the regression line, denoted as Δ, is given by

$$\Delta = \sigma^0 - af - b \tag{4.1}$$

with f the land fraction of the full resolution footprint under consideration, a the regression scaling, and b the regression bias. Note that this is the vertical distance, not the distance perpendicular to the regression line. The vertical distance is sufficient here, as it is also used in deriving the regression formulae in chapter 2. To get the perpendicular distance, (4.1) must be multiplied by a factor $1.0/\sqrt{1+a^2}$.

The weight function may be chosen at will. Here we apply exponential weights,

$$w_{exp} = \exp\left(-\frac{\Delta}{F\sigma_e}\right) \tag{4.2a}$$

or Gaussian weights,

$$w_{gaus} = \exp\left(-\left[\frac{\Delta}{F\sigma_e}\right]^2\right) \tag{4.2b}$$

with σ_e the regression error (2.5) and F an adjustable parameter that determines the amount of weighting. The average radar cross section, denoted as X, is then given by

$$X = \frac{\sum_i w_i \sigma_i^0}{\sum_i w_i} \tag{4.3}$$

where the summation is over all full resolution radar cross sections contributing to a particular beam in a WVC.

Such a weighting not only affects the average radar cross section, but also its K_p, defined as

$$K_p = \frac{std(X)}{X} \tag{4.4}$$

where $std(X)$ stands for the standard deviation in the average cross section. Now K_p is used for quality control and it should not exceed a threshold value K_p^{max}. In fact, the KNMI QC flag is a container QC flag that is set when the MLE exceeds a WVC-dependent threshold of about 18 or when the WVC is above ice, or when $K_p > K_p^{max}$.

Figure 3.3 showed the QC flag settings for the land corrected WVCs, and the K_p flag is set often. In fact, in most cases where the KNMI QC flag is set, it is due to the K_p flag. Note that figure 3.3 might give the wrong impression that the K_p flag is set more often than the KNMI QC flag, but the KNMI QC

flag setting frequencies are split according to whether the VarQC flag is set or not (solid and dot-dot-dashed curves, respectively).

This interferes with a possibility for quality control that has not been mentioned before. Since the regression bias b equals the radar cross section of the sea σ^0_{sea}, a threshold can be imposed upon the quantity σ_b/X which is similar to K_p. In this chapter the effect of using weighted radar cross section values is investigated further.

4.2 Averaging

As an example, figure 4.1 shows a part of the Philippines recorded January 1, 2017, with exponential weights of various strength in the average radar cross sections. The upper left panel is without averaging (corresponding to an infinitely large value for F), and many WVCs close to the land have the KNMI QC flag set (orange arrows and red arrows) because the K_p flag has been set.

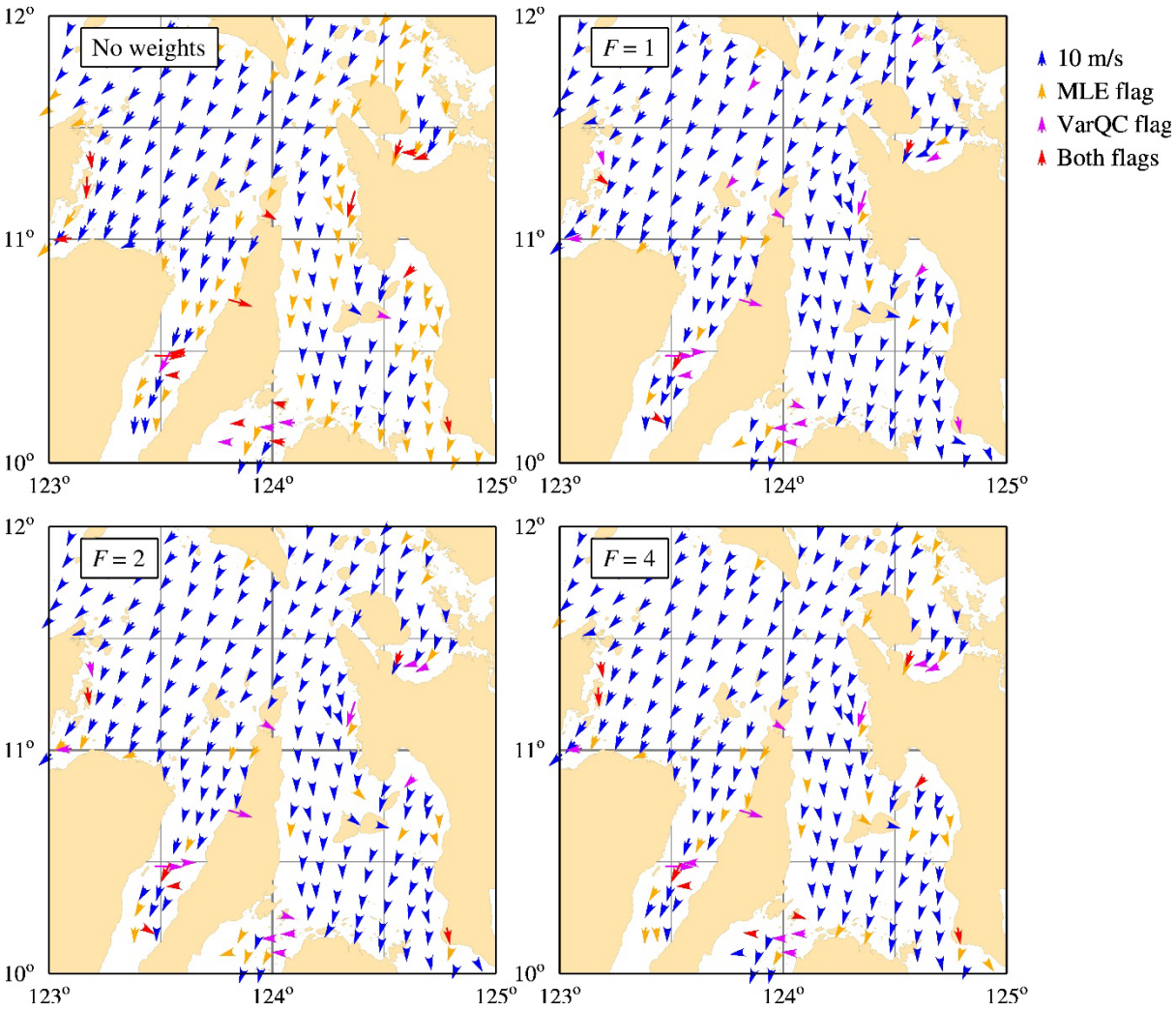

Figure 4.1 Part of the Philippines recorded January 1, 2017, with exponential radar cross section weights of various strength.

Reducing the strength F reduces the number of WVCs in which the KNMI QC flag has been set. A smaller value for F implies a narrower weight function, so full resolution radar cross sections far from the regression line get lower weight. In particular the result for $F = 1$ (upper right panel) looks reliable.

Figure 4.2 is similar to figure 4.1, but now for Gaussian weights. Compared to exponential weights, these give smaller weights to points far away from the regression line and larger weights to points close to the regression line. Again, the result for $F = 1$ (upper right panel) looks reliable, and slightly better than the corresponding exponential weighting in figure 4.1.

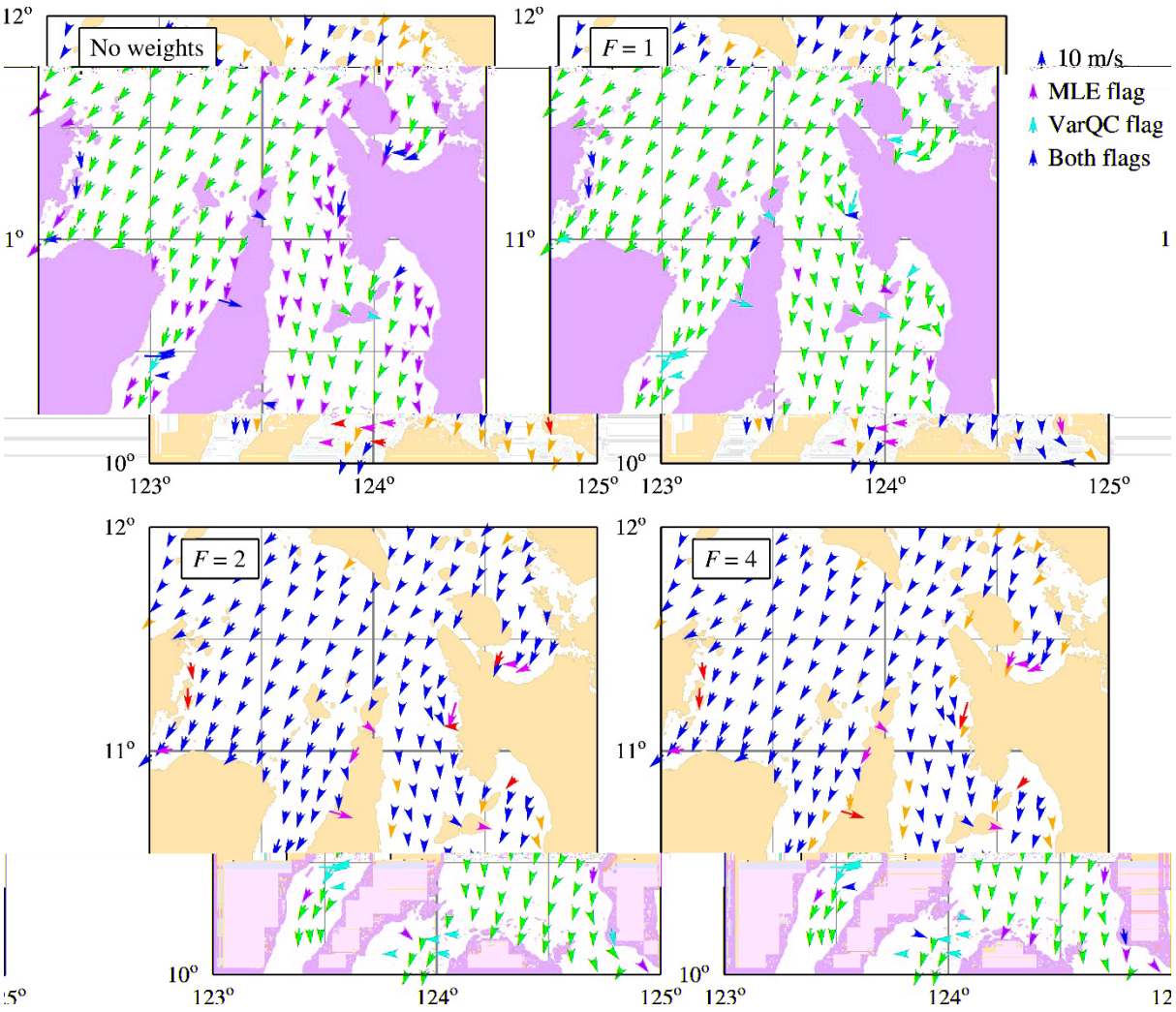

Figure 4.2 As figure 4.1, but with Gaussian weights.

5 Coastal statistics

5.1 Distance and direction to the coast

The averaging with Gaussian weights shown in the previous chapter seems to add many WVCs with good quality winds. In order to further investigate the correctness of these winds we calculate wind statistics as a function of the distance to the coast and of the wind direction relative to the coast. To this end we need to know for each WVC the distance and the direction to the coast.

The distance δ (in radians) between a WVC with geographical longitude and latitude coordinates (λ_1, φ_1) and a GHSSG database point with coordinates (λ_2, φ_2) is given by

$$\delta = 2\sin^{-1}\left(\sqrt{\left(\sin\frac{\varphi_1-\varphi_2}{2}\right)^2 + \cos\varphi_1\cos\varphi_2\left(\sin\frac{\lambda_1-\lambda_2}{2}\right)^2}\right) \tag{5.1}$$

(www.edwilliams.org/avform147.htm). This expression is slightly more complicated than the usual formula based on the spherical cosine rule, but less prone to round-off errors at small distances. The distance in km is obtained as $d = \delta R$ with R the radius of the Earth in km. The direction α (in radians) from the WVC to the coastal point reads

$$\alpha = \pi \quad \text{if } \cos\varphi_1 < 10^{-6} \text{ and } \varphi_1 > 0 \quad \text{(WVC on North pole)} \tag{5.2a}$$

$$\alpha = 2\pi \quad \text{if } \cos\varphi_1 < 10^{-6} \text{ and } \varphi_1 < 0 \quad \text{(WVC on South pole)} \tag{5.2b}$$

$$\alpha = \cos^{-1}\left(\frac{\sin\lambda_2 - \cos d\sin\lambda_1}{\sin\delta\cos\varphi_1}\right) \quad \text{if } \sin(\lambda_2-\lambda_1) < 0 \tag{5.2c}$$

$$\alpha = 2\pi - \cos^{-1}\left(\frac{\sin\lambda_2 - \cos d\sin\lambda_1}{\sin\delta\cos\varphi_1}\right) \quad \text{if } \sin(\lambda_2-\lambda_1) \geq 0 \tag{5.2d}$$

with δ from (5.1). A simple algorithm would be to go through the entire GSHHG data base for each WVC, calculate the distance and direction from (5.1) and (5.2), and retain the values for which δ is smallest. However, this takes too much computation time, in particular when using the full resolution GSHHG data base.

It is more efficient to calculate δ and α in advance on a fine grid and retrieve the values when needed. To this end the following algorithm is designed:

1. Define a global lat/lon grid with grid size 0.01˚ (for statistics calculations) or 0.1˚ (for illustrations). Define two arrays of one-byte integers (to reduce memory requirements) of dimension (N_φ, N_λ), one to contain d and one to contain α for each grid point. Initialize each array element to 127. Values of d will be stored in integer kilometers from 0 to 120, values of α in bins of 3˚ from 0˚ to 360˚ with index ranging from 1 to 120.
2. Go through the full resolution GSHHG database and set each array element containing a coast line point to 0.

3. Go through the array. If a coastal point (value 0) is found, calculate d and α for all surrounding points. Replace the current values of d and α by the new ones if the surrounding cell has not yet been initialized or if the new distance is smaller than the existing one.
4. Write the grids to unformatted files for further use.

An efficient search in step 3 is made if first the latitude index is varied. If the distance between the coastal point and the surrounding point at the same longitude exceeds 125 km, the next latitude is processed; if not the distance is calculated for eastward and westward surrounding points, until the maximum distance is reached. The value of α is calculated and stored only when the distance needs to be updated.

With this scheme the generation of a global map of distance and direction to the coast takes slightly more than 1½ hour on a Linux work station. Figures 5.1 and 5.2 show part of the maps of distance and direction, respectively, on a 0.1˚ grid for the Eastern Mediterranean. It looks like the direction map contains gaps near the coast, but this is an artefact of the plotting software.

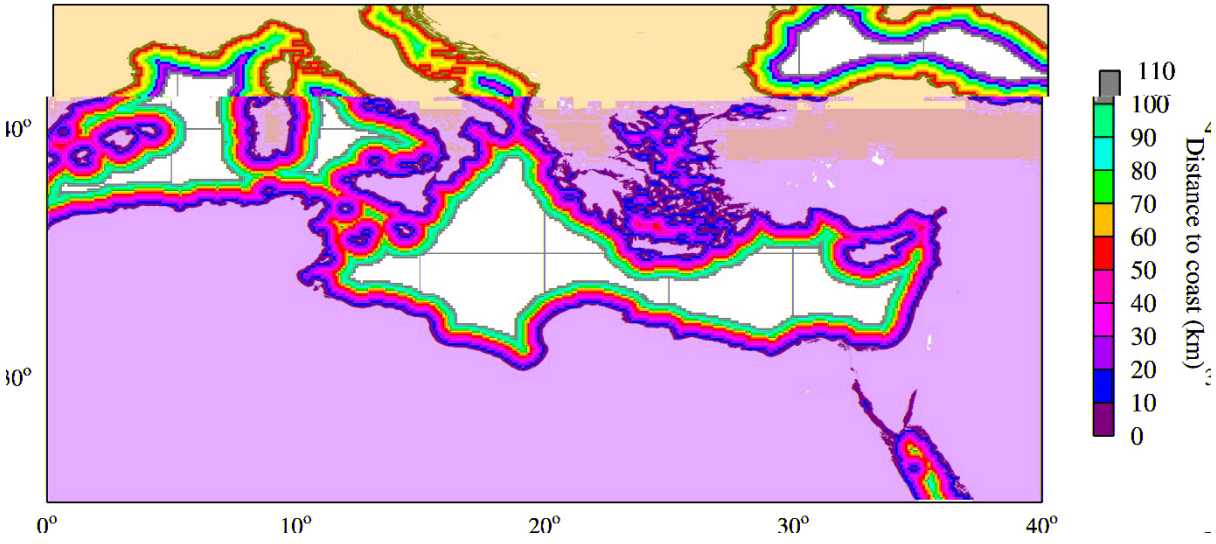

Figure 5.1 Distance to coast on a 0.1 degree grid for the Eastern and Central Mediterranean.

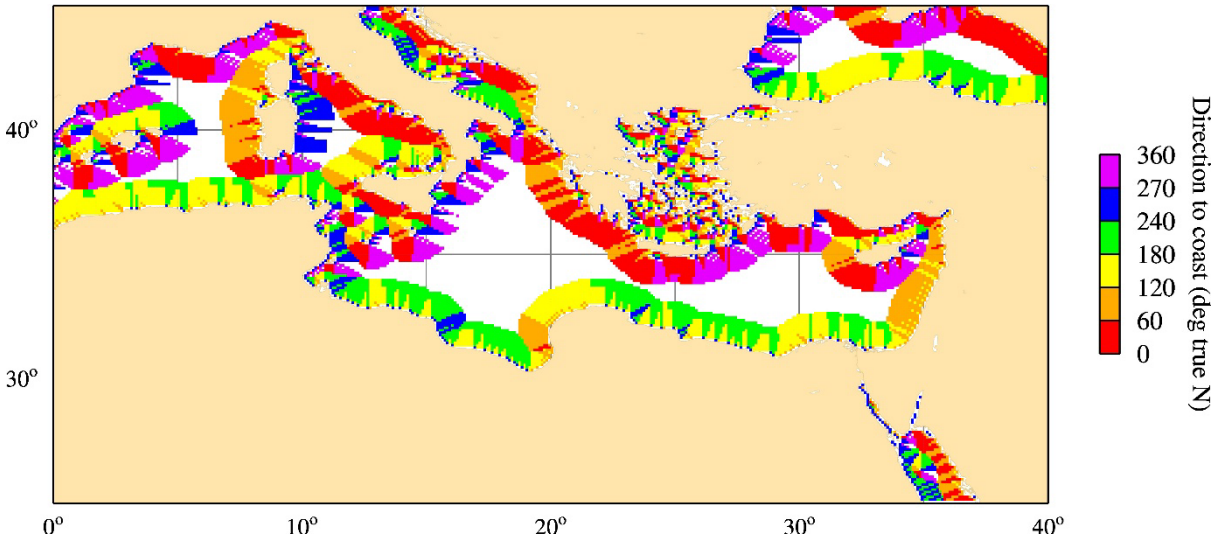

Figure 5.2 Direction to the coast on a 0.1 degree grid for the Eastern and Central Mediterranean.

5.2 Wind speed pdf

In order to obtain good statistics, the land-corrected product with Gaussian weight 1 averaging of the radar cross section was processed for all January 2017 with maximum land fractions 0.2 and 0.5. The pdfs as function of the distance to the coast (in bins of 10 km) are shown in figure 5.3. Three wind products are shown: the land corrected product with maximum land fraction 0.2 (upper left panel), that with maximum land fraction 0.5 (upper right panel), and the operational coastal product (lower left panel). The lower right panel shows the number of contributing WVCs per distance bin for each product.

The peaks of the pdfs of the land-corrected products shift towards lower wind speeds as the distance to the coast decreases, due to the influence of land. Land backscatter is generally higher than the backscatter of low winds and hence land contamination effects may show up as a PDF skewing to higher winds. On the other hand, land roughness is much larger than ocean roughness and hence upstream land influences will reduce the wind speed. The latter effect appears somewhat stronger for the product with maximum land faction 0.50, perhaps due to the relative abundance of samples < 10 km from the coast, but absent from the operational product. It may seem strange that even for the operational product there are WVCs close to the coast, though far less than for the land-corrected products, but this is due to very small isles that have only a limited effect. Note though that without correction and a coarser land-sea mask, relatively less low winds appear for the operational product in the category < 10 km, perhaps indicating remaining land contamination in this category. It is encouraging to note that this effect does not appear for the many data near the coast with maximum land fraction 0.5.

Figure 5.4 is similar to figure 5.3, but now for the ECMWF forecast collocated with each valid wind speed. Figure 5.4 shows that the shift of the pdfs towards lower wind with decreasing distance to the coast is much stronger for the ECMWF forecast than for the land-corrected ASCAT wind products,

even at relatively high wind speeds, which are generally on-shore. There is even some shift in the forecasts at the positions of the operational products. This confirms that the ECMWF model feels the coast too early and too strong, a known weakness of the model.

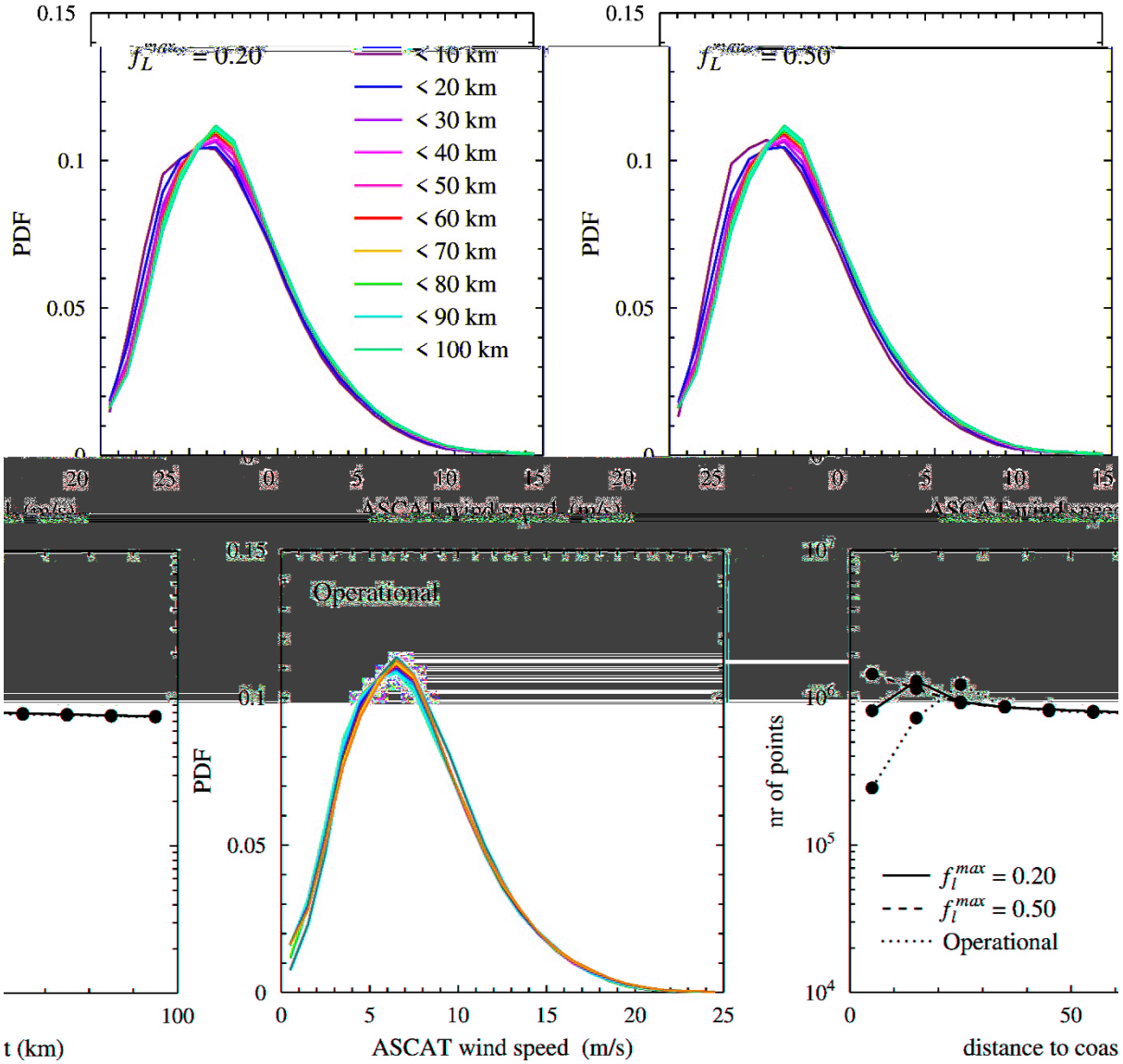

Figure 5.3 Wind speed pdf's per distance bin.

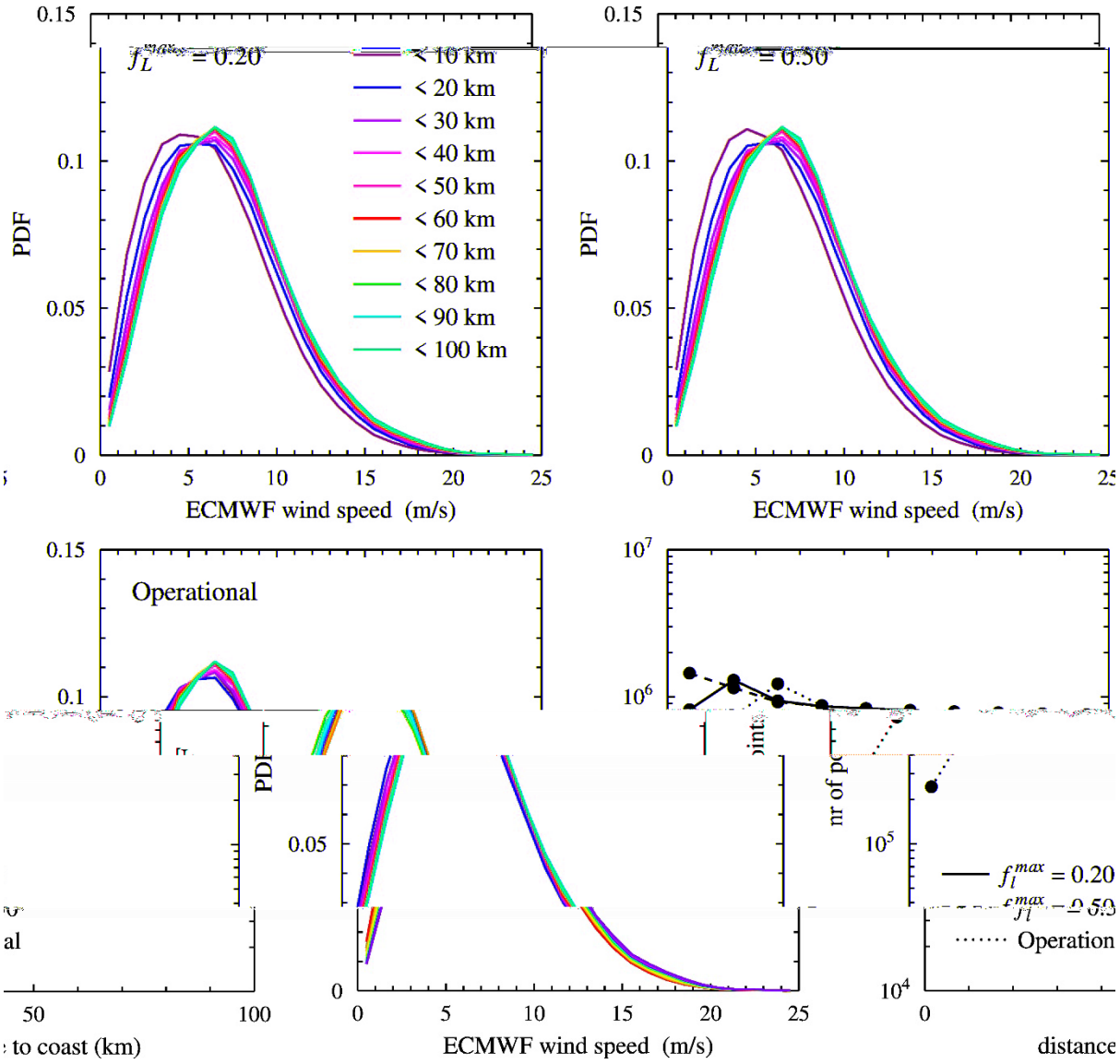

Figure 5.4 As figure 5.3, but for the ECMWF forecasts collocated with the ASCAT winds.

5.3 *Wind direction effects*

The pdfs in the previous section contained all wind directions. Now the effect of wind direction will be addressed. The winds will be split in three classes: on-shore winds where the wind direction makes an angle of 60˚ or less with the direction to the coast, off-shore winds where the wind direction makes an angle of 120˚ or more with the direction to the coast, and parallel winds for all other relative directions.

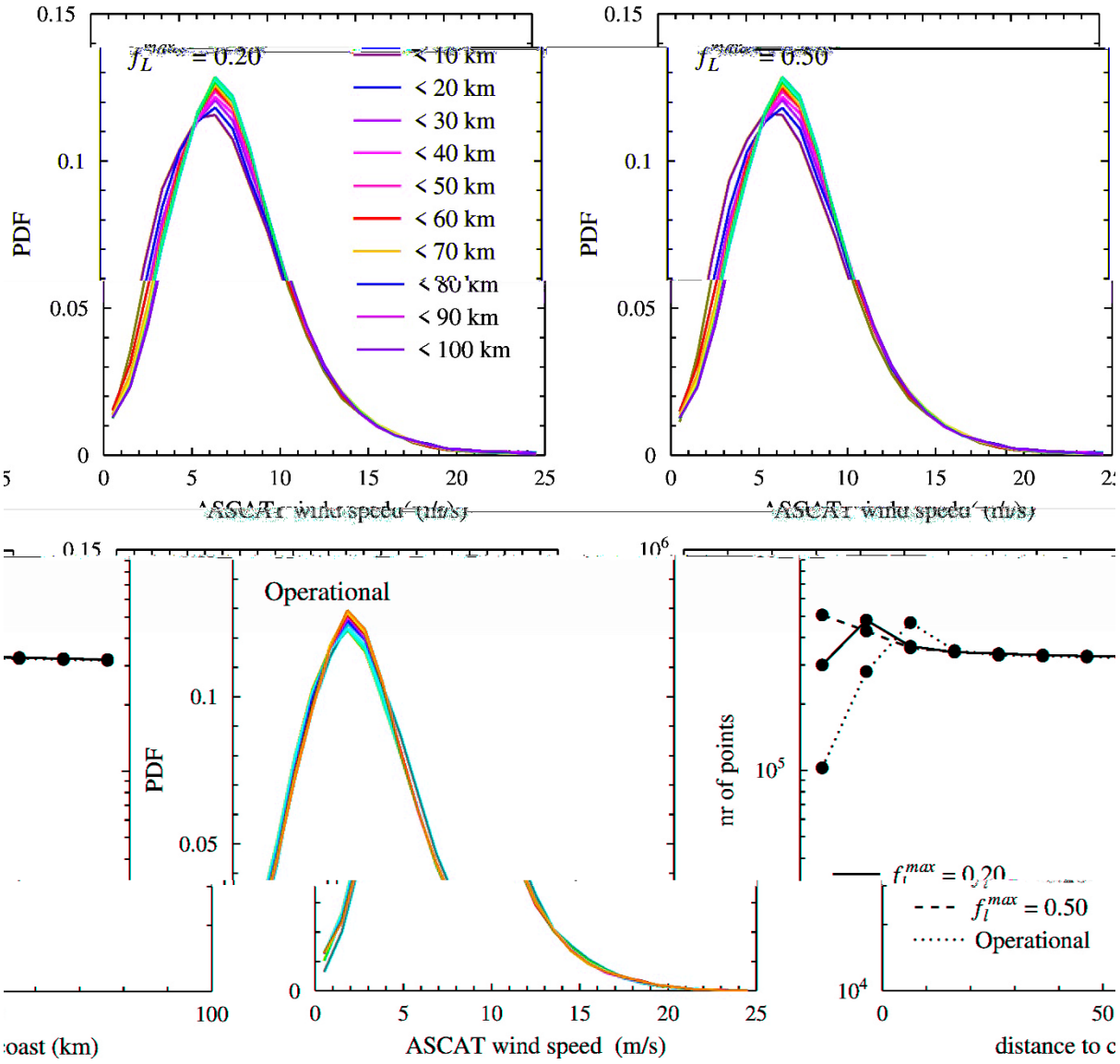

Figure 5.5 Wind speed pdfs as function of the distance to the coast for on-shore ASCAT winds.

Figure 5.5 shows the ASCAT wind speed pdfs as a function of the distance to the coast for onshore winds, figure 5.6 for offshore winds, and figure 5.7 for winds parallel to the coast. Note that most winds are on-shore, then parallel to the coast and least are off-shore. The land effect is clearly stronger for off-shore winds than for on-shore winds. The PDF peaks at lower speeds, while the winds clearly pick up further away from the coast as the internal marine boundary layer grows. . The PDFs of winds parallel to the coast in figure 5.7 lie between those for onshore and offshore winds, as may be expected.

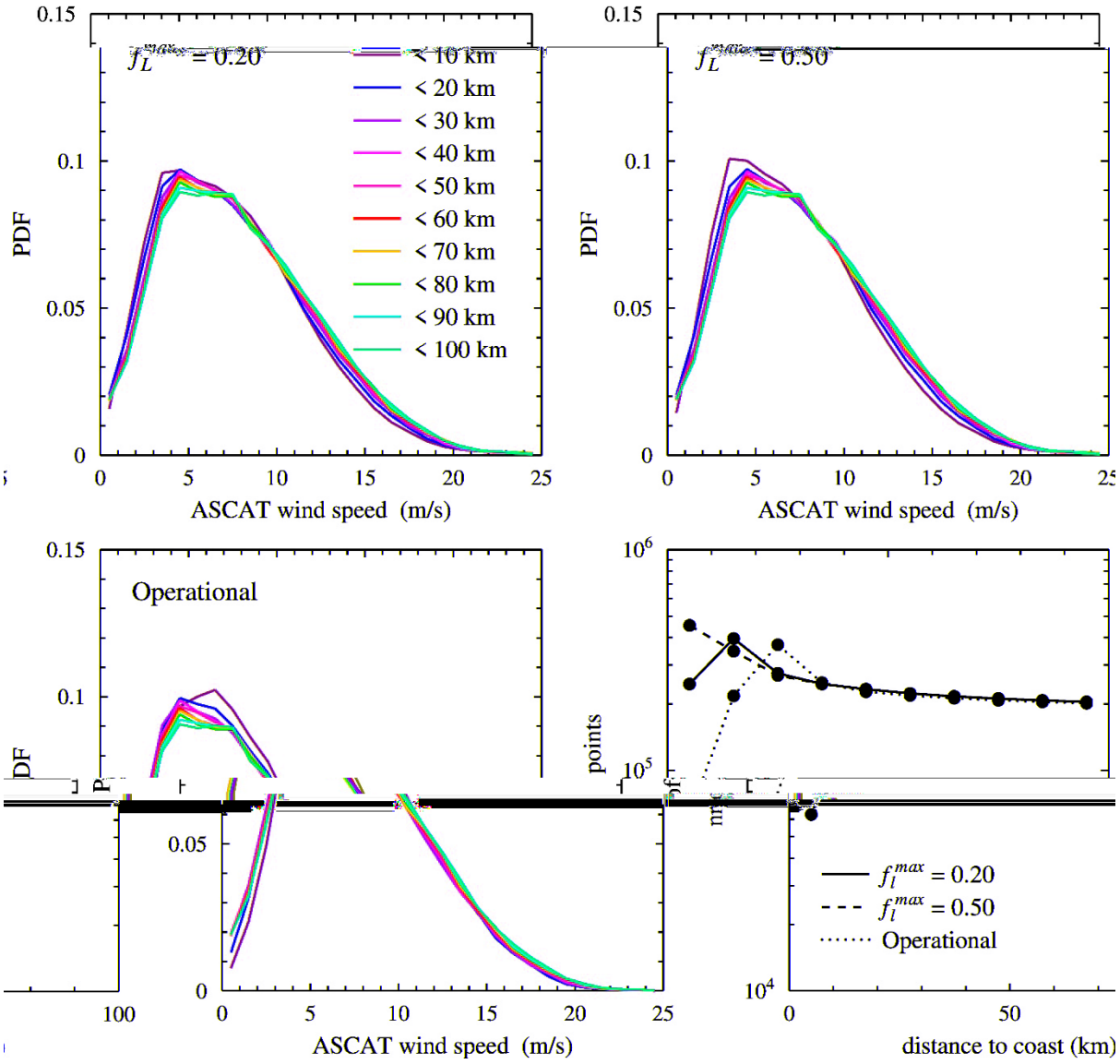

Figure 5.6 As figure 5.5, but for ASCAT off-shore winds.

It is interesting to note, however, that the low winds closer than 10 km to the coast increase with respect to the winds further away from the coast in the operational product. Given the generally negative backscatter corrections due to land, this effect is consistent with the absence of land correction. For a maximum land fractions of 0.5 in particular, low winds closer than 10 km to the coast decrease with respect to the winds further away from the coast, hinting at a successful land contamination correction.

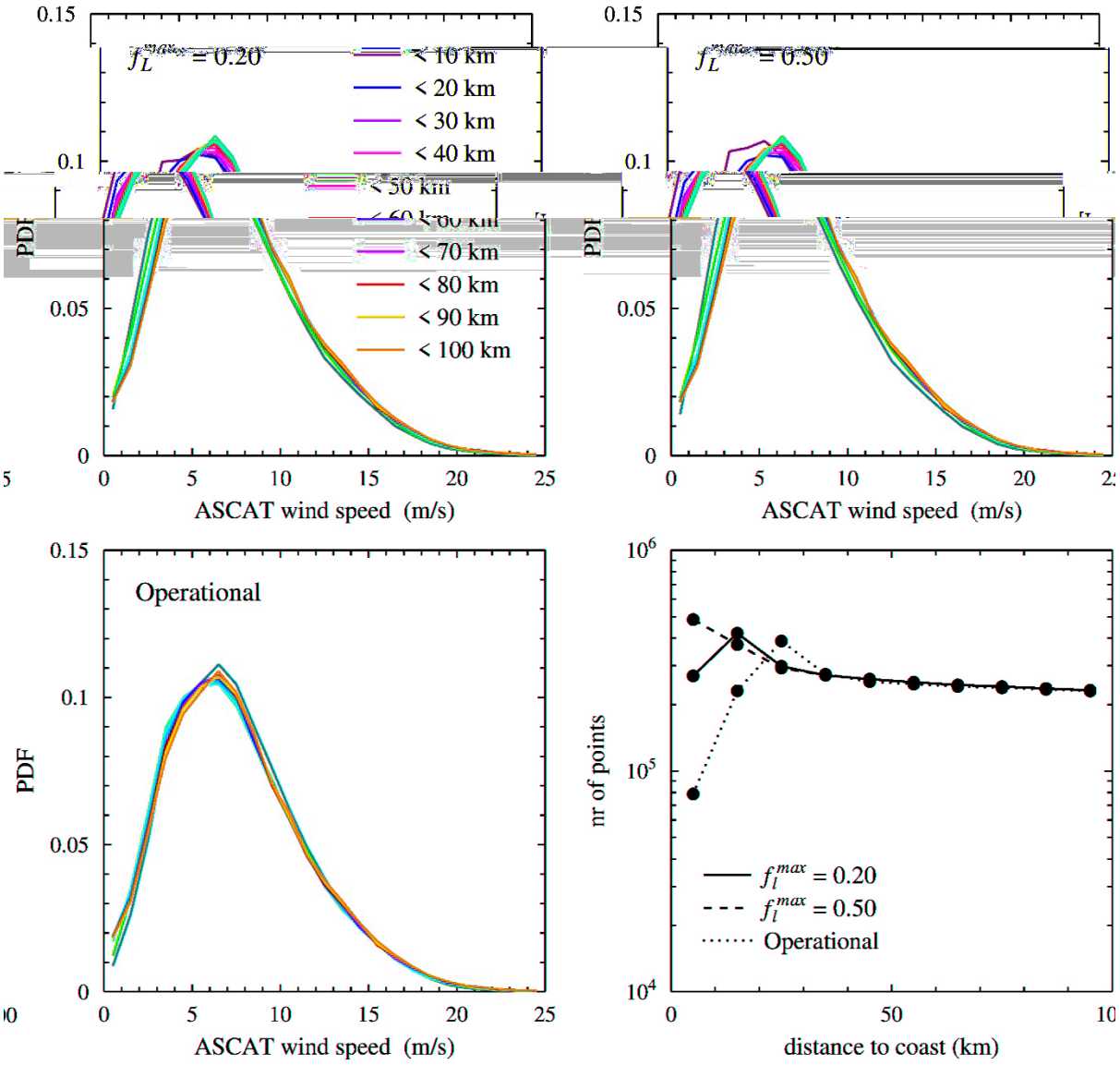

Figure 5.7 As figure 5.5, but for ASCAT winds parallel to the coast.

5.4 MLE

The MLE is related to local wind variability conditions [*Lin et al*., 2015], which may be a function of distance to the coast, due to land-sea effects. In addition, the beam collocation quality will affect the MLE [*Vogelzang and Stoffelen*, 2017], which may also degrade closer to the coast, due to the varying footprint orientations interfering with the local land shapes. Finally, also land contamination effects could affect the MLE. Figure 5.8 shows the MLE pdfs for the various classes of distance to the coast. As in the previous figures, the upper left panel shows the pdfs for the land-corrected product with a maximum land fraction of 0.20, and the upper right panel that for a maximum land fraction of 0.50. The lower left panel shows the pdfs for the operational product, and the lower right panel shows the total number of WVCs as function of the distance to coast class.

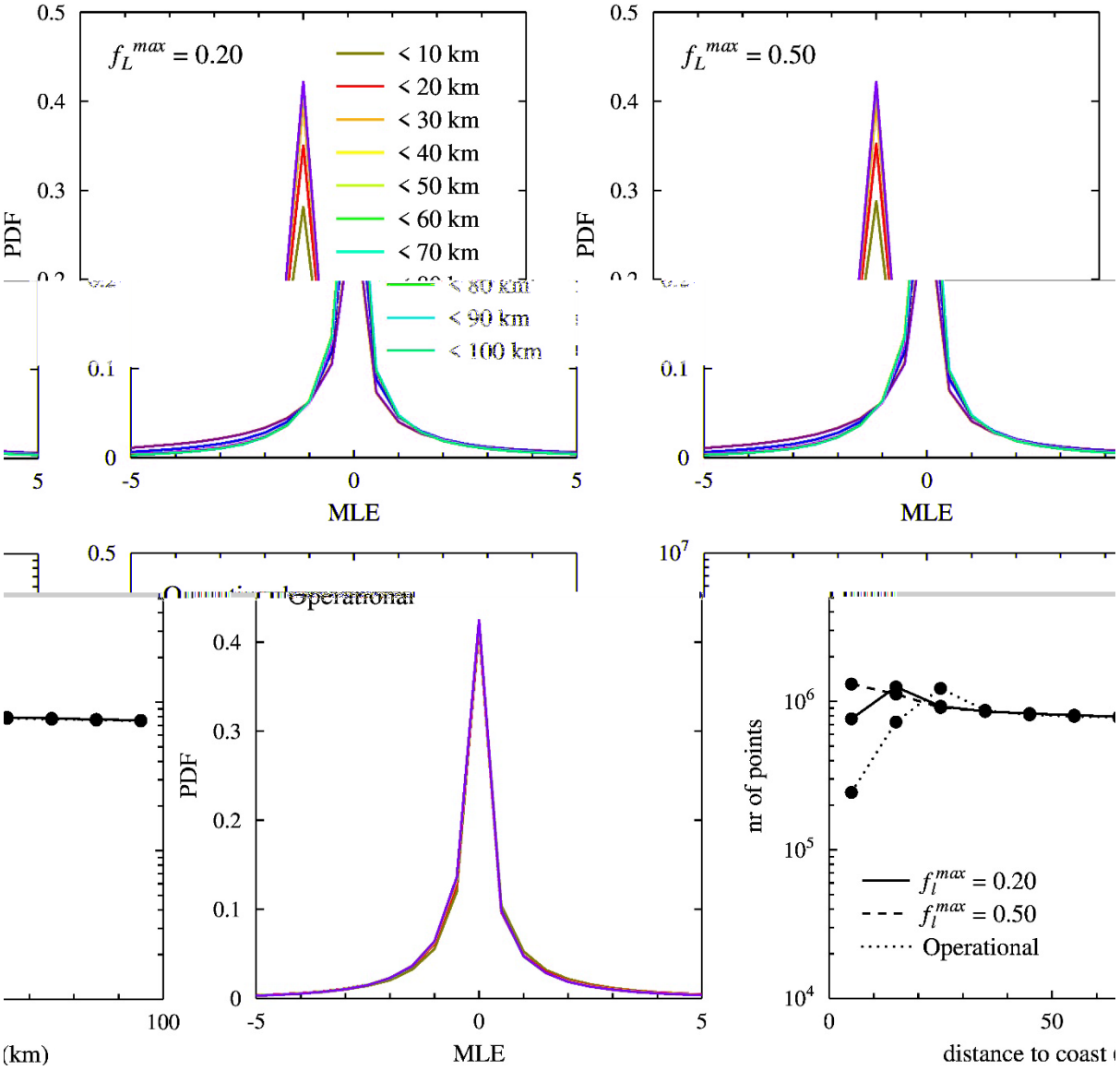

Figure 5.8 MLE histograms for the various distance classes.

Figure 5.8 shows that for distances to the coast of less than 20 km the MLE distribution widens a bit towards the negative MLE values. This means that the σ^0 triplets tend to lie more outside the GMF cone as a result of wind variability reduction or perhaps the land correction procedure. However, the effect is quite modest since the QC threshold for MLE is 18 or slightly more.

5.5 Flag setting frequencies

Figure 5.9 shows the flag setting frequencies for the KNMI QC flag (labeled as MLE), the Variational Quality Control flag (labeled as VarQC), and the Kp flag as a function of distance to the coast. The upper left and right panels are for the land-corrected products with a maximum land fraction of 0.20 and 0.50, respectively. The lower left panel is for the operational product.

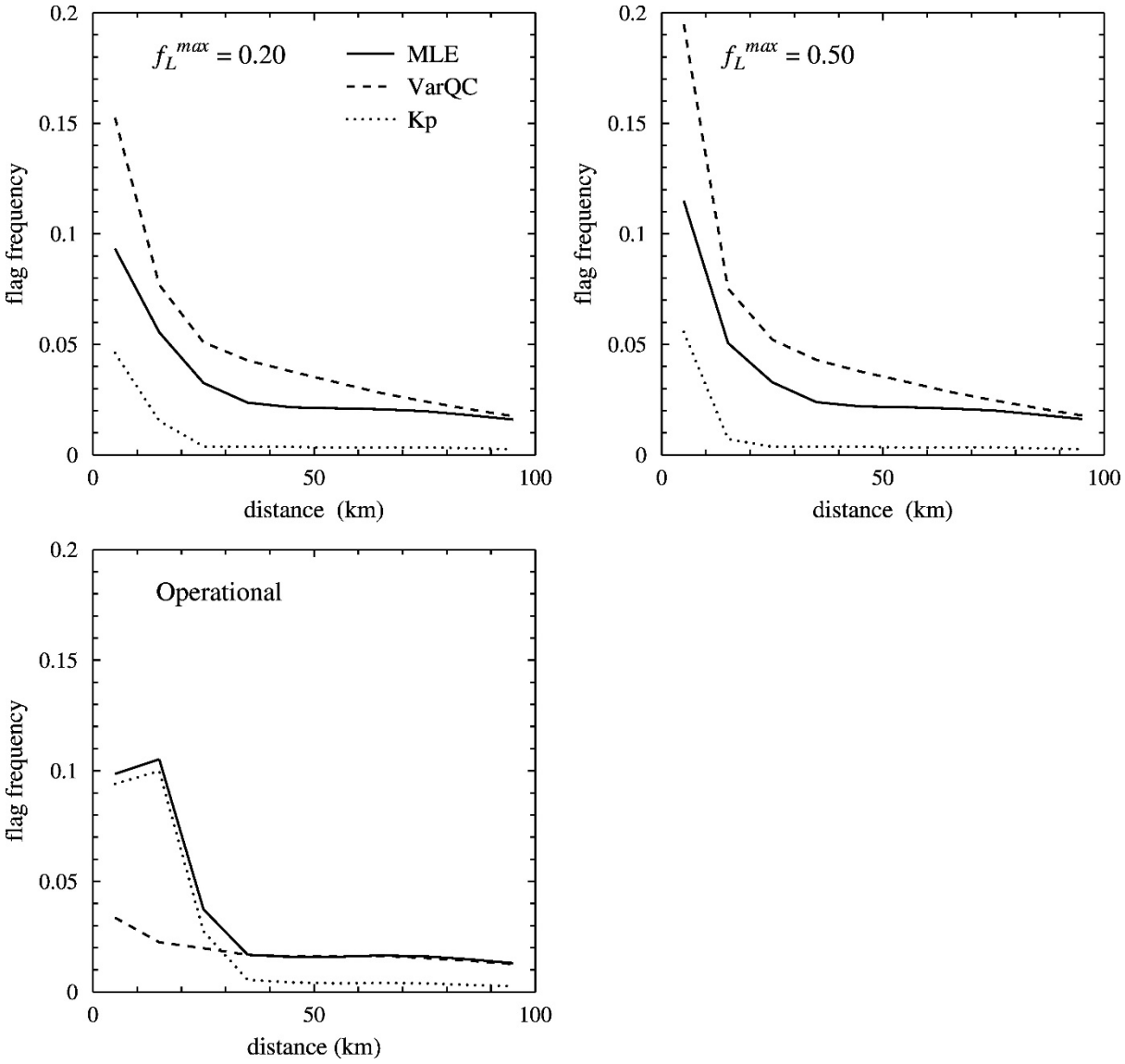

Figure 5.9 Flag setting frequencies as a function of distance to the coast.

For distances to the coast of less than 40 km all flagging frequencies increase for all products, also the operational product. However, for the land-corrected products the increase is stronger, notably for a maximum land fraction of 0.50. The increase in the KNMI QC flag for the land-corrected products is the same or less than for the operational product. Close to the coast, the land-corrected products contain more σ^0 information than the operational product, resulting in better statistics and less outliers due to noise.

Note the increase of the VarQC flag frequency. This flag indicates that the observed wind deviates much from the 2DVAR analysis. The increase in VarQC flag frequency in the land-corrected products starts at about 90 km from the coast and several effects are likely to play a role here. Since wind gradients increase as the distance to the coast decreases, 2DVAR flagging may be enhanced too.

Furthermore, we recall that the model winds used as background are degrading as one moves to the coast, which may degrade the 2DVAR analysis.

Note that the setting frequency of the Kp flag has been greatly reduced by using Gaussian weights in averaging the full resolution radar cross sections contributing to a WVC.

5.6 Radar cross section pdf

Figures 5.10 and 5.11 show the radar cross section pdfs of fore and mid beam, respectively, for the various distance classes and the same wind products as in the previous sections. For the operational product the σ^0 distributions are quite independent from the distance to the coast, except for distances less than 10 km where it is slightly sharper peaked, but also noisier due to a limited number of WVCs. For the land corrected products the distribution shifts slightly towards lower σ^0 values for distances to the coast of 20 km or less. The σ^0 distributions of the aft beam are very similar to those of the fore beam and are not shown here.

Similar to the wind speed PDFs, it is interesting to note that the low NRCSs closer than 10 km to the coast increase with respect to the NRCSs further away from the coast in the operational product. Given the generally negative backscatter corrections due to land, this effect is consistent with the absence of land correction. With land correction, on the other hand, for a maximum land fraction of 0.5 in particular, low NRCSs closer than 10 km to the coast somewhat decrease with respect to the NRCSs further away from the coast, hinting at a successful land contamination correction.

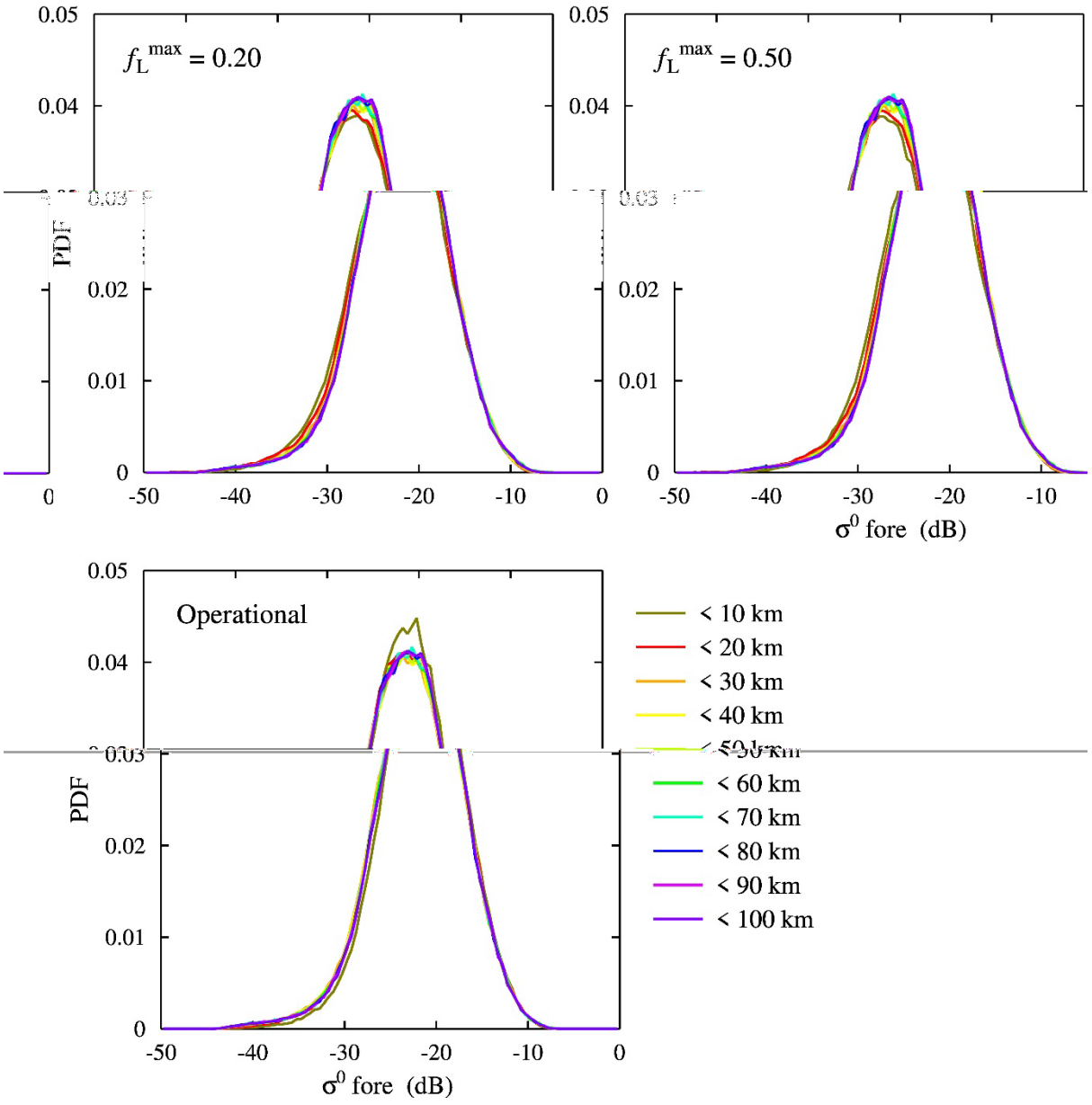

Figure 5.10 Radar cross section histograms for the fore beam.

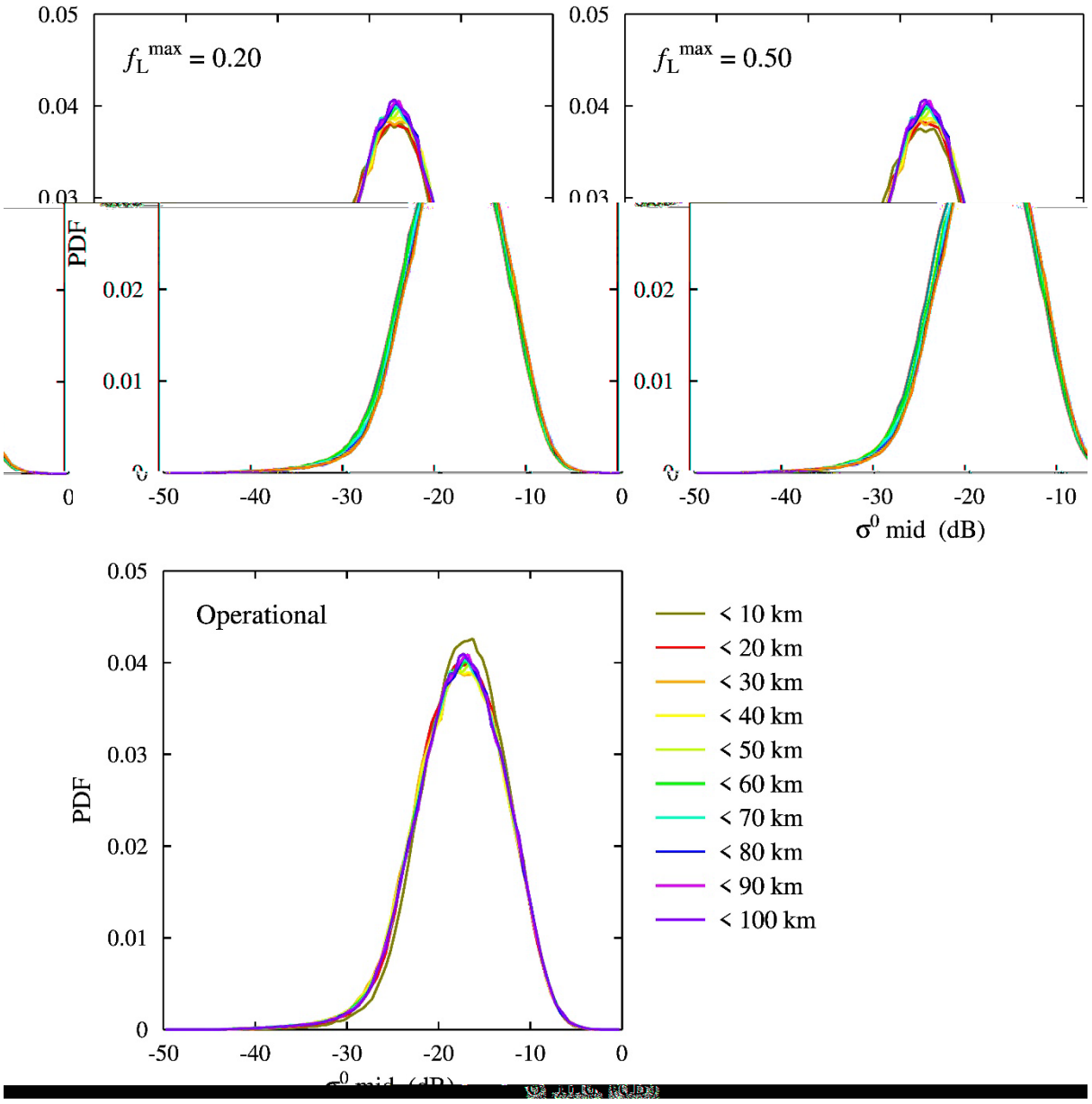

Figure 5.11 As figure 5.10 for the mid beam.

6 Buoy comparison

6.1 Introduction

Another possibility to check the quality of the land-corrected winds is to compare them with measurements from coastal buoys. Again, there are some caveats, as buoys measure very locally so their measurements may not be representative for the larger-scale scatterometer winds. Therefore the ECMWF blacklists were not applied, as these list buoys that differ too much from the model. Statistics of collocated winds will be calculated for each buoy separately. The collocation criteria are as usual: at most 30 minutes difference in time and 17.7 km difference in space.

In this chapter buoy data from three different sources will be considered:

- Copernicus Marine Service In Situ Thematic Assembly Centre (IS TAC), www.marineinsitu.eu, in NetCDF format;
- Meteorological Archival and Retrieval System (MARS) from ECMWF in BUFR format;
- National Data Buoy Centre (NDBC) from NOAA, www.ndbc.noaa.gov, in ASCII.

Buoy comparison was done for the operational product and for the land-corrected products with maximum land fraction f_L^{max} equal to 0.2 or 0.5, both with Gaussian weights with strength $F = 1$ in averaging the radar cross sections.

The IS TAC and NDBC buoy data archives contain hundreds of buoys, but only a limited number of buoys have the required temperature and humidity data together with the required metadata, see next section.

6.2 Metadata

To compare buoy winds with scatterometer winds one needs to convert the buoy measured wind speed to u_{10}^S, the 10-m height stress-equivalent wind (de Kloe *et al.*, 2017). It is obtained from u_{10}^N, the neutral wind speed at 10 m anemometer height using the Coupled Ocean Atmosphere Response Experiment model version 3.6 (further referred to as COARE-3.6) [Fairall *et al.*, 2003; Cronin *et al.*, 2006]. See also www.coaps.fsu.edu/COARE/flux_algor/. The computation of u_{10}^S requires the following additional information:

- Buoy anemometer height;
- Sea temperature
- Air temperature and air temperature sensor height;
- Relative humidity (or wet bulb temperature, or dew point temperature) and humidity sensor height;
- Surface pressure.
- Salinity (though this dependency is very weak and one may safely assume a salinity of 35 psu)

The surface pressure can be retrieved from ECMWF analyses and forecasts by quadratic interpolation in time and bilinear interpolation in space. Temperatures and humidity must be part of the buoy data, and sensor heights must be available from the metadata. Salinity is measured by some of the buoys. When absent, a default value of 35 PSU is used.

The following metadata were available:

- A list of metadata for NDBC buoys prepared by Ethan Wright, FSU;
- A list of metadata from Jean Bidlot of ECMWF;
- The genscat metadata list (which also originates from ECMWF).

The metadata were merged into one list. Also a few IS TAC files contained their own metadata that were not present in any of the lists.

6.3 IS TAC, MARS, and NDBC buoys

Figures 6.1 and 6.2 show the average differences and standard deviations for the MARS buoys located up to 120 km off-shore for the zonal wind component u and the meridional wind component v, respectively. Each figure shows results for the land corrected product with maximum land fraction 0.50 (top panel), with maximum land fraction 0.20 (middle panel), and for the operational product.

Figures 6.1 and 6.2 show the average differences differ less from zero than one standard deviation (or slightly more in a few cases). For distances more than 20 km from the coast the three products give very similar statistics. Closer to the coast the land-corrected products have slightly better statistics than the corresponding operational products. This may indicate that in the operational product some land contamination is present. Overall, the standard deviations are around 2 m/s, with some buoys with larger values.

The land-corrected product with maximum land fraction 0.5 collocates with more buoys very close to the coast than the product with maximum land fraction 0.20, but the latter product has better statistics close to the coast. Note that also the operational product collocates with buoys less than 10 km offshore. In those cases the buoy is located near an island that is so small that there are plenty of full resolution footprints with land fraction less than 0.02 around it.

In order to demonstrate the effect of distance to the coast more clearly, the buoy statistics were binned in 5 km intervals. The results are shown in figures 6.3, 6.4, and 6.5, for the IS TAC buoys, the MARS buoys, and the NDBC buoys, respectively. The bars indicate the number of collocations in each bin, with the scale given on the right hand axis.

Note that the MARS data set contains more collocations than the IS TAC and NDBC sets (see also Figure 6.6). However, the general picture from the three figures is the same: for the operational product the difference between scatterometer wind and buoy wind remains more or less constant as a function

of the distance to the coast while for the land corrected products the difference slightly increases with decreasing distance.

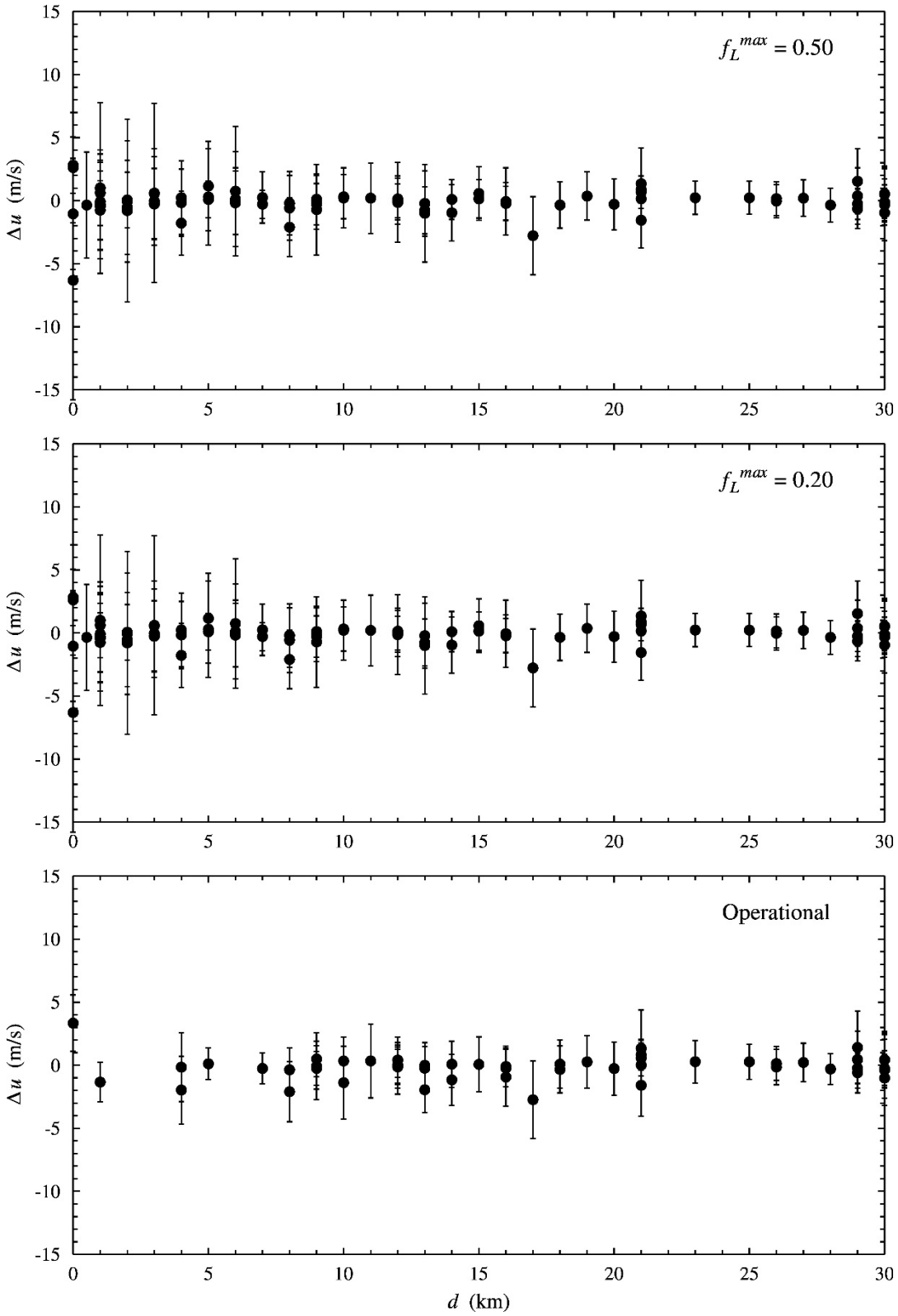

Figure 6.1 Average difference and standard deviation for the zonal wind component $\boldsymbol{u}$.

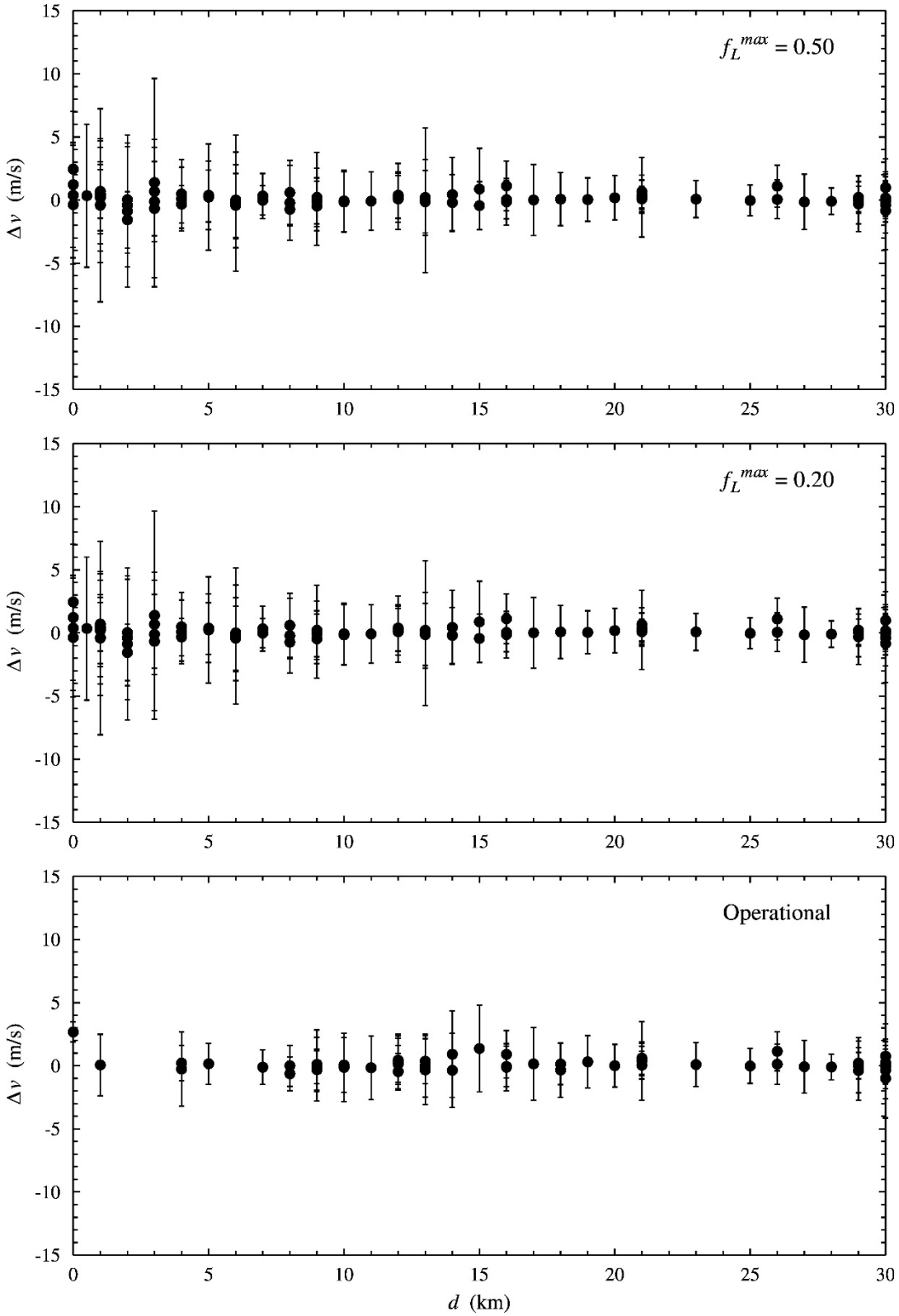

Figure 6.2 Average difference and standard deviation for the zonal wind component $\boldsymbol{v}$.

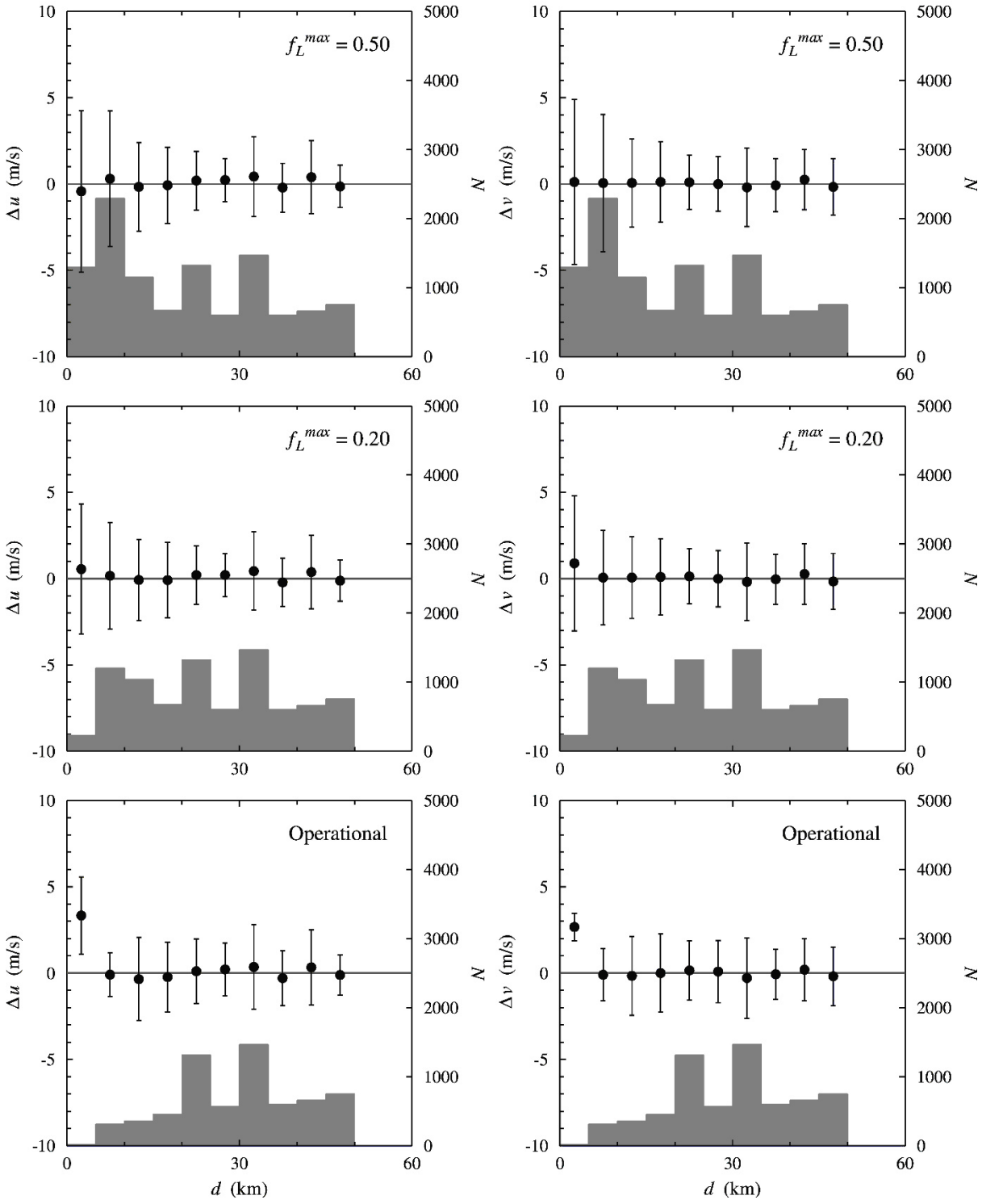

Figure 6.3 Comparison statistics for the IS TAC buoys.

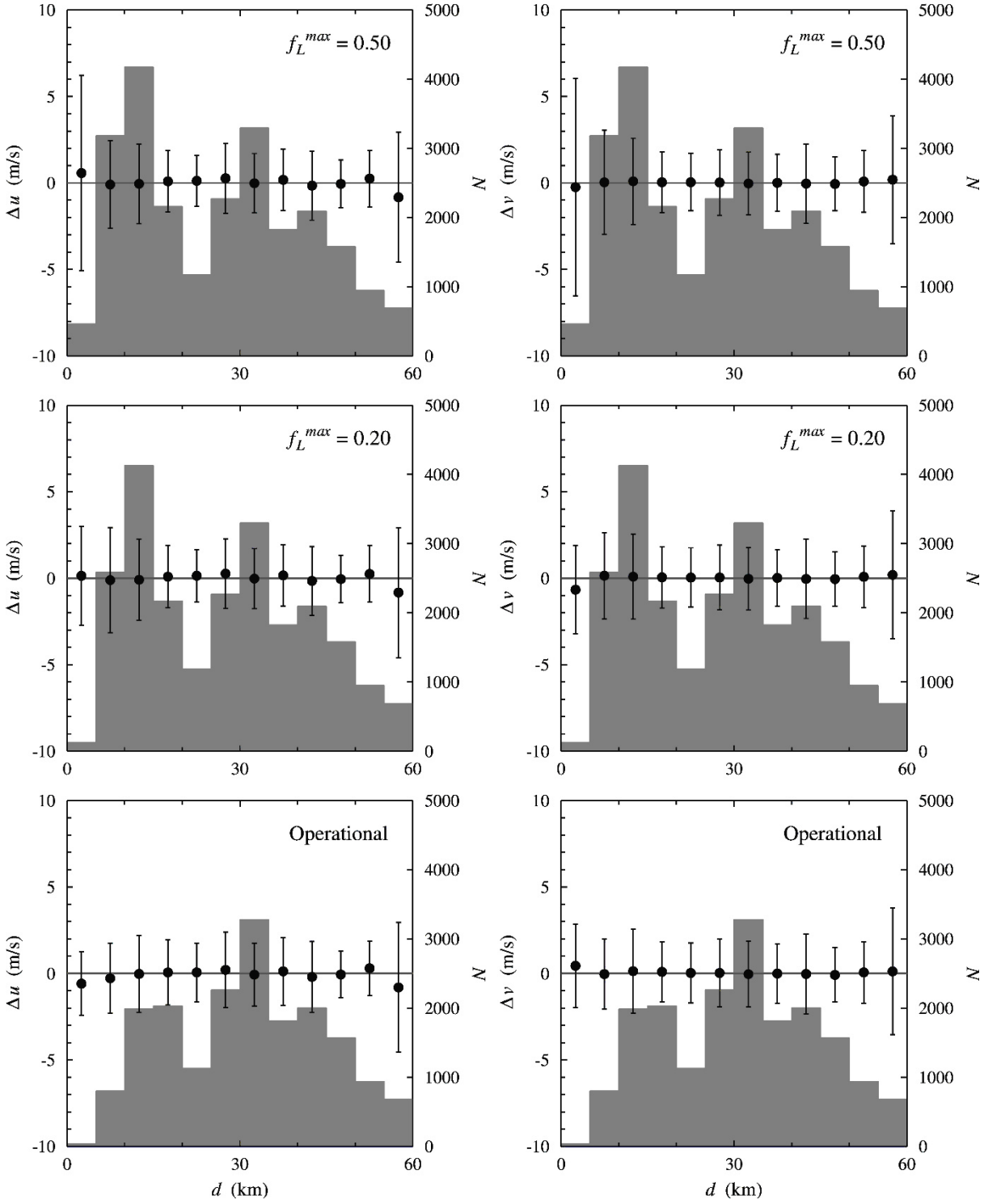

Figure 6.4 Comparison statistics for the MARS buoys.

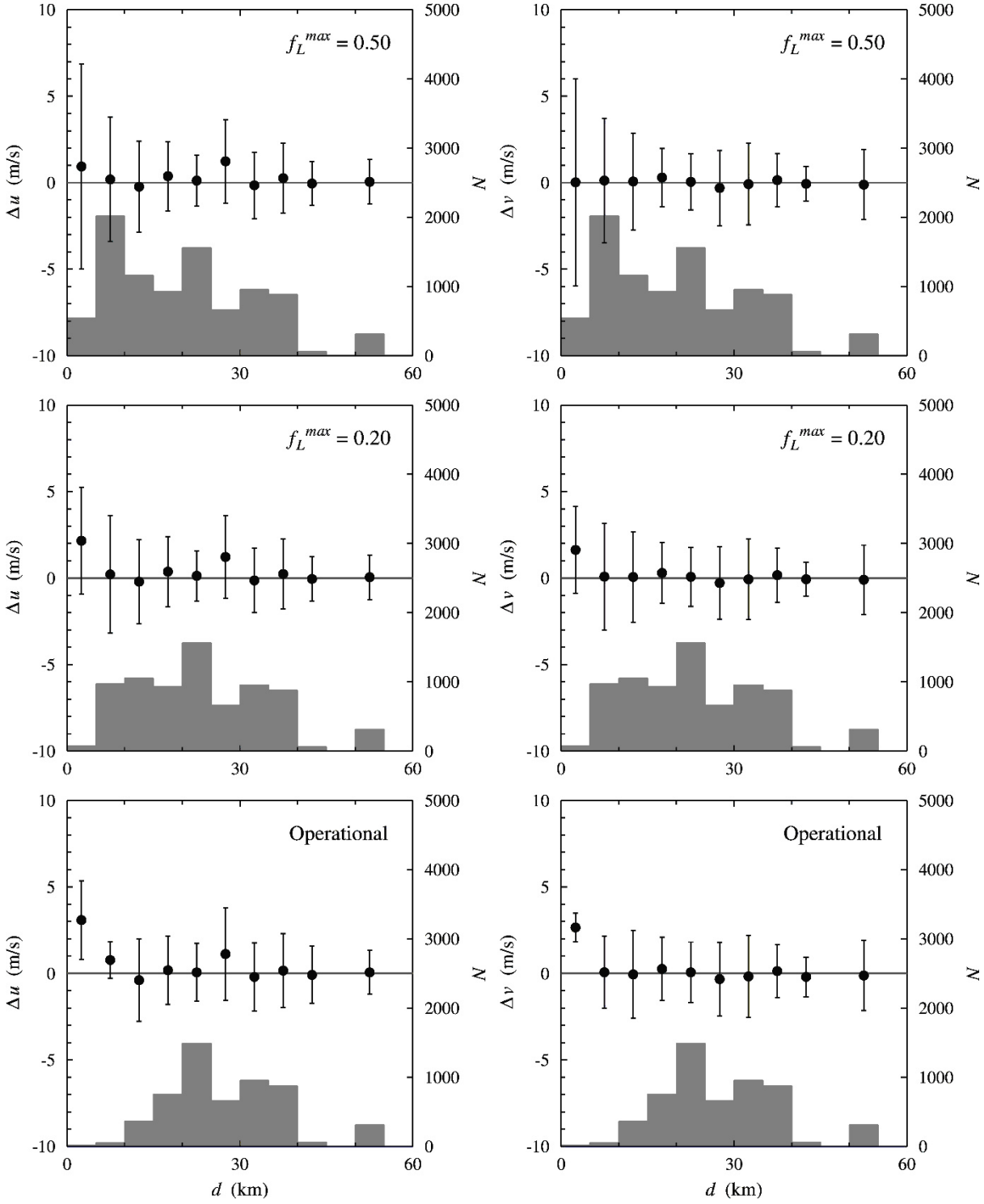

Figure 6.5 Comparison statistics for the NDBC buoys

6.4 *Merged data set*

The IS TAC, MARS, and NDBC collocation data sets were merged into one data set. Figure 6.6 shows the number of buoys in each data set separately and in their intersections for the collocations with the three scatterometer products. The total number of buoys in each merged collocation data set is shown in brackets.

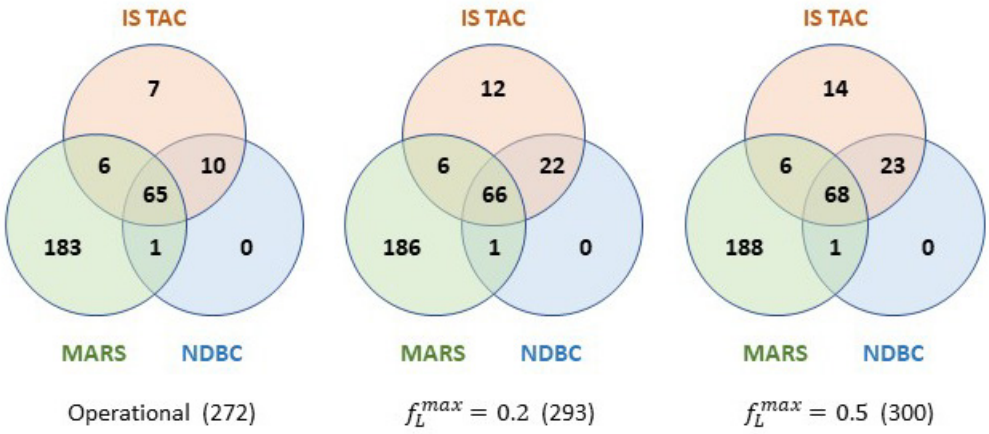

Figure 6.6 Number of buoys in each collocation data set

Figure 6.6 shows that the MARS data set adds most of the buoys, while the NDBC data set has no buoys that are not contained in the IS TAC and MARS data sets. This does not mean that the NDBC data set is worthless, because in a considerable number of cases the NDBC data are more complete than the IS TAC and MARS data. When merging the three data sets the set that adds most collocations is selected.

Figure 6.7 shows the comparison statistics. As for the IS TAC, MARS, and NDBC data sets, the bias between buoy wind and scatterometer wind is close to zero. The standard deviation of the difference between buoy and scatterometer winds is almost constant for the operational product, except for two outliers at 2.5 km and 57.5 km off-shore that may be attributed to poor statistics. The standard deviations for the land-corrected products tend to increase slightly for distances to the coast less than 10 km. Note the large standard deviations at 57.5 km in all three products.

Figure 6.8 shows the locations of the buoys collocating with the land-corrected wind product with $f_L^{max} = 0.50$. The vector root-mean-square (VRMS) of the difference between buoy winds and scatterometer winds is given by the color: green for VRMS op to 2 m/s, orange for VRMS between 2 m/s and 5 m/s, read for VRMS between 5 m/s and 10 m/s, and 10 m/s for VRMS larger than 10 m/s. Note that a considerable number of buoys is located around the Great Lakes in America. There are also some buoys in the European coastal waters and around South Korea.

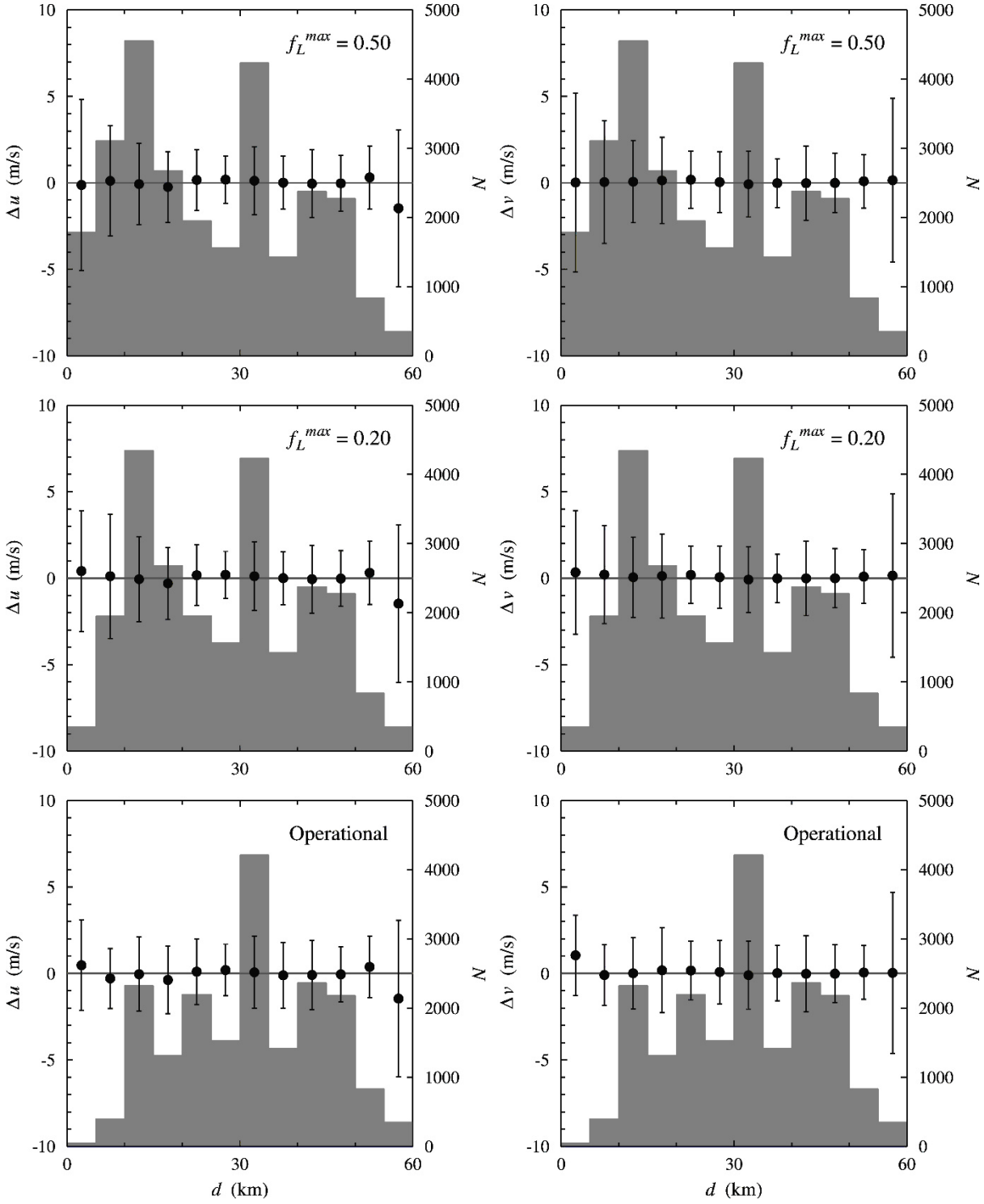

Figure 6.7 Comparison statistics for the merged collocation data set.

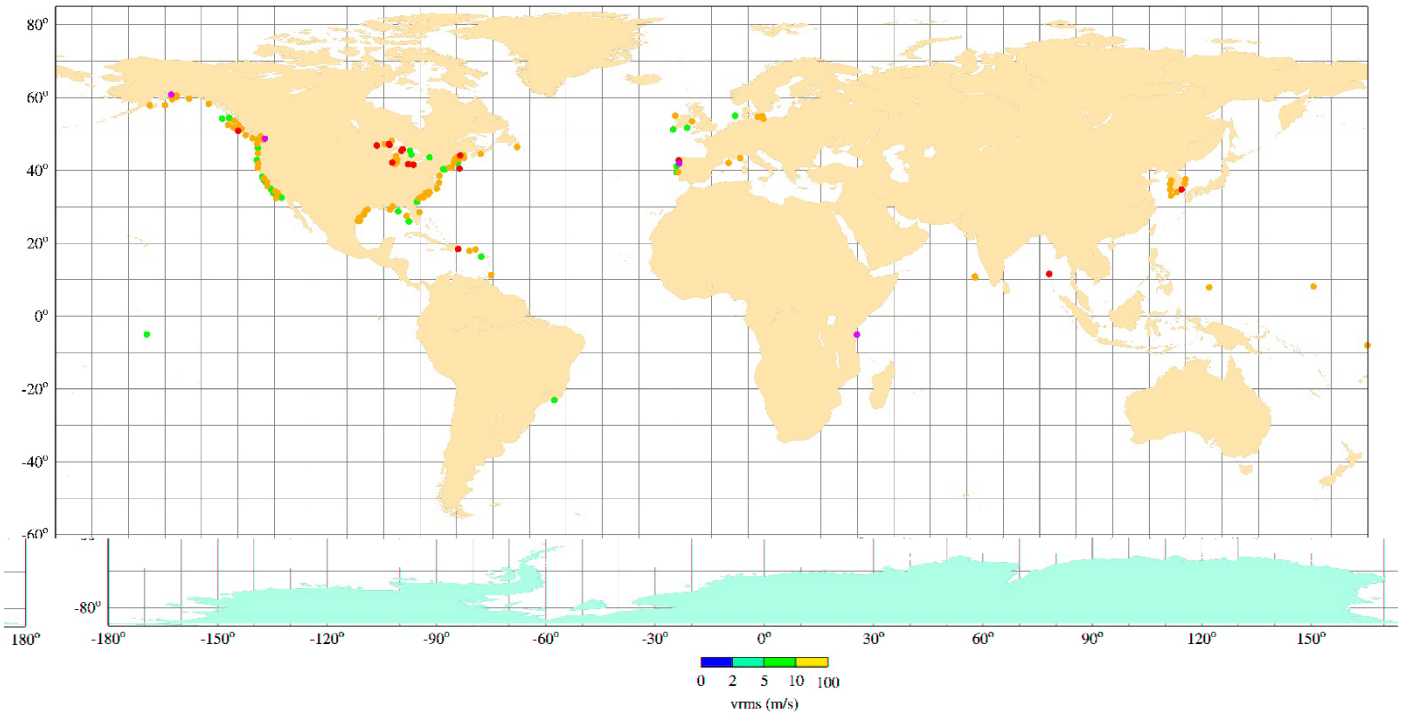

Figure 6.8 Buoys collocating with the land-corrected product with $f_L^{max} = 0.50$. The colors indicate the vector root-mean-square error between the ASCAT land corrected winds and the buoy winds

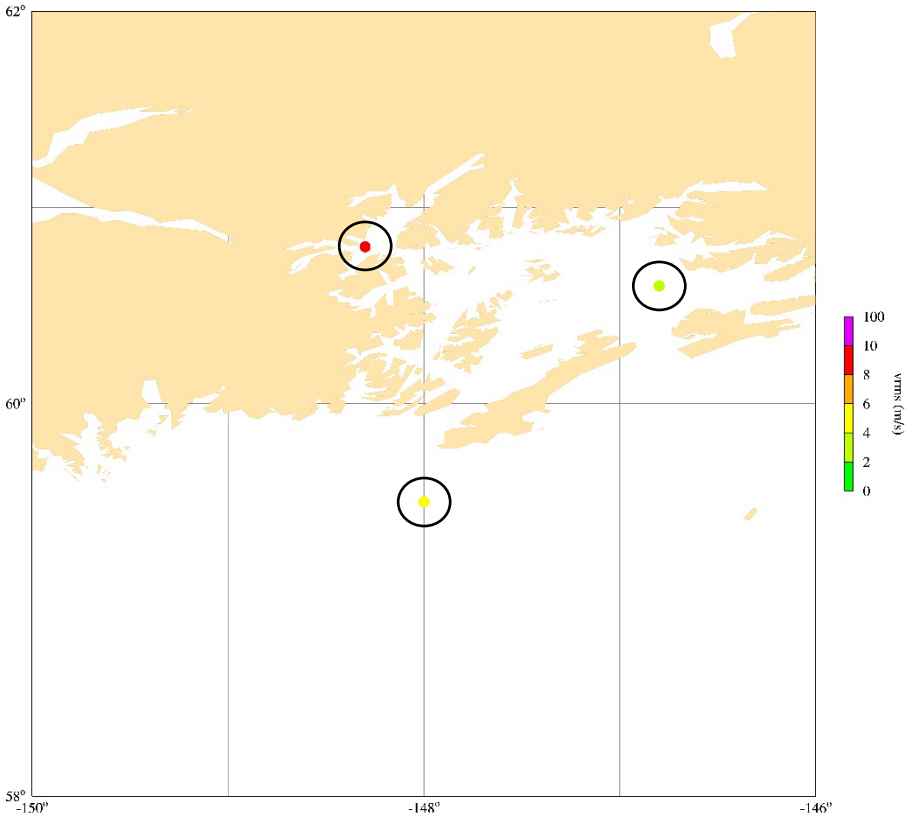

Figure 6.9 Position of buoy number 46081 (red dot).

The increase in the standard deviations close to the coast may be due to flaws in the land correction procedure, to lack of representativeness, or to erroneous buoy measurements (see section 6.6). For instance, the buoy with number 46081 which is located along the South coast of Alaska at 60.80 ˚N, -148.30 ˚E has large standard deviations. This buoy can be seen as the purple dot in figure 6.8. Figure 6.9 shows its location in greater detail: this buoy (red dot) has a distance of 1 km to the coast and lies sheltered between isles. It is not likely that its measurements are representative for a larger area. Other nearby buoys (yellow and green dots) have much smaller standard deviation as indicated by the color scale. Note that the buoy that is most off-shore has the smallest VRMS.

6.5 *Time series*

The results in the previous sections show that some buoys contribute more to the overall difference with scatterometer winds than others. Here we present some time series of buoy winds, land-corrected ASCAT winds, and ECMWF background winds.

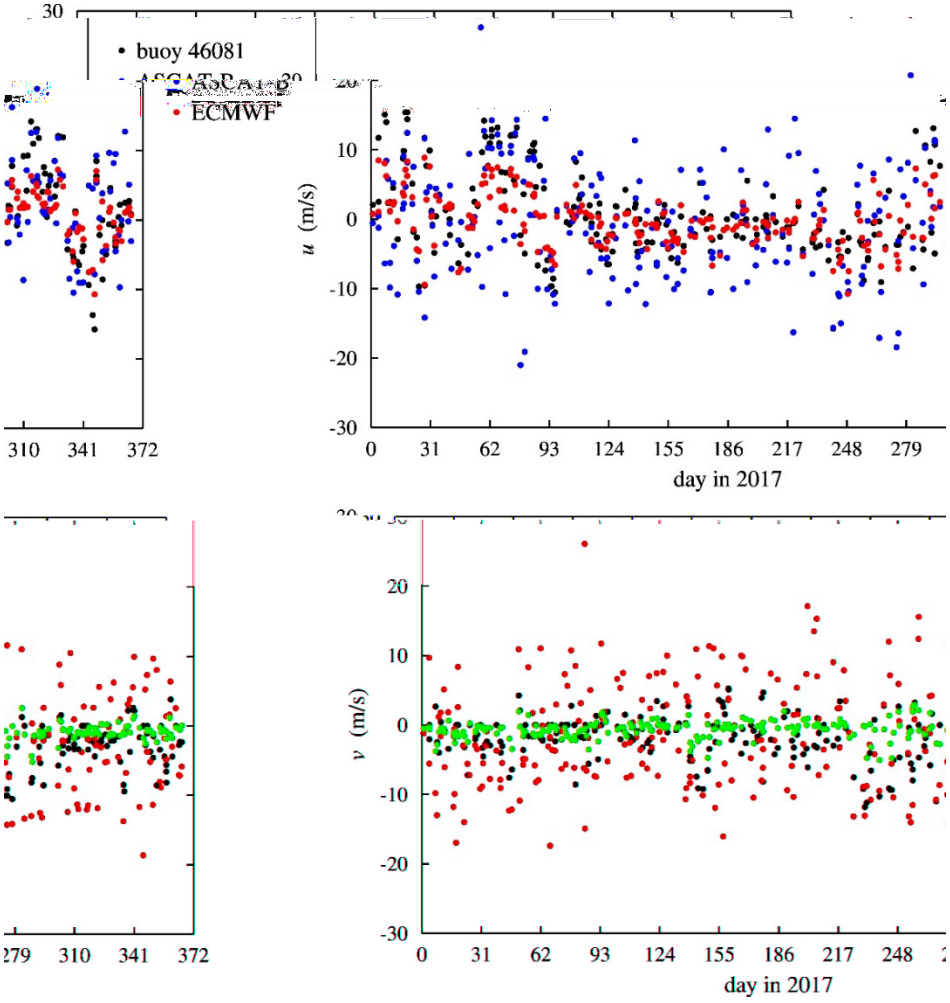

Figure 6.10 Time series for buoy 46081.

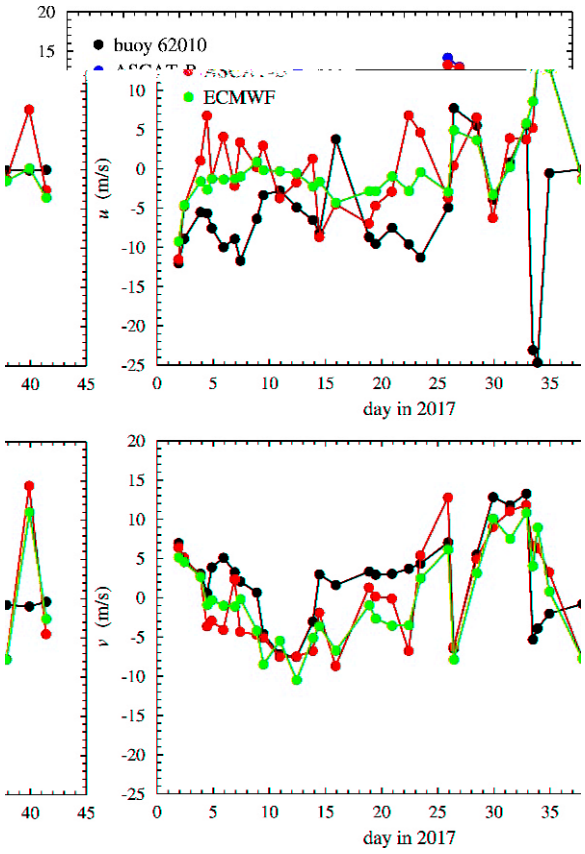

Figure 6.11 Time series for buoy 62010.

Figure 6.10 shows the time series of u (top panel) and v (bottom panel) for buoy 46081, the buoy near the coast of Alaska which location was shown in figure 6.9. It is clear from the figure that the land correction algorithms fails here: the ASCAT winds (maximum land fraction $f_L^{max} = 0.50$) have much larger spreading than the buoy winds and the ECMWF winds. The land-corrected product with $f_L^{max} = 0.20$ has no collocations with this buoy, so the land correction with $f_L^{max} = 0.50$ is based on land fractions more than 0.20, and the radar cross section of the sea is obtained by extrapolating the regression line. This may lead to large errors. Note that the v-component of the ECMWF wind is close to zero throughout the year, while that of the buoy shows more variation. Apparently, the complicated coast line also poses problems to the ECMWF model.

Figure 6.11 shows the time series for buoy 62010, located near the Nantucket shoals at the east coast of the U.S.A. The time series covers only the first part of 2017, as the buoy gets defect in mid-February: it gives a few very large negative u-values and then stops a few days later.

As stated before, large differences between buoy winds and land-corrected scatterometer winds can be due to failure of the land correction algorithm or to malfunctioning of the buoy. Further information on how to proceed is given by figure 6.12. The upper left and upper right panel of figures 6.12 show horizontally the average difference with buoy winds of the operational product and vertically that for the land-corrected product with $f_L^{max} = 0.20$. The left hand panel is for u, the right hand panel for v.

Each dot represents a buoy, and the color indicates its distance class to the coast. The grey line is the line on which the average difference with the buoy winds is the same for both scatterometer products.

The top panels of figure 6.12 shows that there may be considerable differences with the buoy winds, but that these are almost the same for the operational product and the land-corrected product, as the spreading of the points around the grey $y = x$ line is small. The bottom panels of figure 6.12 show horizontally the average difference with buoy winds for the land-corrected product with $f_L^{max} = 0.20$ and vertically that for the land-corrected product with $f_L^{max} = 0.50$. There are now more collocated buoys less than 10 km offshore, but the spread in the values is also larger.

Figure 6.12 suggests to adopt a value of 0.20 for f_L^{max}, but in the next section quality control will be addressed in more detail.

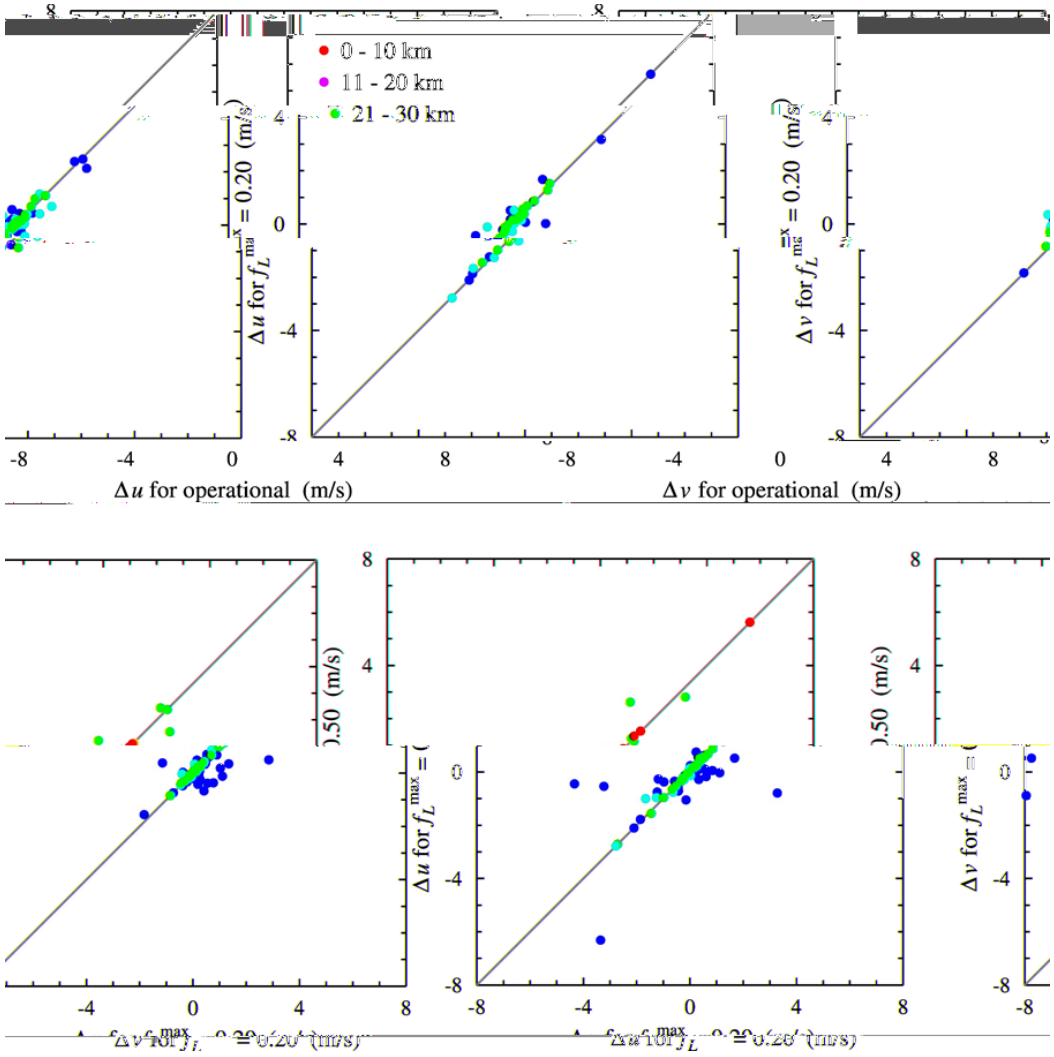

Figure 6.12 Comparison of the difference with buoy winds for the three scatterometer products.

6.6 *Quality control*

One can identify the following quantities that may be used as a quality control parameter:

- the regression error σ_e^2,
- the regression scaling error σ_a^2,
- the regression bias error σ_b^2,
- the minimum land fraction used in the regression analysis, f_L^{min},
- the difference in wind speed between the scatterometer and the ECMWF background, Δs.

These quantities will in this section be referred to as QC parameters. For each of them the maximum allowed value has been varied, and for each value the vector root-mean-square (VRMS) difference with the buoy winds was calculated as well as the number of collocations with buoys. The buoys were split in 6 distance to the coast classes, from 0 to 30 km in steps of 5 km. The results are shown in figures 6.13 to 6.17.

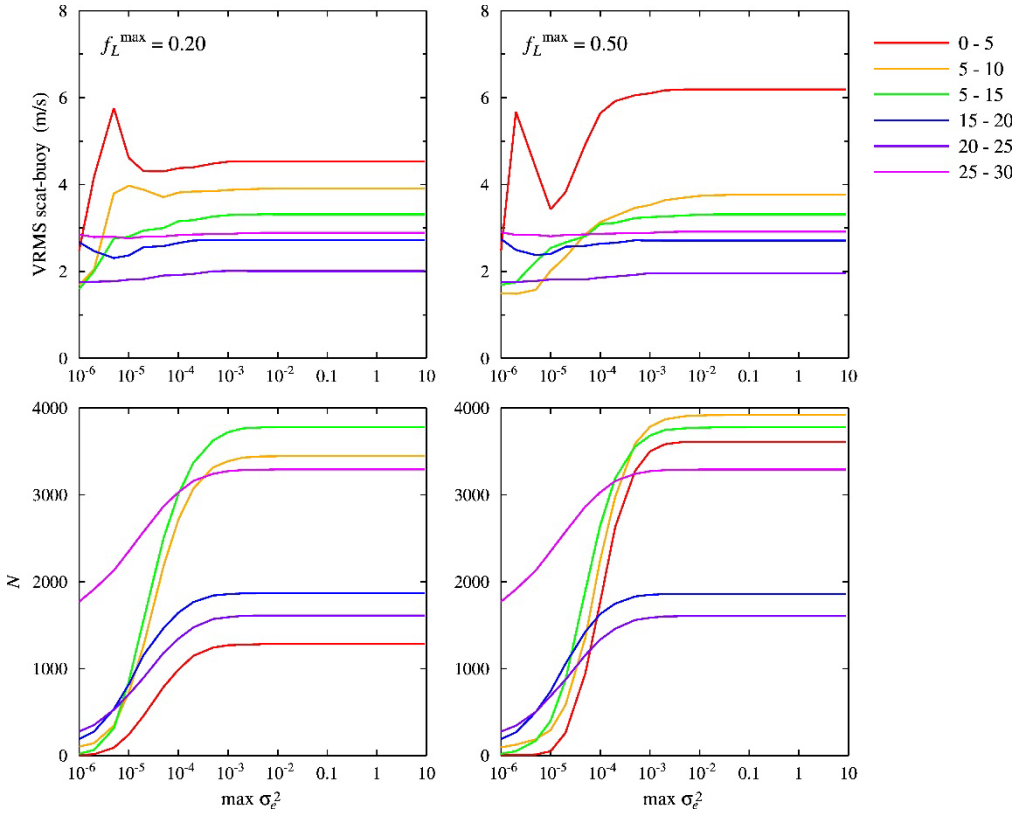

Figure 6.13 VRMS error between buoys and ASCAT-B (upper panels) and number of collocations (lower panels) against the maximum allowed regression error for maximum land fraction 0.20 (left hand panels) and 0.50 (right hand panels).

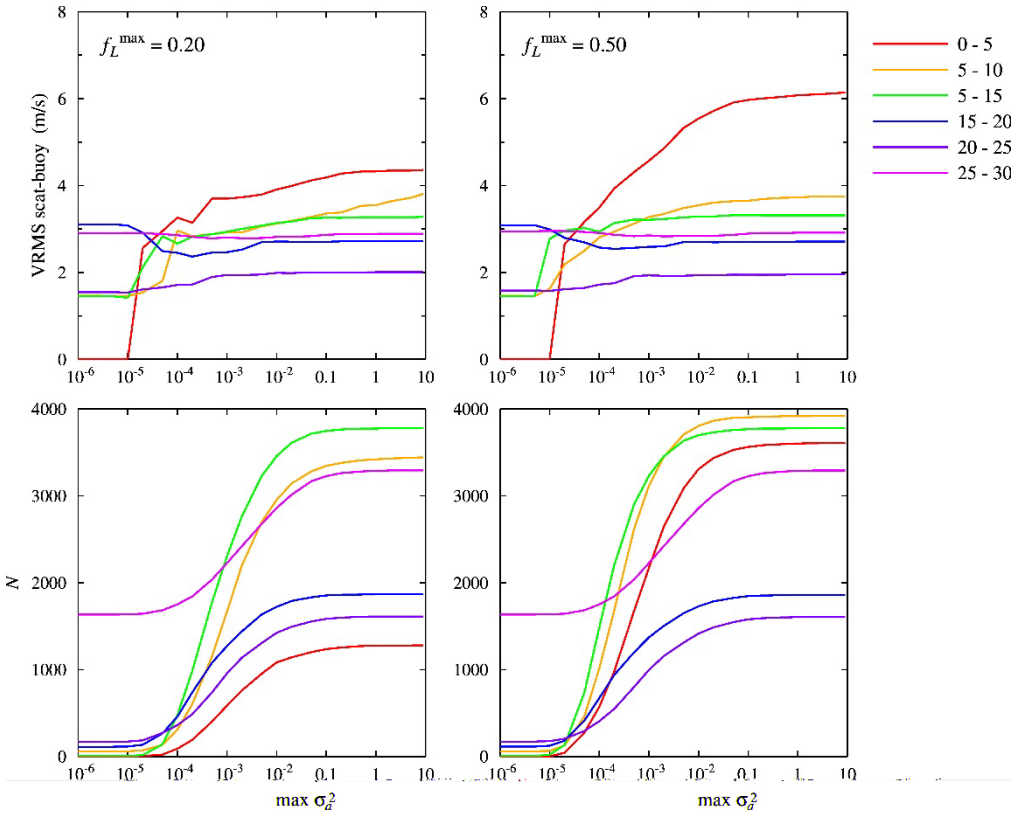

Figure 6.14 As figure 6.13, but for the maximum allowed regression scaling error.

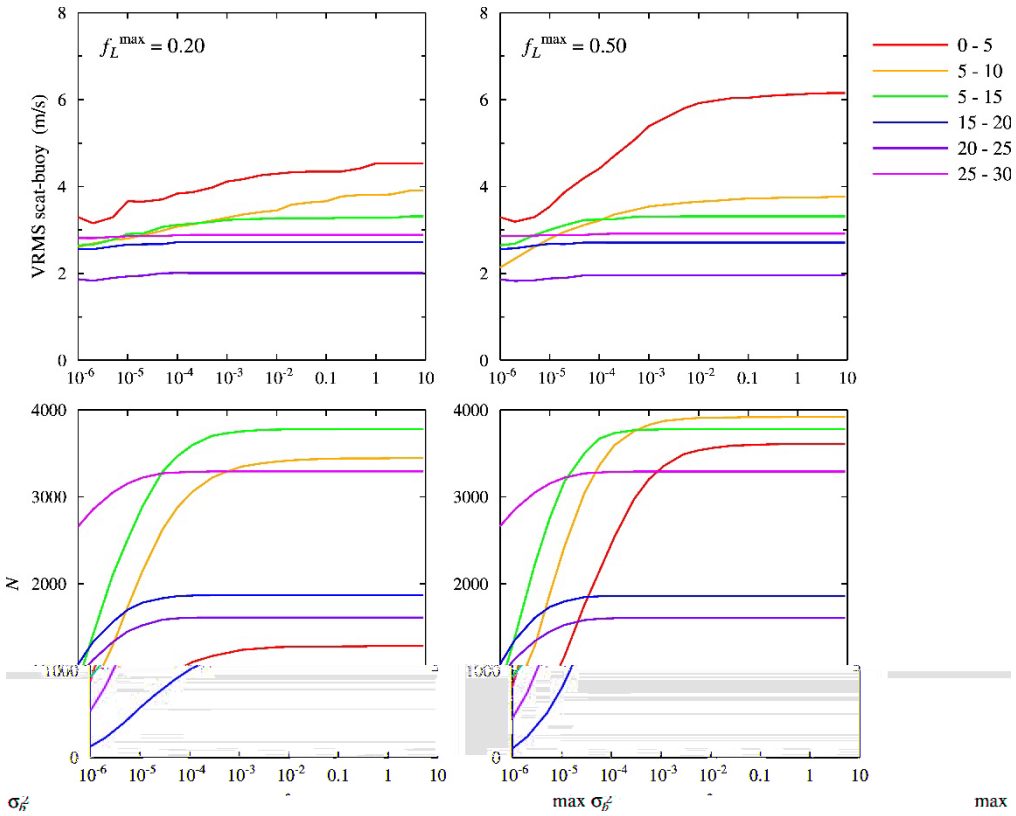

Figure 6.15 As figure 6.13, but for the maximum allowed regression bias error.

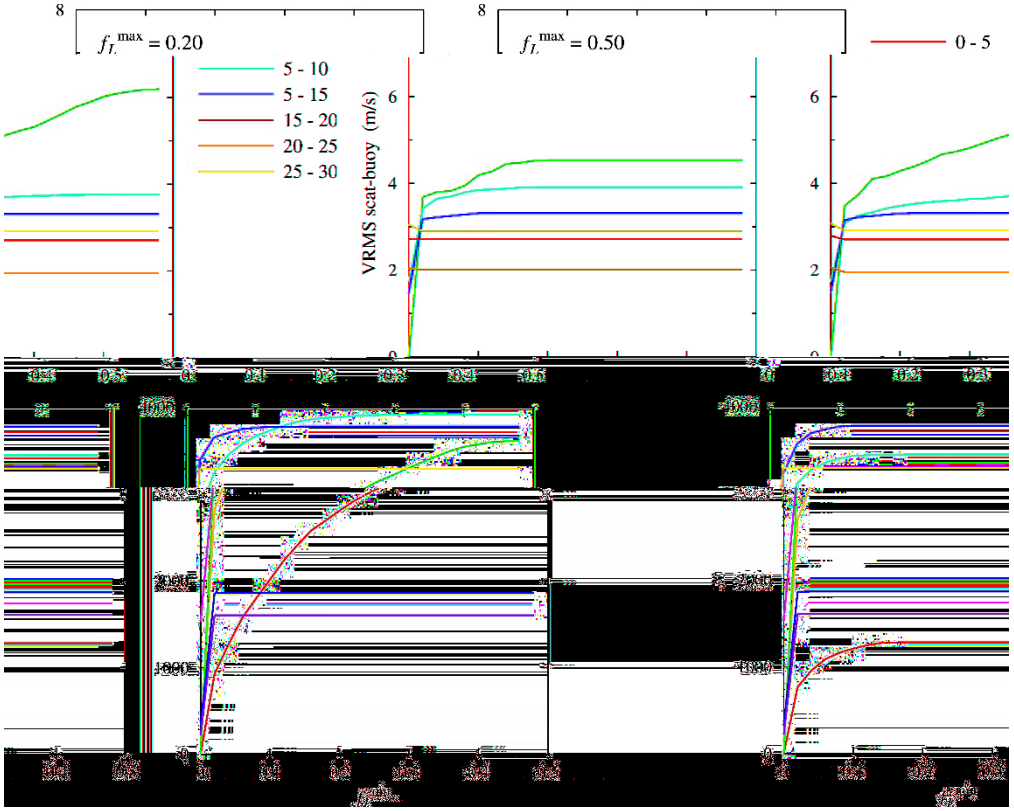

Figure 6.16 As figure 6.13, but for the maximum allowed minimum land fraction.

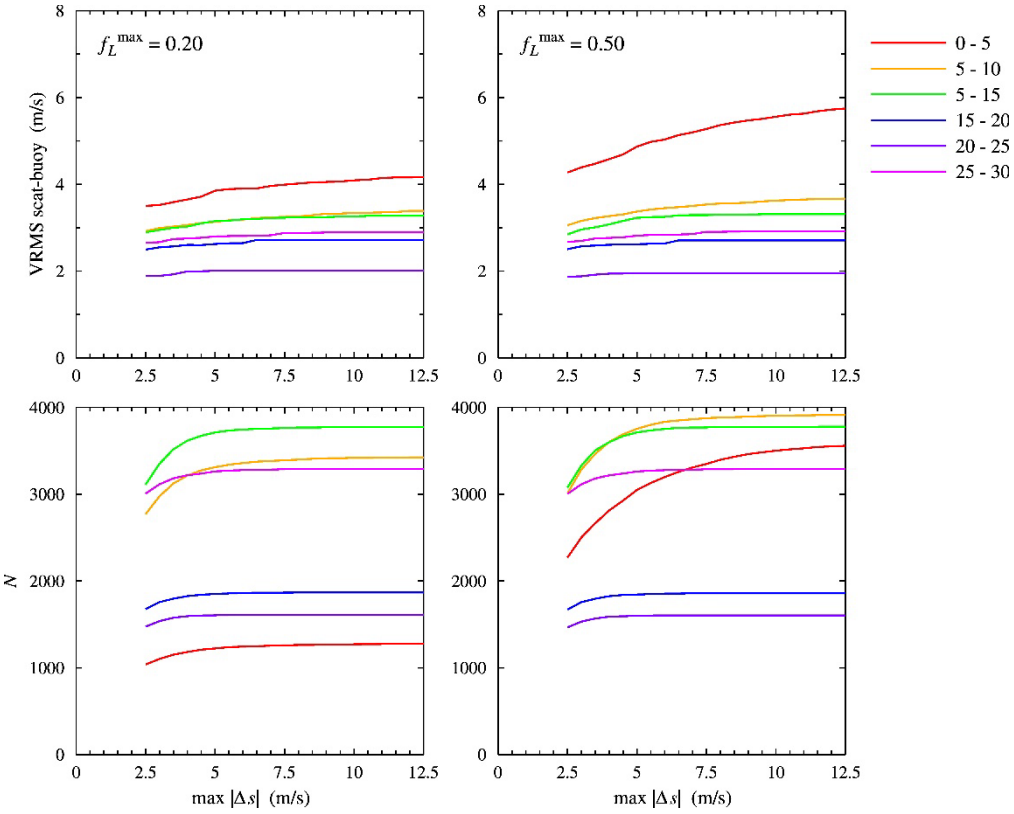

Figure 6.17 As figure 6.13, but for the maximum allowed wind speed difference between scatterometer and ECMWF forecast.

All figures show more or less the same behavior: when decreasing the maximum allowed value of the QC parameter both the VRMS difference with the buoys and the number of collocations decreases. The decrease in the number of collocations starts earlier for the buoys that are 0 - 15 km off the coast than for the buoys 15-30 km off the coast. This behavior is especially clear for the land corrected product with maximum land fraction 0.50. A first rough threshold value for each QC parameter can be found by demanding that for the buoys 15 - 30 km off the coast at most 10% of the collocations is filtered out. This gives the thresholds listed in table 6.1.

QC parameter	Threshold value
σ_e^2	0.00002
σ_a^2	0.02
σ_b^2	0.00001
f_L^{min}	0.02
Δs	2.5

Table 6.1 Threshold values for the QC parameters.

A better view of the influence of the QC parameters can be obtained from figure 6.18. This figure shows the VRMS difference between the buoys and the land corrected product as a function of the distance of the buoys to the coast (in bins of 5 km).

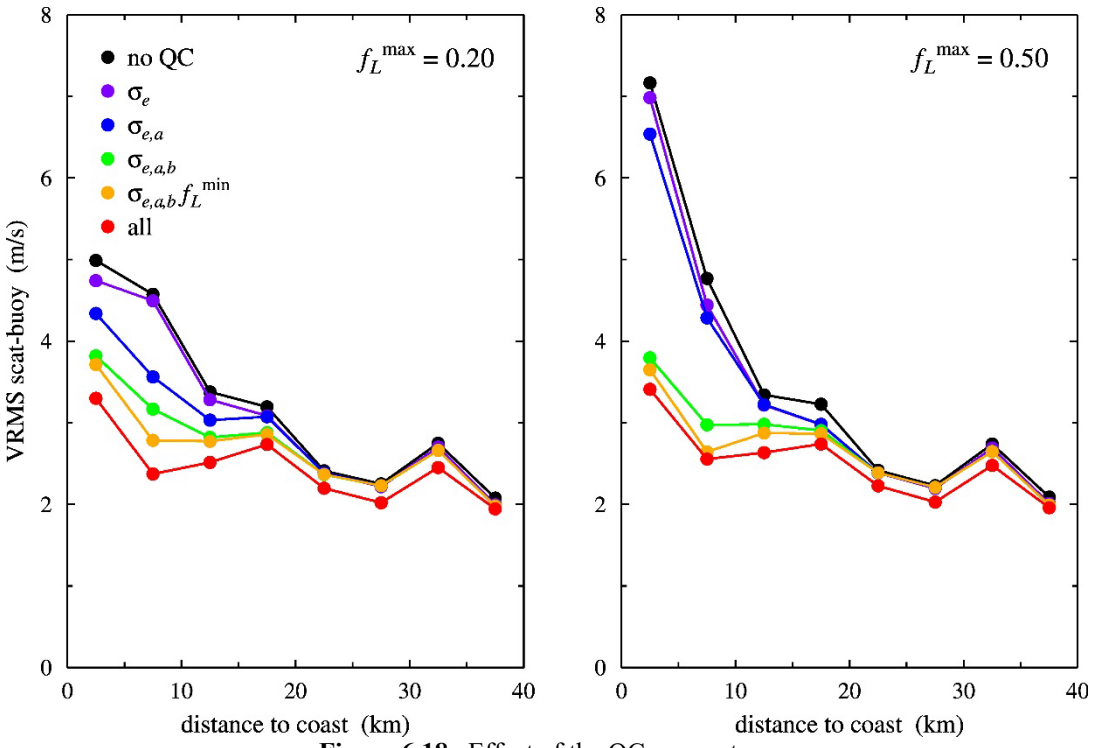

Figure 6.18 Effect of the QC parameters.

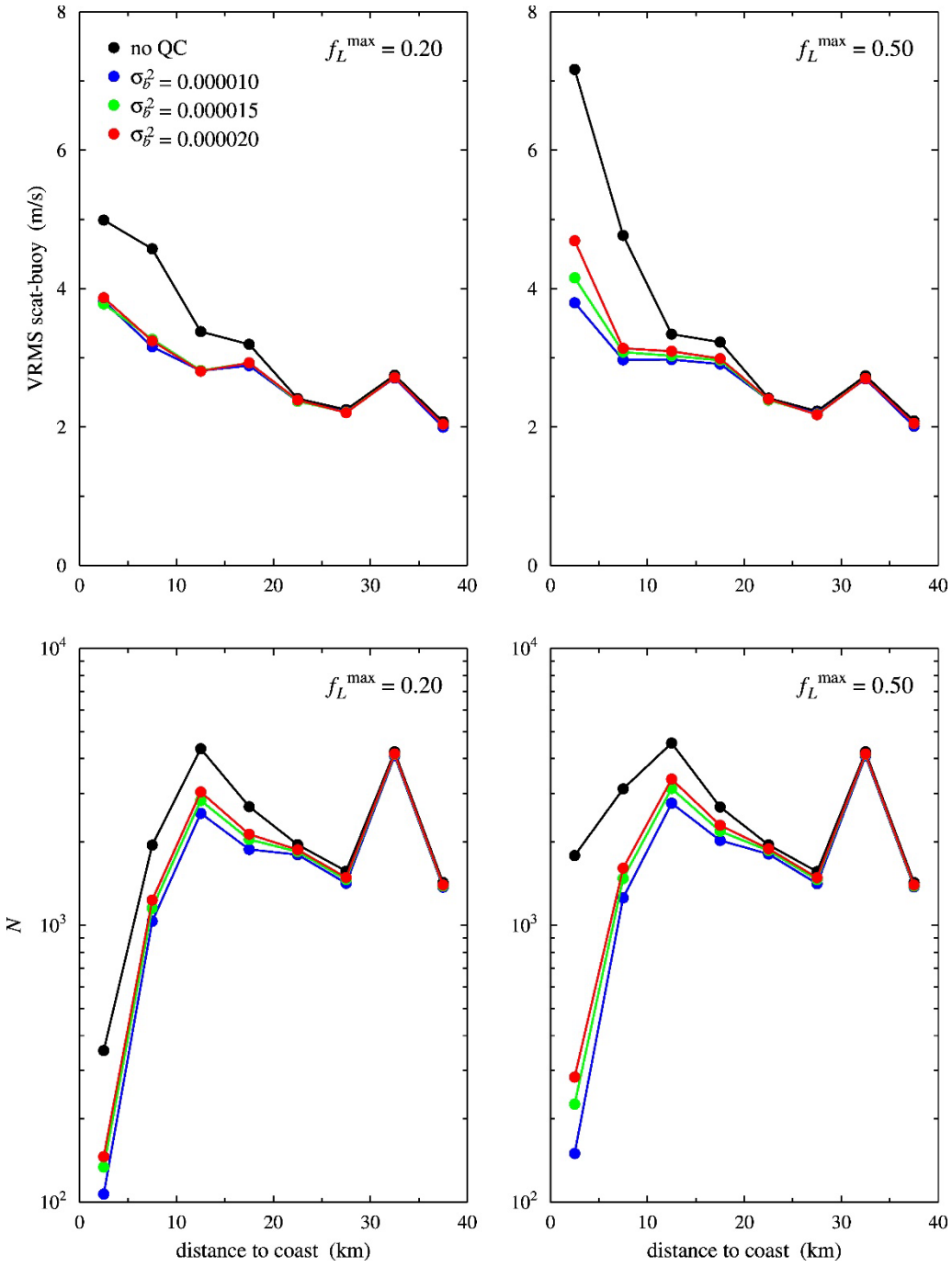

Figure 6.19 Effect of various thresholds for the regression bias error as QC parameter.

The black dots in figure 6.18 are without any QC. The violet dots, labeled σ_e, show the result when the threshold on σ_e is applied. The blue dots, labeled $\sigma_{e,a}$, show the results when also the threshold on σ_a is applied. The green dots, labeled $\sigma_{e,a,b}$ show the results when also the threshold on σ_b is applied. The orange dots are with inclusion of the threshold on f_L^{min} , the red dots, finally, are with inclusion of the threshold on Δs.

The difference in wind speed between the land-corrected product and the ECMWF forecast is not a very good QC parameter, as it is active also for measurements further off the coast. It merely filters out some

extreme cases. The other QC parameters become active at distances less than 20 km off the coast. All decrease the difference with buoys, notably the regression bias error σ_b for maximum land fraction 0.50 (green dots): it makes the VRMS drop by more than 2 m/s for buoys less than 5 km off the coast.

Figure 6.18 shows that the regression bias error is the most effective QC parameter for the land-corrected product with maximum land fraction 0.50, so it is worth to tune it more precisely, taking also the number of remaining collocations into account. This is done in figure 6.19 which shows the VRMS difference for three threshold values of σ_b (upper panels) and the number of remaining collocations (lower panels). The black dots in figure 6.19 show the results without QC as a reference.

Figure 6.19 shows that a threshold value of 0.000015 gives the best balance between the reduction in VRMS difference with the buoys and the number of remaining collocations. For the buoys less than 5 km off the coast the VRMS reduces by 1 m/s for a maximum land fraction of 0.20 and 3 m/s for a maximum land fraction of 0.50. However, the number of remaining collocations drops by about an order of magnitude.

6.7 Discussion

In the previous section it has been shown that imposing a maximum value of 0.000015 to the maximum regression bias error improves the buoy comparison considerably. The numerical results are shown in table 6.2.

Distance to coast (km)	**VRMS scat - buoy (m/s)**			
	$f_L^{max} = 0.20$		$f_L^{max} = 0.50$	
	no QC	**QC**	**no QC**	**QC**
0 - 5	5.0	3.8	7.2	4.2
5 - 10	4.6	3.3	4.8	3.1
10 - 15	3.4	2.8	3.3	3.0
15 - 20	3.2	2.9	3.2	3.0
20 - 25	2.4	2.4	2.4	2.4
25 - 30	2.3	2.2	2.2	2.2
30 - 35	2.7	2.7	2.7	2.7
35 - 40	2.1	2.0	2.1	2.1

Table 6.2 VRMS with buoys for various wind products as a function of the distance to the coast.

Table 6.2 shows that the VRMS difference with buoys increases with decreasing distance to the coast, but that QC keeps the VRMS at about 4 m/s. The reduction in VRMS near the coast is considerable, in particular for the land-corrected product with maximum land fraction 0.50. Moreover, the results for the $f_L^{max} = 0.50$ product differ very little from those of the $f_L^{max} = 0.20$ product.

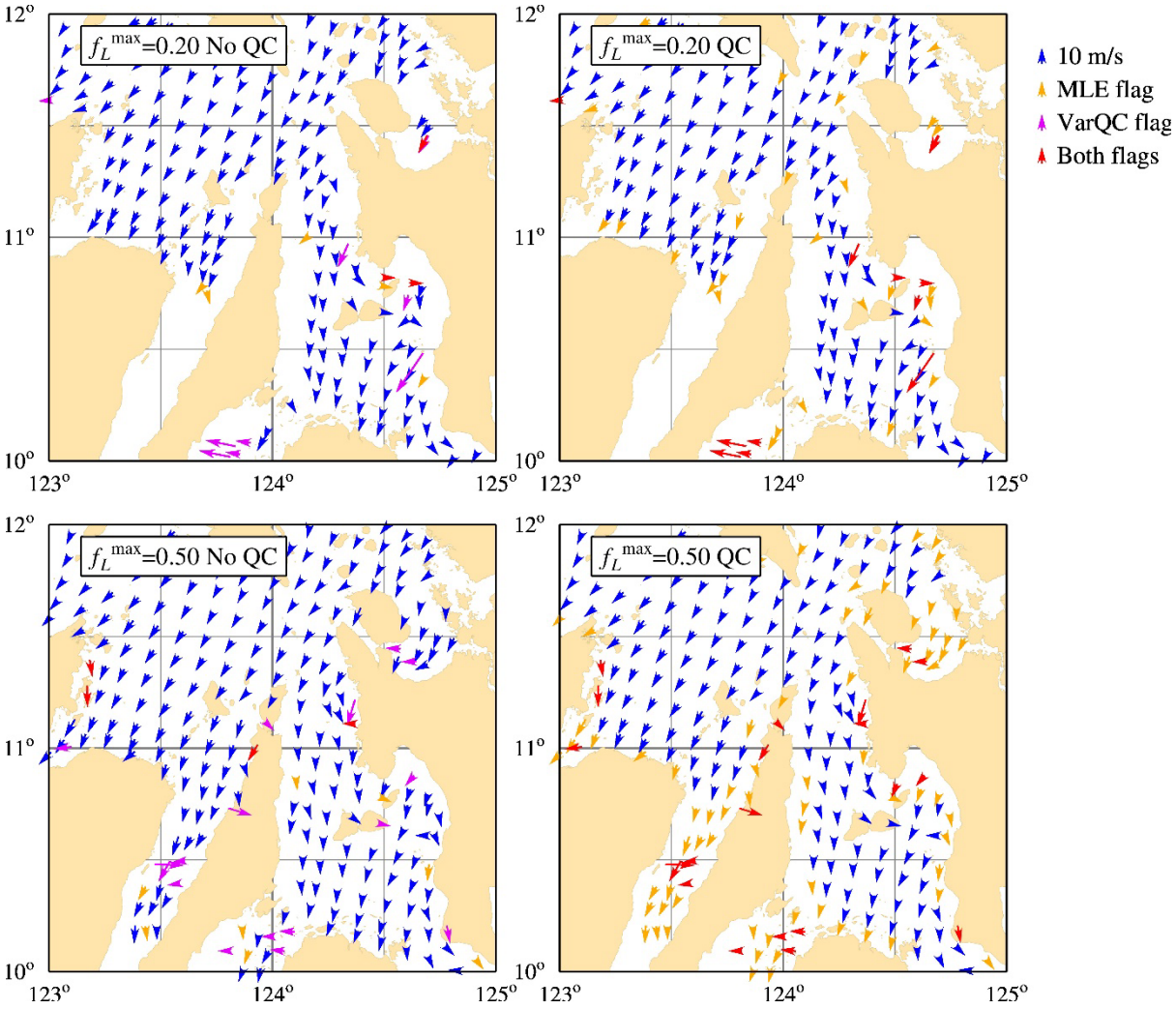

Figure 6.20 Part of the Philippines recorded January 1, 2017,with and without QC based on the regression bias error.

However, QC on the maximum regression bias error filters out many coastal WVCs that look good at first sight. This can be seen in figure 6.20 which shows the same scene of the Philippines as figures 4.1 and 4.2. The QC has been incorporated in the MLE flag, and with QC a lot of coastal WVCs are flagged. Figure 6.21 shows a scatter plot of the maximum regression bias error against the VRMS difference between buoy and scatterometer winds. Indeed the threshold on the regression bias error removes the outliers with a VRMS larger than 20 m/s, but it also removes a lot of points that compare well with buoys. A more specific QC procedure would be welcome.

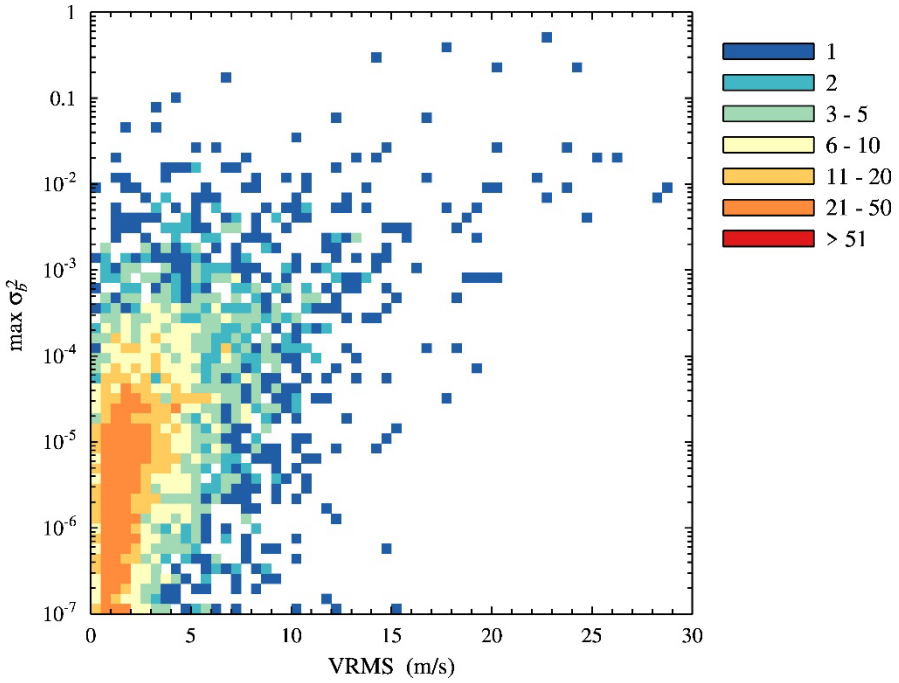

Figure 6.21 Maximum regression bias error against VRMS of the difference between

7 High resolution example

The land-correction scheme can also be applied to the high-resolution products of AWDP as illustrated in Figure 7.1.

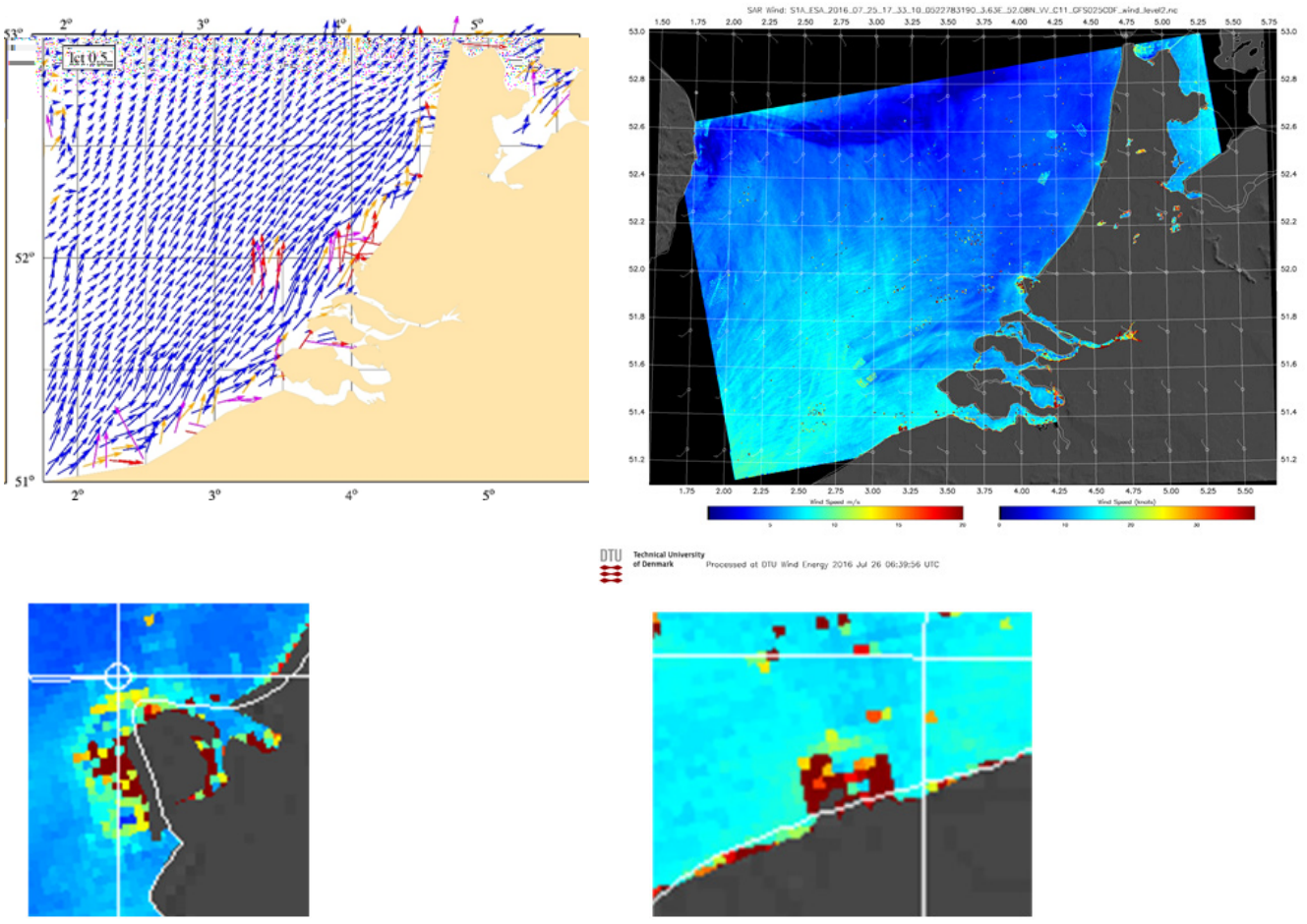

Figure 7.1: Top left: example 6.25-km AWDP product using land correction with . Top right: Sentinel-1 SAR wind image at high resolution. Bottom: zoomed SAR winds, left around the Maasvlakte in the Netherlands and right near Oostende in Belgium.

The coastal processing appears very promising, while the SAR image justifies the need for enhanced QC. The red pixels are winds above 20 m/s, which are really an artefact. Red dots also appear along the sandy coastline, which is due to NRCS variability, which may also not be treated well in our ASCAT land corrections. QC will hence be needed for large structures, breaking waves on shore, Radio Frequency Interference, lakes, shallow waters, (tidal) currents, etc. Investigating collocated ASCAT and SAR images will be useful to better understand these effects in future work.

8 Conclusions and recommendations

The new land fraction that EUMETSAT has developed for the full-resolution radar cross section product has been investigated to improve the ASCAT coastal product.

The new land fraction is based on the GSHHG coast line data base [*Wessel and Smith*, 1996] and a parameterization of the ASCAT Spatial Response Function [*Lindsley et al.*, 2016]. The new land fraction takes the shape and orientation of the Spatial Response Function into account. It differs in some coastal regions from the (rather conservative) land fraction that is currently used in ASCAT operational processing. However, simple replacement of the land fraction does lead to WVCs closer to the coast.

Therefore a simple land-correction model is proposed, based on the assumption that within the aggregation area of a coastal WVC the radar cross sections of land and sea are constant. A simple regression analysis then yields the value of the land-corrected radar cross section.

The resulting wind fields show strongly increased QC flag setting frequency near the coast, mostly due to setting of the K_p flag that indicates a large spread in radar cross section values contributing to a WVC. Nevertheless, from visual inspection the coastal winds look generally good. Therefore the radar cross sections are averaged with a Gaussian weight, the width of which is determined by the regression error. This weighting is also applied in the K_p calculation, leading to smaller K_p values and less frequent QC flag settings in the coastal regions.

Validation of the land-corrected ASCAT winds is not easy. ECMWF model winds are known to be inaccurate near the coast, while buoy measurements may not be representative for the average wind field in a WVC due to the variability of coastal winds.

ECMWF winds show coastal effects at larger distances from the coast than land-corrected ASCAT winds, notably for high winds. This is a good sign, as the ECMWF model is known to 'feel' the coast too far offshore.

The ASCAT land-corrected product was also collocated with buoy winds from three different data sources: ECMWF's MARS archive, the NDBC data base, and the IS TAC data base. Up to 300 buoys contributed to the collocation data set. This could be more, if the proper additional data and metadata to convert the buoy winds to neutral winds using the COARE-3.6 algorithm would be available. The additional data are sea temperature, air temperature, and relative humidity (or wet bulb temperature), the metadata are the heights at which wind, air temperature, and humidity are measured. Surface pressure can be interpolated with sufficient precision from the ECMWF model.

The most important parameter is the maximum land fraction f_L^{max} that determined the maximum land fraction of a single full resolution radar cross section measurement that is accepted for the regression analysis. The optimum value is between 0.20 and 0.50. Small values of f_L^{max} result in less new coastal WVCs but also better buoy comparison. Increasing f_L^{max} increases the number of new coastal WVCs

but also deteriorates the buoy comparison. Moreover, the errors at high values of f_L^{max} may obscure wind features.

The difference between buoy winds and ASCAT land-corrected winds increases at smaller distances to the coast. This increase is partly caused by some outliers for buoys close to an irregular coastline in a mountainous area or buoys in the Great Lakes. Several parameters were tested for their usability as quality control parameters. The best results were obtained by setting a threshold of 0.000015 to the maximum regression bias error. With this choice the VRMS difference with buoys less than 5 km from the coast is 3.8 m/s for the land-corrected product with maximum land fraction 0.20 and 4.2 m/s for that with maximum land fraction 0.50. However, the threshold also filters out a considerable number of WVCs that compare well with buoys.

It is concluded that the land-corrected coastal product adds useful wind information close to the coast, though the differences in the coastal region (less than 10 to 15 km offshore) are larger than those in the open sea, notably in urban areas or mountainous areas with a complicated coast line and enhanced local wind variability. User feedback will be of utmost importance to arrive at a more precise assessment. Furthermore, comparison with additional (non-public) in-situ wind data, HF radar or satellite SAR may be helpful in cases with poor spatial representation.

Glossary

ASCAT	Advanced SCATterometer
AWDP	ASCAT Wind Data Processor
COARE	Coupled Ocean Atmosphere Response Experiment
CMEMS	Copernicus Marine Environment Monitoring Service
IS TAC	In Situ Thematic Assembly Center
ECMWF	European Centre for Medium-Range Weather Forecasts
GMF	Geophysical Model Function
GSHHG	Global Self-consistent, Hierarchical, High-resolution Geography
NDBC	National Data Buoy Center
NRCS	Normalized Radar Cross Section
QC	Quality Control
SAR	Synthetic Aperture Radar
SRF	Spatial Response Function
VRMS	Vector Root Mean Square
WVC	Wind Vector Cell

References

M.F. Cronin, C.W. Fairall, and M.J. McPhaden, 2006
An assessment of buoy derived and numerical weather prediction surface heat fluxes in the tropical Pacific. *J. Geophys. Res*., 111, C06038, doi:10.1029/2005JC003324.

C.W. Fairall, E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson, 2003.
Bulk Parameterization of Air-Sea Fluxes: Updates and Verification for the COARE Algorithm. *J. Climate*, 16, 571-591.

R. D. Lindsley and D. G. Long, 2015.
Enhanced-resolution reconstruction of ASCAT backscatter measurements. *IEEE Trans. Geosci. Rem. Sens*., 54, 2589–2601. doi:10.1109/TGRS.2015.2503762.

R. D. Lindsley, C. Anderson, J. Figa-Saldaña, and D. Long, 2016.
A parameterized ASCAT measurement spatial response function, *IEEE Trans. Geosc. Rem. Sens*., 54, 4570-4579. doi: 10.1109/TGRS.2016.2544835.

Verhoef, A., M. Portabella, and A. Stoffelen, 2012.
High-resolution ASCAT scatterometer winds near the coast. *IEEE Trans. Geosci. Rem. Sens*., 50, 2481 – 2487. doi: 0.1109/TGRS.2011.2175001

Vogelzang, J., and A. Stoffelen, 2017.
Developments in ASCAT wind ambiguity removal. NWP SAF Technical Report NWPSAF-KN-TR-026. Download from www.nwpsaf.eu/site/all-reports-publications.

Vogelzang, J., and A. Stoffelen, 2018.
Two-dimensional variational ambiguity removal (2DVAR), version 1.4. NWP SAF Technical Report NWPSAF-KN-TR-004. Download from www.nwpsaf.eu/site/all-reports-publications.

Vogelzang, J., and A. Stoffelen, 2020.
ASCAT land fraction validation.
OSISAF Technical Report SAF/OSI/CDOP3/KNMI/TEC/TN/???

Wessel, P., and W. H. F. Smith, 1996.
A Global Self-consistent, Hierarchical, High-resolution Shoreline Database, *J. Geophys. Res., 101*, 8741-8743.