The EUMETSAT Network of Satellite Application Facilities

NWP SAF
Numerical Weather Prediction

Document NWPSAF-KN-TR-023

Version 1.0

20-10-2014

# Estimation of background error correlation functions

Jur Vogelzang

**KNMI, de Bilt, the Netherlands**

Met Office

ECMWF

Royal Netherlands Meteorological Institute
*Ministry of Transport, Public Works and Water Management*

METEO FRANCE
Toujours un temps d'avance

# Estimation of background error correlation functions

Jur Vogelzang

KNMI, De Bilt, The Netherlands

This documentation was developed within the context of the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF), under the Cooperation Agreement dated 16 December, 2003, between EUMETSAT and the Met Office, UK, by one or more partners within the NWP SAF. The partners in the NWP SAF are the Met Office, ECMWF, KNMI and Météo France.



| Change record | | | | |
|---|---|---|---|---|
| Version | Date | Author | Approved | Remarks |
| 0.1 | November 2010 | Jur Vogelzang | | First draft |
| 1.0 | October 2014 | Jur Vogelzang | Ad Stoffelen | Final version |
| | | | | |
| | | | | |

# Contents

# 1 Introduction

## *1.1 Background error correlations in 2DVAR*

The NWP SAF scatterometer wind data processors use two-dimensional variational ambiguity removal (2DVAR) as default method for ambiguity removal. 2DVAR consists of two steps [*Vogelzang et al.*, 2009]. In the first step it makes an analysis of the ambiguous scatterometer wind solutions and a background field, usually a prediction by the ECMWF model. In the second step the ambiguous scatterometer solution closest to the analysis is selected as preferred solution. 2DVAR can therefore also be referred to as closest-to-analysis. A detailed description of 2DVAR is given by *Vogelzang* [2009].

In order to make an analysis using methods based on statistical interpolation, accurate knowledge on the background error correlations is needed. 2DVAR defines these in the spatial domain in terms of the wind potential and stream function. These functions are often referred to as error correlation functions, but should not be confused with Kolmogorov error correlation functions, which are averaged velocity differences as a function of separation. In the default 2DVAR settings the error correlation functions are simple Gaussians with parameters derived for the scatterometer carried by ERS.

2DVAR is similar to well known data assimilation systems as 3DVAR and 4DVAR. It can therefore also be seen as a simple, but nontrivial, test bed for data assimilation. From this point of view the analysis is the final product of 2DVAR.

Figure 1.1 shows wind spectra of the zonal component $u$ (left hand panel) and the meridional component $v$ (right hand panel). All spectra were obtained from all available ASCAT-12.5 data of January 2009. The blue curves show the ASCAT-12.5 selected winds spectra, the red curves the ECMWF background spectra, and the green curves the analyses spectra. The black curves show theoretical spectra following from Kolmogorov's turbulence theory at arbitrary scale; the dashed curves $k^{-5/3}$ spectra for three-dimensional turbulence and the dotted curves $k^{-3}$ spectra for two-dimensional turbulence.

Figure 1.1 shows that the ECMWF background wind field contains much less variance than the ASCAT-12.5 field. This is due to numerical cut-off in the ECMWF model that prohibits small structures to grow and degrade the medium range prediction skill. This cut-off starts already at scales of about 800 km. The figure also shows that the 2DVAR analysis does not pick up the small scale information that is present in the ASCAT-12.5 observations. the reason for this is given by *Daley* [1991]: data assimilation schemes like 2DVAR tend to act as low-pass filters. Optimal performance is reached when the background error correlations are described well.

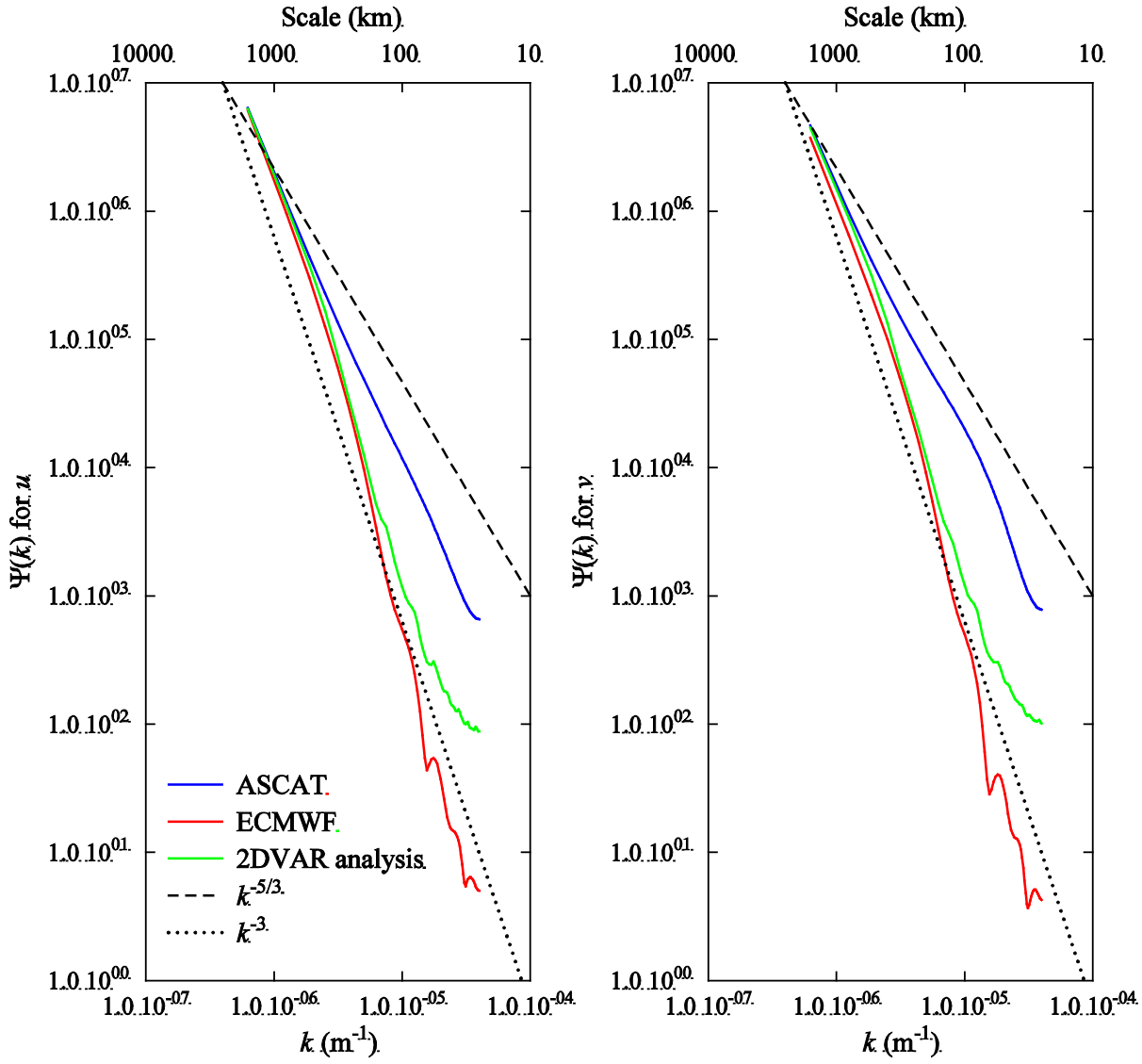


**Figure 1.1** Spectra of ASCAT-12.5 (blue curves), the ECMWF background (read curves), and the 2DVAR analyses (green curves) for all available data of January 2009. The black curves show theoretical spectra at arbitrary scale.

## 1.2 Aims and scope

The aim of this work is to find optimal error correlation functions that enable 2DVAR to arrive at an analysis that contains as much as possible small-scale information. It is expected that such an improved

analysis has little effect on the 2DVAR selection process (step 2), because the ASCAT ambiguities are well defined and limited in number. However, rotating pencil beam scatterometers like SeaWinds and OSCAT have unfavourable measurement geometry at nadir. It has been shown for SeaWinds that the wind product can be improved by using 2DVAR in combination with MSS [*Vogelzang et al.,* 2009]. Under such conditions a more detailed analysis may result in improved ambiguity skill.

For obtaining the background error correlations two methods exist at the moment. The first one is synthetic and employs model predictions at different prediction times to estimate the background error correlations using a procedure similar to Kalman filtering. This method is implemented in a number of NWP models, including that of ECMWF. The second method exploits the fact that the background error correlation equals the correlated differences between observations and background (O-B) when the observation errors are uncorrelated. *Hollingsworth and Lönnberg* [1986] binned the spatial correlations of radiosonde measurements, extrapolated it to $r = 0$ to remove the contribution of the observation error variance, and fitted a Bessel series expansion through the binned values.

In this work an alternative for the second method is given using the fact that scatterometer winds are available on a dense and regular grid. This allows direct solution of the differential equations that relate the autocorrelation of the background errors to the error correlation functions.

## *1.3 Overview*

In this report the term error correlation functions, sometimes shortened to correlation functions or error correlations, will be used for background error correlation functions.

In chapter 2 it is shown how numerical error correlation functions can be derived from background minus observation (O-B) autocorrelations. In the spatial domain each error correlation function is given by a third-order inhomogeneous ordinary differential equation which is remarkably easy to solve. It is also shown that it is possible to calculate the Fourier transform of the error correlation functions directly from the observed O-B spectra. This is mathematically more involving, but leads to a Volterra integral equation of the second kind which always has a solution.

In chapter 3 the algorithms for calculating numerical error correlation functions are tested using Gaussian error correlation functions. For these functions all results can be derived analytically. It is shown that the most precise results are obtained in the spatial domain. It is also shown here how in the spatial domain the remaining parameters (length scales and rotation/divergence ratio) can be obtained. This favours the spatial approach over the frequency approach. Both methods face the same problem: accurate error correlation functions require large sample length. In the spatial domain this is needed to let the

autocorrelations go to zero for large distances and in the frequency domain for getting a fine grid. In the spatial domain it is easier to obtain large sample lengths because missing points are easily handled whereas in the frequency domain missing points must be interpolated. Therefore it is concluded that the error correlation functions are calculated best in the spatial domain.

In chapter 4 error correlation functions obtained from observed O-B statistics are presented. It is shown here that correlations in the observations are negligible and that sampling with exclusion of mission points yields best results. The observed autocorrelations must be forced to zero using a cut-off function. Several choices of the cut-off function are studied. It will be shown that it is not possible to calculate separate error correlation functions for the Tropics and the Extratropics, because of the slow convergence to zero of the observed B-O autocorrelations. However, zonal effects are studies as far as possible.

In chapter 5 the effects on analysis and selection are presented using only global error correlation functions obtained with a brick-wall cut-off. It is shown that a small batch grid size of 25 km is needed in order to handle small-scale information. Combined with a 2DVAR free zone of 6000 km to ensure that the analysis properly goes to zero at the batch edges, this leads to batch grids of about $500 \times 500$ points. As a consequence, 2DVAR requires a lot of computer resources, so the calculations are limited to the first three days of January 2009. It is shown that numerical error correlation functions lead to a significant increase in the analysis spectral density at small scales.

As expected, the effect on the ASCAT-12.5 selection itself is not very large, because in most cases ASCAT's ambiguities are well defined. Differences occur in difficult situations with a limited number of observations due to rain flagging or the presence of the coast, with a large discrepancy between observations and background, or in situations with small-scale features. A case study of ASCAT-12.5 demonstrates the potential of numerical error correlation functions in a case of rapidly varying wind direction over a front.

The report ends with the conclusions and recommendations in chapter 6. A number of derivations is presented in the appendices.

# 2 Numerical correlation functions

## 2.1 Approach

The problem under consideration is how wind spectra from scatterometer measurements and model predictions can be used to derive error correlation functions for 2DVAR that yield analyses as close as possible to the scatterometer observations. In particular, these error correlation functions should retain small scale information in the analysis that is present in scatterometer wind fields but not in model predictions.

Wind spectra are defined in terms of the transverse and longitudal wind components in the frequency domain, $\hat{t}$ and $\hat{l}$, respectively, while error correlation functions are defined for the velocity potential and the stream function in the spatial domain, $\chi$ and $\psi$. These domains are connected by Fourier and Helmholtz transformations as indicated in figure 2.1.

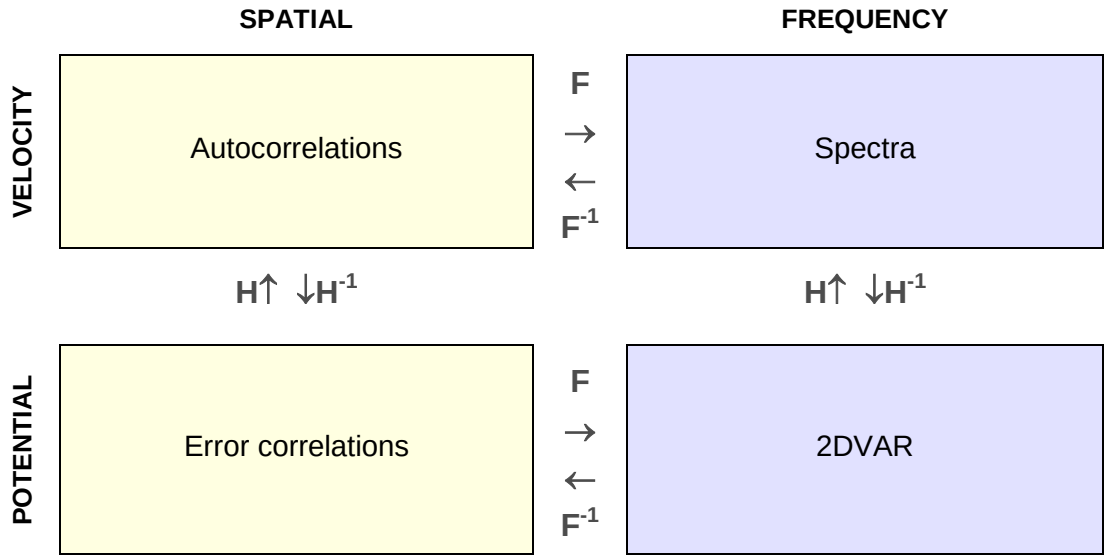


**Figure 2.1** Relations between wind speed and potentials in the two-dimensional spatial and frequency domains. F stands for Fourier transformation, H for Helmholtz transformation. The expressions for forward and inverse Helmholtz transformation are given explicitly.

Figure 2.1 shows that a connection from the upper right to the lower left must be found. This can be achieved by inverting the Helmholz transformation in the spatial domain, as shown in section 2.2, and in the frequency domain, as shown in section 2.3.

The spectra and the error correlation functions are both one-dimensional, the spectra because of the lack of a sufficient number of samples across the satellite track, the error correlation functions because of isotropy. Therefore it is natural to use one-dimensional Fourier and Helmholz transforms to solve the problem.

## 2.2 Spatial domain

### 2.2.1 Decoupling

In the spatial domain the autocorrelations of the longitudinal and transversal wind components, $\rho_{ll}$ and $\rho_{tt}$, are related to the error correlation functions as [*Daley*, 1993, equations 5.2.28-29]

$$\rho_{ll}(r) = -L_\psi^2(1-\nu^2)\frac{1}{r}\frac{d\rho_{\psi\psi}(r)}{dr} - L_\chi^2\nu^2\frac{d^2\rho_{\chi\chi}(r)}{dr^2} \quad , \tag{2.1a}$$

$$\rho_{tt}(r) = -L_\psi^2(1-\nu^2)\frac{d^2\rho_{\psi\psi}(r)}{dr^2} - L_\chi^2\nu^2\frac{1}{r}\frac{d\rho_{\chi\chi}(r)}{dr} \quad . \tag{2.1b}$$

Dropping the argument $r$ and using primes to denote the derivatives, these equations can be rewritten as

$$\rho'_{\psi\psi} = \frac{1}{a_\psi} r\rho_{ll} - \frac{a_\chi}{a_\psi} r\rho''_{\chi\chi} \quad , \tag{2.2a}$$

$$\rho'_{\chi\chi} = \frac{1}{a_\chi} r\rho_{tt} - \frac{a_\psi}{a_\chi} r\rho''_{\psi\psi} \quad , \tag{2.2b}$$

where

$$a_\psi = -L_\psi^2(1-\nu^2) \quad , \quad a_\chi = -L_\chi^2\nu^2 \quad . \tag{2.3}$$

Substitution of (2.2b) in (2.1a) yields

$$\rho_{ll} = a_\psi\frac{1}{r}\rho'_{\psi\psi} + a_\chi\frac{d}{dr}\left[\frac{1}{a_\chi} r\rho_{tt} - \frac{a_\psi}{a_\chi} r\rho''_{\psi\psi}\right] =$$

$$= a_\psi\frac{1}{r}\rho'_{\psi\psi} + \rho_{tt} + r\rho'_{tt} - a_\psi\rho''_{\psi\psi} - a_\psi r\rho'''_{\psi\psi} \quad .$$

This can be rewritten as

$$\rho'''_{\psi\psi} + \frac{1}{r}\rho''_{\psi\psi} - \frac{1}{r^2}\rho'_{\psi\psi} = \frac{1}{a_\psi}\left[\frac{\rho_{tt}-\rho_{ll}}{r} + \rho'_{tt}\right] \quad . \tag{2.4a}$$

In the same manner, substitution of (2.2a) in (2.1b) yields

$$\rho'''_{\chi\chi} + \frac{1}{r}\rho''_{\chi\chi} - \frac{1}{r^2}\rho'_{\chi\chi} = \frac{1}{a_\chi}\left[\frac{\rho_{ll} - \rho_{tt}}{r} + \rho'_{ll}\right] \quad . \tag{2.4b}$$

Equations (2.4a) and (2.4b) are decoupled and give $\rho_{\psi\psi}$ and $\rho_{\chi\chi}$, respectively, in terms of the observables $\rho_{tt}$ and $\rho_{ll}$ as a third-order inhomogeneous linear differential equation. Note that (2.4a) and (2.4b) transform into each other when simultaneously interchanging $\psi \leftrightarrow \chi$, $l \leftrightarrow t$. Therefore only one equation needs to be solved explicitly; the solution of the other follows from interchanging the subscripts.

### 2.2.2 Solution of the third order equation

Equation (2.4a) can be solved by writing

$$X_{\psi\psi} = \frac{d\rho_{\psi\psi}}{dr} \quad \Rightarrow \quad \rho_{\psi\psi}(r) = \int_0^r ds\; X_{\psi\psi}(s) + C_\psi \quad , \tag{2.5a}$$

with $C_\psi$ an integration constant, and

$$X''_{\psi\psi} + \frac{1}{r}X'_{\psi\psi} - \frac{1}{r^2}X_{\psi\psi} = g_\psi \quad , \tag{2.5b}$$

with

$$g_\psi = \frac{1}{a_\psi}\left[\frac{\rho_{tt} - \rho_{ll}}{r} + \rho'_{tt}\right] \quad . \tag{2.5c}$$

Equation (2.5b) is an inhomogeneous second-order Euler differential equation. Its solution is shown in appendix A to be

$$X_{\psi\psi}(r) = r\left[A_\psi + \tfrac{1}{2}K_\psi(r)\right] + \frac{1}{r}\left[B_\psi - \tfrac{1}{2}L_\psi(r)\right] \quad , \tag{2.6a}$$

with

$$K_\psi(r) = \int_0^r ds\; g_\psi(s) = \frac{1}{a_\psi}\int_0^r ds\left[\frac{\rho_{tt}(s) - \rho_{ll}(s)}{s} + \rho'_{tt}(s)\right] \quad , \tag{2.6b}$$

$$L_\psi(r) = \int_0^r ds\; s^2 g_\psi(s) = \frac{1}{a_\psi}\int_0^r ds\left[s(\rho_{tt}(s) - \rho_{ll}(s)) + s^2\rho'_{tt}(s)\right] \quad , \tag{2.6c}$$

and $A_\psi$ and $B_\psi$ integration constants. The terms with $\rho'_{tt}$ in (2.6b) and (2.6c) can be handled with partial integration, resulting in

$$K_\psi(r) = \frac{1}{a_\psi}\int_0^r ds\left[\frac{\rho_{tt}(s) - \rho_{ll}(s)}{s}\right] + \frac{1}{a_\psi}\rho_{tt}\Big|_{s=0}^{s=r} = \frac{1}{a_\psi}\int_0^r ds\left[\frac{\rho_{tt}(s) - \rho_{ll}(s)}{s}\right] + \frac{\rho_{tt}(r) - 1}{a_\psi} \quad , \tag{2.7a}$$

$$L_\psi(r) = \frac{1}{a_\psi}\int_0^r ds\, s[\rho_{tt}(s) - \rho_{ll}(s)] + \left.\frac{s^2\rho_{tt}(s)}{a_\psi}\right|_{s=0}^{s=r} - \frac{2}{a_\psi}\int_0^r ds\, s\rho_{tt}(s) =$$
$$= \frac{-1}{a_\psi}\int_0^r ds\, s[\rho_{tt}(s) + \rho_{ll}(s)] + \frac{r^2\rho_{tt}(r)}{a_\psi} \quad , \tag{2.7b}$$

where we have used that $\rho_{tt}(0) = 1$. Substituting this in (2.6a) gives

$$X_{\psi\psi}(r) = r\left\{A_\psi + \frac{\rho_{tt}(r) - 1}{2a_\psi} + \frac{1}{2a_\psi}\int_0^r ds\left[\frac{\rho_{tt}(s) - \rho_{ll}(s)}{s}\right]\right\} +$$
$$+ \frac{1}{r}\left\{B_\psi - \frac{r^2\rho_{tt}(r)}{2a_\psi} + \frac{1}{2a_\psi}\int_0^r ds\, s[\rho_{tt}(s) + \rho_{ll}(s)]\right\} \quad . \tag{2.8}$$

This can be simplified to

$$X_{\psi\psi}(r) = r\left\{A_\psi - \frac{1}{2a_\psi} + \frac{1}{2a_\psi}\int_0^r ds\left[\frac{\rho_{tt}(s) - \rho_{ll}(s)}{s}\right]\right\} +$$
$$+ \frac{1}{r}\left\{B_\psi + \frac{1}{2a_\psi}\int_0^r ds\, s[\rho_{tt}(s) + \rho_{ll}(s)]\right\} \quad . \tag{2.9}$$

### 2.2.3 Boundary conditions

Now $X_{\psi\psi}$ is the derivative of $\rho_{\psi\psi}$ by (2.5a). We want $\rho_{\psi\psi}$ to become one at $r = 0$ and go to zero as $r \to \infty$. This implies that $X_{\psi\psi}$ should be regular at $r = 0$ (it may become zero, but that is not necessary) and go to zero as $r \to \infty$.

The first term in (2.9) is regular when $r \to 0$, because then the integrand of the integral over $s$ can be expanded in a Taylor series as

$$\frac{\rho_{tt}(s) - \rho_{ll}(s)}{s} \approx \frac{1 + s\rho'_{tt}(s) + O(s^2) - 1 - s\rho'_{ll}(s) + O(s^2)}{s} = \rho'_{tt}(s) - \rho'_{ll}(s) + O(s) \quad . \tag{2.10}$$

This approaches to a constant value when $r \to 0$ so the integral in (2.9) approaches zero at least as fast as $r$.

The integrand of the integral over $s$ in the second term of (2.9) goes to zero as $r$ when $r \to 0$, so the integral itself approaches zero as $r^2$. When multiplied with the factor $r^{-1}$ the second term behaves as

$$\frac{1}{r}\left\{B_\psi + \frac{1}{2a_\psi}\int_0^r ds\, s[\rho_{tt}(s) + \rho_{ll}(s)]\right\} \approx \frac{B_\psi}{r} + \frac{r}{2a_\psi} \quad . \tag{2.11}$$

This will diverge for $r \to 0$ unless $B_\psi = 0$. With this choice we have that

$$\lim_{r\to 0} X_{\psi\psi}(r) = 0 \quad . \tag{2.12}$$

In the limit $r \to \infty$ the first term in (2.9) will diverge, unless

$$A_\psi - \frac{1}{2a_\psi} + \frac{1}{2a_\psi}\int_0^\infty ds\left[\frac{\rho_{tt}(s)-\rho_{ll}(s)}{s}\right] = 0 \quad . \tag{2.13}$$

This fixes the integration constant $A_\psi$ to

$$A_\psi = \frac{1}{2a_\psi} - \frac{1}{2a_\psi}\int_0^\infty ds\left[\frac{\rho_{tt}(s)-\rho_{ll}(s)}{s}\right] \quad . \tag{2.14}$$

With these values for the integration constants the solution for $X_{\psi\psi}$ becomes

$$\begin{aligned} X_{\psi\psi}(r) = r\left\{\frac{1}{2a_\psi}\int_0^r ds\left[\frac{\rho_{tt}(s)-\rho_{ll}(s)}{s}\right] - \frac{1}{2a_\psi}\int_0^\infty ds\left[\frac{\rho_{tt}(s)-\rho_{ll}(s)}{s}\right]\right\} + \\ + \frac{1}{2a_\psi r}\int_0^r ds\, s\left[\rho_{tt}(s)+\rho_{ll}(s)\right] \quad , \end{aligned} \tag{2.15}$$

which can be simplified to

$$X_{\psi\psi}(r) = -\frac{r}{2a_\psi}\int_r^\infty ds\left[\frac{\rho_{tt}(s)-\rho_{ll}(s)}{s}\right] + \frac{1}{2a_\psi r}\int_0^r ds\, s\left[\rho_{tt}(s)+\rho_{ll}(s)\right] \quad . \tag{2.16a}$$

The solution for $X_{\chi\chi}$ immediately follows from the interchange $\psi \leftrightarrow \chi$, $l \leftrightarrow t$ as

$$X_{\chi\chi}(r) = -\frac{r}{2a_\chi}\int_r^\infty ds\left[\frac{\rho_{ll}(s)-\rho_{tt}(s)}{s}\right] + \frac{1}{2a_\chi r}\int_0^r ds\, s\left[\rho_{ll}(s)+\rho_{tt}(s)\right] \quad . \tag{2.16b}$$

Taking the asymmetry and symmetry of the integrand into account, the solution reads

$$X_{\psi\psi}(r) = -\frac{r}{2a_\psi} I(r) + \frac{1}{2a_\psi r} J(r) \quad , \quad X_{\chi\chi}(r) = \frac{r}{2a_\chi} I(r) + \frac{1}{2a_\chi r} J(r) \quad , \tag{2.17a}$$

$$I(r) = \int_r^\infty ds\left[\frac{\rho_{tt}(s)-\rho_{ll}(s)}{s}\right] \quad , \quad J(r) = \int_0^r ds\, s\left[\rho_{tt}(s)+\rho_{ll}(s)\right] \quad . \tag{2.17b}$$

### 2.2.4 Final solution

From (2.5a) and (2.17a) the solutions read

$$\rho_{\psi\psi}(r) = -\frac{1}{2a_\psi}\int_0^r ds\, sI(s) + \frac{1}{2a_\psi}\int_0^r ds\, \frac{J(s)}{s} + C_\psi \quad , \tag{2.18a}$$

$$\rho_{\chi\chi}(r) = \frac{1}{2a_\chi}\int_0^r ds\, sI(s) + \frac{1}{2a_\chi}\int_0^r ds\, \frac{J(s)}{s} + C_\chi \quad . \tag{2.18b}$$

As was pointed out earlier, the integrands in (2.18a-b) go to zero when $r$ approaches zero, and so do the integrals. Since we require that $\rho_{\psi\psi}(0) = \rho_{\chi\chi}(0) = 1$ we must have $C_\psi = C_\chi = 1$, so the final solution reads

$$\rho_{\psi\psi}(r) = 1 + \frac{S(r) - R(r)}{2a_\psi} \quad , \quad \rho_{\chi\chi}(r) = 1 + \frac{S(r) + R(r)}{2a_\chi} \quad , \tag{2.19a}$$

$$R(r) = \int_0^r ds\, sI(s) \quad , \quad S(r) = \int_0^r ds\, \frac{J(s)}{s} \quad , \tag{2.19b}$$

with $I$ and $J$ defined in (2.17b).

## 2.3 Frequency domain

### 2.3.1 Definition of the Fourier transform

The forward and inverse transformation from spatial coordinate $r$ to frequency coordinate $k$ and vice versa are defined as

$$\hat{f}(k) = \int_{-\infty}^{\infty} dr\, e^{2\pi i k r} f(r) \quad , \tag{2.20a}$$

$$f(r) = \int_{-\infty}^{\infty} dk\, e^{-2\pi i k r} \hat{f}(\mathrm{k}) \quad , \tag{2.20b}$$

with orthogonality condition

$$\int_{-\infty}^{\infty} dr\, e^{2\pi i (k-p) r} = \delta(k-p) \quad , \tag{2.21}$$

where $\delta(k-p)$ stands for the Dirac delta function.

In the remainder of this section we will use the caret to denote functions in the frequency domain.

### 2.3.2 Integral equations

The Fourier transform of (2.1a) reads

$$\hat{\rho}_{ll}(k) = \int_{-\infty}^{\infty} dr\, e^{2\pi i k r}\, \rho_{ll}(r) = \int_{-\infty}^{\infty} dr\, e^{2\pi i k r} \left[ \frac{a_\psi}{r} \frac{d\rho_{\psi\psi}(r)}{dr} + a_\chi \frac{d^2 \rho_{\chi\chi}(r)}{dr^2} \right] =$$
$$= a_\psi \int_{-\infty}^{\infty} dr\, e^{2\pi i k r} \frac{d}{dr} \int_{-\infty}^{\infty} dp\, e^{-2\pi i p r}\, \hat{\rho}_{\psi\psi}(p) + a_\chi \int_{-\infty}^{\infty} dr\, e^{2\pi i k r} \frac{d^2}{dr^2} \int_{-\infty}^{\infty} dp\, e^{-2\pi i p r}\, \hat{\rho}_{\chi\chi}(p) \quad . \tag{2.22}$$

Interchanging the order of integrations and differentiations yields

$$\hat{\rho}_{ll}(k) = a_\psi \int_{-\infty}^{\infty} dp\, \hat{\rho}_{\psi\psi}(p) \int_{-\infty}^{\infty} dr \frac{e^{2\pi i k r}}{r} \frac{d}{dr} e^{-2\pi i p r} + a_\chi \int_{-\infty}^{\infty} dp\, \hat{\rho}_{\chi\chi}(p) \int_0^{\infty} dr\, e^{2\pi i k r} \frac{d^2}{dr^2} e^{-2\pi i p r} \quad . \tag{2.23}$$

Evaluating the derivatives gives

$$\hat{\rho}_{ll}(k) = -(2\pi i) a_\psi \int_{-\infty}^{\infty} dp\; p \hat{\rho}_{\psi\psi}(p) \int_{-\infty}^{\infty} dr \frac{e^{2\pi i (k-p) r}}{r} - (2\pi)^2 a_\chi \int_{-\infty}^{\infty} dp\; p^2 \hat{\rho}_{\chi\chi}(p) \int_{-\infty}^{\infty} dr\, e^{2\pi i (k-p) r} \quad . \tag{2.24}$$

The second integral over $r$ equals a delta function, see (2.21), while the first one equals a step function, as shown in appendix B

$$\int_{-\infty}^{\infty} dr \frac{e^{2\pi i (k-p) r}}{r} = \begin{cases} i\pi & , \quad k - p > 0 \Leftrightarrow p < k \\ 0 & , \quad k = p \\ -i\pi & , \quad k - p < 0 \Leftrightarrow p > k \end{cases} \quad . \tag{2.25}$$

Substituting the integrals in (2.25) yields

$$\hat{\rho}_{ll}(k) = 2\pi^2 \int_{-\infty}^{k} dp\; p a_\psi \hat{\rho}_{\psi\psi}(p) - 2\pi^2 \int_{k}^{\infty} dp\; p a_\psi \hat{\rho}_{\psi\psi}(p) - (2\pi)^2 k^2 a_\chi \hat{\rho}_{\chi\chi}(k) \quad . \tag{2.26}$$

Now $\hat{\rho}_{\psi\psi}$ is an even function, because it is the Fourier transform of a real and even function (the error correlation function), so

$$\int_{-\infty}^{\infty} dp\; p \hat{\rho}_{\psi\psi}(p) = 0 \quad . \tag{2.27}$$

Equation (2.26) can therefore be simplified to

$$\hat{\rho}_{ll}(k) = -(2\pi)^2 \int_{k}^{\infty} dp\; p a_\psi \hat{\rho}_{\psi\psi}(p) - (2\pi)^2 k^2 a_\chi \hat{\rho}_{\chi\chi}(k) \quad . \tag{2.28a}$$

Since (2.1b) follows from (2.1a) by the simultaneous interchanges $t \leftrightarrow l$ and $\psi \leftrightarrow \chi$, one immediately finds

$$\hat{\rho}_{tt}(k) = -(2\pi)^2 \int_k^\infty dp\ p a_\chi \hat{\rho}_{\chi\chi}(p) - (2\pi)^2 k^2 a_\psi \hat{\rho}_{\psi\psi}(k) \quad . \tag{2.28b}$$

### 2.3.3 Decoupling

Equations (2.28a-b) constitute a set of two coupled integral equations relating the Fourier transform of the error correlation functions, $\hat{\rho}_{\psi\psi}$ and $\hat{\rho}_{\chi\chi}$, to the spectra $\hat{\rho}_{ll}$ and $\hat{\rho}_{tt}$. They can be decoupled by isolating $a_\chi \hat{\rho}_{\chi\chi}$ from (2.28a) and substituting this into (2.28b) as

$$\begin{aligned}\hat{\rho}_{tt}(k) &= -(2\pi)^2 \int_k^\infty dp\ p \frac{-1}{(2\pi)^2 p^2} \left[ (2\pi)^2 \int_p^\infty dq\, q a_\psi \hat{\rho}_{\psi\psi}(q) + \hat{\rho}_{ll}(p) \right] - (2\pi)^2 k^2 a_\psi \hat{\rho}_{\psi\psi}(k) = \\ &= (2\pi)^2 \int_k^\infty \frac{dp}{p} \int_p^\infty dq\, q a_\psi \hat{\rho}_{\psi\psi}(q) + \int_p^\infty dp \frac{\hat{\rho}_{ll}(p)}{p} - (2\pi)^2 k^2 a_\psi \hat{\rho}_{\psi\psi}(k) \quad .\end{aligned} \tag{2.29}$$

Rearranging terms results in

$$a_\psi \hat{\rho}_{\psi\psi}(k) = g_\psi(k) + \frac{1}{k^2} \int_k^\infty \frac{dp}{p} \int_p^\infty dq\, q a_\psi \hat{\rho}_{\psi\psi}(q) \quad , \tag{2.30a}$$

where

$$g_\psi(k) = \frac{1}{(2\pi)^2 k^2} \left[ -\hat{\rho}_{tt}(k) + \int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p} \right] \quad . \tag{2.30b}$$

The order of the integrations in (2.30a) can be interchanged due to

$$\int_k^\infty dx \int_x^\infty dy\ f(x,y) = \int_k^\infty dy \int_k^y dx\ f(x,y) \quad . \tag{2.31}$$

This relation is easy proved when considering figure 2.1. The left hand side of (2.31) corresponds to the left hand panel of figure 2.1, where for each value of $x$ the integration over $y$ is done from $y = x$ (represented by the dotted line) to $y = \infty$ as indicated by the arrows. The integration area is the infinite triangle bounded by the lines $x = k$ (dashed) and $y = x$. The same area is covered in the right hand panel of figure 2.1, where for each value of $y$ the integration over $x$ is done from $x = k$ to $x = y$, thus establishing (2.31).

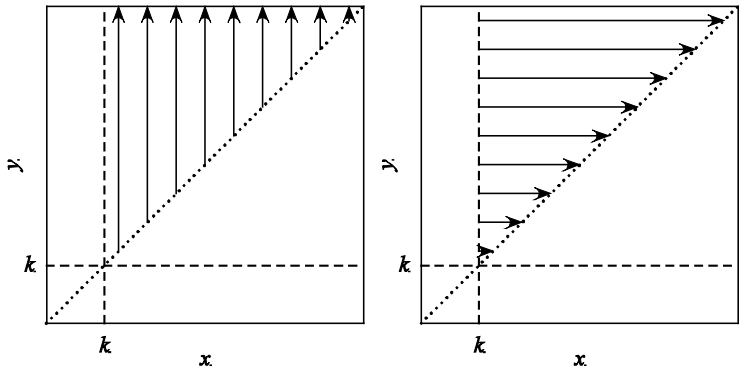


**Figure 2.1** Interchange of the order of integrations

Application to (2.30a) leads to

$$a_\psi \hat{\rho}_{\psi\psi}(k) = g_\psi(k) + \frac{1}{k^2}\int_k^\infty dq\, q a_\psi \hat{\rho}_{\psi\psi}(q) \int_k^q \frac{dp}{p} = g_\psi(k) + \frac{1}{k^2}\int_k^\infty dq\, q \ln\left(\frac{q}{k}\right) a_\psi \hat{\rho}_{\psi\psi}(q) \quad , \qquad (2.32)$$

In the same manner one finds

$$a_\chi \hat{\rho}_{\chi\chi}(k) = g_\chi(k) + \frac{1}{k^2}\int_k^\infty dq\, q \ln\left(\frac{q}{k}\right) a_\chi \hat{\rho}_{\chi\chi}(q) \quad . \qquad (2.33a)$$

where

$$g_\chi(k) = \frac{1}{(2\pi)^2 k^2}\left[ -\hat{\rho}_{ll}(k) + \int_k^\infty dp \frac{\hat{\rho}_{tt}(p)}{p} \right] \quad . \qquad (2.33b)$$

Equations (2.30a) and (2.31a) are Volterra integral equations of the second kind. The integration range is different than in most text book examples, but they can be solved in the same way with the method of successive approximations.

### 2.3.4 Method of successive approximations

Equation (2.32) can be rewritten as

$$\Psi(k) = h(k) + \int_k^\infty dq\, K(k,q)\Psi(q) \quad , \qquad (2.34)$$

with $\Psi(k) = a_\psi \hat{\rho}_{\psi\psi}(k)$ and $h(k) = g_\psi(k)$ to simplify the notation, and kernel

$$K(k,q) = \frac{q}{k^2}\ln\left(\frac{q}{k}\right) \quad . \qquad (2.35)$$

In the method of successive approximations a first solution $\Psi_1 = h$ is substituted in the right hand side of (2.34) to obtain a solution $\Psi_2$. This solution is again substituted in (2.34) to yield $\Psi_3$, etc. By applying the formula for integration order interchange (2.31), the *n*-th iterated solution $\Psi_n$ can be written as

$$\Psi_n(k) = h(k) + \sum_{j=1}^{n} \int_k^\infty dq\, K_n(k,q) h(q) \quad , \tag{2.36}$$

with $K_n$ the *n*-th iterated kernel which satisfies

$$K_n(k,q) = \int_k^q dq_n K_{n-1}(k,q_n) K_1(q_n,k) \quad , \tag{2.37}$$

and $K_1 = K$. The order summation and integration in (2.36) can be changed, leading to the solution

$$\Psi_n(k) = h(k) + \int_k^\infty dq\, \Gamma(k,q) h(q) \quad , \tag{2.38}$$

with the resolvent kernel $\Gamma$ given by

$$\Gamma(k,q) = \sum_{j=1}^{\infty} K_j(k,q) \quad . \tag{2.39}$$

In appendix C it will be shown that

$$K_n(k,q) = \frac{q}{k^2} \frac{\ln^{2n-1}\left(\frac{q}{k}\right)}{(2n-1)!} \quad , \tag{2.40}$$

so the resolvent kernel reads

$$\begin{aligned} \Gamma(k,q) &= \frac{q}{k^2} \sum_{j=1}^{\infty} \frac{\ln^{2j-1}\left(\frac{q}{k}\right)}{(2j-1)!} = \frac{q}{k^2} \sum_{m=0}^{\infty} \frac{\ln^{2m+1}\left(\frac{q}{k}\right)}{(2m+1)!} = \frac{q}{k^2} \sinh\left(\ln\frac{q}{k}\right) = \\ &= \frac{q}{2k^2}\left(\frac{q}{k} - \frac{k}{q}\right) \quad . \end{aligned} \tag{2.41}$$

## 2.3.5 Final solution

Recalling that $\Psi(k) = a_\psi \hat{\rho}_{\psi\psi}(k)$ and $h(k) = g_\psi(k)$, the solution reads

$$a_\psi \hat{\rho}_{\psi\psi}(k) = g_\psi(k) + \int_k^\infty dq \frac{q}{2k^2}\left(\frac{q}{k} - \frac{k}{q}\right) g_\psi(q) \quad , \tag{2.42a}$$

$$a_\chi \hat{\rho}_{\chi\chi}(k) = g_\chi(k) + \int_k^\infty dq \frac{q}{2k^2}\left(\frac{q}{k} - \frac{k}{q}\right) g_\chi(q) \quad , \tag{2.42b}$$

with

$$g_\psi(k) = \frac{1}{(2\pi)^2 k^2}\left[ -\hat{\rho}_{tt}(k) + \int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p} \right] \quad , \tag{2.43a}$$

$$g_\chi(k) = \frac{1}{(2\pi)^2 k^2}\left[ -\hat{\rho}_{ll}(k) + \int_k^\infty dp \frac{\hat{\rho}_{tt}(p)}{p} \right] \quad . \tag{2.43b}$$

Substituting (2.43a) into (2.42a) yields

$$\begin{aligned} a_\psi \hat{\rho}_{\psi\psi}(k) = \frac{1}{(2\pi)^2 k^2}\Bigg[ & -\hat{\rho}_{tt}(k) + \int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p} - \int_k^\infty dq \frac{1}{2q}\left(\frac{q}{k} - \frac{k}{q}\right)\hat{\rho}_{tt}(q) + \\ & + \int_k^\infty dq \frac{1}{2q}\left(\frac{q}{k} - \frac{k}{q}\right)\int_q^\infty dp \frac{\hat{\rho}_{ll}(p)}{p} \Bigg] \quad . \end{aligned} \tag{2.44}$$

The double integral can be simplified by changing once more the order of integration, cf. (2.31), to

$$\begin{aligned} & \int_k^\infty dq \frac{1}{2q}\left(\frac{q}{k} - \frac{k}{q}\right)\int_q^\infty dp \frac{\hat{\rho}_{ll}(p)}{p} = \int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p}\int_k^p dq \frac{1}{2q}\left(\frac{q}{k} - \frac{k}{q}\right) = \int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p}\left[\frac{q}{2k} + \frac{k}{2q}\right]_{q=k}^{q=p} = \\ & = \int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p}\left[\frac{p}{2k} + \frac{k}{2p} - 1\right] \quad . \end{aligned} \tag{2.45}$$

Substitution of (2.45) into (2.44), changing the integration variable in the third term of (2.44) from $q$ to $p$, and rearranging terms gives

$$a_\psi \hat{\rho}_{\psi\psi}(k) = \frac{1}{(2\pi)^2 k^2}\left[ -\hat{\rho}_{tt}(k) + \frac{1}{2}\int_k^\infty dp \frac{\hat{\rho}_{ll}(p)}{p}\left(\frac{p}{k} + \frac{k}{p}\right) - \frac{1}{2}\int_k^\infty dp \frac{\hat{\rho}_{tt}(p)}{p}\left(\frac{p}{k} - \frac{k}{p}\right) \right] \quad , \tag{2.46}$$

with a similar expression for $a_\chi \hat{\rho}_{\chi\chi}$ when simultaneously interchanging $\psi \leftrightarrow \chi$ and $t \leftrightarrow l$. This can also be written as

$$\begin{aligned} a_\psi \hat{\rho}_{\psi\psi}(k) &= \frac{1}{(2\pi)^2 k^2}\left[ -\hat{\rho}_{tt}(k) + \frac{P(k)}{2k} + \frac{k}{2}Q(k) \right] \quad , \\ a_\chi \hat{\rho}_{\chi\chi}(k) &= \frac{1}{(2\pi)^2 k^2}\left[ -\hat{\rho}_{ll}(k) - \frac{P(k)}{2k} + \frac{k}{2}Q(k) \right] \quad , \end{aligned} \tag{2.47a}$$

with

$$P(k) = \int_k^\infty dp \left[\hat{\rho}_{ll}(p) - \hat{\rho}_{tt}(p)\right] \quad ,$$
$$Q(k) = \int_k^\infty dp \frac{\hat{\rho}_{ll}(p) + \hat{\rho}_{tt}(p)}{p^2} \quad . \tag{2.47b}$$

### 2.3.6 Properties of the solution

The solution (2.47a) has a singularity of (at least) order 2 at $k = 0$. Yet, this singularity must be removable, since the solution should be regular at the origin. This is most easily seen by differentiating (2.28a-b) twice to $k$ and evaluating the result at $k = 0$:

$$\hat{\rho}''_{tt}(0) = -(2\pi)^2 \left(a_\chi \hat{\rho}_{\chi\chi}(0) + 2a_\psi \hat{\rho}_{\psi\psi}(0)\right) \quad , \tag{2.48a}$$

$$\hat{\rho}''_{ll}(0) = -(2\pi)^2 \left(a_\psi \hat{\rho}_{\psi\psi}(0) + 2a_\chi \hat{\rho}_{\chi\chi}(0)\right) \quad . \tag{2.48b}$$

Solving for the error correlation function values gives

$$a_\psi \hat{\rho}_{\psi\psi}(0) = \frac{-\hat{\rho}''_{ll}(0) - 2\hat{\rho}''_{tt}(0)}{3(2\pi)^2} \quad , \tag{2.49a}$$

$$a_\chi \hat{\rho}_{\chi\chi}(0) = \frac{-\hat{\rho}''_{tt}(0) - 2\hat{\rho}''_{ll}(0)}{3(2\pi)^2} \quad , \tag{2.49b}$$

which is indeed regular. Equations (2.49a-b) can be used to calculate the error correlation function values at $k = 0$ using the fourth-order formula of appendix D.

# 3 Implementation and test

## *3.1 Determination of the parameters*

### 3.1.1 Spatial domain

The requirement that the error correlation functions $\rho_{\psi\psi}$ and $\rho_{\chi\chi}$ vary between 1 and 0 fixes the constants $a_\psi$ and $a_\chi$ using (2.19a) as

$$a_\psi = -\tfrac{1}{2}\left[S(N\Delta)\text{-}R(N\Delta)\right] \quad , \quad a_\chi = -\tfrac{1}{2}\left[S(N\Delta) + R(N\Delta)\right] \quad , \tag{3.1}$$

where, cf. equation (2.3)

$$a_\psi = -L_\psi^2(1-\nu^2) \quad , \quad a_\chi = -L_\chi^2\nu^2 \quad . \tag{3.2}$$

The length scales (or strength parameters) $L_\psi$ and $L_\chi$ can also be obtained from the error correlation functions themselves [*Daley*, 1991, equation (4.3.12)] as

$$L_\psi^2 = \left.\frac{\rho_{\psi\psi}(r)}{\nabla^2\rho_{\psi\psi}(r)}\right|_{r=0} \quad , \quad L_\chi^2 = \left.\frac{\rho_{\chi\chi}(r)}{\nabla^2\rho_{\chi\chi}(r)}\right|_{r=0} \quad . \tag{3.3}$$

Note that Daley's equation contains an additional factor 2 which is attributed to the fact that the problem is two-dimensional. The Gaussian test described in the next paragraph shows that (3.3) yields the correct values.

Equation (3.3) can be handled in two ways. It can be cast in second-order finite difference form to yield

$$L_\psi^2 = \frac{\Delta^2}{2\left[\rho_{\psi\psi}(\Delta)\text{-}1\right]} \quad , \quad L_\chi^2 = \frac{\Delta^2}{2\left[\rho_{\chi\chi}(\Delta)\text{-}1\right]} \quad , \tag{3.4}$$

where we have used that the error correlation functions that are symmetric around the origin, $\rho(-\Delta) = \rho(\Delta)$ with $\rho(0) = 1$. Now (3.2) yields two independent estimates for $\nu^2$. In the ideal case these are equal, but in practice differences may occur, most likely due to the fact that numerical differentiation is notoriously prone to numerical errors.

A more consistent and numerically stable approach is obtained by substituting (2.19a-b) and (2.17b) in (3.3). The Laplacian operator in polar coordinates reads [*Daley*, 1991]

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2} \quad . \tag{3.5}$$

When there is no angular dependency and if the first derivative vanishes, as is the case here, this reduces to

$$\nabla^2 = \frac{1}{r}\frac{d}{dr} r \frac{d}{dr} = \frac{1}{r}\left( r\frac{d^2}{dr^2} + \frac{d}{dr} \right) = \frac{d^2}{dr^2} \quad . \tag{3.6}$$

Then we have

$$\begin{aligned} \rho''_{\psi\psi}(r) &= \frac{d^2}{dr^2}\left( 1 + \frac{S(r) - R(r)}{2a_\psi} \right) = \frac{1}{2a_\psi}\left[ \left( \frac{d}{dr}\frac{J(r)}{r} \right) - \left( \frac{d}{dr} rI(r) \right) \right] = \\ &= \frac{1}{2a_\psi}\left[ -\frac{J(r)}{r^2} + \frac{J'(r)}{r} - I(r) - rI'(r) \right] \quad , \end{aligned} \tag{3.7a}$$

$$\begin{aligned} \rho''_{\chi\chi}(r) &= \frac{d^2}{dr^2}\left( 1 + \frac{S(r) + R(r)}{2a_\psi} \right) = \frac{1}{2a_\chi}\left[ \left( \frac{d}{dr}\frac{J(r)}{r} \right) + \left( \frac{d}{dr} rI(r) \right) \right] = \\ &= \frac{1}{2a_\chi}\left[ -\frac{J(r)}{r^2} + \frac{J'(r)}{r} + I(r) + rI'(r) \right] \quad , \end{aligned} \tag{3.7b}$$

since $S'(r) = J(r)/r$ and $R'(r) = rI(r)$ by (2.19b). Equations (3.7a-b) must be evaluated at $r = 0$. Now

$$\begin{aligned} \lim_{r\to 0} \frac{J(r)}{r^2} &= \lim_{r\to 0} \frac{1}{r^2} \int_0^r ds\, s\left[ \rho_{tt}(s) + \rho_{ll}(s) \right] = \lim_{r\to 0} \frac{1}{r^2} \int_0^r ds\, s\left( 1 + \tfrac{1}{2}\rho''_{tt} s^2 + 1 + \tfrac{1}{2}\rho''_{ll} s^2 + O(s^4) \right) = \\ &= \lim_{r\to 0} \frac{1}{r^2}\left[ s^2 + \tfrac{1}{8}\left( \rho''_{tt} + \rho''_{ll} \right) s^4 \right]_{s=0}^{s=r} = 1 \quad , \end{aligned} \tag{3.8a}$$

$$\lim_{r\to 0} \frac{J'(r)}{r} = \lim_{r\to 0} \frac{1}{r} r\left[ \rho_{tt}(r) + \rho_{ll}(r) \right] = 2 \quad , \tag{3.8b}$$

$$\lim_{r\to 0} I(r) = I(0) = \int_0^\infty ds \left[ \frac{\rho_{tt}(s) - \rho_{ll}(s)}{s} \right] \quad , \tag{3.8c}$$

$$\lim_{r\to 0} rI'(r) = -\lim_{r\to 0} r\left[ \frac{\rho_{tt}(r) - \rho_{ll}(r)}{r} \right] = 0 \quad , \tag{3.8d}$$

where we used that $\rho_{tt}$ and $\rho_{ll}$ are even functions. Substituting (3.8) into (3.7) yields

$$\rho''_{\psi\psi}(0) = \frac{1}{2a_\psi}\left[ -1 + 2 - I(0) \right] = \frac{1 - I(0)}{2a_\psi} \quad , \tag{3.9a}$$

$$\rho''_{\chi\chi}(0) = \frac{1}{2a_\chi}\left[ -1 + 2 + I(0) \right] = \frac{1 + I(0)}{2a_\chi} \quad , \tag{3.9b}$$

and so

$$L_\psi^2 = -\frac{2a_\psi}{1-I(0)} \quad , \quad L_\chi^2 = -\frac{2a_\chi}{1+I(0)} \tag{3.10}$$

Both expressions in (3.2) lead with (3.10) to the same expression for $\nu^2$

$$\nu^2 = \tfrac{1}{2}\left[1+I(0)\right] \quad . \tag{3.11}$$

Equations (3.10) and (3.11) together with (3.1) fix all parameters in the error correlation functions.

### 3.1.2 Frequency domain

In the spatial frequency domain the parameters $L_\psi$ and $L_\chi$ can also be obtained from (3.3). Using (2.20b) one readily finds

$$\rho_{\psi\psi}(r)\Big|_{r=0} = \int_{-\infty}^{\infty} dk\, e^{-2\pi i k r}\, \hat{\rho}_{\psi\psi}(k)\Big|_{r=0} = \int_{-\infty}^{\infty} dk\, \hat{\rho}_{\psi\psi}(k) = 2\int_0^{\infty} dk\, \hat{\rho}_{\psi\psi}(k) \quad , \tag{3.12}$$

where use has been made of the fact that $\hat{\rho}_{\psi\psi}$ is an even function. Using (3.6) one finds

$$\begin{aligned} \nabla^2 \rho_{\psi\psi}(r)\Big|_{r=0} &= \frac{d^2}{dr^2} \int_{-\infty}^{\infty} dk\, e^{-2\pi i k r}\, \hat{\rho}_{\psi\psi}(k)\Big|_{r=0} = -(2\pi)^2 \int_{-\infty}^{\infty} dk\, k^2 \hat{\rho}_{\psi\psi}(k) = \\ &= -2(2\pi)^2 \int_0^{\infty} dk\, k^2 \hat{\rho}_{\psi\psi}(k) \quad . \end{aligned} \tag{3.13}$$

Note that similar relations hold for $\rho_{\chi\chi}$ . Substituting (3.12) and (3.13) into (3.3) yields

$$\begin{aligned} L_\psi^2 &= \frac{\rho_{\psi\psi}(r)}{\nabla^2 \rho_{\psi\psi}(r)}\Big|_{r=0} = -\frac{\int_0^{\infty} dk\, \hat{\rho}_{\psi\psi}(k)}{(2\pi)^2 \int_0^{\infty} dk\, k^2 \hat{\rho}_{\psi\psi}(k)} \quad , \\ L_\chi^2 &= \frac{\rho_{\chi\chi}(r)}{\nabla^2 \rho_{\chi\chi}(r)}\Big|_{r=0} = -\frac{\int_0^{\infty} dk\, \hat{\rho}_{\chi\chi}(k)}{(2\pi)^2 \int_0^{\infty} dk\, k^2 \hat{\rho}_{\chi\chi}(k)} \quad . \end{aligned} \tag{3.14}$$

Writing (2.47a) as

$$a_\psi \hat{\rho}_{\psi\psi}(k) = V(k) \quad , \quad a_\chi \hat{\rho}_{\chi\chi}(k) = W(k) \quad , \tag{3.15a}$$

where

$$V(k)=\frac{1}{(2\pi)^2k^2}\left[-\hat{\rho}_{tt}(k)+\frac{P(k)}{2k}+\frac{k}{2}Q(k)\right] \quad ,$$
$$W(k)=\frac{1}{(2\pi)^2k^2}\left[-\hat{\rho}_{ll}(k)-\frac{P(k)}{2k}+\frac{k}{2}Q(k)\right] \quad , \tag{3.15b}$$

simplifies (3.14) to

$$L_\psi^2=-\frac{\int_0^\infty dk\,V(k)}{(2\pi)^2\int_0^\infty dk\,k^2V(k)} \quad , \quad L_\chi^2=-\frac{\int_0^\infty dk\,W(k)}{(2\pi)^2\int_0^\infty dk\,k^2W(k)} \quad . \tag{3.16}$$

Note that $V(k)$ and $W(k)$ can be calculated directly from the spectra. The parameters $a_\psi$ and $a_\chi$ cancel out in (3.16) but they can be determined from the boundary conditions $\rho_{\psi\psi}(0)=\rho_{\chi\chi}(0)=1$. Their Fourier transforms read, c.f. (3.12)

$$\int_0^\infty dk\,\hat{\rho}_{\psi\psi}(k)=\frac{1}{2} \quad , \quad \int_0^\infty dk\,\hat{\rho}_{\chi\chi}(k)=\frac{1}{2} \quad . \tag{3.17}$$

Substituting (3.15a) in (3.17) yields

$$a_\psi=2\int_0^\infty dk\,V(k) \quad , \quad a_\chi=2\int_0^\infty dk\,W(k) \quad . \tag{3.18}$$

From (3.16) and (3.18) one can extract the following two equations for $\nu^2$

$$\nu^2=1+\frac{a_\psi}{L_\psi^2} \quad , \quad \nu^2=-\frac{a_\chi}{L_\chi^2} \quad . \tag{3.19}$$

In most cases the two values for $\nu^2$ will not be the same because $V(0)$ and $W(0)$ are determined from the numerical second derivative of the spectra according to (2.49). However, the difference between the two values is small, so their average comprises a good estimate.

## 3.2 Numerical implementation

The autocorrelations $\rho_{ll}$ and $\rho_{tt}$ can be obtained on the same grid size $\Delta$ as the scatterometer and background winds are given on: $\Delta=25$ km (SeaWinds-KNMI and ASCAT-25) or $\Delta=12.5$ km (ASCAT-12.5). The maximum correlation distance, $N\Delta$, must be large enough to let the autocorrelations approach zero. The integrals $I, J, R,$ and $S$ are calculated on the scatterometer grid using the trapezium rule. In paragraph 3.3 it will be shown that this yields enough accuracy.

The spectra $\hat{\rho}_{ll}$ and $\hat{\rho}_{tt}$ are obtained on a grid with size $\Delta k = (N\Delta)^{-1}$ with $\Delta$ the spatial grid size (25 km or 12.5 km) and $N$ the number of points. The maximum spatial frequency is $k_{\max} = (2\Delta)^{-1}$. The integrals $P$, $Q$, $V$, and $W$, as well as the integrals over $V$ and $W$ are calculated with the trapezium rule. For small values of $k$ the integral $Q$ is split into a regular and a singular part as

$$\begin{aligned} Q(k) &= \int_k^\infty dp \frac{\hat{\rho}_{ll}(p) + \hat{\rho}_{tt}(p)}{p^2} = \int_k^\infty dp \frac{\hat{\rho}_{ll}(p) - \hat{\rho}_{ll}(0) + \hat{\rho}_{tt}(p) - \hat{\rho}_{tt}(0)}{p^2} + \left(\hat{\rho}_{ll}(0) + \hat{\rho}_{tt}(0)\right)\int_k^\infty \frac{dp}{p^2} = \\ &= \int_k^\infty dp \frac{\hat{\rho}_{ll}(p) - \hat{\rho}_{ll}(0) + \hat{\rho}_{tt}(p) - \hat{\rho}_{tt}(0)}{p^2} - \frac{\hat{\rho}_{ll}(0) + \hat{\rho}_{tt}(0)}{k} \quad . \end{aligned} \tag{3.20}$$

The values $V(0)$ and $W(0)$ are obtained from (2.49a-b) using the 4th order formula of appendix D. The Gaussian test described in the next section shows that the accuracy in the values of $V$ and $W$ decreases for small values of $k$, but that the values of $V(0)$ and $W(0)$ are quite accurate. This suggests to calculate the value of $V(\Delta k)$ from 4th order interpolation between $V(0)$ and $V(2\Delta k)$, and similar for $W(\Delta k)$.

## 3.3 Gaussian test

### 3.3.1 Spatial domain

The solution scheme outlined in the previous paragraph was tested using Gaussian error correlation functions

$$\rho_{\psi\psi}(r) = e^{-\frac{r^2}{R_\psi^2}} \quad , \quad \rho_{\chi\chi}(r) = e^{-\frac{r^2}{R_\chi^2}} \quad . \tag{3.21}$$

From (2.1a-b) one readily finds

$$\rho_{ll}(r) = (1-\nu^2) e^{-\frac{r^2}{R_\psi^2}} + \nu^2 \left(1 - 2\frac{r^2}{R_\chi^2}\right) e^{-\frac{r^2}{R_\chi^2}} \quad , \tag{3.22a}$$

$$\rho_{tt}(r) = (1-\nu^2)\left(1 - 2\frac{r^2}{R_\psi^2}\right) e^{-\frac{r^2}{R_\psi^2}} + \nu^2 e^{-\frac{r^2}{R_\chi^2}} \quad , \tag{3.22b}$$

where we have used that for a Gaussian function $L_\psi^2 = \frac{1}{2} R_\psi^2$ and $L_\chi^2 = \frac{1}{2} R_\chi^2$.

In Appendix E it is shown that (2.17b) and (3.22a-b) yield

$$I(0) = 2\nu^2 - 1 \quad , \tag{3.23a}$$

$$J(\infty) = 0 \quad . \tag{3.23b}$$

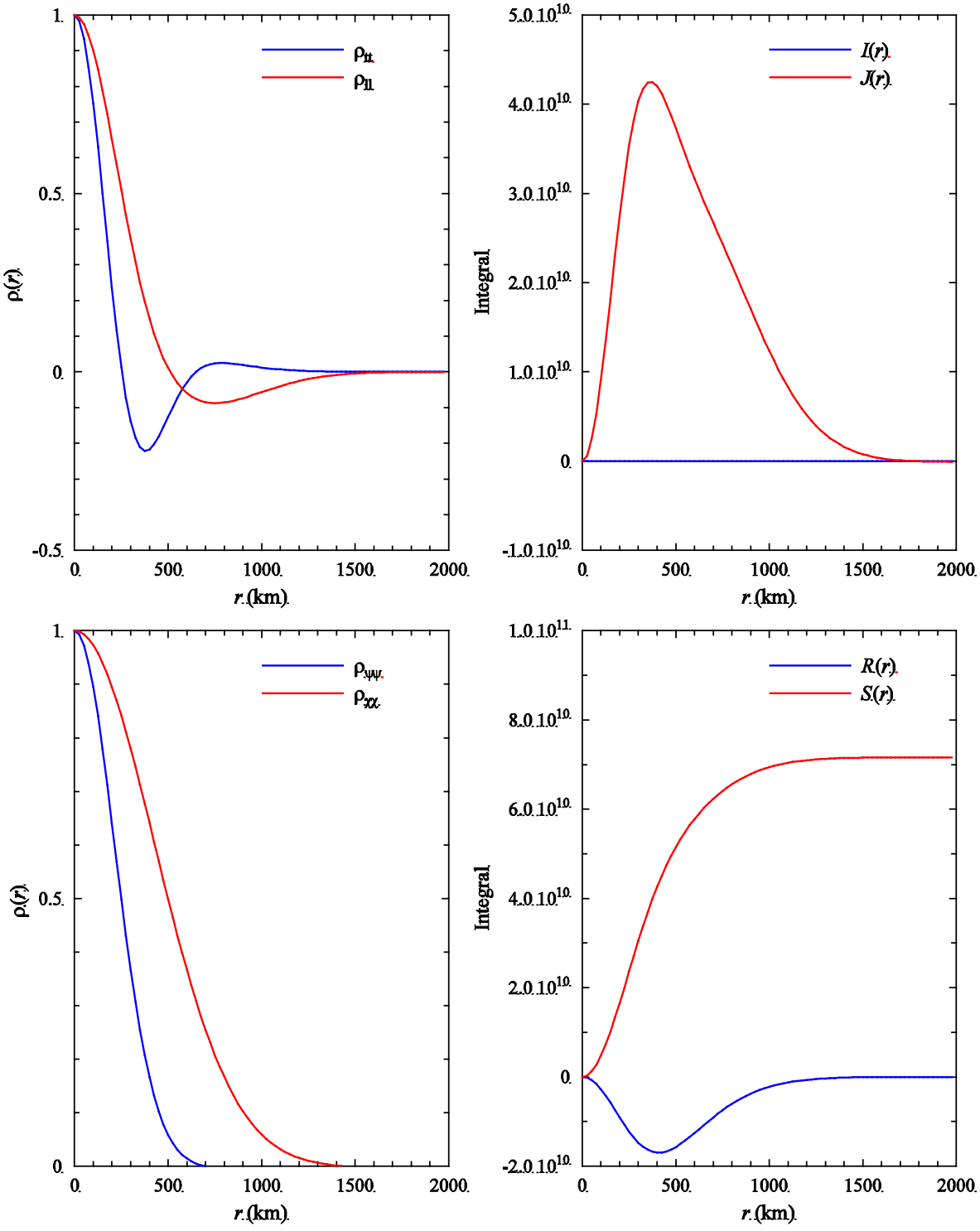


**Figure 3.1** Results for the Gaussian test.

Equations (3.22a-b) were calculated on a 25 km grid of 512 points and on a 12.5 km grid of 1024 points, both with $R_\psi = 300\,\text{km}$, $R_\chi = 600\,\text{km}$ and $\nu^2 = 0.2$. These values were then used as input for calculating back the error correlation functions. The results for the 25 km grid are shown in figure 3.1.

The upper left panel of figure 3.1 shows the input autocorrelations. The upper right panel shows the integrals $I$ and $J$. Note that the value of $J$ is much larger than that of $I$, and that both integrals approach zero for large $r$. The lower right panel shows the integrals $R$ and $S$. Now the two integrals are of the same order of magnitude and they become constant for large $r$. The lower left panel shows the resulting error correlation functions, reproduced with an accuracy of three decimals or better which is sufficient for our purposes.

The error correlation function parameters retrieved for both grids are listed in table 3.1. Also their values are retrieved with sufficient accuracy. Note that the values for $J(\infty)$ look very high and different from zero, but compared to the maximum value of 4.2 $10^{10}$ the 25 km grid result is close enough to zero (five decimals) while the 12.5 km grid result is still accurate to three decimal places.

| Parameter | Input value | Retrieved value | |
|---|---|---|---|
| | | 25 km | 12.5 km |
| $L_\psi$ | 212.13 km | 210.77 km | 211.92 |
| $L_\chi$ | 424.26 km | 423.47 km | 423.70 |
| $\nu^2$ | 0.2 | 0.20044 | 0.20011 |
| $I(0)$ | -0.6 | -0.59913 | -0.59978 |
| $J(\infty)$ | 0 | $-0.25152\ 10^5$ | $-0.26047\ 10^8$ |

**Table 3.1** Input values and retrieved values of the error correlation function parameters in the Gaussian test. The input values for *I* and *J* are the analytical values.

### 3.3.2 Frequency domain

Equations (3.21) and (3.22a-b) are easily Fourier transformed, resulting in

$$\hat{\rho}_{\psi\psi}(k) = R_\psi \sqrt{\pi} e^{-\pi^2 R_\psi^2 k^2} \quad , \tag{3.24a}$$

$$\hat{\rho}_{\chi\chi}(k) = R_\chi \sqrt{\pi} e^{-\pi^2 R_\chi^2 k^2} \quad , \tag{3.24b}$$

for the Fourier transform of the error correlation functions, and

$$\hat{\rho}_{ll}(k) = \sqrt{\pi}\left[ (1-\nu^2) R_\psi e^{-\pi^2 R_\psi^2 k^2} + 2\nu^2 R_\chi \pi^2 R_\chi^2 k^2 e^{-\pi^2 R_\chi^2 k^2} \right] \quad , \tag{3.25a}$$

$$\hat{\rho}_{tt}(k) = \sqrt{\pi}\left[ 2(1-\nu^2) R_\psi \, \pi^2 R_\psi^2 k^2 e^{-\pi^2 R_\psi^2 k^2} + \nu^2 R_\chi \, e^{-\pi^2 R_\chi^2 k^2} \right] , \qquad (3.25b)$$

for the spectra.

Figure 3.2 shows the result for the Gaussian test in the frequency domain using the same error correlation function parameters as listed in table 3.1 and a 12.5 km grid in the spatial domain.

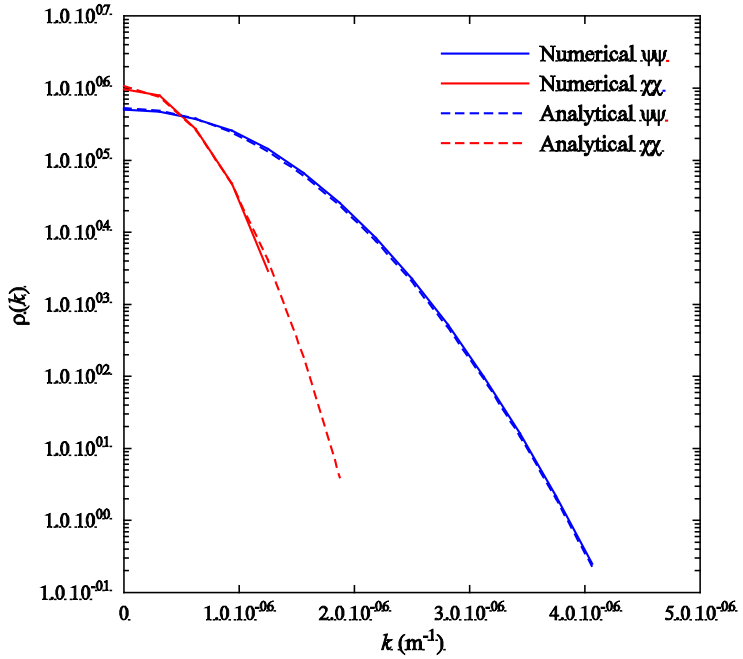


**Figure 3.2** Numerical and analytical error correlation functions in the frequency domain using a 12.5 km grid.

As can be seen from figure 3.2, the numerical and analytical error correlation functions are close to each other over a large range of spatial frequencies. Note, however, that the vertical scale in figure 3.2 covers eight orders of magnitude. The relative error in the numerical error correlation function is 10% at most. Note that this is less accurate than in the spatial domain.

The lower accuracy is also reflected in the comparison of input and retrieved values of the error correlation function parameters shown in table 3.2.

| Parameter | Input value | Retrieved value |
|---|---|---|
| $L_\psi$ | 212.13 km | 206.30 |
| $L_\chi$ | 424.26 km | 423.17 |
| $\nu^2$ | 0.2 | 0.21 |

**Table 3.2** Input values and retrieved values of the error correlation function parameters for the Gaussian test in the frequency domain.

### 3.3.3 Transformation between the domains

In order to be able to compare error correlation functions calculated in both the spatial and the frequency domain, program `A2S` in directory `genscat/tools/structure_functions` contains the possibility to transform error correlation functions from one domain to another. Figures 3.3 and 3.4 show the results for the Gaussian tests using a 12.5 km spatial grid and the error correlation function parameters of table 3.2.

Figure 3.3 shows the results for the Gaussian test in the frequency domain transformed to the spatial domain. As can be seen from figure 3.3, the reduced accuracy in the frequency domain remains visible when transformed to the spatial domain. In particular $\rho_{\chi\chi}$ deviates from its analytical value for large distances, going to a value of -0.0065. Note that the agreement between numerical and analytical results could be improved by linear scaling of the numerical error correlation functions to the interval [0.1].

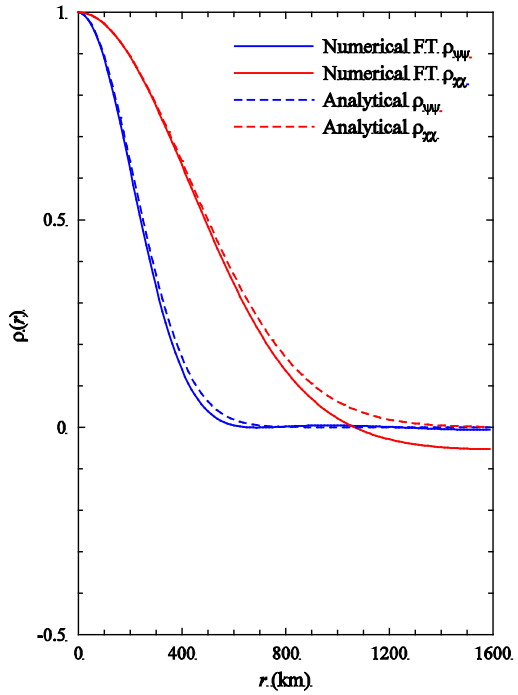


**Figure 3.3** Results for the Gaussian test in the frequency domain transformed to the spatial domain.

Figure 3.4 shows the results for the Gaussian test in the spatial domain transformed to the frequency domain. For large spatial frequencies the Fourier transformed error correlation functions reach a noise level determined by the numerical FFT operation. The noise level is about five orders of magnitude below the maximum value at zero spatial frequency, which is acceptable.

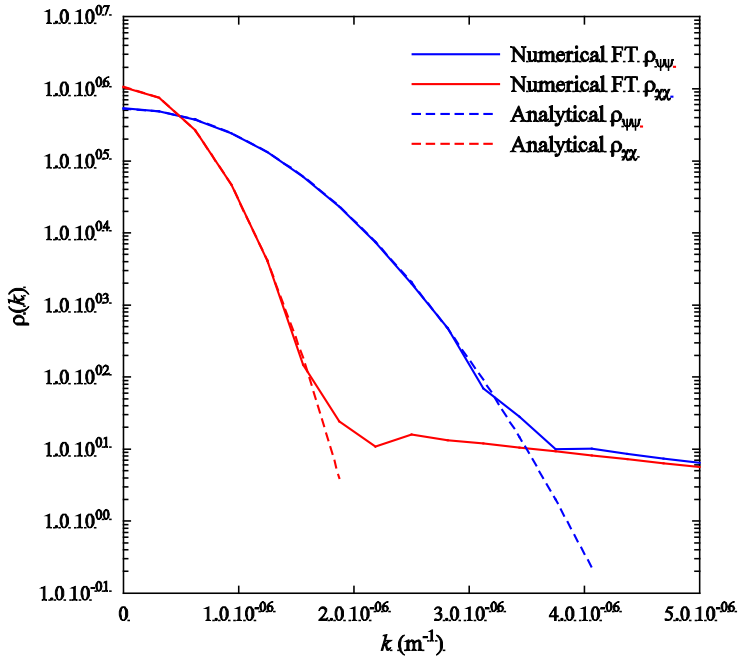


Figure 3.4 Results for the Gaussian test in the spatial domain transformed to the frequency domain.

# 4 Results

In this chapter the algorithm for obtaining error correlation functions from scatterometer and background spectra will be applied to real scatterometer data. Since the error correlation functions are the background error correlations in the potential domain, the choice of scatterometer is not important unless it introduces sampling artefacts. It is clear that from a numerical point of view that a small grid size leads to more accurate integrations in the spatial domain, whereas a large sample length leads to a smaller grid size in the spatial frequency domain and, hence, more accurate integrals there.

The calculation of error correlation functions starts from observed O-B correlations in the spatial domain or O-B spectra in the spatial frequency domain. First the effect of the observations on the O_B correlations will be investigated. Then the effect of sampling strategy, sampling grid size, and sampling length will be studied.

All autocorrelations and spectra were calculated by program `WSC` in `genscat/tools/sac`, using all data available in January, 2009.

## *4.1 Effect of the observations on O-B statistics*

The basic assumptions which enables calculation of error correlation functions from O-B correlations is that the error correlations in the observations are uncorrelated and that observations and background are well calibrated relative to each other. In that case, both observation and background yield some true value $t$ with error $\varepsilon_o$ and $\varepsilon_b$ , respectively. Then O-B equals $\varepsilon_o + \varepsilon_b$. If $\varepsilon_o$ is uncorrelated, the autocorrelation of O-B, $\rho_{o-b}$, is equal to the autocorrelation of B, except for $r = 0$ where

$$\rho_{o-b}(0) = \sigma_o^2 + \sigma_b^2 \quad , \tag{4.1}$$

where $\sigma_o$ and $\sigma_b$ stand for the error standard deviation in observation and background, respectively. As a result, $\rho_{o-b}$ plotted as function of $r$ will show a discontinuity at $r = 0$. The size of this discontinuity can be estimated by extrapolating the autocorrelation to zero and is directly related to the observation error standard deviation. The extrapolated autocorrelation then represents the background error autocorrelation.

Errors in scatterometer winds are in general correlated due to the procedure to for calculating radar cross section multiplets: basic radar cross sections are averaged over an area representing a WVC. This is done to average out speckle noise that is inherent in radar measurements. The area representing a WVC is often larger than the WVC. For the ASCAT-25 and ASCAT-12.5 products the averaging area is a square centered on the WVC with a linear size four times that of the WVC, while the averaging is done with a

Hamming window. For the ASCAT coastal product (WVC size 12.5 km) the averaging area is a circle with a radius of 15 km.

The size of the averaging radius can be varied in AWDP in order to study its effect. Figure 4.1 shows the results for the autocovariance (unnormalized autocorrelation) of the transversal wind component $t$ (left hand panel) and the longitudal wind component $l$ (right hand panel). Results are shown for averaging radii $R$ = 25 km (blue curves), $R$ = 12.5 km (green curves), and $R$ = 6.25 km (red curves).

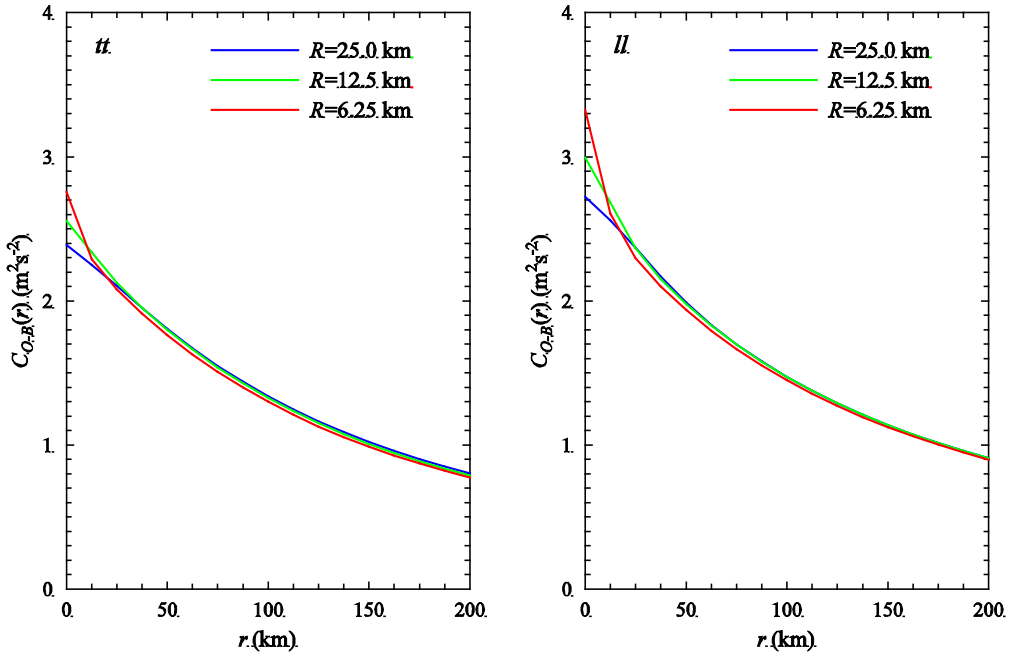


**Figure 4.1** Autocovariances of the ASCAT coastal product for various values of the averaging radius.

For $R$ equal to 6.25 km there is no overlap between the averaging area's. This does not mean that the radar cross sections on the WVC level are uncorrelated, because the underlying basic cross sections cover an area larger than the WVC. Nevertheless, the covariances in figure 4 show a discontinuity caused by the observation error. Note that the observation error only has a small effect, despite the fact that the averaging area is small, so the radar cross section in a WVC is rather noisy. The effect disappears when $R$ is increased and saturates for $R$ around 25 km: the observation error discontinuity becomes invisible and seems to be smeared out over the smaller scales, making it impossible to correct for the observation error by simple extrapolation. However, the difference between the various curves is small, so it seems justified to neglect the contribution of the observation error to the O-B statistics.

## 4.2 Error correlation functions in the spatial domain

### 4.2.1 Sampling strategy

There are two ways for calculating the autocorrelation:

- From samples. In this approach samples of a fixed length are collected for each WVC number. Isolated missing or flagged points may be interpolated. When a valid sample is found, i.e., a sample without missing values, the autocorrelation is calculated and averaged with that of earlier samples.
- From all winds. In this approach the complete dataset is used for updating the autocorrelation per WVC. Missing or flagged WVC's do not contribute.

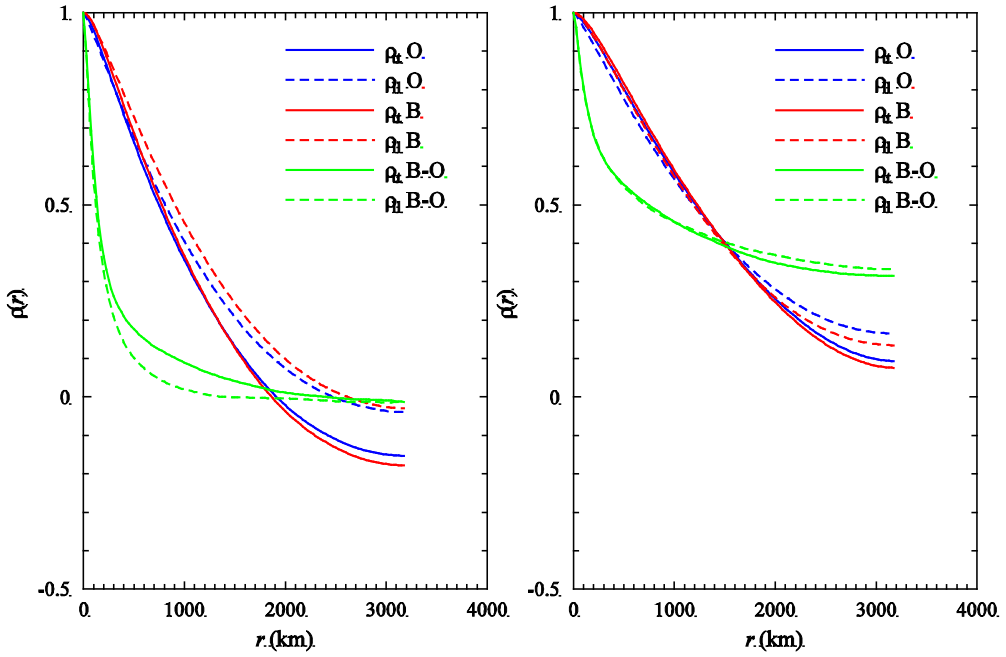


**Figure 4.2** Spatial autocorrelations for ASCAT-25 obtained with the all winds method (left hand panel) and the sample method (right hand panel).

The first method is best suited for calculating spectra when combined with some detrending method to get rid of wind variations larger than the sample size [*Vogelzang*, 2010]. Figure 4.2 shows the results of both methods (all winds: left panel; samples: right panel) for the spatial autocorrelation of the ASCAT-25 observations (O), the ECMWF forecast (B), and their difference (B-O).

ASCAT-25 is used as an example here, but the remarks in this subsection pertain also to other scatterometer wind products. Figure 4.2 shows that the all winds method yields more realistic

autocorrelations that go to zero for large $r$ more or less like the Gaussian test results in figure 3.1. The sample method yields autocorrelations that approach a finite value, notably those for B-O (green curves). This is because the samples are considered periodic, so the last point of a sample is a direct neighbor of the first point. To prevent large differences, the large scale information of the sample is filtered out by detrending [*Vogelzang*, 2010]. Therefore the all winds method will be used in the remainder of this document.

Figure 4.2 also shows that the autocorrelations for B-O in the all winds method (green curves in the left hand panel) seem to contain less rotation that those in the Gaussian test. This is visible in the curves for $\rho_{tt}$ (solid green curve in figure 4.2; solid blue curve in figure 3.1) which in figure 4.2 exhibit no clear negative minimum value. Apparently, the amounts of rotation and divergence in the observed B-O data are about equal, with perhaps slightly more divergence, see also figure 5.2 in *Daley* [1993].

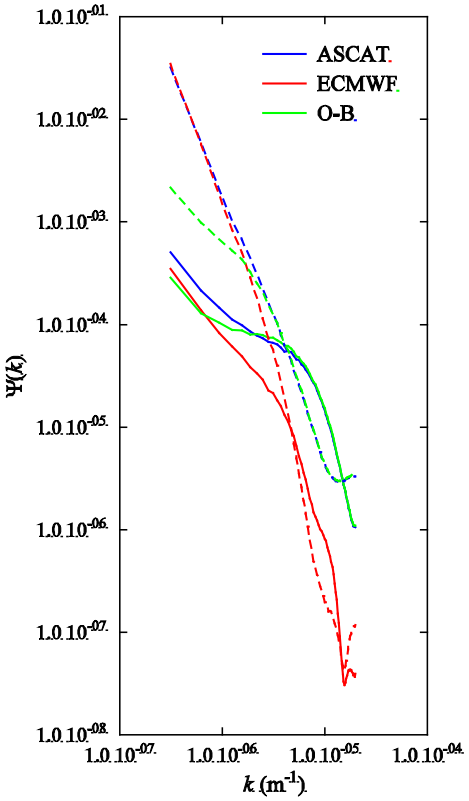


**Figure 4.3** Spectra for divergence (solid curves) and rotation (dashed curves) for O, B, and O-B.

This is corroborated by figure 4.3 which shows the divergence and rotation spectra of O, B, and O-B. The O-B divergence spectrum (solid green curve) lies slightly above the O-B rotation spectrum (dashed green

curve) for small scales, implying that the autocorrelation should be more divergence-like. The divergence and rotation were calculated for each WVC using second order finite differences, see also appendix F.

Finally it should be noted that the autocorrelations in figure 4.2 do not exactly go to zero for large $r$, because the scatterometer samples the ocean surface from a fixed orbit, crossing stationary wind structures like the trades each orbit. This point will be addressed in the next section.

### 4.2.2 Cut-off functions

The accuracy of the numerical integrations in the spatial domain depends on the grid size, so it is logical to use ASCAT-12.5 O-B data. Figure 4.4 shows the autocorrelations for all ASCAT-12.5 data from January 2009 together with the error correlation functions derived from them as well as the intermediate integrals. The FFT size was 1024, so the autocorrelations cover a range of 6400 km of which the last 400 km is not shown in figure 4.4. The error correlation functions do not show a smooth approach to zero for large $r$. $\rho_{\chi\chi}$ even crosses zero for $r \approx 2000\,\mathrm{km}$. Comparison with figure 3.1 shows that this is due to the fact that $J(r)$ does not go to zero in order to fulfill (3.14b), but becomes negative. Such behavior can be forced by multiplying the observed autocorrelations with a suitable cut-off function.

There are many possible forms for such a cut-off function: any form that starts at one for $r = 0$ and goes to zero for $r \to \infty$ will do. In this work two forms will be considered: a cosine cut-off

$$C_c(r) = \begin{cases} 1 & , \quad r < a \\ \frac{1}{2} + \frac{1}{2}\cos\left(\dfrac{\pi r}{b-a}\right) & , \quad a \le r \le b \\ 0 & , \quad r > b \end{cases} \quad , \tag{4.2}$$

and a brick-wall cut-off

$$C_b(r) = \begin{cases} 1 & , \quad r < a \\ 0 & , \quad r \ge a \end{cases} \quad , \tag{4.3}$$

where $a$ and $b$ are adjustable parameters. For the brick-wall cut-off there is only one value for $a$ that has the desired effect, for the cosine cut-off there is a range of $(a,b)$ values.

Table 4.1 lists the choices made for the cut-off functions and some parameters of the resulting error correlation functions.

| **Name** | $a$ **(km)** | $b$ **(km)** | $L_\psi$ **(km)** | $L_\chi$ **(km)** | $\nu^2$ |
|---|---|---|---|---|---|
| Cos-3000 | 3000 | 6282 | 673 | 564 | 0.632 |
| Cos-4000 | 4000 | 5375 | 660 | 562 | 0.632 |
| Brick-wall | 4700 | -- | 658 | 562 | 0.632 |

**Table 4.1** Parameters for the cut-off functions and the error correlation functions retrieved from them.

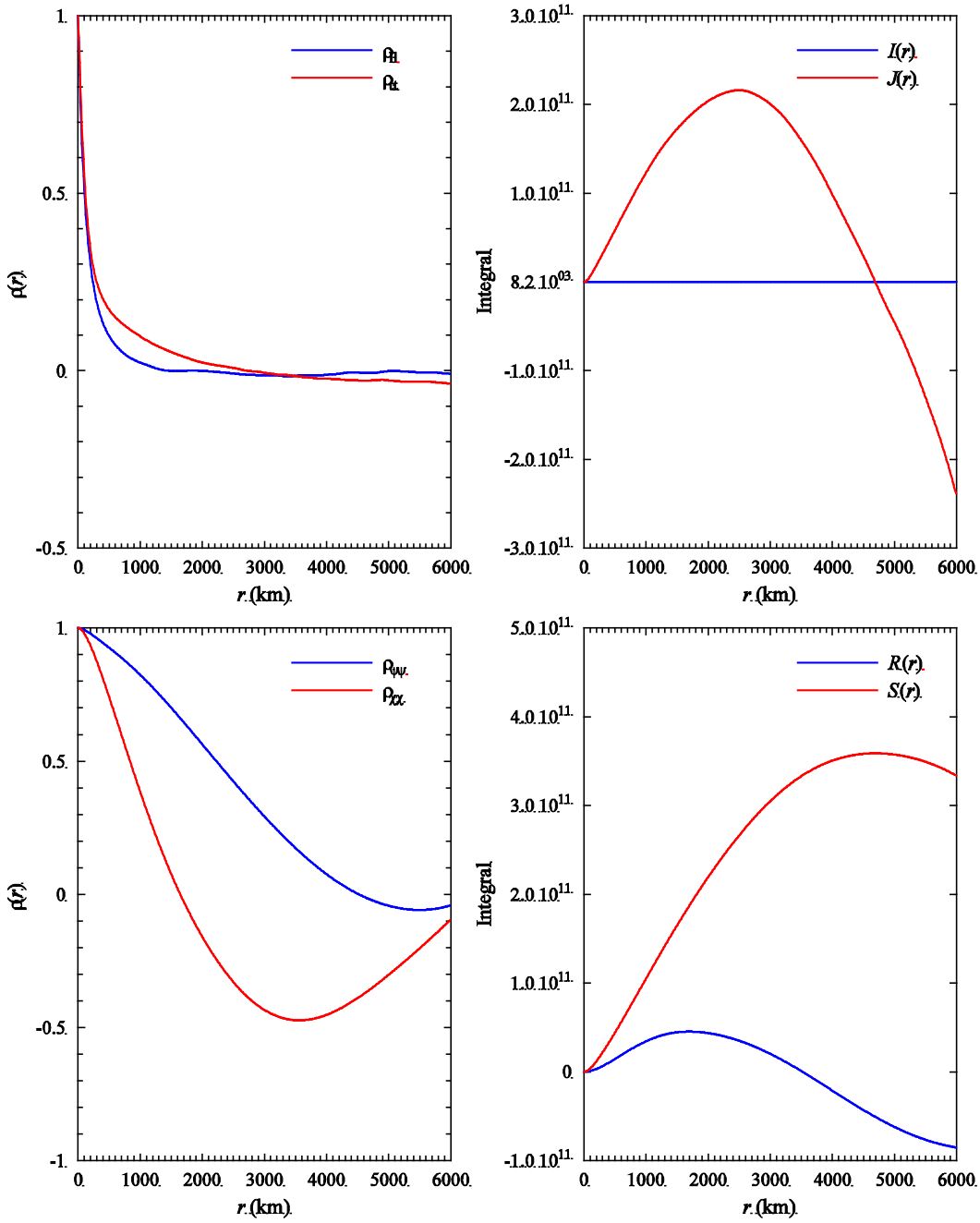


**Figure 4.4** Autocorrelations (upper left panel), error correlation functions (lower left panel) and intermediate results (right hand panels) for ASCAT-12.5 B-O.

The transition region for the cosine cut-off lies between $a$ and $b$. From table 4.1 it can be inferred that the Cos-3000 function has the softest cut-off. That of the Cos-4000 function is slightly harder, while the Brick-wall has, of course, the hardest cut-off. The values for the length scales $L_\psi$ and $L_\chi$ show little variation for the various cut-off functions. Note that $L_\psi$ is a lot larger than $L_\chi$ and that the estimates for $\nu^2$ are all the same.

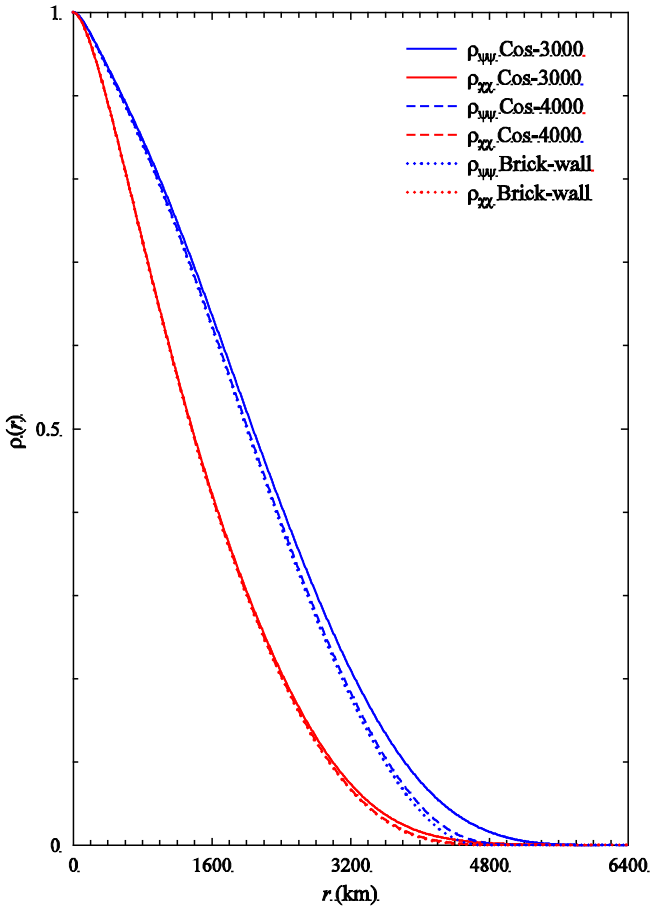


**Figure 4.5** Numerical error correlation functions for ASCAT-12.5 with various cut-off functions.

Figure 4.5 shows the error correlation functions obtained with the cut-off functions from table 4.1. The error correlation functions now approach zero smoothly as $r \to \infty$. Moreover, the choice of cut-off has little effect on the error correlation function.

### 4.2.3 Zonal effects

The results in figure 4.5 were obtained with autocorrelations obtained from over the whole globe. One expects that zonal effects play a role, and that error correlation functions for the tropics have different properties than those for the extratropics. In 2DVAR the tropics are defined as the region with latitude between 20°N and 20°S. With this choice the tropics part of the swath has a length less than 6400 km, so the autocorrelations obtained for this region extend not far enough for calculating error correlation functions with proper asymptotic properties as those shown in figure 4.5. However, this is feasible for the region with latitude between 30°N and 30°S, referred to as the extended tropics. The results are shown in figure 4.6.

Figure 4.6 shows the error correlation functions obtained for all Earth (upper left panel), the extratropics (upper right panel), the tropics (lower left panel), and the extended tropics (lower right panel). The sample size of 1024 appears to be too small for the extratropics: the integral $J(r)$ does not cross zero for $r < 6400$ km, so no cut-off can be defined. Therefore only results without cut-off are shown, as for the tropics, but these error correlation functions do not have the required asymptotic behaviour. The cut-off parameters for the extratropics had to be determined again, but the minimum cut-off scale for the cosine cut-off functions were taken the same. The cut-off parameters are listed in table 4.2. The results for all Earth, upper left panel of figure 4.6, are identical to those in figure 4.5.

| **Name** | $a$ **(km)** | $b$ **(km)** | $L_\psi$ **(km)** | $L_\chi$ **(km)** | $\nu^2$ |
|---|---|---|---|---|---|
| Cos-3000 | 3000 | 5414 | 555 | 470 | 0.659 |
| Cos-4000 | 4000 | 4625 | 544 | 470 | 0.659 |
| Brick-wall | 4325 | -- | 543 | 471 | 0.659 |

**Table 4.2** Parameters for the cut-off functions and the error correlation functions retrieved from them for the extended tropics.

Comparison of the error correlation functions for all Earth and for the extended tropics shows that the error correlation functions for the extended tropics are narrower and contain more divergence than those for the whole Earth. This agrees with the cut-off and error correlation function parameters in table 4.2.

As a reference, the solid green curves show the default Gaussian error correlation function in the tropics ($R_\psi = R_\chi = 600$ km) and the extratropics ($R_\psi = R_\chi = 300$ km). The dotted green curves show the TOAR error correlation functions for the same parameters. The Gaussian error correlation functions are much narrower than the numerical error correlation functions. The TOAR error correlation functions are wider, and the one for the tropics has a range comparable with that of the numerical error correlation functions.

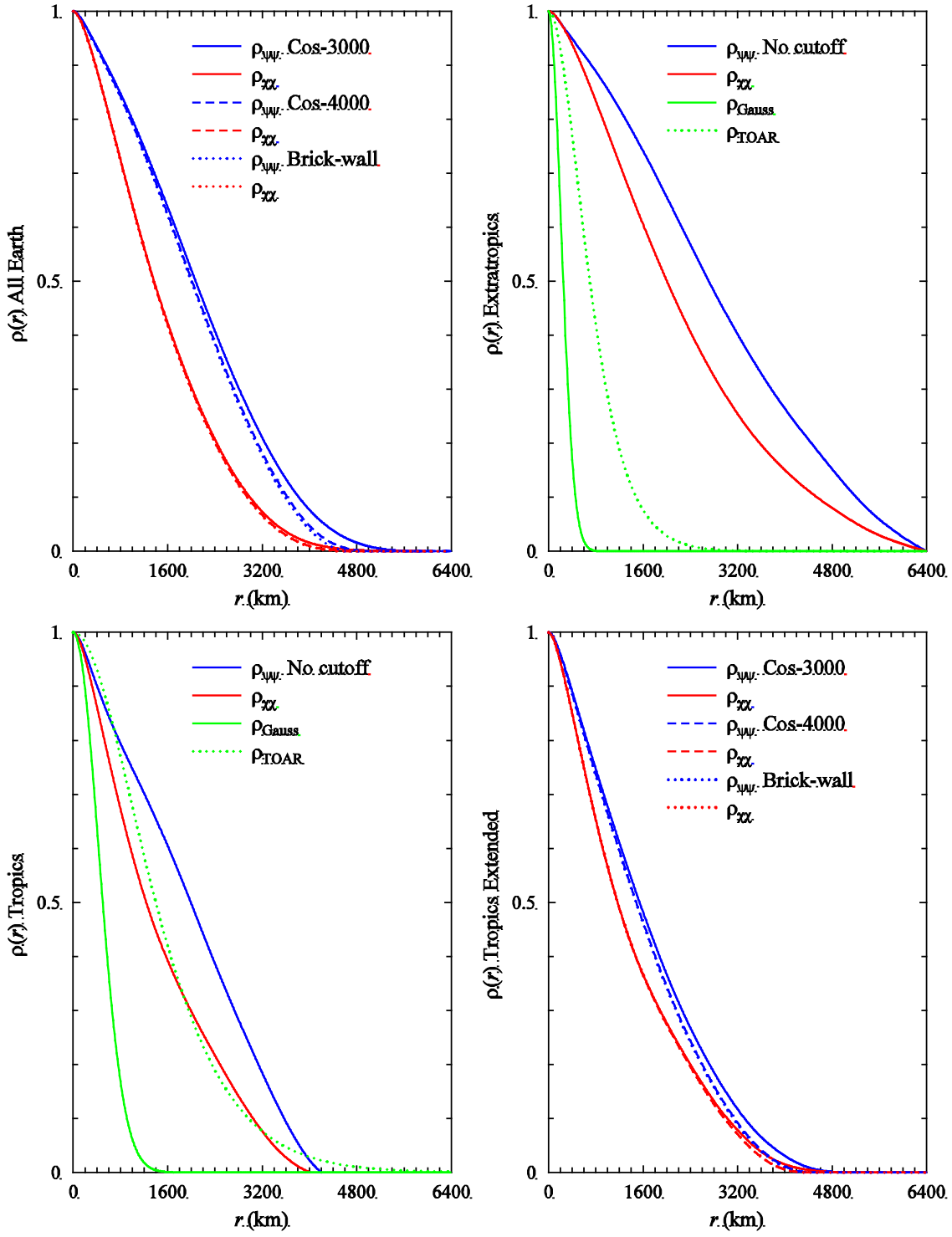


**Figure 4.6** Error correlation functions for ASCAT-12.5 in various regions.

### 4.2.4 Effect of wind calibration

When using O-B statistics to retrieve error correlation functions, one assumes that observations and background are well calibrated, so that their difference contains only the errors. In practice, this is not precisely the case. To first order (linear calibration) one has

$$x = \alpha t + \beta \quad , \tag{4.4}$$

where $x$ stands for the measured wind component, $x = \{u, v, t, l\}$ and $t$ for the calibrated or true wind. The calibration trend and bias coefficients are given by $\alpha$ and $\beta$, respectively. So far, all calculations were performed with $x$ rather than $t$.

*Stoffelen* [1998] has introduced the triple collocation method for simultaneous calculation of the error variances and linear calibration coefficients from collocated triplets (scatterometer wind, buoy wind, and ECMWF background field). *Vogelzang et al.* [2011] applied the triple collocation technique to the operational ASCAT-12.5, ASCAT-25, SeaWinds-KNMI, and Seawinds-NOAA products. Since the error correlation functions describe the background error covariance, they are a property of the NWP model. The calibration coefficients for ASCAT-12.5 and ECMWF with respect to buoy measurements are given in table 4.3.

| **Wind** | **ASCAT-12.5** | | **ECMWF** | |
|---|---|---|---|---|
| **component** | $\alpha$ | $\beta$ | $\alpha$ | $\beta$ |
| $u$ | 1.012 | 0.19 | 1.032 | 0.32 |
| $v$ | 1.008 | -0.01 | 1.043 | 0.09 |

**Table 4.3** Triple collocation calibration coefficients.

Table 4.3 shows that the calibration biases and trends are small. Figure 4.7 shows the effect of calibration on the spatial autocorrelations by calibrating the wind components using the inverse of (4.4) with the values from table 4.3. The autocorrelations of uncalibrated wind components are given by the solid curves, those of the calibrated ones by the dashed curves. results are given for all Earth (bottom panel), tropics (middle panel) and extratropics (top panel)

Figure 4.3 shows that calibration has a significant effect on the autocorrelations: for all zones there is a decrease for distances between 400 km and 2400 km. In the tropics there is an increase in $\rho_{tt}$ for scales larger than about 300 km. Note that the autocorrelations for the tropics go to zero abruptly at distances of more than 4600 km due to the finite size of this zone, as mentioned before. Therefore the increase at large distances could be caused by poor sampling.

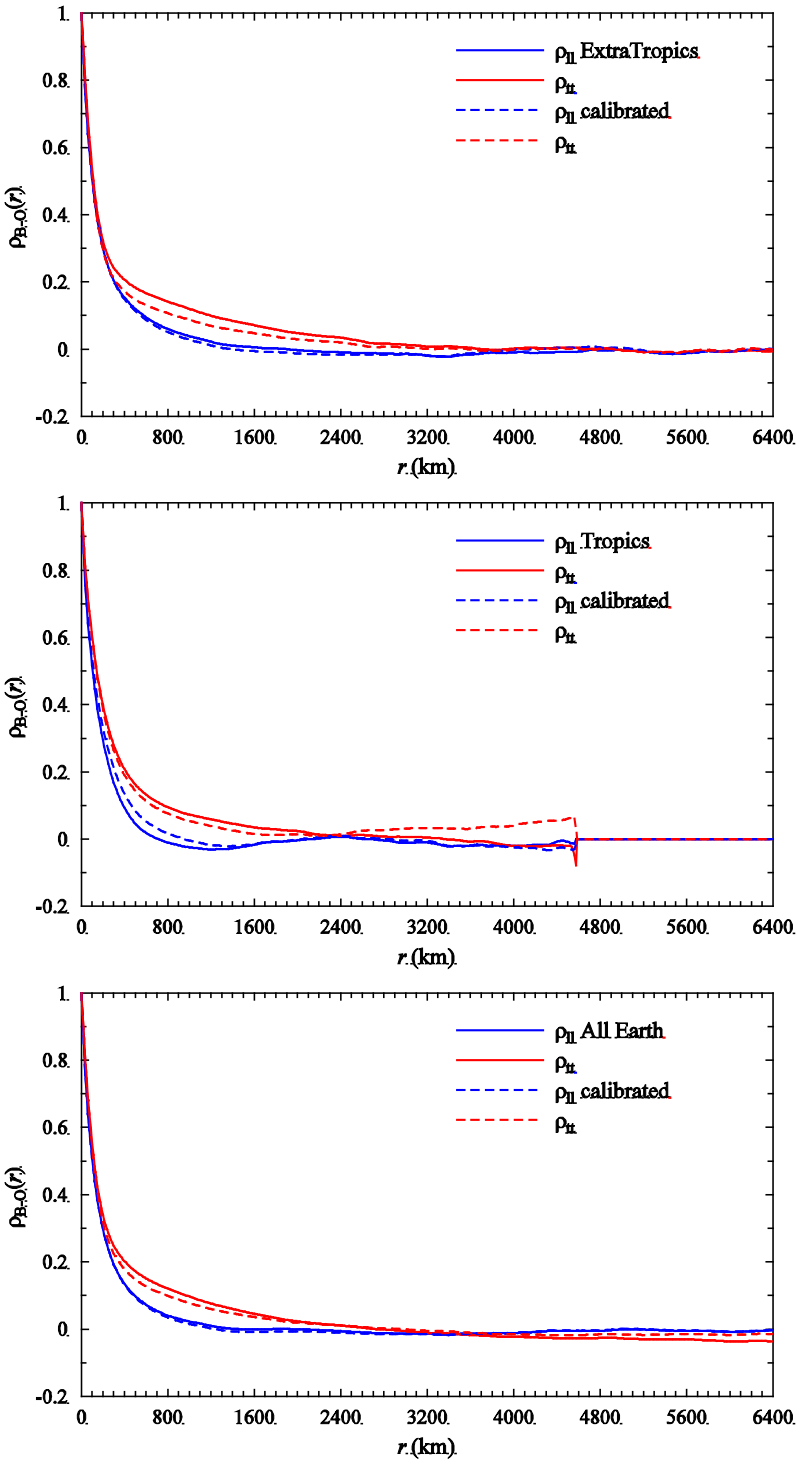


**Figure 4.7** Effect of calibration on the autocorrelations.

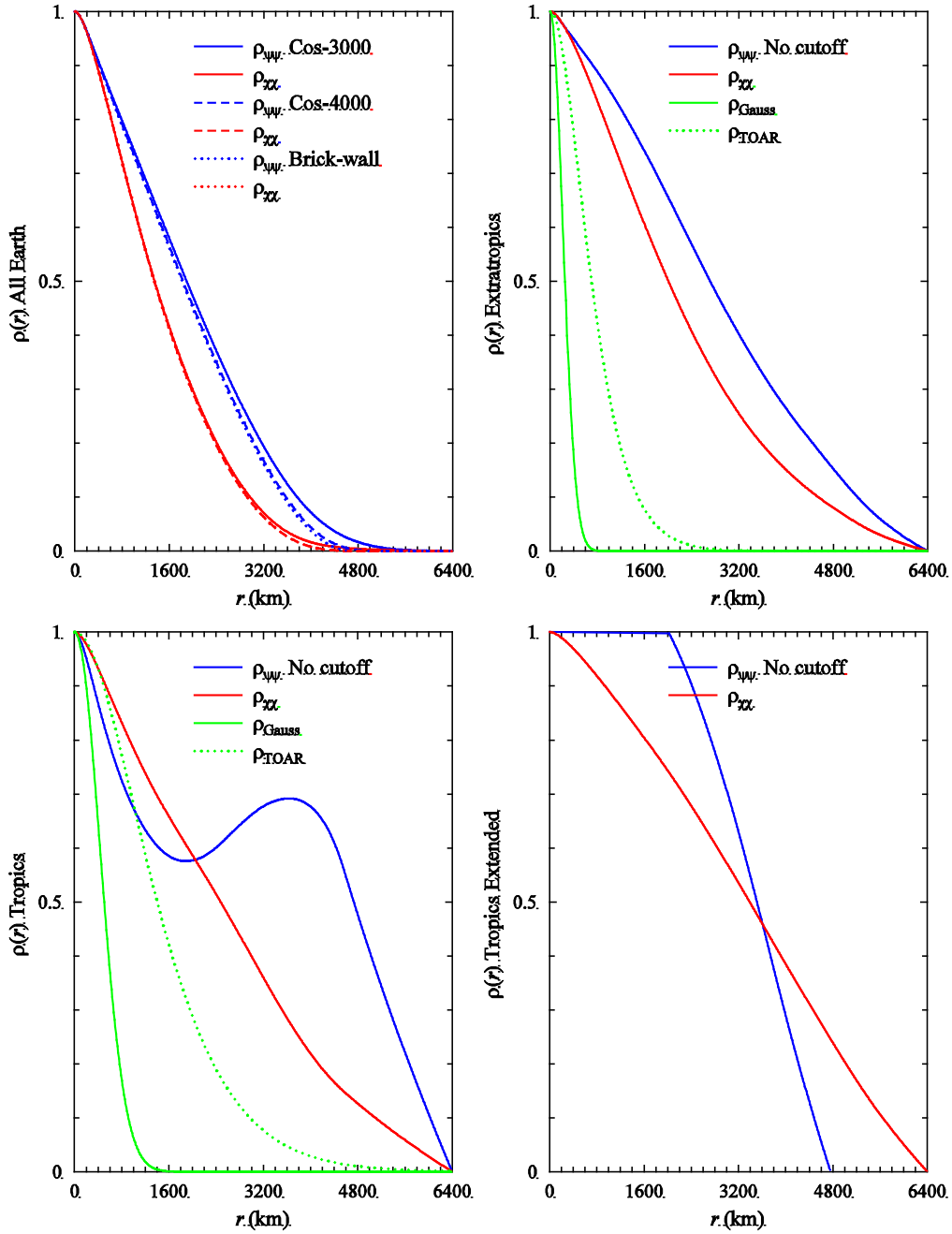


**Figure 4.8** Effect of calibration on the error correlation functions

Figure 4.8 shows the error correlation functions obtained with calibrated wind data. A cut-off function can only be defined for the global error correlation functions; for the others the range is too limited. Comparison with figure 4.6 shows that the global and extratropics error correlation functions are quite

similar, but that the use of calibrated wind data has an adverse effect for the error correlation functions in the Tropics and the Extended Tropics.

## 4.3 Error correlation functions in the frequency domain

Figure 4.9 shows the error correlation functions in the frequency domain using all ASCAT-12.5 data from January 2009 collocated with ECMWF predictions. The error correlation functions were obtained from spectra with an FFT size of 256 (solid curves) and 512 (dashed curves) over all of the earth. Since calculation of the spectra requires samples without missing points (only isolated missing points are interpolated), the maximum sample length is restricted to the size of the oceans. An FFT size of 1024 would require samples of 12800 km length, leading to too few valid samples in one month of data.

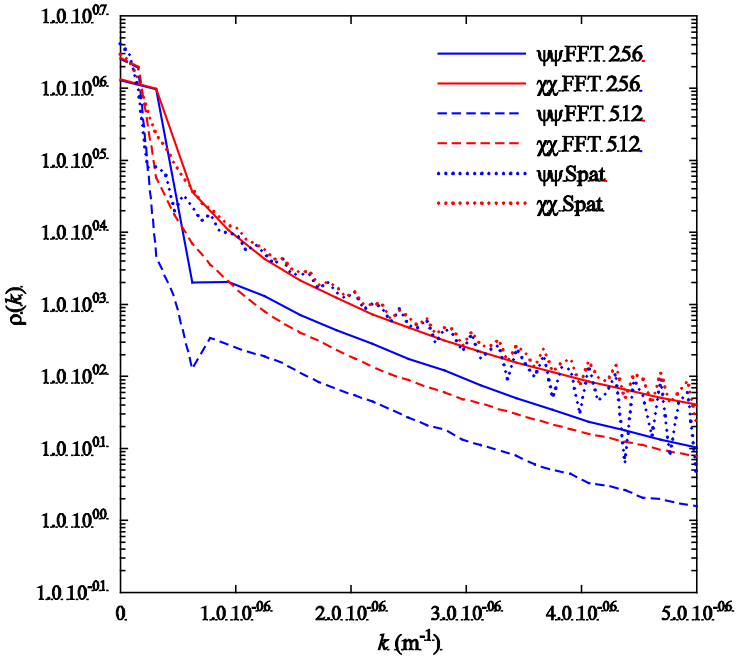


**Figure 4.9** Error correlation functions in the frequency domain.

Figure 4.9 shows that the FFT size has great impact on the shape of the error correlation functions, especially for small spatial frequencies. An FFT size of 512 is needed to resolve the shape of the error correlation functions near $k = 0$. As a reference, the dotted curves show the Fourier transform of the spatial error correlation functions obtained from autocorrelations with a lag size of 1024 and a Brick-wall cut-off function. The oscillations for high frequencies are numeric effects of the FFT operation. They are not relevant, since the error correlation functions have fallen off by several orders of magnitude. Note that

there is a considerable difference between the frequency domain error correlation functions $\hat{\rho}_{\psi\psi}$ and $\hat{\rho}_{\chi\chi}$ for larger values of $k$, whereas the Fourier transformed spatial error correlation functions are almost the same, both coinciding with $\hat{\rho}_{\chi\chi}$ for FFT size 256.

In chapter 3 it has been shown from the Gaussian tests that the error correlation functions in the frequency domain are less accurate than those in the spatial domain. Combining this with the results above leads to the conclusion that the error correlation functions can better be calculated in the spatial domain, because this leads to more detailed and accurate results.

# 5 Effect on analysis and selection

In this chapter the effect of the error correlation functions on the analysis will be studied. Since there are three scatterometer wind products (ASCAT-12.5, ASCAT-25, and SeaWinds-KNMI), five error correlation functions (Gauss, TOAR, C3000, C4000, and Brick-wall), and at least three geographical zones (all Earth, tropics, and extratropics), the number of combinations is very large. Moreover, as will be shown in the next section, TOAR and numerical error correlation functions require a large free edge around the observations, leading to large batch grid dimensions. In order to reduce computational load, all results shown in this chapter are based on the data of only three days, January 1-3, 2009.

As was mentioned in the introduction, there is little freedom in adjusting the error variances of observations and background within 2DVAR. However, the default error variances in 2DVAR do not agree with values found by *Vogelzang et al.* [2011] from triple collocation, so the effect of more realistic error variances will be investigated too in this section.

For ASCAT-12.5, the wind component spectra from this dataset are an average over 31342 samples, while the divergenge and rotation spectra are an average over 14623 samples, so the accuracy in the spectral densities is better than 1%.

## *5.1 Batch grid*

The batch grid size and dimension can be determined quickly using the Single Observation Aanalysis (SOA) test [*Vogelzang*, 2007]. In this test a batch grid with given size and dimension is generated, together with zero background and a single observation in the centre of the batch grid. 2DVAR will generate an analysis increment, the shape of which depends on the error correlation function parameter. The analysis wind speed at the location of the observation, however, solely depends on the error standard deviations of observation and background, $\sigma_o$ and $\sigma_b$, respectively, as

$$u_a = f u_o \quad , \quad v_a = f v_o \quad , \quad f = \frac{\sigma_o^2 \sigma_b^2}{(\sigma_o^2 + \sigma_b^2)^2} \quad . \tag{5.1}$$

Taking the standard choices $\sigma_o$ =1.8 m/s and $\sigma_b$ =2.0m/s and putting $(u_o, v_o) = (1,0)$ this readily leads to $u_a$ =0.5525 at the location of the observation. Table 5.1 shows the results of the SOA test.

The various error correlation functions were tested for different batch grid sizes $\Delta$. The precision (Prec.) gives the deviation in percent from the expected value for $u_a$ = 0.5525; a plus sign indicating too high retrieved values and a minus sign too low values. Both the Gaussian and the TOAR error correlation functions had the standard 2DVAR parameters: $R_\psi = R_\chi = 600$ km with $\nu^2$=0.5 in the tropics and

$R_\psi = R_\chi = 300$ km with $\nu^2 = 0.2$ in the extratropics. The batch grid was constructed for ASCAT-12.5, assuming a WVC size of 12.5 km with 41 WVC's across track and 96 WVC's along track. The Gaussian error correlation functions were tested with a free edge around the batch of 1800 km wide; the others with a free edge of 6000 km. The batch grid parameters are listed in table 5.2.

| Structure Function | $\Delta$ =12.5 km | | $\Delta$ =25 km | | $\Delta$ =50 km | | $\Delta$ =100 km | |
|---|---|---|---|---|---|---|---|---|
| | $u_o$ | Prec. | $u_o$ | Prec. | $u_o$ | Prec. | $u_o$ | Prec. |
| Gauss | -- | -- | 0.5528 | +0.1 | 0.5525 | 0.0 | 0.5525 | 0.0 |
| TOAR | -- | -- | 0.5767 | +4.4 | 0.5626 | +1.8 | 0.5566 | +0.8 |
| C-3000 | 0.6756 | +22.3 | 0.5627 | +1.8 | 0.5309 | -3.9 | 0.5101 | -7.7 |
| C-4000 | 0.6797 | +23.0 | 0.5640 | +2.1 | 0.5311 | -3.9 | 0.5100 | -7.7 |
| Brick-wall | 0.6950 | +23.0 | 0.5778 | +4.6 | 0.5326 | -2.5 | 0.5114 | -7.4 |

**Table 5.1** Results for the single observation test.

| $\Delta$ (km) | Free edge 1800 km | Free edge 6000 km |
|---|---|---|
| 100 | 42 × 48 | 126 × 132 |
| 50 | 84 × 96 | 252 × 264 |
| 25 | 168 × 192 | 504 × 528 |
| 12.5 | -- | 1008 × 1056 |

**Table 5.2** Batch grid sizes and dimensions.

The Gaussian error correlation function gives analysis wind speeds very close to the analytical solution. The TOAR error correlation function requires a large free edge, as the numerical error correlation functions. The TOAR error correlation function yields test results closest to the analytical results for a rather large batch grid size $\Delta$ = 100 km. This keeps the batch grid dimension within reasonable limits, as can be seen from table 5.2. The numerical error correlation functions perform best for $\Delta$ = 25 km. Here the error correlation function with soft cut-off, C-3000, seems a bit better than those with hard cut-off, C-4000 and Brick-wall. Note the large differences (more than 20%) between the test results and the analytical solution for numerical error correlation functions with minimum batch grid size $\Delta$ = 12.5 km. The reason for this is not clear.

## *5.2 ASCAT-12.5*

### 5.2.1 Effect batch grid size

Figure 5.1 shows the spectra of ASCAT-12.5 (blue curves), ECMWF background (red curves), and 2DVAR analysis (green curves) for $u$ (left hand panel) and $v$ (right hand panel) with batch grid sizes of 100 km (solid), 50 km (dashed), and 25 km (dotted). The spectra were obtained over all Earth with the

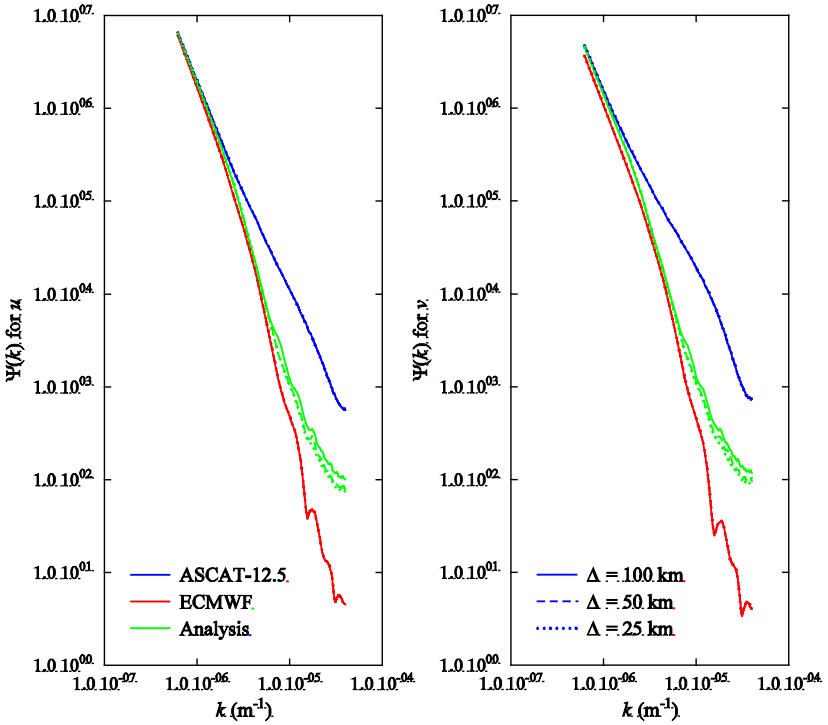


**Figure 5.1** Spectra for Gaussian error correlation functions with different batch grid sizes.

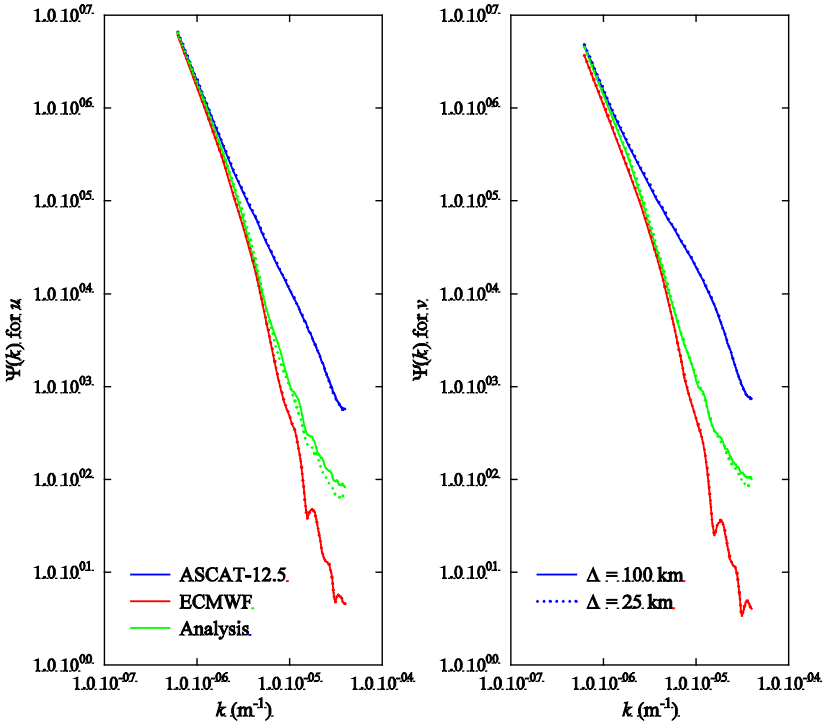


**Figure 5.2** As figure 5.1, for TOAR error correlation functions.

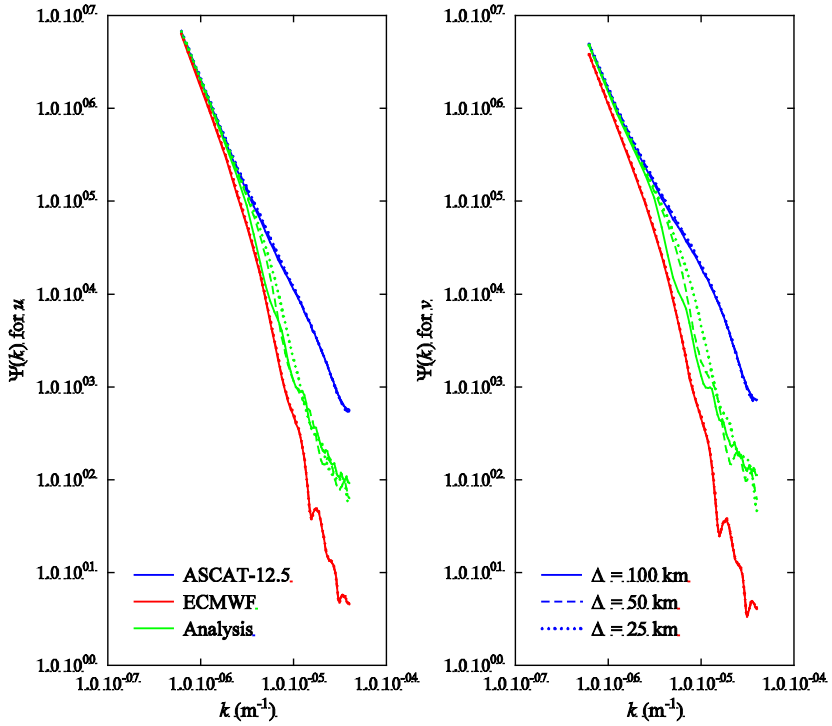


**Figure 5.3** As figure 5.1, for Cos-4000 error correlation functions.

Gaussian error correlation function parameters as used in operational processing: $R_\psi = R_\chi = 600$ km, $\nu^2 = 0.5$ in the tropics, and $R_\psi = R_\chi = 300$ km, $\nu^2 = 0.2$ in the extratropics. The free edge around the scatterometer observations in the 2DVAR batch grid was 1800 km.

Figure 5.1 shows that for Gaussian error correlation functions the observation and background spectra show no effect of the batch grid size, but that the analysis contains more information for a batch grid size of 100 km.

Figure 5.2 is similar to figure 5.1, but now using TOAR error correlation functions with the same error correlation function parameters as the Gaussian error correlation functions and a free edge of 6000 km. Only results for batch grid sizes of 100 km and 25 km are shown. The results are very similar to that for the Gaussian error correlation function.

Figure 5.3 is also similar to figure 5.1, but now for the numerical Cos-4000 error correlation function (see table 4.1 for its cut-off parameters). Here the situation is less clear than before: with a batch grid size of 25 km the analysis improves at intermediate scales, with some effect on the selection, but at the smallest scales the information content of the analysis decreases rapidly. This is in particular clear for $\nu$ (right hand panel).

In the remainder of this chapter the batch grid size for Gaussian error correlation functions will be 100 km, unless explicitly stated otherwise. The batch grid size for TOAR and numerical error correlation functions results for both 25 km and 100 km batch grid size will be shown.

### 5.2.2 Effect on the analysis

Figure 5.4 shows the effect of the various error correlation functions on the analysis spectrum. This figure shows the spectral ratio for various error correlation functions relative to the Gaussian error correlation function with default parameters and 25 km batch grid size (dotted curves in figure 5.1). Since the TOAR and numerical error correlation functions require a large free edge around the batch grid of 6000 km and a small batch grid size, 2DVAR consumes a lot of time: processing of a single three-minutes ASCAT file takes of the order of three minutes. Therefore the spectra on which figure 5.4 is based are calculated for the period January 1-3, 2009. The results in figure 5.4 pertain to all Earth.

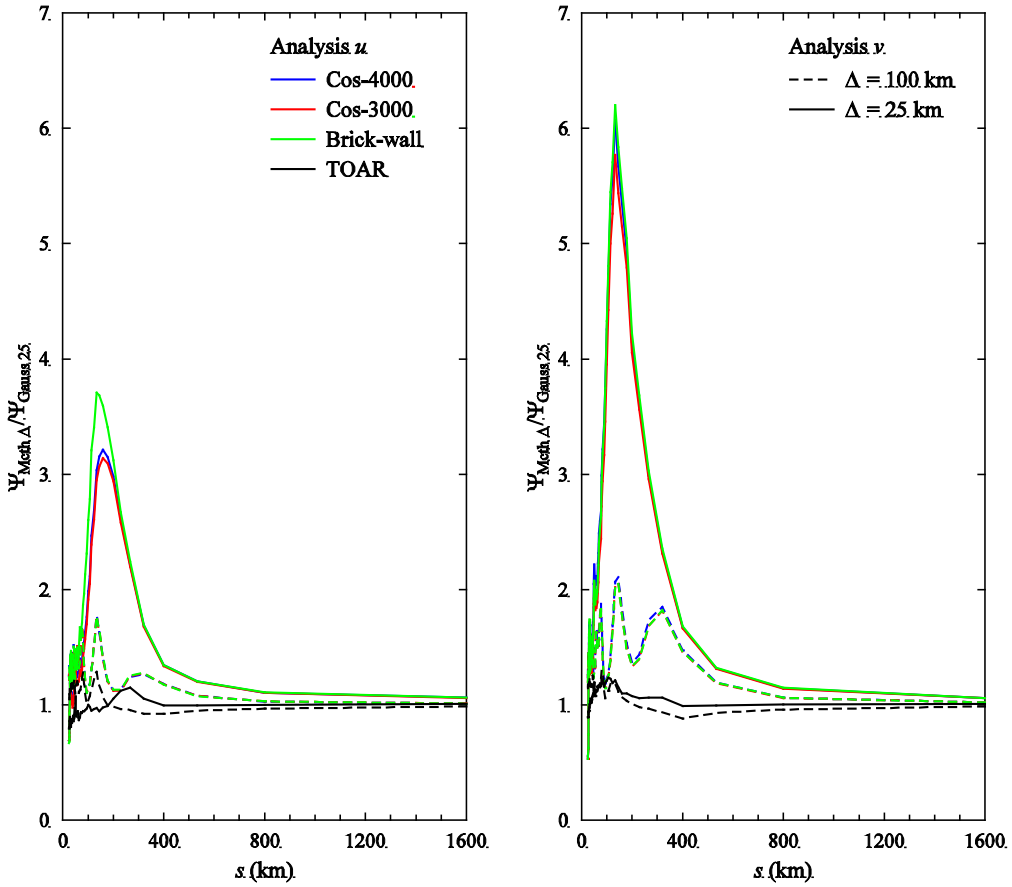


**Figure 5.4** Spectral ratio of analysis spectra for various error correlation functions relative to the default Gaussian as a function of scale.

The spectral ratio is shown as a function of the scale $s$ (in km) defined as

$$s = \frac{1}{k} \quad , \tag{5.2}$$

with $k$ the spatial frequency.

The numerical error correlation functions with cosine cut-off (Cos-4000 and Cos-3000, blue and red curves, respectively) and brick-wall cut-off (Brick-wall, green curves) have the largest effect on the analysis when the batch grid size $\Delta$ = 25 km (solid curves). For $\Delta$ = 100 km the effect is much smaller, though the increase in spectral density may be more than 2 for $v$ at scales of 200 km. The TOAR error correlation functions have little effect.

For the meridional wind component $v$ the spectral density increases with a factor of more than 7 when $s$ is slightly larger than 100 km. The effect for the zonal wind component $u$ is smaller, but still sizeable: an increase by a factor of more than 3.5 for $s$ around 150 km. For both wind components the Brick-wall error correlation function give the largest increase in spectral density.

The numerical error correlation functions have the effect that they were intended to have: they increase the analysis spectral density, notably at small scales of the order of 100 km. However, for very small scales of less than 30 km the numerical error correlation functions with 25 km batch grid size show a decrease in spectral density relative to the default Gaussian error correlation function. This decrease is strongest for $v$ when using the Brick-wall function with 25 km batch grid size. This decrease does not occur when the batch grid size is 100 km.

### 5.2.3 Effect on the selection

One expects that when 2DVAR makes an analysis that contains more variance this will also affect the selection. Figure 5.5 shows the results for the 2DVAR selected ASCAT 12.5 km wind using the same data as used for obtaining figure 5.4.

As was the case for the analysis, the numerical error correlation functions at 25 km batch grid size have the largest effect on the selection. The variance in the selected wind increases in both wind components for all scales except the smallest ones (smaller than 40 km). The increase in selection spectral density is 15% for scales between 200 and 300 km, but drops by 10% for the smallest scales. This may well be an indication that the use of numerical error correlation functions decreases the noise in the scatterometer wind fields.

The TOAR error correlation function at 25 km batch grid size shows a slight increase over the whole range for both $u$ and $v$; those at 100 km batch grid size show a slight decrease.

The results in figures 5.4 and 5.5 are a bit noisy, due to the limited period for which the spectra were calculated. The noise can be reduced by calculating the spectra with a multitaper method [*Vogelzang*, 2010]. This does not affect the results significantly (no figures shown).

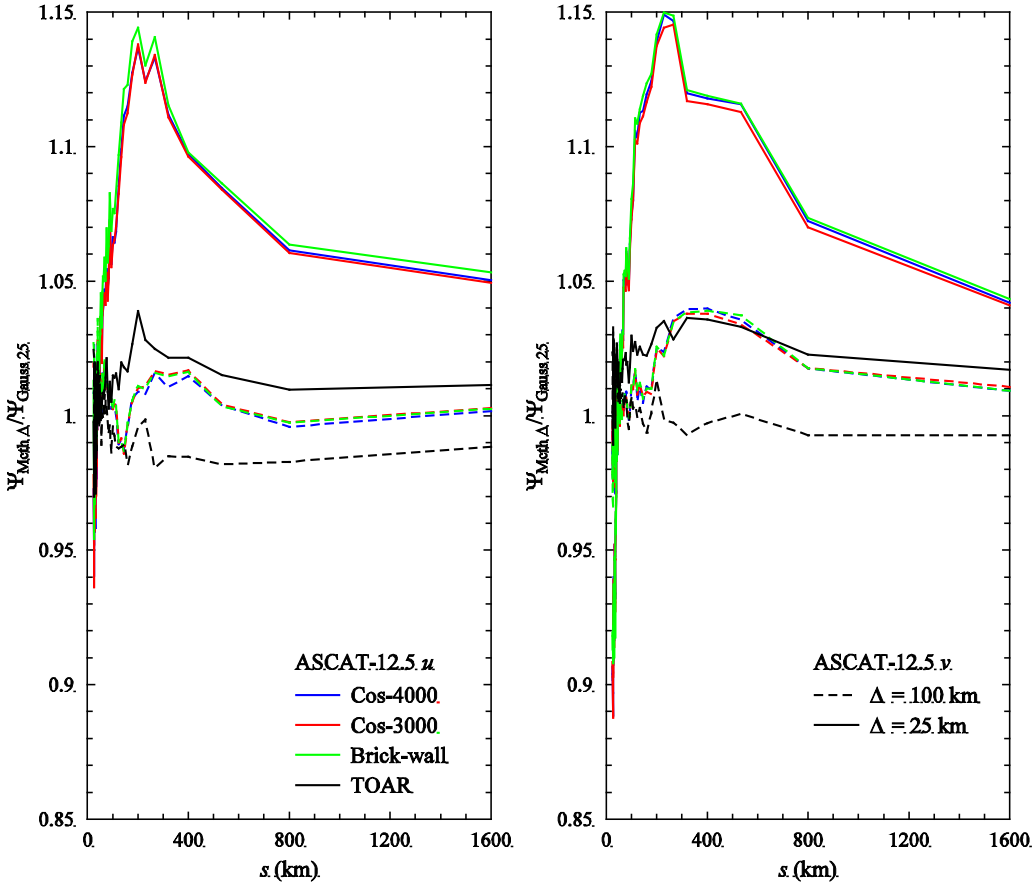


**Figure 5.5** As figure 5.4, but for the ASCAT-12.5 selected wind.

## 5.2.4 Effect of calibrated winds and error variances

Figure 5.6 shows the effect of various error correlation functions on the spectrum of selected winds. The solid green curves show the spectra of ASCAT-12.5 for the period January 1-3, 2009 obtained with the default Gaussian error correlation functions. The dashed green curves show the results with the numerical error correlation function with brick-wall cut-off. These results were already discussed in the previous subsections.

The dot-dot-dashed blue curve shows the results for the global brick-wall cut-off numerical error correlation function obtained from calibrated winds using the calibration coefficients of *Vogelzang et al.* [2011]. This curve coincides with the dashed green curve that was obtained from uncalibrated winds. This

leads to the conclusion that wind calibration has negligible effect on the numerical error correlation functions.

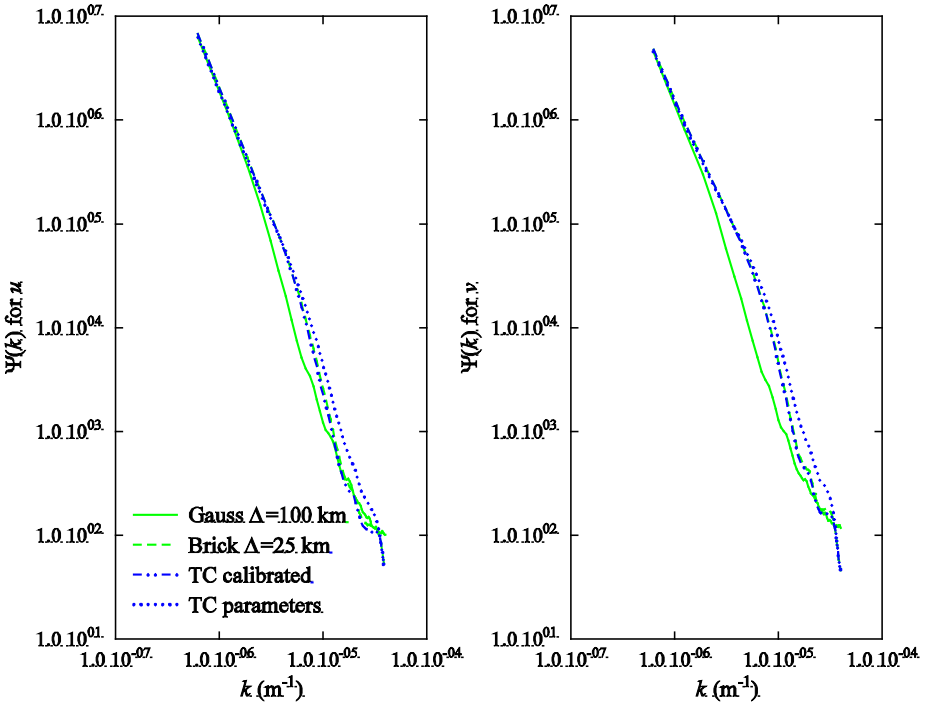


**Figure 5.6** Spectrum of selected winds for various numerical error correlation functions.

Finally, the dotted blue curves show the spectrum obtained using global numerical error correlation functions with brick-wall cut-off obtained from calibrated winds, but now using error variances for observations and background as found by *Vogelzang et al.* [2011] instead of the default values, see table 5.3. This has a significant effect on the spectrum for scales between about 40 km up to more than 100 km.

| **Error variances** | **Observation** | | **Background** | |
|---|---|---|---|---|
| | ***u*** | ***v*** | ***u*** | ***v*** |
| Default | 1.8 | 1.8 | 2.0 | 2.0 |
| Triple collocation | 0.69 | 0.82 | 1.54 | 1.55 |

**Table 5.3** Values for observation and background error variances in 2DVAR.

### 5.2.5 Case January 2, 2010

Figure 5.7 shows a detail of the ASCAT-12.5 wind field recorded on January 2, 2009 around 6:06 UT off the Pacific coast of Canada. The figure covers an area of 4° by 4° centered at 43°N, 129° W, and shows a frontal zone with a change in wind direction. The purple arrows are set by the variational quality control (VQC) and indicate a large difference between selected ambiguity and analysis. The left hand panel of figure 5.7 is obtained with the default Gaussian error correlation functions and shows a large area with VQC flagged cells associated with the front. There are also some ambiguity removal errors, visible as ↑↓↑ flow patterns. The right hand panel of figure 5.7 is obtained with the global brick-wall cut-off error correlation functions (and default error variances in 2DVAR). The ambiguity removal errors have disappeared and VQC flagging is restricted to a much narrower zone. The right hand panel of figure 5.7 also contains two orange arrows. These are set by quality control and most probably indicate rain. They were overridden by the VQC flag in the left hand panel.

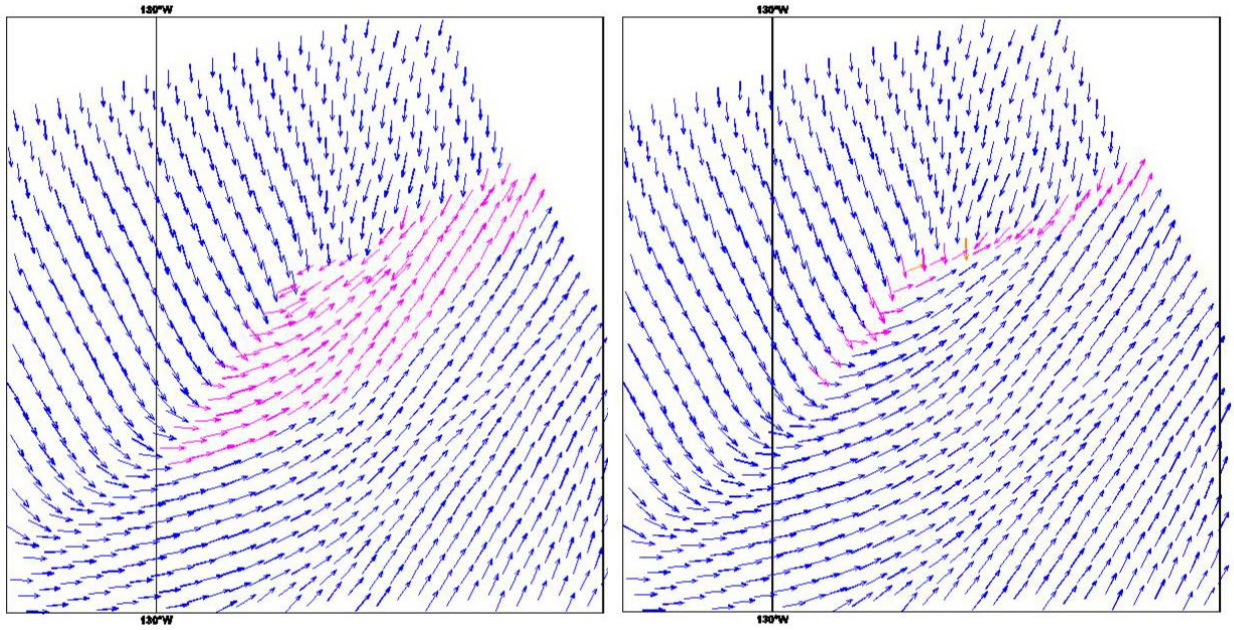

**Figure 5.7** ASCAT-12.5 wind fields off the Pacific coast of Canada on January 2, 2009 around 6:06 UT. Left hand panel: with default Gaussian error correlation functions; right hand panel: with global brick-wall cut-off functions:

The ECMWF background wind field is shown in figure 5.8. As expected, the frontal zone is represented as a wide area with gradually changing wind direction. The ECMWF background resembles the ASCAT wind field obtained with the standard Gaussian error correlation functions in figure 5.7. This is because the analysis, which is shown in the left hand panel of figure 5.9, does not contain much small-scale

information, while the analysis obtained with numerical error correlation functions, shown in the right hand panel of figure 5.9, exhibits much more sharp detail. Without additional information it is very hard, if not impossible, to decide which wind field is correct. Nevertheless, this example demonstrates the potential of numerical error correlation functions for ambiguity removal in complicated situations.

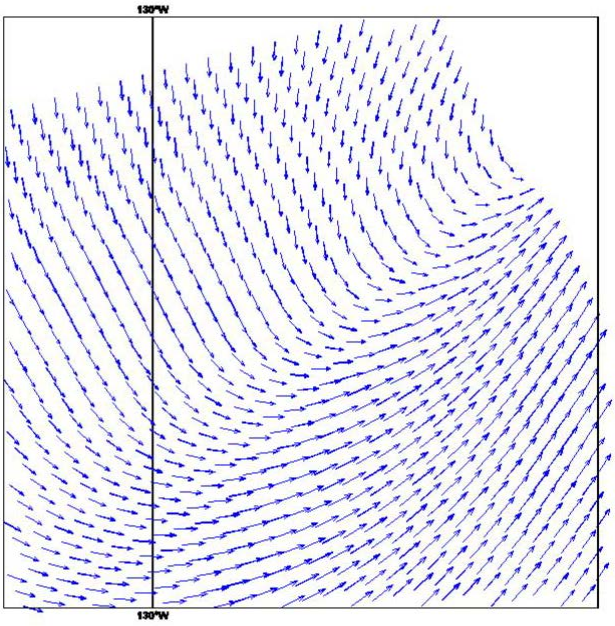

**Figure 5.8** ECMWF background corresponding to figure 5.7

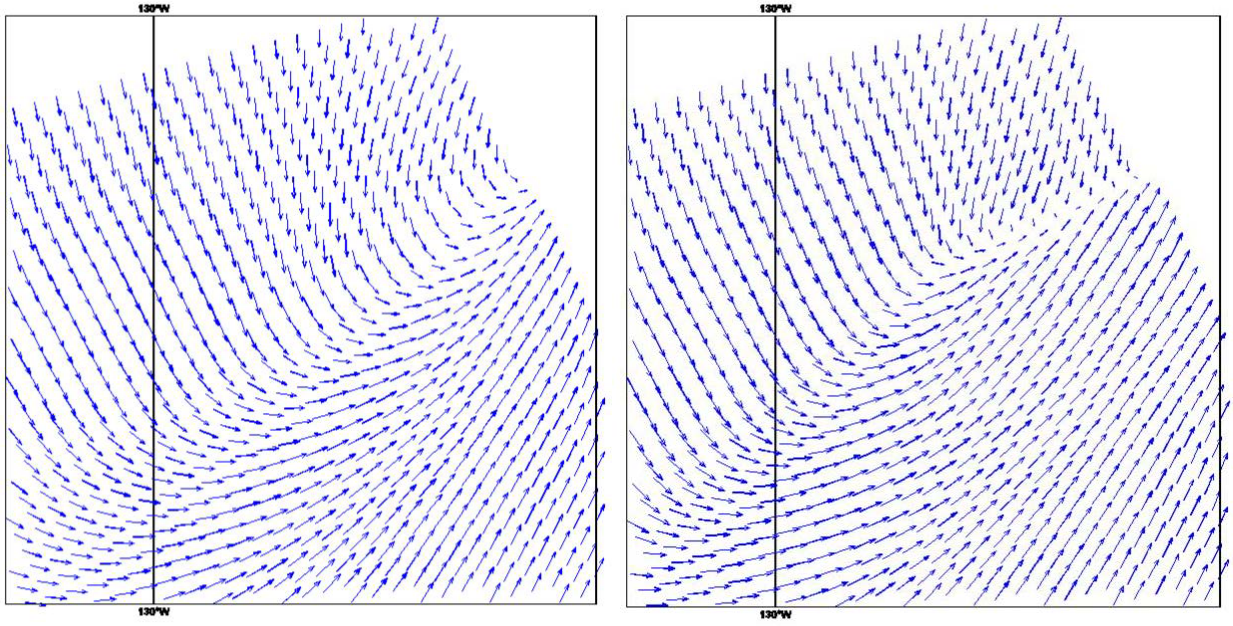


**Figure 5.9** Analyses corresponding to figure 5.7.

## *5.3 SeaWinds-KNMI*

### 5.3.1 Spectra

The Brick-wall error correlation function that was applied to ASCAT-12.5 data in the previous sections has also been used for SeaWinds data processed with the SeaWinds Data Processor (SDP). This wind product is further referred to as SeaWinds-KNMI. Figure 5.10 shows the resulting wind spectra, together with some ASCAT results for comparison. The SeaWinds spectra were obtained using all data from January 2009.

The blue curves in figure 5.10 show the scatterometer winds selected by 2DVAR. The green curves show the analyses, and the red curves the ECMWF background spectra. Solid curves pertain to ASCAT-12.5 with default Gaussian error correlation functions (G), and dashed curves to ASCAT-12.5 with Brick-wall numerical error correlation functions (NSF). Similarly, dot-dot-dashed curves are for SeaWinds-KNMI with default Gaussian error correlation functions, and dotted curves for SeaWinds-KNMI with numerical error correlation functions.

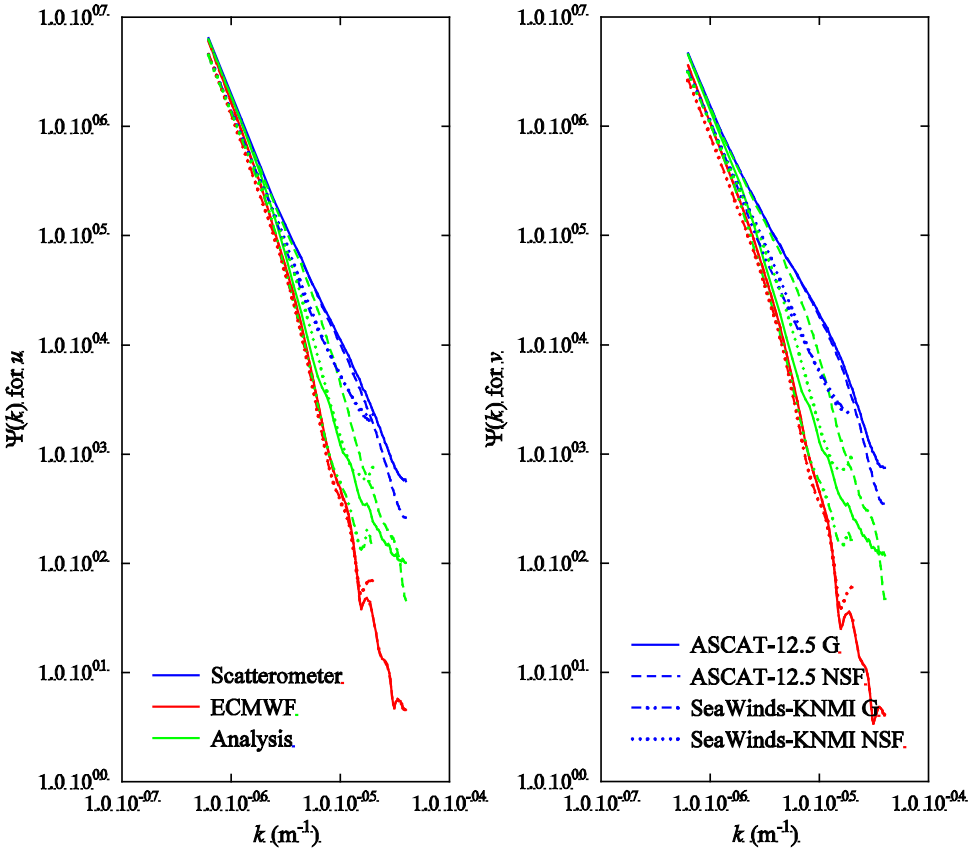


**Figure 5.10** Wind component spectra for ASCAT-12.5 and SeaWinds-KNMI

Figure 5.10 shows that the spectral level of the SeaWinds-KNMI analysis improves considerably when applying numerical error correlation functions: it has the same or even higher level than that of ASCAT-12.5 with default Gaussian error correlation functions. At the smallest spatial scales the SeaWinds-KNMI analysis shows some indication of noise.

As with ASCAT-12.5, the effect of numerical error correlation functions on the SeaWinds-KNMI selection is smaller than that on the analysis. This is particularly well visible in figure 5.11 which shows the spectral ratio of numerical error correlation function results over default Gaussian ones for ASCAT-12.5 and SeaWinds-KNMI.

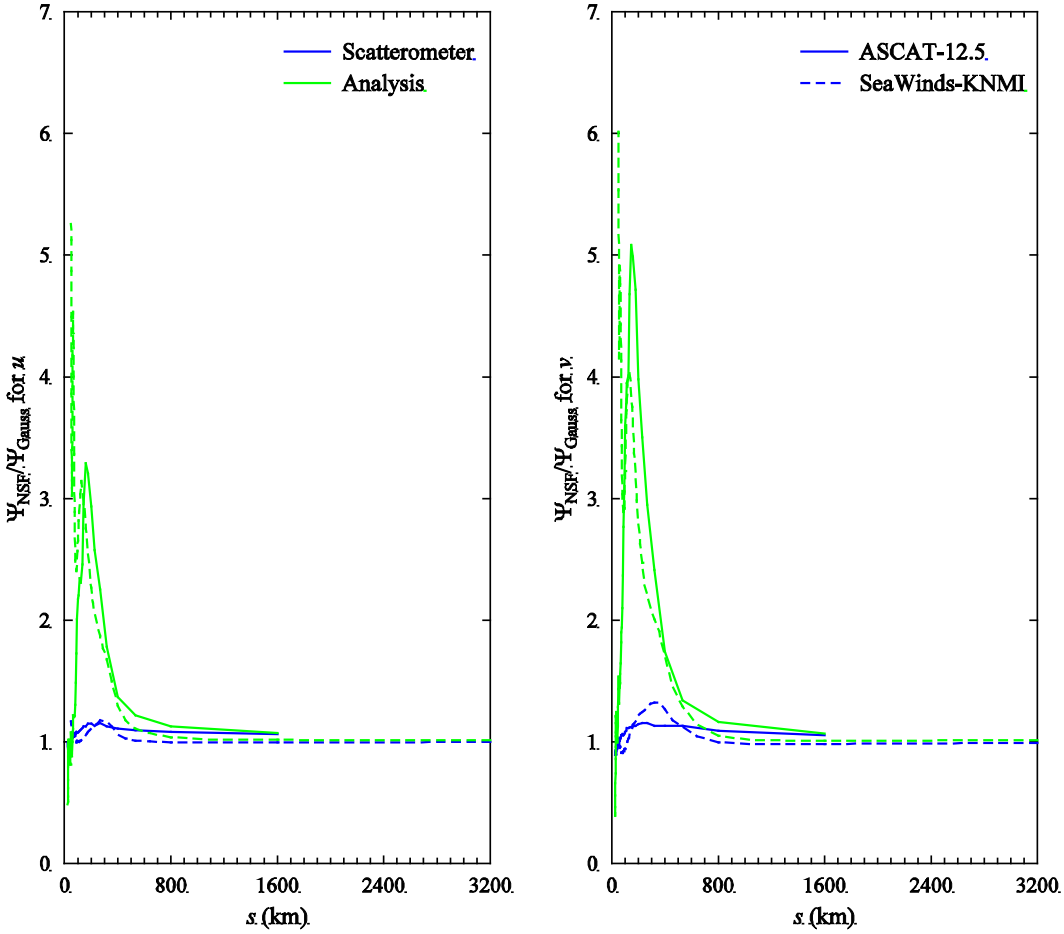


**Figure 5.11** Ratio of spectra obtained with numerical error correlation functions over spectra obtained with default Gaussian error correlation functions for ASCAT-12.5 (solid curves) and SeaWinds-KNMI (dashed curves). results are shown for the 2DVAR selected winds (blue curves) and the 2DVAR analysis (green curves).

Figure 5.11 shows that the increase in spectral density of the analysis (green curves) is of the same order of magnitude for SeaWinds-KNMI and ASCAST-12.5, though the increase for SeaWinds-KNMI is at smaller spatial scales. The increase in spectral density of the 2DVAR selected winds (blue curves) is of the same order of magnitude for SeaWinds-KNMI and for ASCAT-12.5 at the zonal wind component $u$. At the meridional wind component $v$ the ratio for SeaWinds-KNMI is more than twice as large as for ASCAT-12.5. The maximum increase for SeaWinds-KNMI is for scales around 300 km. Note that the increase for SeaWinds-KNMI is restricted to small scales, whereas for ASCAT-12.5 the increase extends over all scales.

These results show that indeed the use of numerical error correlation functions have more impact on SeaWinds than on ASCAT, at least in a statistical sense.

### 5.3.2 Statistics

Since the SeaWinds data were processed for one month, it is possible to collocate the data with buoy data from that month and apply statistics to the differences. The SeaWinds data were processed in three different ways:

1. default processing with Gaussian error correlation functions;
2. NSF processing with Brick-wall numerical error correlation functions;
3. NSF+TC processing with Brick-wall numerical error correlation functions as well as calibration using triple collocation results.

The results are given in table 5.4. The number of buoy collocations is 3047, leading to a precision in the standard deviations of 2% (about 0.03 m/s for the values in table 5.4)

.

| Processing | Bias | | | | Standard deviation | | | |
|---|---|---|---|---|---|---|---|---|
| | Scat-Buoy | | Scat-Back | | Scat-Buoy | | Scat-Back | |
| | *u* | *v* | *u* | *v* | *u* | *v* | *u* | *v* |
| Default | -0.36 | -0.02 | -0.06 | 0.01 | 1.71 | 1.69 | 1.56 | 1.47 |
| NSF | -0.34 | -0.03 | -0.08 | 0.02 | 1.75 | 1.63 | 1.72 | 1.60 |
| NSF + TC | -0.33 | -0.07 | -0.09 | 0.02 | 1.69 | 1.62 | 1.65 | 1.53 |

**Table 5.4** Statistics of the differences between scatterometer winds and buoy or background winds for January 2009.

Table 5.4 shows that the bias in $u$ varies little for the various datasets, while that in $v$ increases slightly for the difference between scatterometer and buoy when applying numerical error correlation functions and triple collocation calibration. However, the bias in $v$ is very small. Note that the bias in $u$ is much larger for the difference between scatterometer and buoy than for that between scatterometer and background.

The standard deviations for the difference between scatterometer and buoy wind component lie close together, almost within the precision of 0.03 m/s. The results for $v$ show a decrease when successively applying numerical error correlation functions and triple collocation calibration. The results for $u$ are less clear, though the smallest standard deviation is obtained when applying both numerical error correlation functions and triple collocation calibration. The standard deviations of the difference between scatterometer and ECMWF background wind components show clear and statistically significant effects: they increase when applying numerical error correlation functions and decrease when also applying triple collocation calibration.

These results indicate that application of numerical error correlation functions cause the resulting wind field to become closer to the high resolution buoy measurements and further away from the low resolution background fields. This is a further indication that 2DVAR with numerical error correlation functions leads to more detailed wind fields. Application of the triple collocation calibration removes (in first order) scaling differences between the various wind fields, leading to better agreement.

### 5.3.3 Case Central America, January 31, 2009

Figure 5.12 shows the SeaWinds-KNMI wind field recorded on January 31, 2009 off the Pacific coast of Central America, using default Gaussian error correlation functions. Two strong southward air flows from land are visible, one around 95° W and one around 87° W, the latter being on the eastern edge of the figure. between the two outflows is a calm area. Figure 5.12 shows a line of divergence around 7° N with a clear separation between flow towards the northwest and the southwest.

Figure 5.13 is the same as figure 5.12, except that the wind field now was obtained with Brick-wall numerical cutoff functions and that first-order calibration from triple collocation has been applied. The line of divergence has disappeared, and the flow pattern appears more realistic than that in figure 5.12.

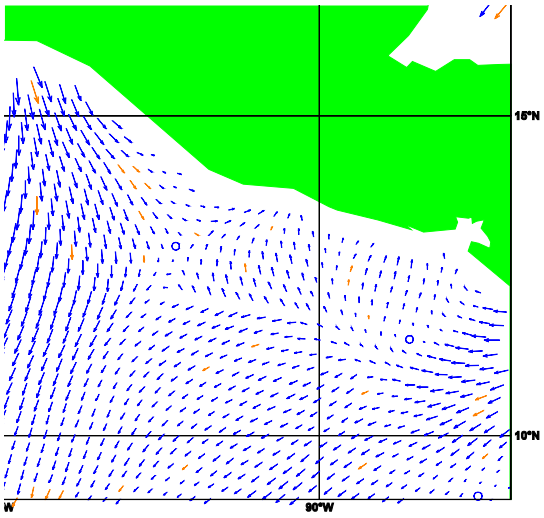


**Figure 5.12** SeaWinds-KNMI wind field recorded Jan 31, 2009 off the Pacific coast of Central America obtained with default Gaussian error correlation functions.

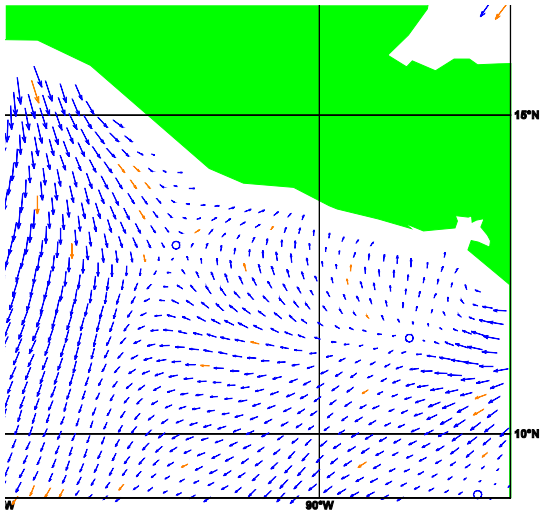


**Figure 5.13** As figure 5.12, but with Brick-wall numerical error correlation functions

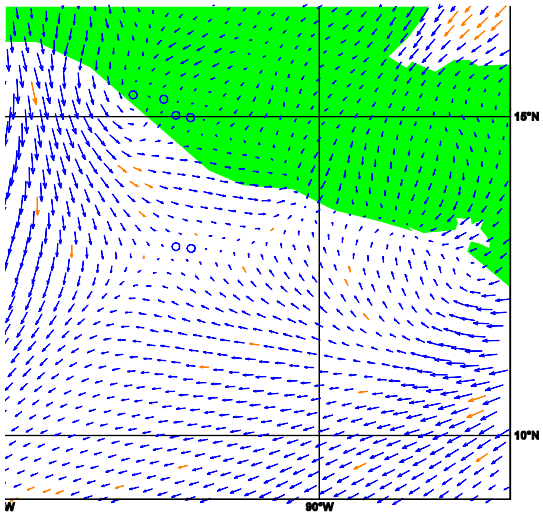


**Figure 5.14** ECMWF background corresponding to figures 5.12 and 5.13.

Figure 5.14 shows the ECMWF background corresponding to figures 5.12 and 5.13. The overall features of the background agree with those of the scatterometer wind fields, except that the high pressure area between the outflows (visible as the small circles that indicate very low wind speeds) is located slightly more to the east in the background than in the scatterometer.

### 5.3.4 Case Madagascar, January 22

Figure 5.15 shows a SeaWinds-KNMI wind field recorded on January 22, 2009, around 15:00 UT south-east of Madagascar. The figure shows a small tropical cyclone with wind speeds up to 20 m/s at the southwestern flank. Closer to the centre wind speeds up to 22 m/s are found, but those are VarQC flagged and drawn in purple in figure 5.13

Figure 5.16 is the same as figure 5.15, except that Brick-wall numerical error correlation functions have been used and first-order triple collocation calibration has been applied. Note that the centre of the cyclone is more to the northeast in figure 5.16. This happens quite often in the SeaWinds dataset of January 2009: in many cases application of numerical error correlation functions slightly shifts the position of small cyclones. In general the shift is small, but figures 5.15 and 5.16 show a case where the shift is quite large.

Figure 5.17 shows the ECMWF background wind field. Note that the SeaWinds-KNMI wind field with numerical error correlation functions has the centre of the cyclone close to the background prediction, whereas the wind field with default Gaussian error correlation functions has the centre further away.

Figure 5.18 shows two Meteosat images of the area southeast of Madagascar, the left one recorded at 12:00 and the right one at 18:00. The figure shows that the cyclone moves towards the southeast. The SeaWinds data were recorded right between those two images. Though it is hard to get an exact positions of the cyclone from figure 5.18, it seems like its centre is at 50° E or slightly more to the east. The white crosses in figure 5.18, located every 10 degrees, may help to estimate the position of the cyclone centre. This result favours the default Gaussian error correlation functions over the numerical ones.

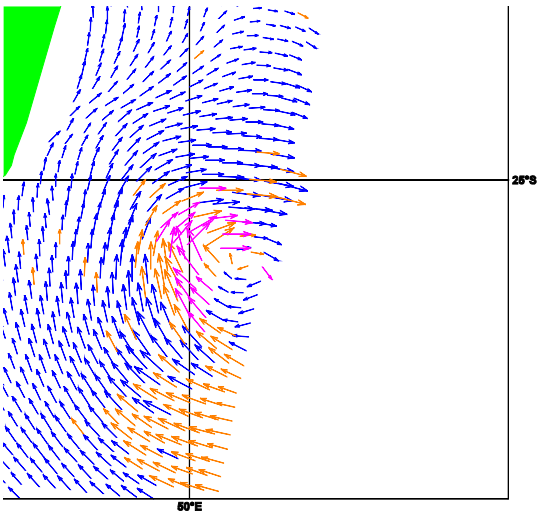


**Figure 5.15** SeaWinds-KNMI wind field on January 22, 2009 southeast of Madagascar (green) obtained with default Gaussian error correlation functions.

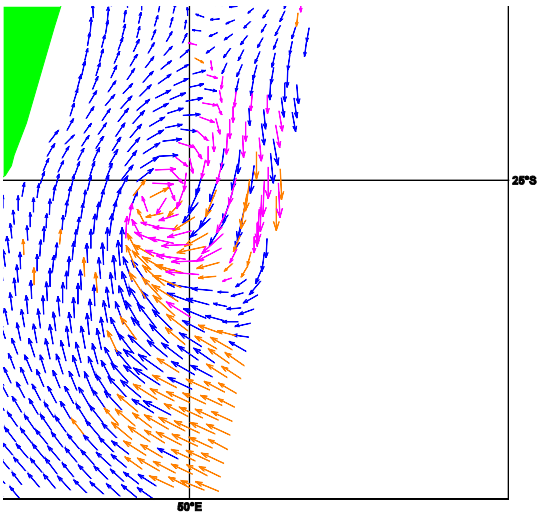


**Figure 5.16** As figure 5.15, but with Brick-wall numerical error correlation functions and first-order triple collocation calibration.

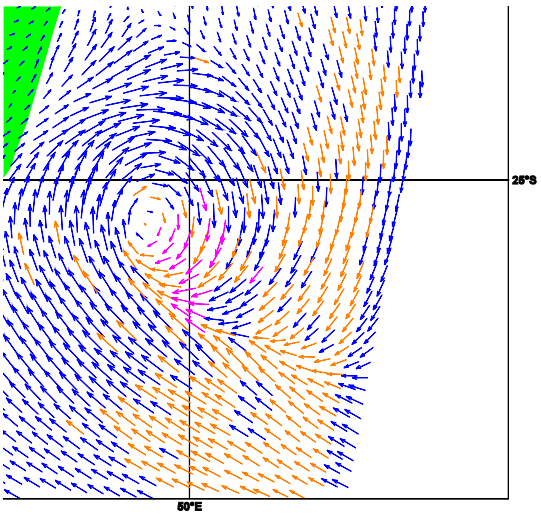


**Figure 5.17** ECMWF background corresponding to figures 5.15 and 5.16.

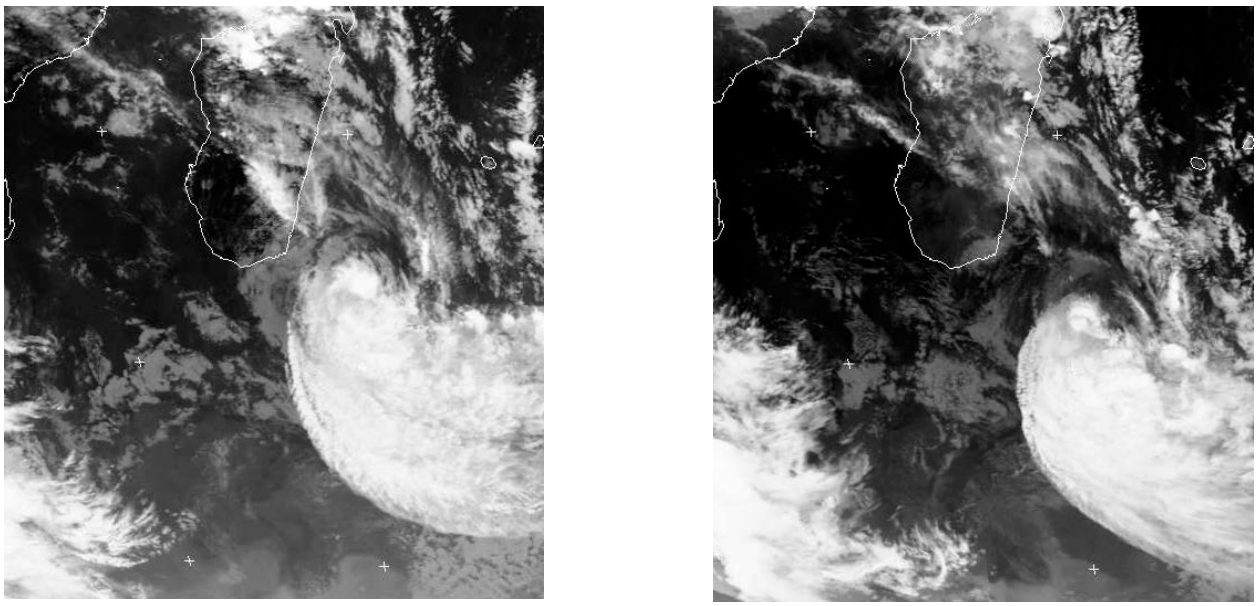

**Figure 5.18** Meteosat 7 images of the area southeast of Madagascar at 12:00 (left) and at 18:00 (right).

# 6 Conclusions

Error correlation functions were calculated from spatial correlation statistics of the difference of scatterometer winds and ECNWF background. It has been shown that the error correlation functions can be calculated both in the spatial domain (from autocorrelations) and in the spatial frequency domain (from spectra). Calculation in the spatial domain yields the best results: the numerical implementation of this method is more accurate and the autocorrelation method can easily handle missing points, thus allowing calculation of correlations over large distances. Moreover, the values of the length scales and the rotation/divergence ratio directly follow from the boundary conditions.

Nevertheless, the autocorrelations do not converge nicely to zero for large distances because of the finite size of the Earth. As a consequence, it is not possible to derive error correlation functions for specific zones (tropics and Extratropics). Only global error correlation functions were constructed. Even then, a numerical cutoff at large distances must be introduced in order to obtain error correlation functions that approach zero with zero derivative at large distances. Several types of cutoff functions have been investigated, and best results were obtained by a Brick-wall function with cutoff of 4700 km.

Error correlation functions were calculated using all ASCAT-12.5 data from January 2009. These functions have a much longer range than the default Gaussian ones employed now operationally. As a consequence, 2DVAR needs a large batch grid with a free edge large enough to let the error correlation functions go to zero. In this work a free edge of 6000 km was used. As a result, 2DVAR requires much computer time.

The ASCAT-12.5 data of January 1-3, 2009, were reprocessed with numerical error correlation functions with Brick-wall cutoff at 4700 km. A small batch grid size of 25 km is needed to introduce small scale features in the analysis and, subsequently, in the selected wind field. The effect on the analysis is large: the spectral density of $u$ increases by a factor up to 3.5, and that in $v$ by a factor up to 7. The increase in the spectral density of the selected winds is much smaller, up to 15% in both $u$ and $v$. The maximum increase is for spatial scales between 200 km and 300 km. At the smallest scales the spectral density decreases, but this may be a sign of more effective noise filtering by 2DVAR.

It can be expected that 2DVAR is less sensitive to the details of the analysis for ASCAT data, because the ambiguities here are well defined. Nevertheless, a case of cold air outflow from Canada into the Pacific shows a frontal zone that is resolved much better with numerical error correlation functions than with the default Gaussian ones. Compared to the default results, numerical error correlation functions cause different wind fields in cases with small scale features or in cases with a limited number of observations caused by rain or the vicinity of the coast.

More effects can be expected for SeaWinds. The SeaWinds-KNMI data of January 2009 were reprocessed with Brick-wall cutoff error correlation functions on a batch grid with 25 km grid size and 6000 km free edge. The spectral density in the analysis increases by a factor 5 to 6, which is the same order of magnitude as for ASCAT. The spectral density in the selected wind increases up to17% for $u$ (which is almost the same as for ASCAT) and up to 32% for $v$ (which is twice as much as for ASCAT). Moreover, the increase in spectral density for the SeaWinds-KNMI selected winds is restricted to small scales below 500 km, whereas for ASCAT the effect extends over all spatial scales.

Two case studies give mixed results. In a case off the Pacific coast of Central America the use of numerical error correlation functions in 2DVAR solves some ambiguity removal issues. In a case southeast of Madagascar numerical error correlation functions places the centre of a small tropical cyclone at a different position than default Gaussian ones. The result obtained with numerical error correlation functions is closer to the ECMWF background, whereas Meteosat-7 imagery suggests that the default Gaussian result is more likely to give the right position.

The SeaWinds-KNMI results were also collocated with buoy winds. Application of Brick-wall numerical error correlation functions decreases the standard deviation of the difference between scatterometer and buoy wind components (except for $u$), though the differences are statistically not significant due to the small data set. It also increases the standard deviation of the difference between scatterometer and background wind components, and these increases are significant. This indicates that the increase in small details in the 2DVAR analysis leads to wind fields with more spatial variation: closer to the buoy winds and further away from the background. Application of first-order calibration decreases all standard deviations, because scaling effects are mitigated. Though the differences in standard deviation are smaller, the overall picture is clearer: better agreement with the buoys and less with the background for both $u$ and $v$.

At the moment 2DVAR with numerical error correlation functions is too time consuming to apply operationally. It may be used for reprocessing purposes, but optimalisation of 2DVAR in terms of code efficiency and 2DVAR batch grid tuning is desirable.

# References

Bourassa, M.A., and K. McBeth Ford, 2010,
Uncertainty in scatterometer-derived vorticity.
*J. Atm. Ocean. Tech.* **27**, 594-603

Conway, J.B., 1973,
*Functions of one complex variable.*
2nd printing, Springer-Verlag, New York, U.S.A.

Daley, R., 1993,
*Atmospheric data analysis.*
Cambridge University Press, Cambridge, USA.

Dawkins, P., 2003,
Paul's online math notes.
http://tutorial.math.lamar.edu/Classes/DE/VariationofParameters.aspx

Hollingsworth, A. and P. Lönnberg, 1986,
The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field
*Tellus*, 38A, 111-136.

Stoffelen, A., 1998,
Towards the true near-surface wind speed: Error modeling and calibration using triple collocation.
*J. Geophys. Res.* 103C3, 7755-7766.

Vogelzang, J., 2007,
*Two dimensional variational ambiguity removal (2DVAR).*
Report NWPSAF-KN-TR-004, EUMETSAT, Darmstadt, Germany.
http://research.metoffice.gov.uk/research/interproj/nwpsaf/scatterometer

Vogelzang, J., A. Stoffelen, A. Verhoef, J. de Vries, and H. Bonekamp, 2009,
Validation of two-dimensional variational ambiguity removal on SeaWinds scatterometer data.
*J. Atm. Ocean. Tech.* 26, 1229-1245.

Vogelzang, J., 2010,
*How to calculate wind spectra.*
Report NWPSAF-KN-TR-007, EUMETSAT, Darmstadt, Germany.
http://research.metoffice.gov.uk/research/interproj/nwpsaf/ scatterometer

Vogelzang, J., A. Stoffelen, A. Verhoef, and J. Figa-Saldaña, 2011,
On the quality of high-resolution scatterometer winds.
Accepted for publication in *J. Geophys. Res.*

# Appendix A Euler equation

The second order inhomogeneous Euler equation reads

$$X'' + \frac{1}{r} X' - \frac{1}{r^2} X = g \quad . \tag{A.1}$$

The solution of this equation is $X = X_h + X_p$, the sum of the solutions of the homogeneous equation and a particular solution of the inhomogeneous equation.

The homogeneous Euler equation reads

$$X'' + \frac{1}{r} X' - \frac{1}{r^2} X = 0 \quad . \tag{A.2}$$

By trying a solution of the form $X = r^j$ one readily finds that the solutions of the homogeneous equation are

$$X_1 = Ar \quad , \quad X_2 = \frac{B}{r} \quad , \tag{A.3}$$

where $A$ and $B$ are integration constants.

A particular solution can be found from the homogeneous solutions with the method of variation of parameters [*Dawkins*, 2003] as

$$X_p = -X_1 \int \frac{X_2 g}{W} + X_2 \int \frac{X_1 g}{W} \quad , \tag{A.4}$$

where $W$ is the Wronskian defined as

$$W = X_1 X_2' - X_1' X_2 \quad . \tag{A.5}$$

Substituting the solutions of the homogeneous equation in (A.5) yields $W = -2/r$, so the particular solution (A.4) becomes

$$X_p(r) = \frac{r}{2} \int_0^r ds\, g(s) - \frac{1}{2r} \int_0^r ds\, s^2 g(s) \quad . \tag{A.6}$$

From (A.3) and (A.6) the final solution reads

$$X(r) = r \left[ A + \tfrac{1}{2} \int_0^r ds\, g(s) \right] + \frac{1}{r} \left[ B - \tfrac{1}{2} \int_0^r ds\, s^2 g(s) \right] \quad . \tag{A.7}$$

# Appendix B The integral $S(a)$

Consider the integral

$$S(a) = \int_{-\infty}^{\infty} dx \frac{e^{2\pi i a x}}{x} \quad . \tag{B.1}$$

This can be written as

$$S(a) = \int_{-\infty}^{\infty} dx \frac{\cos(2\pi a x)}{x} + i \int_{-\infty}^{\infty} dx \frac{\sin(2\pi a x)}{x} \quad . \tag{B.2}$$

The first integral is zero because the integrand is an odd function of $x$, so

$$S(a) = i \int_{-\infty}^{\infty} dx \frac{\sin(2\pi a x)}{x} \quad . \tag{B.3}$$

The result of (B.3) depends on the value of $a$. When $a = 0$ the integrand vanishes, so $S(0) = 0$.

When $a > 0$ one can apply the change of variables $y = 2\pi a x$ to obtain

$$S(a) = i \int_{-\infty}^{\infty} dy \frac{\sin y}{y} \quad . \tag{B.4}$$

This integral can be evaluated by contour integration using Cauchy's theorem [*Conway,* 1973], resulting in a value of $\pi$.

When $a < 0$ one can apply the same change of variables $y = 2\pi a x$, but now the integration limits change sign, so

$$S(a) = i \int_{\infty}^{-\infty} dy \frac{\sin y}{y} = -i \int_{-\infty}^{\infty} dy \frac{\sin y}{y} = -i\pi \quad , \quad a < 0 \quad . \tag{B.5}$$

Collecting results one arrives at the final answer

$$S(a) = \int_{-\infty}^{\infty} dx \frac{e^{2\pi i a x}}{x} = \begin{cases} i\pi & , \quad a > 0 \\ 0 & , \quad a = 0 \\ -i\pi & , \quad a < 0 \end{cases} \quad . \tag{B.6}$$

# Appendix C Iterated kernel

It will be proven by induction that for $n > 0$

$$K_n(k,q) = \frac{q}{k^2}\frac{\ln^{2n-1}\left(\frac{q}{k}\right)}{(2n-1)!} \quad , \tag{C.1}$$

with

$$K_1(k,q) = \frac{q}{k^2}\ln\left(\frac{q}{k}\right) \quad . \tag{C.2}$$

It is easy to see that (C.2) follows from (C.1) for $n=1$. Suppose now that (C.1) is valid up to a certain value of $n$. Then

$$K_{n+1}(k,q) = \int_k^q dq_n K_n(k,q_n)K_1(q_n,k) = \frac{q}{k^2}\int_k^q \frac{dq_n}{q_n}\frac{\ln^{2n-1}\left(\frac{q_n}{k}\right)}{(2n-1)!}\ln\left(\frac{q}{q_n}\right) \quad . \tag{C.3}$$

Writing $\ln\left(\frac{q}{q_n}\right) = \ln\left(\frac{q}{k}\frac{k}{q_n}\right) = \ln\left(\frac{q}{k}\right) - \ln\left(\frac{q_n}{k}\right)$, (C.3) reads

$$K_{n+1}(k,q) = \frac{q}{k^2}\frac{1}{(2n-1)!}\left[-\int_k^q dq_n \frac{\ln^{2n}\left(\frac{q_n}{k}\right)}{q_n} + \ln\left(\frac{q}{k}\right)\int_k^q dq_n \frac{\ln^{2n-1}\left(\frac{q_n}{k}\right)}{q_n}\right] \quad . \tag{C.4}$$

Changing integration variables to $z = \frac{q_n}{k}$ with $k > 0$, (C.4) can be written as

$$K_{n+1}(k,q) = \frac{q}{k^2}\frac{1}{(2n-1)!}\left[-\int_1^{\frac{q}{k}} dz \frac{\ln^{2n} z}{z} + \ln\left(\frac{q}{k}\right)\int_1^{\frac{q}{k}} dz \frac{\ln^{2n-1} z}{z}\right] =$$

$$= \frac{q}{k^2}\frac{1}{(2n-1)!}\left[-\frac{\ln^{2n+1} z}{2n+1} + \ln\left(\frac{q}{k}\right)\frac{\ln^{2n} z}{2n}\right]_{z=1}^{z=\frac{q}{k}} = \tag{C.5}$$

$$= \frac{q}{k^2}\frac{1}{(2n-1)!}\ln^{2n+1}\left(\frac{q}{k}\right)\left[\frac{1}{2n} - \frac{1}{2n+1}\right] \quad .$$

The integrals follow from the relation

$$\frac{d}{dz}\frac{ln^{m+1} z}{m+1} = \frac{ln^m z}{z} \quad , \tag{C.6}$$

which holds for all $m \geq 0$ .The term in square brackets in (C.5) reduces to $(2n)^{-1}(2n+1)^{-1}$ so

$$K_{n+1}(k,q) = \frac{q}{k^2}\frac{\ln^{2n+1}\left(\frac{q}{k}\right)}{(2n+1)!} \quad . \tag{C.7}$$

By induction it follows that (C.1) holds for all values of $n$.

# Appendix D Double derivative

Suppose we want to evaluate the double derivative of an even function $f$ for zero argument. Since $f$ is even, its Taylor expansion around $x = 0$ contains only even powers of $x$. Up to 4th order one has

$$f(x) = a_0 + a_2 x^2 + a_4 x^4 \quad , \tag{D.1}$$

and the double derivative at $x = 0$ is

$$f''(0) = 2a_2 \quad . \tag{D.2}$$

If $f$ is given for $x = j\Delta,\ j = 0,1,...,N$, and $f_j = f(j\Delta)$

$$\begin{aligned} f_0 &= a_0 \quad , \\ f_1 &= a_0 + a_2\Delta^2 + a_4\Delta^4 \quad , \\ f_2 &= a_0 + 4a_2\Delta^2 + 16a_4\Delta^4 \quad . \end{aligned} \tag{D.3}$$

Solving for $a_2$ yields

$$a_2 = \frac{-15f_0 + 16f_1 - f_2}{12\Delta^2} \quad . \tag{D.4}$$

So $f''(0)$ follows from (D.2) and (D.4).

# Appendix E Gaussian test

Substituting (3.13a-b) into (2.17b) yields

$$
\begin{aligned}
I(0) &= \int_0^\infty dr\,\frac{\rho_{tt}(r)-\rho_{ll}(r)}{r} = \\
&= \int_0^\infty dr\,\frac{1}{\mathrm{r}}\left[(1-\nu^2)\left(1-2\frac{r^2}{R_\psi^2}\right)e^{-\frac{r^2}{R_\psi^2}}+\nu^2 e^{-\frac{r^2}{R_\chi^2}}-(1-\nu^2)e^{-\frac{r^2}{R_\psi^2}}-\nu^2\left(1-2\frac{r^2}{R_\chi^2}\right)e^{-\frac{r^2}{R_\chi^2}}\right] = \\
&= \int_0^\infty dr\,\frac{1}{\mathrm{r}}\left[(1-\nu^2)\left(1-2\frac{r^2}{R_\psi^2}-1\right)e^{-\frac{r^2}{R_\psi^2}}-\nu^2\left(1-2\frac{r^2}{R_\chi^2}-1\right)e^{-\frac{r^2}{R_\chi^2}}\right] = \\
&= -\frac{2(1-\nu^2)}{R_\psi^2}\int_0^\infty dr\,re^{-\frac{r^2}{R_\psi^2}}+\frac{2\nu^2}{R_\chi^2}\int_0^\infty dr\,re^{-\frac{r^2}{R_\chi^2}} \quad .
\end{aligned}
\tag{E.1}
$$

Using

$$
\int_0^\infty dx\,xe^{-a^2x^2} = \frac{1}{2a^2} \quad , \tag{E.2}
$$

this reduces to

$$
I(0) = -\frac{2(1-\nu^2)}{R_\psi^2}\frac{R_\psi^2}{2}+\frac{2\nu^2}{R_\chi^2}\frac{R_\chi^2}{2} = 2\nu^2-1 \quad . \tag{E.3}
$$

In the same way one obtains

$$
\begin{aligned}
J(\infty) &= \int_0^\infty dr\,r\left(\rho_{tt}(r)+\rho_{ll}(r)\right) = \\
&= \int_0^\infty dr\,\mathrm{r}\left[(1-\nu^2)\left(1-2\frac{r^2}{R_\psi^2}+1\right)e^{-\frac{r^2}{R_\psi^2}}+\nu^2\left(1+1-2\frac{r^2}{R_\chi^2}\right)e^{-\frac{r^2}{R_\chi^2}}\right] = \\
&= 2(1-\nu^2)\int_0^\infty dr\,re^{-\frac{r^2}{R_\psi^2}}-\frac{2(1-\nu^2)}{R_\psi^2}\int_0^\infty dr\,r^3e^{-\frac{r^2}{R_\psi^2}}+2\nu^2\int_0^\infty dr\,re^{-\frac{r^2}{R_\chi^2}}-\frac{2\nu^2}{R_\chi^2}\int_0^\infty dr\,r^3e^{-\frac{r^2}{R_\chi^2}} \quad .
\end{aligned}
\tag{E.4}
$$

With (E.3) and

$$\int_0^\infty dx\, x^3 e^{-a^2x^2} = \frac{1}{2a^4} \quad , \tag{E.5}$$

one finds

$$J(\infty) = 0 \quad . \tag{E.6}$$

# Appendix F Rotation

The rotation of a given wind cell reads

$$R = \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} \quad . \tag{F.1}$$

For gridded WVC's the rotation $R_{i,j}$ of the cell at position $(x_i, y_j)$ can be obtained from (F.1) using second order finite differences as

$$R_{i,j} = \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta} + \frac{v_{i+1,j} - v_{i-1,j}}{2\Delta} \quad , \tag{F.2}$$

with $\Delta$ the scatterometer grid size which is assumed to be the same in both directions. The neighbouring WVC's should, of course, contain valid winds.

*Bourassa and McBeth Ford* [2010] apply the circulation theorem to arrive at

$$R_{i,j} = \frac{1}{\Delta^2} \oint d\mathbf{l} \cdot \mathbf{v} \quad , \tag{F.3}$$

where $\mathbf{v} = (u, v)$ and $d\mathbf{l}$ the line element tangent to the contour ABCDEFGH sketched in figure F.1.

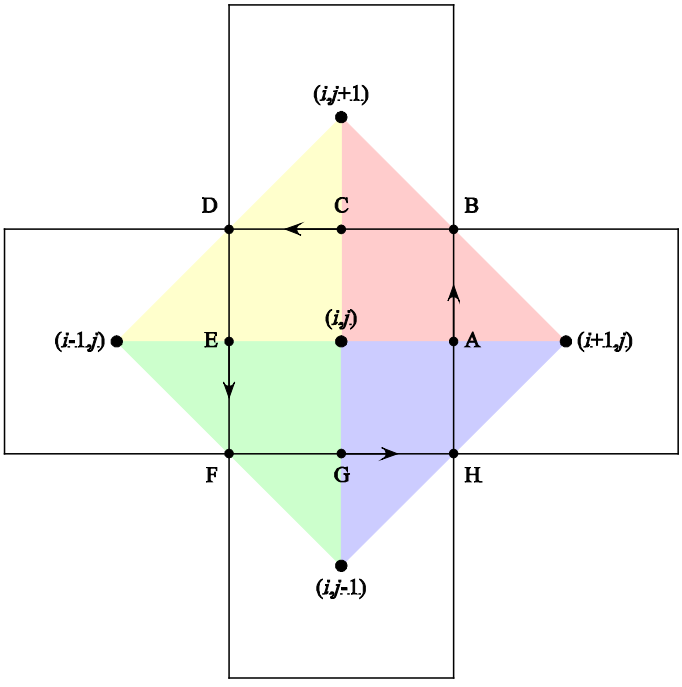


**Figure F.1** Sketch of the WVC's (solid black squares) with their centres (large dots) and integration contour points (small dots)

Equation (F.3) can easily be generalized for any contour. The factor $\Delta^2$ is then replaced by the area enclosed by the contour.

It will be shown now that (F.2) and (F.3) are identical for a single WVC when using linear interpolation. Figure F.1 shows that the wind components for every point on the integration contour ABCDEFGH can be obtained by linear interpolation between the wind components at the central cell with indices $(i, j)$ and its four neighbours with indices $(i+1, j)$, $(i, j+1)$, $(i-1, j)$, and $(i, j-1)$. The interpolation regions are indicated by the coloured triangles, and the average wind in each cell is attributed to its centre.

Since the wind may be linearly interpolated in each coloured triangle, the wind halfway two points with known winds is simply the average of these known values. Now each of the integration contour points A-H lies halfway two grid points, so the wind components at the contour points are easily obtained.

When written in components, the integrand $d\mathbf{l} \cdot \mathbf{v}$ reduces to $dx\, u + dy\, v$. The first term contributes only along FGH and, with a minus sign, along BDE; the second term along HAB and, with a minus sign, along DEF.

Finally, if $f(x)$ is a linear function in $[x_1, x_2]$ then

$$\int_{x_1}^{x_2} dx\, f(x) = \int_{x_1}^{x_2} dx \left( \frac{f_2 - f_1}{x_2 - x_1} x + f_1 - \frac{f_2 - f_1}{x_2 - x_1} x_1 \right) = \tfrac{1}{2}(f_1 + f_2)(x_2 - x_1) \quad , \tag{F.4}$$

where $f_i = f(x_i)$. Equation (F.4) simply states that the integral of a linear function over an interval equals the average of this function times the size of the interval.

Now the integral (F.3) over the contour sketched in figure (F1) can be evaluated as

$$\begin{aligned}
R_{i,j} = {} & \frac{1}{\Delta^2}\left[ -\tfrac{1}{2}(u_B + u_C)\tfrac{1}{2}\Delta - \tfrac{1}{2}(u_C + u_D)\tfrac{1}{2}\Delta + \tfrac{1}{2}(u_F + u_G)\tfrac{1}{2}\Delta + \tfrac{1}{2}(u_G + u_H)\tfrac{1}{2}\Delta \right] + \\
& + \frac{1}{\Delta^2}\left[ \tfrac{1}{2}(v_H + v_A)\tfrac{1}{2}\Delta + \tfrac{1}{2}(v_A + v_B)\tfrac{1}{2}\Delta - \tfrac{1}{2}(v_D + v_E)\tfrac{1}{2}\Delta - \tfrac{1}{2}(v_E + v_F)\tfrac{1}{2}\Delta \right] = \\
& = \frac{1}{2\Delta}\left[ -\frac{u_B + 2u_C + u_D}{2} + \frac{u_F + 2u_G + u_H}{2} + \frac{v_H + 2v_A + v_B}{2} - \frac{v_D + 2v_E + v_F}{2} \right] = \\
& = \frac{1}{2\Delta}\left[ -\frac{4u_{i,j+1} + u_{i+1,j} + 2u_{i,j} + u_{i-1,j}}{4} + \frac{4u_{i,j-1} + u_{i+1,j} + 2u_{i,j} + u_{i-1,j}}{4} \right] + \\
& + \frac{1}{2\Delta}\left[ \frac{4v_{i+1,j} + v_{i,j-1} + 2v_{i,j} + v_{i,j+1}}{4} - \frac{4v_{i-1,j} + v_{i,j-1} + 2v_{i,j} + v_{i,j+1}}{4} \right] \quad .
\end{aligned} \tag{F.5}$$

This simplifies to

$$R_{i,j} = \frac{1}{2\Delta}\left[ u_{i,j+1} - u_{i,j-1} + v_{i+1,j} - v_{i-1,j} \right] \quad , \tag{F.6}$$

which is exactly equal to (F.2).