DANISH METEOROLOGICAL INSTITUTE

SCIENTIFIC REPORT

98-15

# Variational analysis using spatial filters

Xiang-Yu Huang

DMI

COPENHAGEN 1998

ISSN-Nr. 0905-3263
ISBN-Nr. 87-7478-383-1

# Variational analysis using spatial filters

Xiang-Yu Huang*
Danish Meteorological Institute, Denmark

December 8, 1998

**Abstract**

In this note the standard variational analysis scheme is modified, through a simple transform, to avoid the inversion of the background error covariance matrix. A close inspection of the modified scheme reveals that it is possible to use a filter to replace the multiplication of the covariance matrix. A variational analysis scheme using a filter is then formulated, which does not explicitly involve the covariance matrix. The modified scheme and the filter scheme have the advantage of avoiding the inversion or any usage of the large matrix for analyses using gridpoint representation. To illustrate the use of these schemes a small-sized analysis problem, for which the best linear unbiased estimate is easily obtainable as a reference to solutions of different schemes, is considered. It is shown that both the modified scheme and the filter scheme work well. Compared with the standard and modified schemes the storage and computational requirements of the filter scheme could be reduced by several orders of magnitude for realistic atmospheric applications.

*Corresponding author address: Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen Ø, Denmark. Email: xyh@dmi.dk

# 1 Introduction

Objective atmospheric data analysis is widely used in observational studies, numerical weather prediction (NWP) and for verifying new models and theories. In many analysis schemes, observations are optimally used together with a background (also called firstguess) field in a statistical sense. At most NWP centers, a particular implementation of an Optimal Interpolation (OI) (Eliassen, 1954; Gandin, 1963) or a VARiational data assimilation (VAR) (Lorenc, 1986) is used. In the OI and VAR formulations, the inverse of the observation and background error covariance matrices, $\mathbf{R}$ and $\mathbf{B}$, is required. Direct inversions are difficult or even impossible due to the sizes of $\mathbf{R}$ and $\mathbf{B}$, which are of the order $10^5 \times 10^5$ and $10^7 \times 10^7$ respectively for realistic atmospheric applications. Special considerations are needed in each implementation.

In most OI implementations $\mathbf{B}$ is computed at the observation locations, often denoted as $\mathbf{P}$. Furthermore, a data selection algorithm, *e.g.* the box structures in the OI implementation at the European Centre for Medium range Weather Forecast, is needed to reduce $\mathbf{P}$ and $\mathbf{R}$ to a manageable size.

In most VAR implementations there is no OI-type data selection. The most common assumption on $\mathbf{R}$ is that the observation errors are not correlated (for most variables), which makes $\mathbf{R}$ diagonal [Courtier *et al.* (1998)]. Using spectral representations, $\mathbf{B}$ is also diagonal (Parrish and Derber, 1992; Courtier *et al.*, 1998; Gustafsson *et al.*, 1998). The inversion of $\mathbf{R}$ and $\mathbf{B}$ is thus a trivial issue.

However, if an implementation of the VAR scheme using a gridpoint formulation is desirable, the inversion of $\mathbf{B}$ needs to be considered. There are at least two potential problems due to the size of $\mathbf{B}$. The first is how to invert $\mathbf{B}$. This is impossible for a realistic atmospheric implementation. The second is how to handle the storage and computation related to $\mathbf{B}$. This is a difficult issue in operational implementations.

In this note, we address the first problem by modifying the standard variational analysis scheme so that the inversion is not needed, and then address the second problem by formulating a new scheme which uses a filter to replace the use of the matrix so that $\mathbf{B}$ itself does not enter the analysis. This note is organized as follows. The standard, the modified and the new schemes are described in Section 2 together with comparisons to other related schemes. An illustrative example is given in Section 3 using real surface temperature observations. The conclusions are summarized in Section 4.

# 2 Analysis methods

## 2.1 VAR: variational analysis

Briefly speaking, a variational analysis can be formulated to minimize the following cost function (Lorenc, 1986):

$$J(\mathbf{x}) = J_b(\mathbf{x}) + J_o(\mathbf{x}) = \frac{1}{2}\left[\left(\mathbf{x}-\mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) + (H\mathbf{x}-\mathbf{y})^T \mathbf{R}^{-1}(H\mathbf{x}-\mathbf{y})\right],$$

where, $\mathbf{x}$ is the analysis vector (on model gridpoints), $\mathbf{x}^b$ is the background vector (on model gridpoints), $\mathbf{y}$ is the observation vector (on observation locations), $H$ is the (linear) observation operator which converts a model variable from model gridpoints to observation locations, $\mathbf{B}$ is the background error covariance matrix, $\mathbf{R}$ is the observation error covariance matrix which includes the effects of both instrument and representativeness, $(.)^T$ indicates transpose, and $(.)^{-1}$ indicates inversion. The notations closely follow Ide *et al.* (1997).

Starting with an initial guess field $\mathbf{x}^0$, normally $\mathbf{x}^0 = \mathbf{x}^b$, the final analysis $\mathbf{x}^{VAR} = \mathbf{x}^\infty$ can be obtained using a minimization algorithm. The simplest method is the so-called steepest descent, which can be expressed as

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \alpha \nabla_{\mathbf{x}} J, \tag{1}$$

where, the superscript denotes the iteration cycle number, $\alpha$ is a constant ($\alpha > 0$), and $\nabla_{\mathbf{x}} J$ is the gradient of the cost function $J$ with respect to $\mathbf{x}$:

$$\nabla_{\mathbf{x}} J = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) + H^T \mathbf{R}^{-1}(H\mathbf{x}-\mathbf{y}). \tag{2}$$

Assuming $N$ model gridpoints and $M$ observation locations, then $\mathbf{x}$ and $\mathbf{x}^b$ are vectors of length $N$, $\mathbf{y}$ is a vector of length $M$, $\mathbf{B}$ is a $N \times N$ matrix, and $\mathbf{R}$ is a $M \times M$ matrix.

## 2.2 VAN: Variational analysis with No inversion of B

To avoid the inversion of $\mathbf{B}$, a new vector $\mathbf{z}$ is introduced, defined as

$$\mathbf{z} = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) \tag{3}$$

and the cost function $J$ can now be written as

$$J(\mathbf{z}) = J_b(\mathbf{z}) + J_o(\mathbf{z}) = \frac{1}{2}\left[\mathbf{z}^T\mathbf{B}^T\mathbf{z} + \left(H\mathbf{B}\mathbf{z} + \delta\mathbf{y}^b\right)^T \mathbf{R}^{-1}\left(H\mathbf{B}\mathbf{z} + \delta\mathbf{y}^b\right)\right],$$

where, $\delta\mathbf{y}^b$ is the background innovation vector defined as

$$\delta\mathbf{y}^b = H\mathbf{x}^b - \mathbf{y} \tag{4}$$

which remains unchanged during the iterations. Using $\mathbf{z}$ as the control variable, the minimization scheme (1) is now re-written as,

$$\mathbf{z}^{n+1} = \mathbf{z}^n - \alpha \nabla_{\mathbf{z}} J, \tag{5}$$

where $\nabla_{\mathbf{z}}$ is the gradient of $J$ with respect to $\mathbf{z}$:

$$\nabla_{\mathbf{z}} J = \mathbf{B}^T \left[ \mathbf{z} + H^T \mathbf{R}^{-1} \left( H\mathbf{B}\mathbf{z} + \delta \mathbf{y}^b \right) \right].$$

Starting with an initial guess field $\mathbf{z}^0$,

$$\mathbf{z}^0 = \mathbf{B}^{-1} \left( \mathbf{x}^0 - \mathbf{x}^b \right),$$

the minimum of $J$ is reached at $\mathbf{z}^\infty$ and the final analysis $\mathbf{x}^{VAN}$ can easily be calculated as

$$\mathbf{x}^{VAN} = \mathbf{x}^b + \mathbf{B}\mathbf{z}^\infty.$$

In this scheme $\mathbf{B}^{-1}$ is only needed to provide the initial guess value for the control variable. Consequently, if we assume

$$\mathbf{z}^0 = 0,$$

we have formulated a variational analysis scheme which needs no inversion of $\mathbf{B}$. This assumption is certainly a reasonable one. It simply uses the background vector as the initial analysis vector as done in most existing analysis schemes.

The incremental approach used in (3) is necessary. Otherwise, $\mathbf{B}^{-1}$ is needed to transform the background vector $\mathbf{x}^b$ to $\mathbf{B}^{-1}\mathbf{x}^b$ in the same way as the analysis vector. Furthermore, it would be unreasonable to assume $\mathbf{B}^{-1}\mathbf{x}^0 = 0$.

The modified scheme is named VAN, Variational Analysis with No inversion of $\mathbf{B}$, and is summarized as:

$$\begin{aligned} \mathbf{z}^0 &= 0 \\ \nabla_{\mathbf{z}} J &= \mathbf{B}^T \left[ \mathbf{z} + H^T \mathbf{R}^{-1} \left( H\mathbf{B}\mathbf{z} + \delta \mathbf{y}^b \right) \right] \\ \mathbf{z}^{n+1} &= \mathbf{z}^n - \alpha \nabla_{\mathbf{z}} J \\ \mathbf{x}^{VAN} &= \mathbf{x}^b + \mathbf{B}\mathbf{z}^\infty. \end{aligned} \tag{6}$$

The transform from $\mathbf{x}$ to $\mathbf{z}$ using (3) and the redefinition of the cost function with respect to $\mathbf{z}$ were used by Lorenc (1988) to speed up the convergence of descent algorithms such as the steepest descent and conjugate-gradient methods. In this note, VAN is used to provide guidance to formulate a new scheme which does not explicitly involve $\mathbf{B}$. VAN is also used as a reference when the new scheme is tested.

## 2.3 VAF: Variational Analysis using a Filter

In NWP applications, the matrix $\mathbf{B}$ ($N \times N \sim 10^{14}$) is too large to be handled by present computers and therefore has to be calculated every time it is needed. In VAN, $\mathbf{B}$ is required during each iteration loop. This makes an operational implementation of VAN difficult. However, a close inspection of VAN, in particular of the gradient of the cost function $\nabla_{\mathbf{z}} J$, reveals that $\mathbf{B}$ may be parameterized by a spatial filter.

For example, consider a two-dimensional uni-variate problem with homogeneous and isotropic background error correlations. The following Gaussian function could be used to calculate $\mathbf{B}$ (Daley, 1991):

$$b_{ij} = \sigma_b^2 \exp\left[-\left(\frac{r_{ij}}{L}\right)^2\right], \tag{7}$$

where, $b_{ij}$ are the elements of $\mathbf{B}$ ($i = 1, N; j = 1, N$), $\sigma_b$ is the standard deviation of the background error, $r_{ij}$ is the distance between model gridpoints $i$ and $j$, and $L$ is a length scale which can be determined theoretically or by observations. Multiplying a vector by the $\mathbf{B}$ matrix constructed by (7) is equivalent to applying a Gaussian filter to the vector. Therefore the VAN scheme can be written by using a filter to replace $\mathbf{B}$. The scheme is named VAF, Variational Analysis using a Filter, and is summarized as:

$$\begin{aligned} \mathbf{z}^0 &= 0 \\ \nabla_{\mathbf{z}} J &= G\left[\mathbf{z} + H^T \mathbf{R}^{-1}\left(HG\mathbf{z} + \delta\mathbf{y}^b\right)\right] \\ \mathbf{z}^{n+1} &= \mathbf{z}^n - \alpha \nabla_{\mathbf{z}} J \\ \mathbf{x}^{VAF} &= \mathbf{x}^b + G\mathbf{z}^\infty \end{aligned} \tag{8}$$

where $G$ is a spatial filter which should be designed based on the *a priori* knowledge of the covariance matrix $\mathbf{B}$.

Re-write $\mathbf{x}(N)$ as $\mathbf{x}(N_\lambda, N_\phi)$, where $N_\lambda \times N_\phi = N$, $\lambda$ and $\phi$ are coordinates in the west-east and south-north directions, respectively. A two-dimensional (Gaussian) filter is chosen for $G$, with the filter coefficients calculated by (7), and a Lanczos window is included to reduce the truncation related Gibbs oscillations:

$$\begin{aligned} \mathbf{x}^* &= G(\mathbf{x}) \\ x^*(n_\lambda, n_\phi) &= \sum_{i=-I/2}^{I/2} \sum_{j=-J/2}^{J/2} g(i,j) x(n_\lambda + i, n_\phi + j) \\ g(i,j) &= w(i,j)\sigma^2 \exp\left[-\frac{(i\Delta s_\lambda)^2 + (j\Delta s_\phi)^2}{L^2}\right] \\ w(i,j) &= \frac{\sin\left[i\pi/(I/2+1)\right]}{i\pi/(I/2+1)} \frac{\sin\left[j\pi/(J/2+1)\right]}{j\pi/(J/2+1)}, \end{aligned} \tag{9}$$

where, $\mathbf{x}^*$ is a filtered matrix with its elements $x^*(n_\lambda, n_\phi)$, $g(i,j)$ is the (Gaussian) filter coefficient, $w(i,j)$ is the Lanczos window, $I$ and $J$ are the filter orders in $\lambda$- and $\phi$-direction, $\Delta s_\lambda$ and $\Delta s_\phi$ are grid distance in $\lambda$- and $\phi$-direction.

Note that the index pair $(i, j)$ for $g(i, j)$ in (9) has a different meaning to $ij$ for $b_{ij}$ in (7). Here $(i, j)$ indicates one model gridpoint while $ij$ indicates two model gridpoints, $i$ and $j$.

The motivation for using a filter to parameterize $\mathbf{B}$ is to reduce the memory consumption of storing $\mathbf{B}$ $(N \times N)$ and the computational burden caused by multiplying $\mathbf{B}$ $(N \times N)$. The storage of the filter coefficients is $(I + 1)(J + 1)$ and the computation required is $< (I + 1)(J + 1) \times N$ ($<$ is used here because filter does not take values outside of the analysis domain). The choice of $I$ and $J$ depends on the characteristics of $\mathbf{B}$. Intuitively, the following approximations should hold:

$$\frac{1}{2} I \Delta\lambda \approx \frac{1}{2} J \Delta\phi \approx L.$$

Take a typical two-dimensional analysis domain, $200 \times 200$ gridpoints, with the grid distance 50 km and the error covariance scale $L = 200km$. The dimension of $\mathbf{B}$ and also the multiplications of $\mathbf{B}$ required by each VAN iteration are $N \times N = 2 \times 10^9$ ($10^{14}$ for 3-dimensional multi-variate analysis). The required size of the filter is approximately $(I + 1)(J + 1) \approx 10^2$ and the filter multiplications required by each VAF iteration is $(I+1)(J+1) \times N \approx 4 \times 10^6$. [For a VAR scheme using the spectral technique, the storage required for $b^*_{ii}$ and the multiplications in each iteration are $N = 4 \times 10^4$.]

When the background error at every analysis gridpoint is considered to be correlated to each other, $N_\lambda - 1$ and $N_\phi - 1$ should be chosen for $I/2$ and $J/2$. In this special case, the VAF scheme without the Lanczos window should give exactly the same solution as the VAN scheme (this particular case is useful for checking the VAF implementation).

Lorenc (1992) also proposed an analysis scheme which uses recursive filters to replace $\mathbf{B}$. Using our notations, his scheme may be derived as the following. Starting from (1)

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \alpha \nabla_{\mathbf{x}} J = \mathbf{x}^n - \alpha \mathbf{B}^{-1} \mathbf{B} \nabla_{\mathbf{x}} J,$$

denoting,

$$\mathbf{Q} = \alpha \mathbf{B}^{-1},$$

and using (2), the basic analysis scheme of Lorenc (1992) is obtained:

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \mathbf{Q} \left[ \left( \mathbf{x}^n - \mathbf{x}^b \right) + \mathbf{B} H^T \mathbf{R}^{-1} \left( H \mathbf{x}^n - \mathbf{y} \right) \right].$$

In this scheme, $\mathbf{B}^{-1}$ is only needed in $\mathbf{Q}$, which in practical implementations needs simplifications as outlined by Lorenc (1992). In the remaining scheme, $\mathbf{B}$ only occurs in front of the innovation vector, $H\mathbf{x}^n - \mathbf{y}$. Replacing $\mathbf{B}$ by a recursive filter in the above equation, it becomes the filter scheme of Lorenc (1992). The major difference between VAF and the Lorenc filter scheme is the control variable. The advantage of using $\mathbf{z}$ as proposed in VAN and VAF is in the calculation of the cost function gradient, which is required by most minimization algorithms. As has been discussed already, $\nabla_{\mathbf{x}} J$ requires the inverse of $\mathbf{B}$ while $\nabla_{\mathbf{z}} J$ does not.

## 2.4 BLUE: the Best Linear Unbiased Estimation

By setting the gradient of the cost function, (2), to zero:

$$\mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right)+H^T\mathbf{R}^{-1}\left(H\mathbf{x}-\mathbf{y}\right)=0.$$

the Best Linear Unbiased Estimation (BLUE) can be derived:

$$\mathbf{x}^{BLUE}=\mathbf{x}^b+\mathbf{B}H^T\left(H\mathbf{B}H^T+\mathbf{R}\right)^{-1}\left(\mathbf{y}-H\mathbf{x}^b\right) \tag{10}$$

where, $\mathbf{x}^{BLUE}$ is the BLUE solution, which should be the final solution of variational schemes and therefore can be used as a reference.

The numerical solution for BLUE depends on the accuracy of $(H\mathbf{B}H^T+\mathbf{R})^{-1}$, which could be problematic especially when $M$ is large. The size of the matrix $\mathbf{B}H^T$ is $N\times M$. The elements of this matrix are the background error covariance between all analysis gridpoints and observation locations. The size of $H\mathbf{B}H^T$ is $M\times M$. The elements of this matrix is the background error covariance between all observation locations. [Note that although observation locations are involved, the two matrices are for background error covariances.] In real NWP and atmospheric data analysis applications, $\mathbf{B}$ is a very large matrix. Simplifications to $\mathbf{B}H^T$ (dimension $N\times M$) and $(H\mathbf{B}H^T+\mathbf{R})^{-1}$ lead to the widely used OI scheme. In most OI scheme implementations [*e.g.* Lönnberg and Shaw (1987)], different data selection strategies are used to reduce $M$ and $N$. With the reduced $M$, the multiplication by $\mathbf{B}H^T$ is coded directly to avoid memory allocation of the large matrix. The matrix $H\mathbf{B}H^T$, often named $\mathbf{P}$, is calculated using the background error covariances at the reduced (selected) observation locations ($\mathbf{B}$ is not involved).

## 2.5 PSAS: Physical-space Statistical Analysis Scheme

The Physical-space Statistical Analysis Scheme (PSAS) described by Cohn *et al.* (1998) is another variational analysis scheme avoiding the inversion of $\mathbf{B}$. The BLUE solution can be re-written as:

$$\mathbf{x}^{BLUE}=\mathbf{x}^b+\mathbf{B}H^T\mathbf{w}$$
$$\mathbf{w}=\left(H\mathbf{B}H^T+\mathbf{R}\right)^{-1}\left(\mathbf{y}-H\mathbf{x}^b\right).$$

Instead of solving $\mathbf{w}$ exactly, which needs to invert $\left(H\mathbf{B}H^T+\mathbf{R}\right)$ and has the same difficulties in practical applications as discussed in the previous subsection, PSAS obtains $\mathbf{w}$ through minimizing the following cost function:

$$J(\mathbf{w})=\frac{1}{2}\mathbf{w}^T\left(H\mathbf{B}H^T+\mathbf{R}\right)\mathbf{w}-\mathbf{w}^T\left(\mathbf{y}-H\mathbf{x}^b\right).$$

The PSAS solution is given as:

$$\mathbf{x}^{PSAS}=\mathbf{x}^b+\mathbf{B}H^T\mathbf{w}^\infty.$$

where $\mathbf{w}^\infty$ is the minimization result.

Although PSAS does not invert $\left(H\mathbf{B}H^T+\mathbf{R}\right)$, this $M\times M$ matrix could still cause significant computational expenses. In the implementation of Cohn *et al.* (1998), simplifications are used to make the matrix sparse, which in a way mimic the OI data selection.

The major difference between PSAS and VAN (VAF) is where the control variable is based. In PSAS, the minimization is performed at the observation locations. The dimension of the control variable, $\mathbf{w}$, is $M$ ($\sim 10^5$). The research efforts have been devoted in simplifying $\left(H\mathbf{B}H^T+\mathbf{R}\right)$, which is also in the observation space. In VAN, the minimization is performed at the model gridpoints. The dimension of the control variable, $\mathbf{z}$, is $N$ ($\sim 10^7$). Research efforts are also required in simplifying $\mathbf{B}$. As discussed in previous subsections, a simple spatial filter could be used to mimic $\mathbf{B}$.

Comparing PSAS and VAN, the former has a smaller sized minimization problem while the latter has a more regular matrix to be modeled or simplified.

# 3 An illustrative example

To demonstrate the VAN and VAF schemes described in Section 2, a small-sized two-dimensional uni-variate analysis problem is considered, for which the BLUE solutions are easily obtained as references to evaluate the new schemes. (A numerical lab, ANALAB, for VAR, VAN, VAF, OI and PSAS schemes, written in FORTRAN-77, has been developed by the author for education purposes at the Danish Meteorological Institute. The complete ANALAB is available from the author. ANALAB is based on the OILAB which has been used in undergraduate courses for many years at the Department of Meteorology, Stockholm University.)

## 3.1 Observations

Real surface temperature observations are used in this illustrative example. The observations are shown in Figure 1. They are collected mainly in Denmark and Southern Sweden at 0300 UTC 3 March 1992. Totally 38 observations are in the database ($M = 38$). The observation locations are indicated by stars. The observed values in °C are printed to the right of their locations.

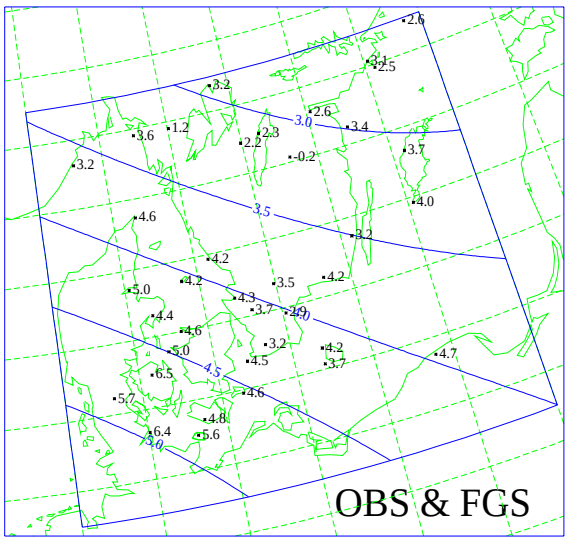


Figure 1: Surface temperature observations (locations are indicated by stars and values in °C) at 0300 UTC 3 March 1992 and the background (firstguess) temperature field, indicated by the contour lines, used by all analyses discussed in this note.

## 3.2 Analysis grid

A regular latitude-longitude grid is chosen. The number of analysis gridpoints is $21 \times 21$ ($N = 441$), which is much larger than the number of observations. The grid resolution

is 0.3° in latitude and 0.6° in longitude. The analysis domain is also shown in Figure 1. The center of the model domain is located at (56.5,14.0). Using this coordinate, the grid resolution is about 37 km in the east-west direction and 33 km in the north-south direction.

## 3.3 Background (firstguess)

The background (firstguess) temperature field is also shown in Figure 1. It is obtained by the BLUE scheme using the average of all the observations as its background. In order to have a smooth background field for later experiments $L = 1000$ km is used here (while 200 km and 100 km are used in the experiments). From the figure, it is clear that the large scale features of the temperature field are captured by the background. It is the purpose of the analysis with a smaller error covariance scale to extract the small scale information in the observations in a statistically optimal way.

## 3.4 R, B and $G$

In the following experiments, the observation errors are assumed non-correlated. The matrix $\mathbf{R}$ is therefore a diagonal matrix with all diagonal elements equal to the square of the observation standard deviation, $\sigma_o^2$, and we choose $\sigma_o = 1$°C.

The background errors are assumed correlated and the Gaussian function, defined in (7), is used to calculate the elements in $\mathbf{B}$. The background standard deviation, $\sigma_b$, is also chosen to be 1 °C. Using the analysis grid defined earlier, the error covariance between one arbitrary gridpoint and the neighboring gridpoints is calculated using (7) and plotted against distance in Figure 2. In the figure, b200 and b100 are covariances using $L = 200$ km and $L = 100$km, respectively. In the testing of the VAF code, the 40th order Gaussian filter without the Lanczos window is used and the filter coefficients, g200_40_nowin, are identical to b200 plotted in Figure 2.

From the figure, it is clear that the covariance become less than 10% only when two gridpoints are separated by more than 10 gridpoints when $L = 200$ km and by more than 5 gridpoints when $L = 100$ km. Using these numbers, a 20th order filter and a 10th order filter are chosen in the VAF experiments.

Due to the Lanczos window, the effective scale of the Gaussian filter is reduced. To compensate this scale reduction, $L = 300$ km is chosen for the 20th order filter and $L = 130$ km is chosen for the 10th order filter. The Gaussian filter coefficients, calculated using (9), as functions of distance are also plotted in Figure 2, where g300_20 is for a 20th order filter and g130_10 is for a 10th order filter. As can be seen from the figure, these two filters closely resemble the two covariance functions.

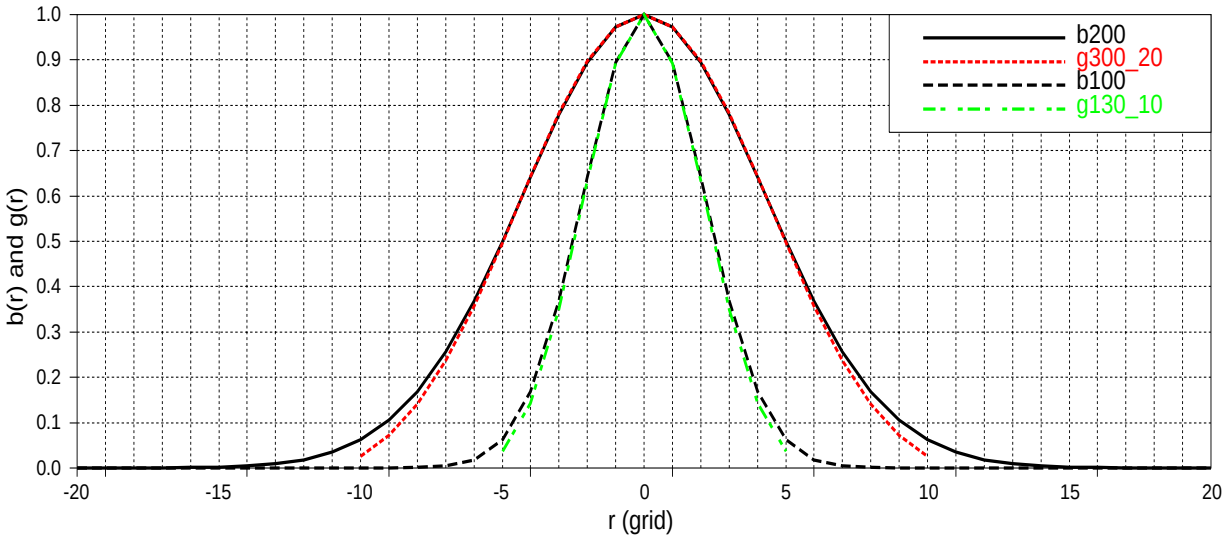


Figure 2: The background error covariance $[b(r)]$ and the Gaussian filter coefficients $[g(r)]$, as functions of distance $(r)$.

## 3.5 The BLUE solution

The BLUE analyses are shown in Figure 3. They are obtained by using $L = 200$ km and $L = 100$ km, respectively. As has been discussed in section 2.4, the BLUE solution is the final solution of the variational schemes assuming the inversion of the $M \times M$ matrix, $(H\mathbf{B}H^T + \mathbf{R})^{-1}$, to be exact. This solution is used as a reference to compare the results in the following subsections.

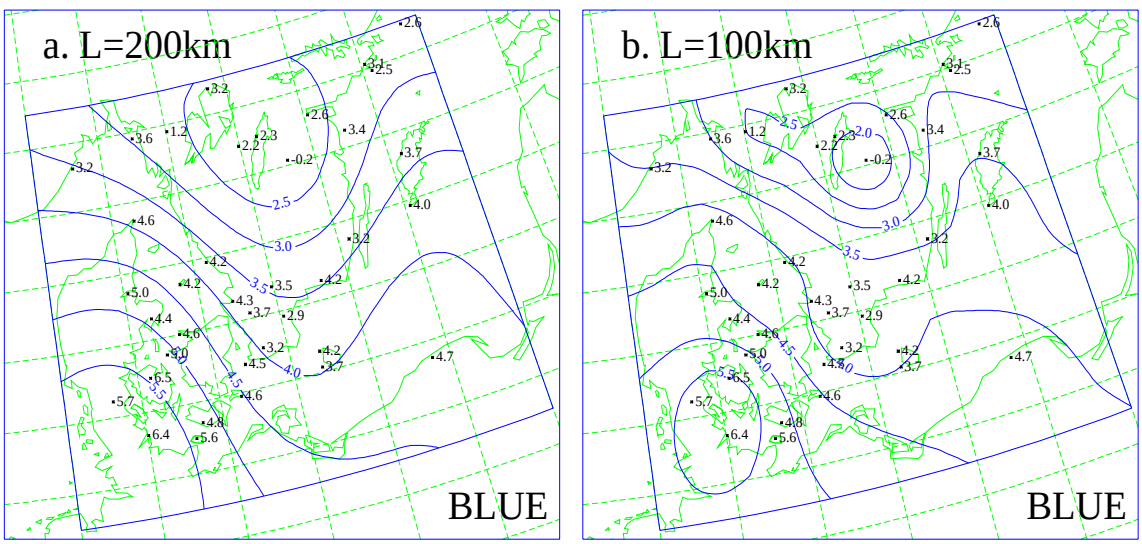


Figure 3: BLUE analyses using a) $L = 200$ km and b) $L = 200$km.

### 3.6 Minimization

To solve the variational problems discussed in Section 2 an algorithm for minimizing the cost function is needed. In this study, for illustration purposes, the steepest descent minimization algorithm is used. There is no general rule for the choice of the minimization step size, $\alpha$ used in (1) and (5). In the experiments described in this note, $\alpha = 0.002$ is chosen, which is approximately the largest value with which the steepest descent minimization method converges for all the experiments.

### 3.7 VAN solutions

In the experiment using the VAN scheme, the parameters are set to be the same as those used to obtain the BLUE solution with $L = 200$ km. The cost functions $J_o$, $J_b$ and $J$ at each iteration (up to 100) are plotted in Figure 4. Note that a logarithmic scale is used.

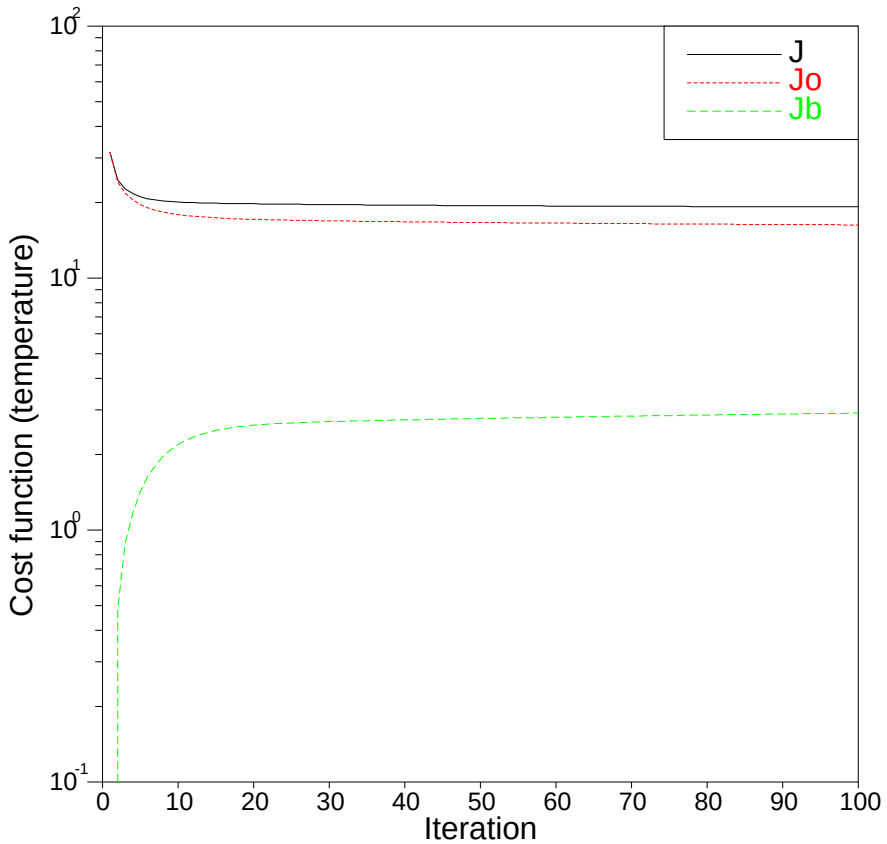


Figure 4: The cost functions $J_o$, $J_b$ and $J$ in logarithmic scale at each iteration.

The choice of the analysis grid, observation and background error covariances leads to $J_o/J_b \sim 5$, *i.e.* the difference between the cost functions of the analysis and observations is about 5 times larger than that between analysis and background. With different parameters this ratio can change.

As can be seen from Figure 4, most changes in cost function occur during the first iteration. Even 10 iterations are enough for practical use. At the end of the iterations (10000 iterations are performed although only 1000 iterations are shown), the VAN solution is actually indistinguishable to the BLUE solution.

The VAN analyses after 1, 10, 100 and 1000 iterations are shown in Figure 5. The VAN analysis obtained after 1 iteration is already quite different to the background and has many features of the BLUE solution. A detailed comparison with the BLUE solution (Figure 3) reveals that the VAN analysis converges to the expected one. The VAN analyses after 10, 100 and 1000 iterations could be considered the same using any subjective judgment, although minor differences, for instance close to the northern boundary, do exist.

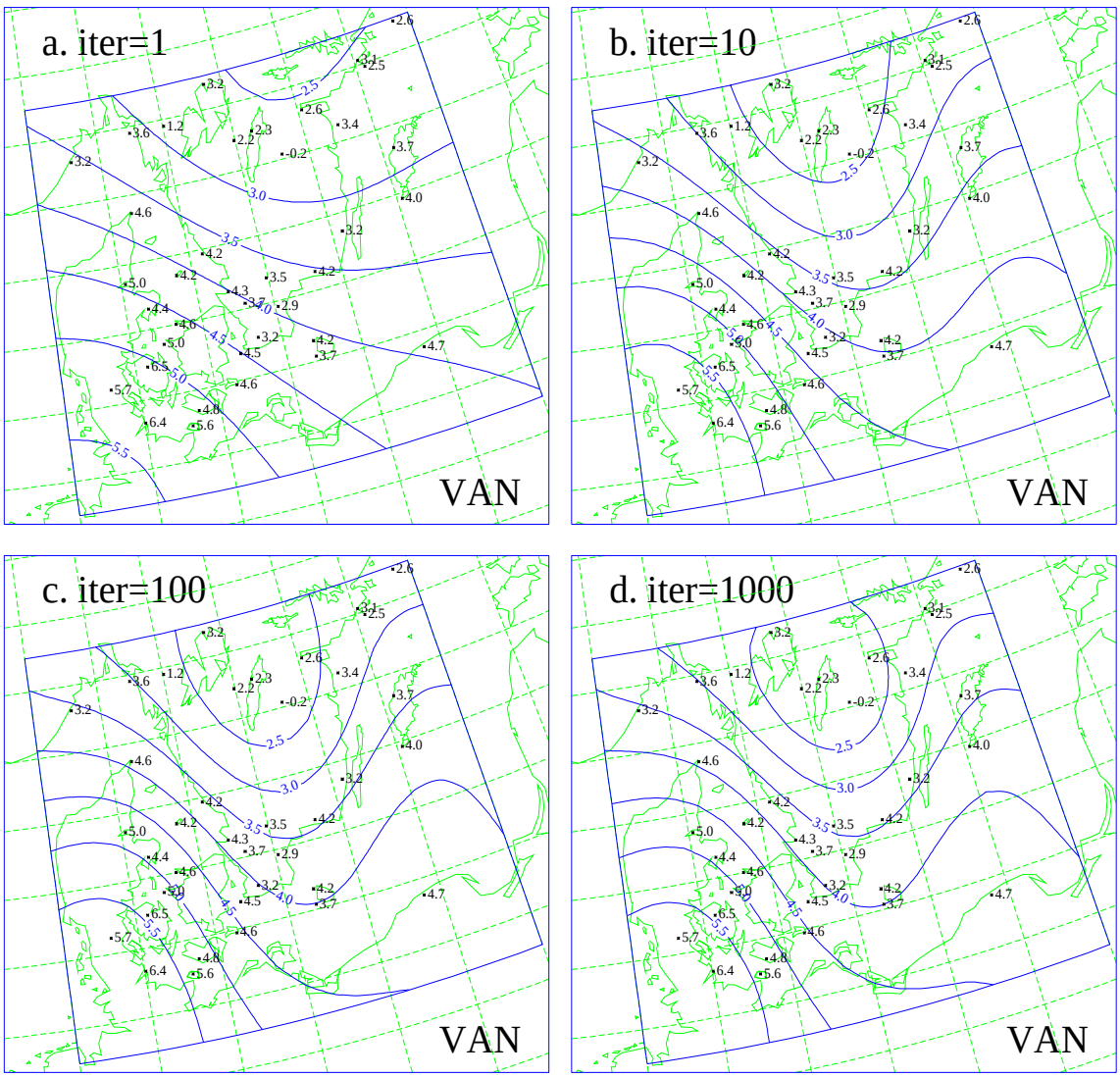


Figure 5: VAN analyses obtained after (a) 1 iteration; (b) 10 iterations; (c) 100 iterations and (d) 1000 iterations.

## 3.8 VAF solutions

Using the Gaussian filter (9) with different orders, a series of experiments using the VAF scheme are conducted. First, a 40th order filter without the Lanczos window is used as a check of the coding, since the VAF scheme is equivalent to the VAN scheme when the filter parameter $(I/2+1) \times (J/2+1)$ is equal to the number of gridpoints of the analysis domain $N_\lambda \times N_\phi$, which are all $21 \times 21$. The VAF results (not shown) are identical to the VAN results as expected.

Reducing the filter order by a factor of 2 ($I = J = 20$), the solutions (Figure 6) looks practically the same as the VAN solutions (Figure 5). The solutions after 10 iterations are actually very close to the BLUE solution using $L = 200$ km (Figure 3).

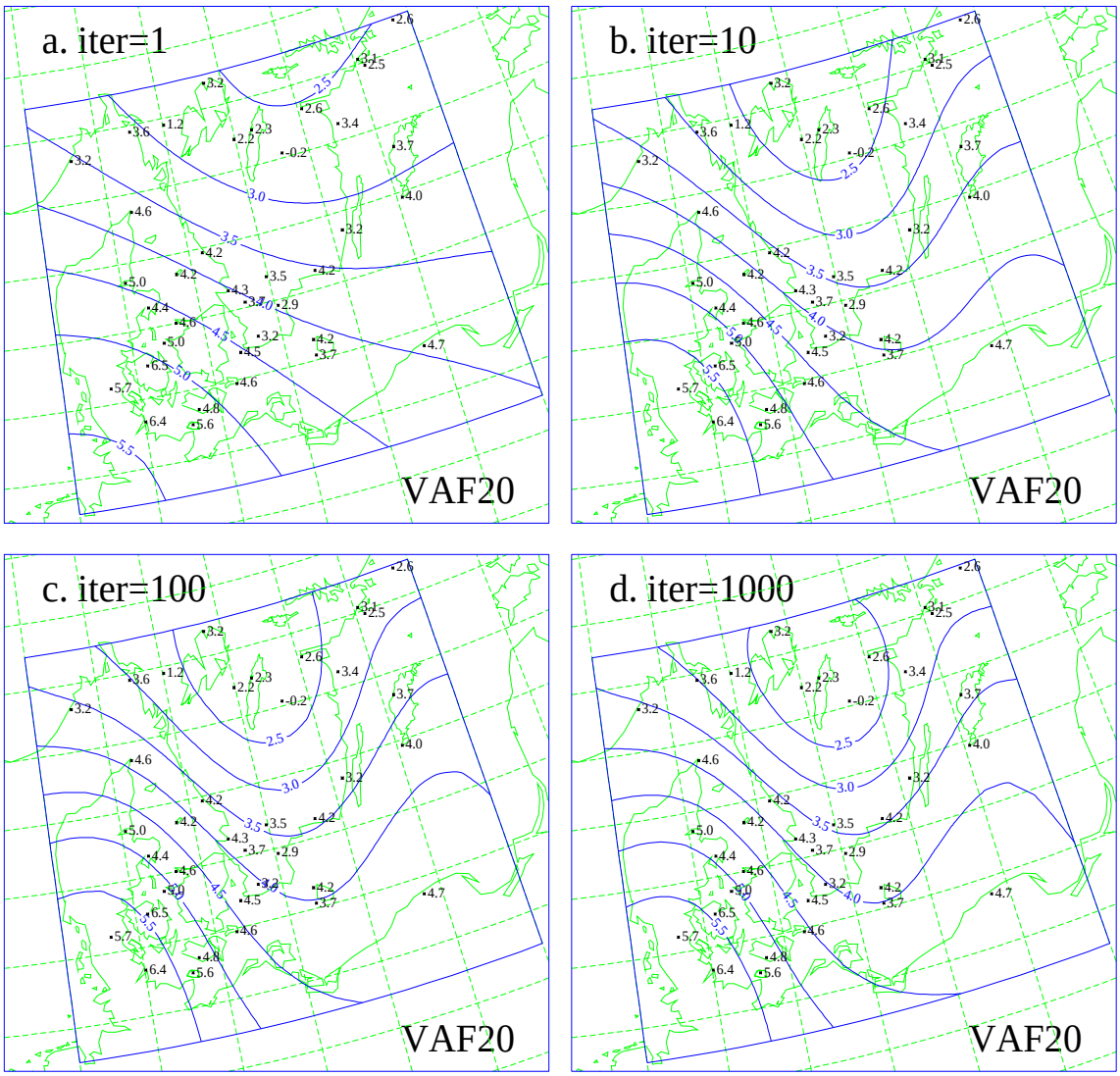


Figure 6: VAF analyses using a 20th order Gaussian filter after (1) 1 iteration; (b) 10 iterations; (c) 100 iterations and (d) 1000 iterations.

A further halving of the filter order ($I = J = 10$) leads to a VAF solution with a smaller spatial scale (Figure 7). As expected from the characteristics of the filter, the solutions converge to the BLUE solution using $L = 100$ km.

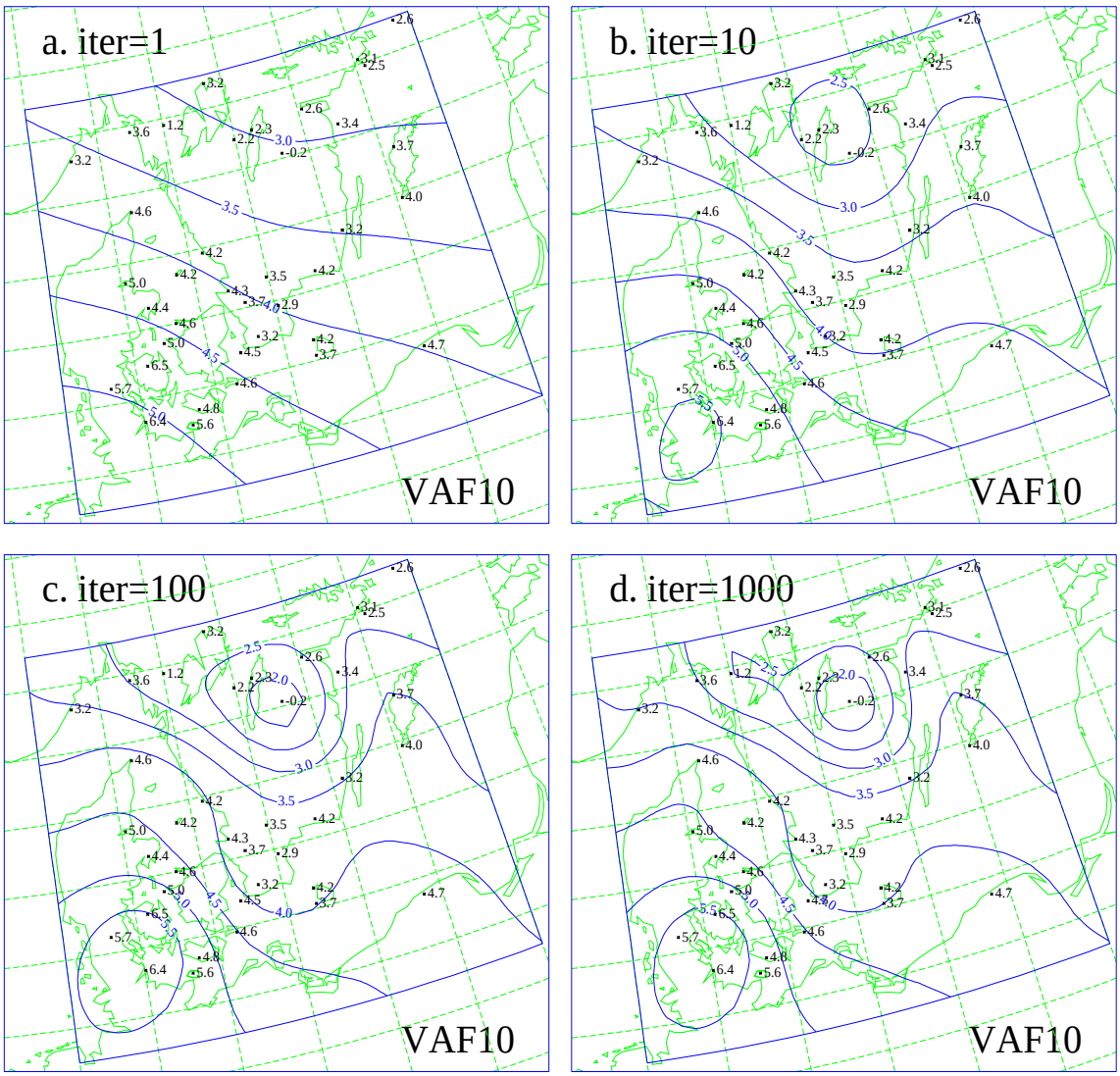


Figure 7: VAF analyses using a 10th order Gaussian filter after (a) 1 iteration; (b) 10 iterations; (c) 100 iterations and (d) 1000 iterations.

# 4 Conclusions

In this note, we have developed two schemes for variational data analysis. The first one, named VAN - Variational Analysis with No inversion of the background error covariance matrix - is a modified scheme of the standard variational analysis, VAR, as formulated by Lorenc (1986). The second one, named VAF - Variational Analysis using a Filter - is a new scheme based on the VAN scheme.

In the VAN scheme, a simple transform is used to avoid the inversion of the background error covariance matrix, which could be impossible for a realistic implementation of the VAR scheme using gridpoint representations. The remaining storage and computation difficulties in VAN is mainly due to the presence of the matrix itself. Although a full-sized implementation of the VAN scheme is still difficult, the VAN scheme is considered as a step forward from VAR. The problems related to the inversion of a very large and often ill-conditioned matrix are completely avoided.

In the VAF scheme, a spatial filter is used to replace the multiplications of the background error covariance matrix necessary in the VAN scheme. This replacement is based on the characteristics of the covariance matrix, which in many practical applications are assumed to have an analytical form. Using a truncated filter the spatial scale of the analysis is changed slightly. However, a re-definition of the length scale in the filter could be used to adjust the effective scale of the analysis. The storage and computation required by the VAF scheme could be orders of magnitude smaller than the VAR and VAN schemes. A full-sized implementation of the VAF scheme is possible.

To illustrate how these schemes work, a small-sized two-dimensional uni-variate analysis problem is taken with real surface temperature observations. Using the best linear unbiased estimate as the reference, it is shown that both VAN and VAF work satisfactorily.

For more realistic implementations, a full-sized three-dimensional multi-variant problem should be considered. A more sophisticated minimization algorithm may be necessary in such a situation. A non-uniform grid version of VAF needs to be derived as shown by Lorenc (1992). In order to make efficient use of asynoptic data and to obtain a coherent 4-dimensional analysis, a time dependence and a numerical weather prediction model should also be included by the VAN and VAF schemes. These will be research subjects in the future.

# 5 Acknowledgments

The author would like to thank Nils Gustafsson, Erland Källén, Leif Laursen, Andrew Lorenc, Peter Lynch and Henrik Vedel for comments on an earlier version of the manuscript.

# References

Cohn, S. E., Da Silva, A., Guo, J., Sienkiewicz, M. and Lamich, D. 1998. Assessing the effects of data selection with the DAO physical-space statistical analysis system. *Mon. Wea. Rev.*, **126**, 2913–2926.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. 1998. The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1808.

Daley, R. 1991. *Atmospheric Data Assimilation*. Cambridge University Press. 457pp.

Eliassen, A. 1954. *Provisional report on calculation of spatial covariance and autocorrelation of the pressure field*. Report 5. Videnskaps-Akademiet, Institut for Vaer och Klimaforskning (Norwegian Academy of Sciences, Institute of Weather and Climate Research), Oslo, Norway.

Gandin, L. 1963. *Objective Analysis of Meteorological Fields*. Leningrad: Gidromet; English translation Jerusalem: Israel Program for Scientific Translations, 1965.

Gustafsson, N., Hörnquist, S., Lindskog, M., Rantakokko, J., Berre, L., Navascues, B., Huang, X.-Y. and Thorsteinsson, S. 1998. *Three-dimensional variational data assimilation for a high resolution limited area model (HIRLAM)*. HIRLAM Tech. Rep. Available from Met Éireann, Glasnevin Hill, Dublin 9, Ireland.

Ide, K., Courtier, P., Ghil, M. and Lorenc, A.C. 1997. Unified Notation for data assimilation: Operational, Sequential and Variational. *J. Meteorol. Soc. Japan*, **75**, 181–189.

Lönnberg, P. and Shaw, D. (Eds.). 1987. *ECMWF Data Assimilation - Scientific Documentation. E.* ECMWF Research Manual 1, 2nd Revised Edition.

Lorenc, A.C. 1986. Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194.

Lorenc, A.C. 1988. Optimal nonlinear objective analysis. *Quart. J. Roy. Meteor. Soc.*, **114**, 205–240.

Lorenc, A.C. 1992. Interactive analysis using covariance functions and filters. *Quart. J. Roy. Meteor. Soc.*, **118**, 569–591.

Parrish, D.F. and Derber, J.C. 1992. The national meteorological center's global spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

## DANISH METEOROLOGICAL INSTITUTE

### Scientific Reports

Scientific reports from the Danish Meteorological Institute cover a variety of geophysical fields, i.e. meteorology (including climatology), oceanography, subjects on air and sea pollution, geomagnetism, solar-terrestrial physics, and physics of the middle and upper atmosphere.

Reports in the series within the last five years:

No. 94-1
**Bjørn M. Knudsen**: Dynamical processes in the ozone layer.

No. 94-2
**J. K. Olesen and K. E. Jacobsen**: On the atmospheric jet stream with clear air turbulences (CAT) and the possible relationship to other phenomena including HF radar echoes, electric fields and radio noise.

No. 94-3
**Ole Bøssing Christensen and Bent Hansen Sass**: A description of the DMI evaporation forecast project.

No. 94-4
**I.S. Mikkelsen, B. Knudsen**, E. Kyrö and M. Rummukainen: Tropospheric ozone over Finland and Greenland, 1988-94.

No. 94-5
**Jens Hesselbjerg Christensen, Eigil Kaas, Leif Laursen**: The contribution of the Danish Meteorological Institute (DMI) to the EPOCH project "The climate of the 21st century" No. EPOC-003-C (MB).

No. 95-1
**Peter Stauning** and T.J. Rosenberg: High-Latitude, Day-time Absorption Spike Events
1. Morphology and Occurrence Statistics.
Not Published.

No. 95-2
**Niels Larsen**: Modelling of changes in stratospheric ozone and other trace gases due to the emission changes : CEC Environment Program Contract No. EV5V-CT92-0079. Contribution to the final report.

No. 95-3
**Niels Larsen, Bjørn Knudsen, Paul Eriksen, Ib Steen Mikkelsen, Signe Bech Andersen and Torben Stockflet Jørgensen**: Investigations of ozone, aerosols, and clouds in the arctic stratosphere : CEC Environment Program Contract No. EV5V-CT92-0074. Contribution to the final report.

No. 95-4
**Per Høeg and Stig Syndergaard**: Study of the derivation of atmospheric properties using radio-occultation technique.

No. 95-5
Xiao-Ding Yu, **Xiang-Yu Huang** and **Leif Laursen** and Erik Rasmussen: Application of the HIRLAM system in China: Heavy rain forecast experiments in Yangtze River Region.

No. 95-6
**Bent Hansen Sass**: A numerical forecasting system for the prediction of slippery roads.

No. 95-7
**Per Høeg**: Proceeding of URSI International Conference, Working Group AFG1 Copenhagen, June 1995. Atmospheric Research and Applications Using Observations Based on the GPS/GLONASS System.

No. 95-8
**Julie D. Pietrzak**: A Comparison of Advection Schemes for Ocean Modelling.

No. 96-1
**Poul Frich** (co-ordinator), H. Alexandersson, J. Ashcroft, B. Dahlström, G.R. Demarée, A. Drebs, A.F.V. van Engelen, E.J. Førland, I. Hanssen-Bauer, R. Heino, T. Jónsson, K. Jonasson, L. Keegan, P.Ø. Nordli, **T. Schmith**, **P. Steffensen**, H. Tuomenvirta, O.E. Tveito: North Atlantic Climatological Dataset (NACD Version 1) - Final Report.

No. 96-2
**Georg Kjærgaard Andreasen**: Daily Response of High-latitude Current Systems to Solar Wind Variations: Application of Robust Multiple Regression. Methods on Godhavn magnetometer Data.

No. 96-3

**Jacob Woge Nielsen, Karsten Bolding Kristensen, Lonny Hansen**: Extreme sea level highs: A statistical tide gauge data study.

No. 96-4
**Jens Hesselbjerg Christensen, Ole Bøssing Christensen, Philippe Lopez**, Erik van Meijgaard, Michael Botzet: The HIRLAM4 Regional Atmospheric Climate Model.

No. 96-5
**Xiang-Yu Huang**: Horizontal Diffusion and Filtering in a Mesoscale Numerical Weather Prediction Model.

No. 96-6
**Henrik Svensmark and Eigil Friis-Christensen**:
Variation of Cosmic Ray Flux and Global Cloud Coverage - A Missing Link in Solar-Climate Relationships.

No. 96-7
**Jens Havskov Sørensen and Christian Ødum Jensen**: A Computer System for the Management of Epidemiological Data and Prediction of Risk and Economic Consequences During Outbreaks of Foot-and-Mouth Disease. CEC AIR Programme. Contract No. AIR3 - CT92-0652.

No. 96-8
**Jens Havskov Sørensen**: Quasi-Automatic of Input for LINCOM and RIMPUFF, and Output Conversion. CEC AIR Programme. Contract No. AIR3 - CT92-0652.

No. 96-9
**Rashpal S. Gill and Hans H. Valeur**:
Evaluation of the radarsat imagery for the operational mapping of sea ice around Greenland.

No. 96-10
**Jens Hesselbjerg Christensen**, Bennert Machenhauer, Richard G. Jones, Christoph Schär, Paolo Michele Ruti, Manuel Castro and Guido Visconti:
Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions.

No. 96-11
**Niels Larsen, Bjørn Knudsen, Paul Eriksen, Ib Steen Mikkelsen, Signe Bech Andersen and Torben Stockflet Jørgensen**: European Stratospheric Monitoring Stations in the Artic: An European contribution to the Network for Detection of Stratospheric Change (NDSC): CEC Environment Programme Contract EV5V-CT93-0333: DMI contribution to the final report.

No. 96-12
**Niels Larsen**: Effects of heterogeneous chemistry on the composition of the stratosphere: CEC Environment Programme Contract EV5V-CT93-0349: DMI contribution to the final report.

No. 97-1
**E. Friis Christensen og C. Skøtt:** Contributions from the International Science Team. The Ørsted Mission - A Pre-Launch Compendium.

No. 97-2
**Alix Rasmussen, Sissi Kiilsholm, Jens Havskov Sørensen, Ib Steen Mikkelsen:** Analysis of Tropospheric Ozone Measurements in Greenland: Contract No. EV5V-CT93-0318 (DG 12 DTEE): DMI's contribution to CEC Final Report Arctic Trophospheric Ozone Chemistry ARCTOC.

No. 97-3
**Peter Thejll:** A Search for Effects of External Events on Terrestrial Atmospheric Pressure: Cosmic Rays

No. 97-4
**Peter Thejll:** A Search for Effects of External Events on Terrestrial Atmospheric Pressure: Sector Boundary Crossings

No. 97-5
**Knud Lassen:** Twentieth Century Retreat of Sea-Ice in the Greenland Sea

No. 98-1
**Niels Woetman Nielsen, Bjarne Amstrup, Jess U. Jørgensen:**
HIRLAM 2.5 parallel tests at DMI:
Sensitivity to type of schemes for turbulence, moist processes and advection

No. 98-2
**Per Høeg, Georg Bergeton Larsen, Hans-Henrik Benzon, Stig Syndergaard, Mette Dahl Mortensen:**
The GPSOS project
Algorithm Functional Design and Analysis of ionosphere, Stratosphere and Troposphere Observations

No. 98-3
**Dahl Mortensen, Mette; Per Høeg:**
Satellite Atmosphere Profiling Retrieval in a Nonlinear Troposphere
Previously entitled: Limitations Induced by Multipath

No. 98-4

**Dahl Mortensen, Mette Per Høeg:**
Resolution Properties in Atmospheric Profiling with GPS

No. 98-5
**Gill, R. S. and M. K. Rosengren**
Evaluation of the Radarsat Imagery for the Operational Mapping of Sea Ice around Greenland in 1997.

No. 98-6
**Gill, R. S., Valeur, H. H., Nielsen, P. and Hansen, K. Q.:** Using ERS SAR images in the operational mapping of sea ice in the Greenland waters: final report for ESA-ESRIN's: pilot projekt no. PP2.PP2.DK2 and 2nd Announcement of opportunity for the exploitation of ERS data projekt No. AO2..DK 102.

No. 98-7
**Høeg, Per et al.:** GPS Atmosphere Profiling Methods and Error Assessments.

No. 98-8
**Svensmark, H., N. Woetmann Nielsen and A.M. Sempreviva:** Large Scale Soft and Hard Turbulent States of the Atmosphere.

No. 98-9
**Lopez, Philippe, Eigil Kaas and Annette Guldberg:** The Full Particle-In-Cell advection scheme in spherical geometry**.**

No. 98-10
**Svensmark, H.:** Influence of Cosmic Rays on Earth's Climate.

No. 98-11
**Thejll, Peter and Henrik Svensmark:** Notes on the method of normalized multivariate regression.

No. 98-12
**Lassen. K:** Extent of sea ice in the Greenland Sea 1977-1997: an extension of DMI Scientific Report 97-5

No. 98-13
**Larsen, Niels; Adriani, Alberto and Donfrancesco, Guido Di:** Microphysical analysis oof plar stratospheric clouds observed by Lidar at McMurco, Antarctica

No.98-14
**Dahl Mortensen, Mette:** The Back-Propagation Method for Inversion of Radio Occultation Data.