Proceedings of the 2014 Federated Conference on Computer Science and Information Systems pp. 587–595
DOI: 10.15439/2014F279
ACSIS, Vol. 2

An error estimate of Gaussian Recursive Filter in 3Dvar problem

Salvatore Cuomo
University of Naples Federico II
Department of Mathematics and Applications "R. Caccioppoli", Italy
Email: salvatore.cuomo@unina.itl

Ardelio Galletti
University of Naples "Parthenope"
Department of Science and Technology, Italy
Email: ardelio.galletti@uniparthenope.it

Raffaele Farina
Centro Euro-Mediterraneo sui Cambiamenti Climatici
CMCC, Italy
Email: raffaele.farina@cmcc.it

Livia Marcellino
University of Naples "Parthenope"
Department of Science and Technology, Italy
Email: livia.marcellino@uniparthenope.it

***Abstract*—Computational kernel of the three-dimensional variational data assimilation (3D-Var) problem is a linear system, generally solved by means of an iterative method. The most costly part of each iterative step is a matrix-vector product with a very large covariance matrix having Gaussian correlation structure. This operation may be interpreted as a Gaussian convolution, that is a very expensive numerical kernel. Recursive Filters (RFs) are a well known way to approximate the Gaussian convolution and are intensively applied in the meteorology, in the oceanography and in forecast models. In this paper, we deal with an oceanographic 3D-Var data assimilation scheme, named OceanVar, where the linear system is solved by using the Conjugate Gradient (GC) method by replacing, at each step, the Gaussian convolution with RFs. Here we give theoretical issues on the discrete convolution approximation with a first order (1st-RF) and a third order (3rd-RF) recursive filters. Numerical experiments confirm given error bounds and show the benefits, in terms of accuracy and performance, of the 3-rd RF.**

I. Introduction

In recent years, Gaussian filters have assumed a central role in image filtering and techniques for accurate measurement [26]. The implementation of the Gaussian filter in one or more dimensions has typically been done as a convolution with a Gaussian kernel, that leads to a high computational cost in its practical application. Computational efforts to reduce the Gaussian convolution complexity are discussed in [16], [24]. More advantages may be gained by employing a *spatially recursive filter*, carefully constructed to mimic the Gaussian convolution operator.

Recursive filters (RFs) are an efficient way of achieving a long impulse response, without having to perform a long convolution. Initially developed in the context of time series analysis [5], they are extensively used as computational kernels for numerical weather analysis, forecasts [17], [20], [25], digital image processing [8], [23]. Recursive filters with higher order accuracy are very able to accurately approximate a Gaussian convolution, but they require more operations.

In this paper, we investigate how the RF mimics the Gaussian convolution in the context of variational data assimilation analysis. Variational data assimilation (Var-DA) is popularly used to combine observations with a model forecast in order to produce a *best* estimate of the current state of a system and enable accurate prediction of future states. Here we deal with the three-dimensional data assimilation scheme (3D-Var), where the estimate minimizes a weighted nonlinear least-squares measure of the error between the model forecast and the available observations. The numerical problem is to minimize a cost function by means of an iterative optimization algorithm. The most costly part of each step is the multiplication of some grid-space vector by a covariance matrix that defines the error on the forecast model and observations. More precisely, in 3D-Var problem this operation may be interpreted as the convolution of a covariance function of background error with the given forcing terms.

Here we deal with numerical aspects of an oceanographic 3D-Var scheme, in the real scenario of OceanVar. Ocean data assimilation is a crucial task in operational oceanography and the computational kernel of OceanVar software is a linear system resolution by means of the Conjugate Gradient (GC) method, where the iteration matrix is relate to an errors covariance matrix, having a Gaussian correlation structure.

In [9], it is shown that a computational advantage can be gained by employing a first order RF that mimics the required Gaussian convolution. Instead, we use the 3rd-RF to compute numerically the Gaussian convolution, as how far is only used in signal processing [27], but only recently used in the field of Var-DA problems.

In this paper we highlight the main sources of error, introduced by these new numerical operators. We also investigate the real benefits, obtained by using 1-st and 3rd-RFs, through a careful error analysis. Theoretical aspects are confirmed by some numerical experiments. Finally, we report results in the case study of the OceanVar software.

The rest of the paper is organized as follows. In the next section we recall the three-dimensional variational data assimilation problem and we remark some properties on the

978-83-60810-58-3/$25.00

conditioning for this problem. Besides, we describe our case study: the OceanVar problem and its numerical solution with CG method. In section III, we introduce the n-th order recursive filter and how it can be applied to approximate the discrete Gaussian convolution. In section IV, we estimate the effective error, introduced at each iteration of the CG method, by using 1st-RF and 3rd-RF instead of the Gaussian convolution. In section V, we report some experiments to confirm our theoretical study, while the section VI concludes the paper.

II. Mathematical Background

The aim of a generic variational problem (VAR problem) is to find a best estimate x, given a previous estimate x_b and a measured value y. With these notations, the VAR problem is based on the following regularized constrained least-squared problem:

$$\min_x J(x)$$

where x is defined in a grid domain D. The objective function $J(x)$ is defined as follows:

$$J(x) = \|y - \mathcal{H}(x)\|^2 + \lambda R(x, x_b) \tag{1}$$

where measured data are compared with the solution obtained from a nonlinear model given by $\mathcal{H}(x)$.
In (1), we can recognize a quadratic data-fidelity term, the first term and the general regularization term (or penalty term), the second one. When $\lambda = 1$ and the regularization term can be write as:

$$R(x, x_b) = \|x - x_b\|^2$$

we deal with a three-dimensional variational data assimilation problem (3D-Var DA problem). The purpose is to find an optimal estimate for a vector of states x_t (called the analysis) of a generic system S, at each time $t \in T = \{0, .., n\}$ given:

- a prior estimate vector x_t^b (called the background) achieved by numerical solution of a forecasting model $\mathcal{L}_{t-1,t}(x_{t-1}) = x_t^b$, with error $\delta x_t = x_t^b - x_t$;
- a vector y_t of observations, related to the nonlinear model by δy_t that is an effective measurement error:

$$y_t = H(x_t) + \delta y_t.$$

At each time t, the errors δx_t in the background and the errors δy_t in the observations are assumed to be random with mean zero and covariance matrices $\mathbf{B}$ and $\mathbf{R}$, respectively. More precisely, the covariance $\mathbf{R} =< \delta y_t, \delta y_t^T >$ of observational error is assumed to be diagonal, (observational errors statistically independent). The covariance $\mathbf{B} =< \delta x_t, \delta x_t^T >$ of background error is never assumed to be diagonal as justified in the follow. To minimize, with respect to x_t and for each $t \in T$, the problem becomes:

$$\min_{x_t \in D} J(x_t) = \min_{x_t \in D} \{\frac{1}{2}\|y_t - H(x_t)\|^2_{\mathbf{R}} + \frac{1}{2}\|x_t - x_t^b\|^2_{\mathbf{B}}\} \tag{2}$$

In explicit form, the functional cost of (2) problem can be written as:

$$\begin{aligned} J(x_t) = \tfrac{1}{2}(y_t - H(x_t))^T\mathbf{R}^{-1}(y_t - H(x_t)) + \\ + \tfrac{1}{2}(x_t - x_t^b)^T\mathbf{B}^{-1}(x_t - x_t^b) \end{aligned} \tag{3}$$

It is often numerically convenient to approximate the effects on $H(x_t)$ of small increments of x_t, using the linearization of H. For small increments δx_t, follows [18], it is:

$$H(x_t) \simeq H(x_t^b) + \mathbf{H}\delta x_t$$

where the linear operator $\mathbf{H}$ is the matrix obtained by the first order approximation of the Jacobian of H evaluated at x_t^b.
Now let $d_t = y_t - H(x_t^b)$ be the *misfit*. Then the function J in (3) takes the following form in the increment space:

$$J(\delta x_t) = \tfrac{1}{2}(d_t - \mathbf{H}\delta x_t)^T\mathbf{R}^{-1}(d_t - \mathbf{H}\delta x_t) + \tfrac{1}{2}\delta x_t^T\mathbf{B}^{-1}\delta x_t \tag{4}$$

At this point, at each time t, the minimum of (4) is obtained by requiring $\nabla J = 0$. This gives rise to the linear system:

$$(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\delta x_t = \mathbf{H}^T\mathbf{R}^{-1}d_t$$

or equivalently:

$$(I + \mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\delta x_t = \mathbf{B}\mathbf{H}^T\mathbf{R}^{-1}d_t \tag{5}$$

For each time $t = 0, ..., n$, iterative methods, able to converge toward a practical solution, are needed to solve the linear system (5). However this problem, so as formulated, is generally very ill conditioned. More precisely, by following [15], and assuming that

$$\mathbf{\Psi} = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{6}$$

is a diagonal matrix, it can be proved that the conditioning of $I + \mathbf{B}\mathbf{\Psi}$ is strictly related to the conditioning of the matrix $\mathbf{B}$ (the covariance matrix). In general, the matrix $\mathbf{B}$ is a block-diagonal matrix, where each block is related to a single state of vector x_t and it is ill conditioned.
This assertion is exposed in [14] starting from the expression of $\mathbf{B}$ for one-state vectors as:

$$\mathbf{B} = \sigma_b^2\mathbf{C}$$

where σ_b^2 is the background error variance and $\mathbf{C}$ is a matrix that denotes the correlation structure of the background error. Assuming that the correlation structure of matrix $\mathbf{C}$ is homogeneous and depends only on the distance between states and not on positions, an expression of $\mathbf{C}$ as a symmetric matrix with a circulant form is given; i. e. as a Toeplitz matrix. By means of a spectral analysis of its eigenvalues, the ill-conditioning of the matrix $\mathbf{C}$ is checked. As in [7], it follows that $\mathbf{B}$ is ill-conditioned and the matrix $I + \mathbf{B}\mathbf{\Psi}$, of the linear system (5), too. A well-known technique for improving the convergence of iterative methods for solving linear systems is to *preconditioning* the system and thus reduce the condition number of the problem.
In order to precondition the system in (5), it is assumed that $\mathbf{B}$ can be written in the form $\mathbf{B} = \mathbf{V}\mathbf{V}^T$, where $\mathbf{V} = \mathbf{B}^{1/2}$ is the square root of the background error covariance matrix $\mathbf{B}$.

Because $\mathbf{B}$ is symmetric Gaussian, $\mathbf{V}$ is uniquely defined as the symmetric ($\mathbf{V^T} = \mathbf{V}$) Gaussian matrix such that $\mathbf{V^2} = \mathbf{B}$. As explained in [18], the cost function (4) becomes:

$$J(\delta x_t) = \tfrac{1}{2}(d_t - \mathbf{H}\delta x_t)^T \mathbf{R}^{-1}(d_t - \mathbf{H}\delta x_t) + \tfrac{1}{2}\delta x_t^T (\mathbf{V}\mathbf{V^T})^{-1}\delta x_t$$
$$= \tfrac{1}{2}(d_t - \mathbf{H}\delta x_t)^T \mathbf{R}^{-1}(d_t - \mathbf{H}\delta x_t) + \tfrac{1}{2}\delta x_t^T (\mathbf{V^T})^{-1}\mathbf{V}^{-1}\delta x_t$$

Now, by using a new control variable v_t, defined as $v_t = \mathbf{V}^{-1}\delta x_t$, at each time $t \in T$ and observing that $\delta x_t = \mathbf{V}v_t$ we obtain a new cost function:

$$\widetilde{J}(v_t) = \frac{1}{2}(d_t - \mathbf{H}\mathbf{V}v_t)^T \mathbf{R}^{-1}(d_t - \mathbf{H}\mathbf{V}v_t) + \frac{1}{2}v_t^T v_t. \quad (7)$$

Equation (7) is said the *dual problem* of equation (4). Finally, to minimize the cost function $\widetilde{J}(v_t)$ in (7) leads to the new linear system:

$$(I + \mathbf{V}\mathbf{\Psi}\mathbf{V})v_t = \mathbf{V}\mathbf{H}^T\mathbf{R}^{-1}d_t \quad (8)$$

Upper and lower bounds on the condition number of the matrix $I + \mathbf{V}\mathbf{\Psi}\mathbf{V}$ are shown in [14]. In particular it holds that:

$$\mu(I + \mathbf{V}\mathbf{\Psi}\mathbf{V}) << \mu(I + \mathbf{B}\mathbf{\Psi}).$$

Moreover, under some special assumptions, it can be proved that $I + \mathbf{V}\mathbf{\Psi}\mathbf{V}$ is very well-conditioned ($\mu(I + \mathbf{V}\mathbf{\Psi}\mathbf{V}) < 4$).

The OceanVar model

As described in [9], at each time $t \in T$, OceanVar software implements an oceanographic three-dimensional variational DA scheme (3D Var-DA) to produce forecasts of ocean currents for the Mediterranean Sea. The computational kernel is based on the resolution of the linear system defined in (8). To solve it, the Conjugate Gradient (CG) method is used and a basic outline is described in **Algorithm 1**.

Algorithm 1 CG Algorithm

1: $k = 0$; $\mathbf{x}_0$, the initial guess;
2: $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
3: $\rho_\mathbf{0} = \mathbf{r_0}$;
4: **while** $\left(\|\mathbf{r}_k\|/\|\mathbf{b}\| > \epsilon \ .and. \ k \leq n\right)$ **do**
5: $\mathbf{q}_k = \mathbf{A}\rho_k$;
6: $\alpha_k = (\mathbf{r}_k, \mathbf{r}_k)/(\rho_k, \mathbf{q}_k)$; $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\rho_\mathbf{k}$;
7: $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{q}_k$; $\beta_k = (\mathbf{r}_{k+1}, \mathbf{r}_{k+1})/(\mathbf{r}_k, \mathbf{r}_k)$;
8: $\rho_{k+1} = \mathbf{r}_{k+1} + \beta_k\rho_k$; $k = k + 1$;
9: **end while**

We focus our attention on step 5.: at each iterative step, a matrix-vector product $\mathbf{A}\,\rho_k$ is required, where

$$\mathbf{A} = \mathbf{I} + \mathbf{V}\mathbf{\Psi}\mathbf{V},$$

$\rho_\mathbf{k}$ is the residual at step k and $\mathbf{\Psi}$ depends on the number of observations and is characterized by a bounded norm (see [15] for details). More precisely, we look to the matrix-vector product

$$\mathbf{q}_k = (I + \mathbf{V}\mathbf{\Psi}\mathbf{V})\rho_\mathbf{k}$$

which can be schematized as shown in **Algorithm 2**.

Algorithm 2 $(I + \mathbf{V}\mathbf{\Psi}\mathbf{V})\rho_k$ Algorithm

1: $z_1 = \mathbf{V}\rho_k$;
2: $z_2 = \mathbf{\Psi}z_1$;
3: $z_3 = \mathbf{V}z_2$;
4: $\mathbf{q}_k = \rho_k + z_3$;

The steps 1. and 3. in **Algorithm 2** consist in a matrix-vector product. These products, as detailed in next section, can be considered discrete Gaussian convolutions and the matrix $\mathbf{V}$, for one-dimensional state vectors, has Gaussian structure. Even for state vectors defined on two (or more) dimensions, the matrix $\mathbf{V}$ can be represented as product of two (or more) Gaussian matrices. Since a single matrix-vector product of this form becomes prohibitively expensive if carried out explicitly, a computational advantage is gained by employing Gaussian RFs to mimic the required Gaussian convolution operators.
In the previous OceanVar scheme, it was implemented a 1st-RF algorithm, as described in [21], [20]. Here, we study the 3rd-RF introduction, based on [27], [23].
The aim of the following sections is to precisely reveal how the n-th order recursive filters are defined and, through the error analysis, to investigate on their effect in terms of error estimate and perfomences.

III. Gaussian recursive filters

In this section we describe Gaussian recursive filters as approximations of the discrete Gaussian convolution used in steps 1. and 3. of **Algorithm 2**. Let denote by

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

the normalized Gaussian function and by $\mathbf{V}$ the square matrix whose entries are given by

$$\mathbf{V}_{i,j} = g(i-j) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(i-j)^2}{2\sigma^2}\right). \quad (9)$$

Now let be $s^0 = (s_1^0, \ldots, s_m^0)^T$ a vector; the discrete Gaussian convolution of s^0 is a new vector $s = (s_1, \ldots, s_m)^T$ defined by means of the matrix-vector product

$$s = \mathbf{V} \otimes s^0 \equiv \mathbf{V}\ s^0. \quad (10)$$

The discrete Gaussian convolution can be considered as a discrete representation of the continuous Gaussian convolution. As is well known, the continuous Gaussian convolution of a function s^0 with the normalized Gaussian function g is a new function s defined as follows:

$$s(x) = [g \otimes s^0](x) = \int_{-\infty}^{+\infty} g(x-\tau)s^0(\tau)d\tau. \quad (11)$$

Discrete and continuous Gaussian convolutions are strictly related. This fact could be seen as follows. Let assume that

$$I = \{x_1 < x_2 < \ldots < x_{m+1}\}$$

is a grid of evaluation points and let set for $i = 1, \ldots, m$

$$s_i \equiv s(x_i), \qquad s_i^0 \equiv s^0(x_i) \quad \text{and} \quad \Delta x_i = x_{i+1} - x_i = 1.$$

By assuming that s^0 is 0 outside of $[x_1, x_{m+1}]$ and by discretizing the integral (11) with a rectangular rule, we obtain

$$s_i = \int_{-\infty}^{+\infty} g(x_i-\tau)s^0(\tau)d\tau = \int_{x_1}^{x_{m+1}} g(x_i-\tau)s^0(\tau)d\tau = \\ = \sum_{j=1}^{m}\int_{x_j}^{x_{j+1}} g(x_i-\tau)s^0(\tau)d\tau \approx \sum_{j=1}^{m}\Delta x_j g(x_i-x_j)s_j^0 = \\ = \sum_{j=1}^{m} g(i-j)s_j^0 = \sum_{j=1}^{m}\mathbf{V}_{i,j}s_j^0 = (\mathbf{V}s^0)_i. \quad (12)$$

An optimal way for approximating the values s_i is given by Gaussian recursive filters. The n-order RF filter computes the vector $s^K = (s_1^K, \ldots, s_m^K)^T$ as follows:

$$\begin{cases} p_i^k = \beta_i s_i^{k-1} + \sum_{j=1}^{n}\alpha_{i,j}p_{i-j}^k \quad i = 1,\ldots,m \\ s_i^k = \beta_i p_i^k + \sum_{j=1}^{n}\alpha_{i,j}s_{i+j}^k \quad i = m,\ldots,1 \end{cases}. \quad (13)$$

The iteration counter k goes from 1 to K, where K is the total number of filter iterations. Observe that values $p_1^k, \ldots p_n^k$ are computed taking in the sums terms $\alpha_{i,j}p_{i-j}^k$ provided that $i-j \geq 1$. Analogously values $s_m^k, \ldots s_{m-n+1}^k$ are computed taking in the sums terms $\alpha_{i,j}s_{i+j}^k$ provided that $i+j \leq m$. The values $\alpha_{i,j}$ and β_i, at each grid point x_i, are often called *smoothing coefficients* and they obey to the constraint

$$\beta_i = 1 - \sum_{j=1}^{n}\alpha_{i,j}.$$

In this paper we deal with first-order and third-order RFs. The first-order RF expression ($n = 1$) becomes:

$$\begin{cases} p_1^k = \beta_1 s_1^{k-1}, \\ p_i^k = \beta_i s_i^{k-1} + \alpha_i p_{i-1}^k \quad i = 2,\ldots,m \\ s_m^k = \beta_m p_m^k, \\ s_i^k = \beta_i p_i^k + \alpha_i s_{i+1}^k \qquad i = m-1,\ldots,1. \end{cases} \quad (14)$$

If R_i is the correlation radius at x_i, by setting

$$\sigma_i = \frac{R_i}{\Delta x_i} \quad \text{and} \quad E_i = \frac{K\Delta x_i^2}{R_i^2} = \frac{K}{\sigma_i^2},$$

coefficients α_i e β_i are given by [21]:

$$\alpha_i = 1 + E_i - \sqrt{E_i(E_i+2)}, \qquad \beta_i = \sqrt{E_i(E_i+2)} - E_i. \quad (15)$$

The third-order RF expression ($n = 3$) becomes:

$$\begin{cases} p_i^k = \beta_i s_i^{k-1} + \sum_{j=1}^{3}\alpha_{i,j}p_{i-j}^k \quad i = 1,...,m \\ s_i^k = \beta_i p_i^k + \sum_{j=1}^{3}\alpha_{i,j}s_{i+j}^k \quad i = m,\ldots,1. \end{cases} \quad (16)$$

Third-order RF coefficients $\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}$ and β_i, for one only filter iteration ($K = 1$), are computed in [11]. If

$$a_i = 3.738128 + 5.788982\sigma_i + 3.382473\sigma_i^2 + \sigma_i^3.$$

the coefficients expressions are:

$$\alpha_{i,1} = (5.788982\sigma_i + 6.764946\sigma_i^2 + 3\sigma_i^3)/a_i$$
$$\alpha_{i,2} = -(3.382473\sigma_i^2 + 3\sigma_i^3)/a_i$$
$$\alpha_{i,3} = \sigma_i^3/a_i$$
$$\beta_i = 1 - (\alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3}) = 3.738128/a_i.$$

In [23] the use of a value $q = q(\sigma_i)$ instead of σ_i is proposed. The q value is:

$$q(\sigma_i) = \begin{cases} 0.98711\sigma_i - 0.96330 & \text{if} \quad \sigma_i > 2.5 \\ 3.97156 - 4.14554\sqrt{1 - 0.26891\sigma_i} & \text{oth.} \end{cases} \quad (17)$$

In order to understand how Gaussian RFs approximate the discrete Gaussian convolution it is useful to represent them in terms of matrix formulation. As explained in [5], the n-order recursive filter computes s^K from s^0 as the solution of the linear system

$$(LU)^K s^K = s^0, \quad (18)$$

where matrices L and U are respectively lower and upper band triangular with nonzero entries

$$U_{i,i} = L_{i,i} = \frac{1}{\beta_i}, \qquad L_{i,i-j} = U_{i,i+j} = -\frac{\alpha_{i,j}}{\beta_i}. \quad (19)$$

By formally inverting the linear system (18) it results

$$s^K = \mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}\, s^0, \quad (20)$$

where $\mathbf{F}_{\mathbf{n}}^{(\mathbf{K})} \equiv (LU)^{-K}$. A direct expression of $\mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}$ and its norm could be obtained, for instance, for the first order recursive filter in the homogenus case ($\sigma_i = \sigma$). However, in the following, it will be shown that $\mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}$ has always bounded norm, i.e.

$$\|\mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}\|_\infty \leq 1. \quad (21)$$

Observe that $\mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}$ is the matrix operator that substitutes the Gaussian operator $\mathbf{V}$ in (10), then a measure of how well s^K approximates s can be derived in terms of the operator distance

$$\|\mathbf{V} - \mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}\|_\infty.$$

Ideally one would expect that $\|\mathbf{V} - \mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}\|$ goes to 0 (and $s^K \to s$) as K approaches to ∞, yet this does not happen due to the presence of edge effects. In the next sections we show the numerical behaviour of the distance $\|\mathbf{V} - \mathbf{F}_{\mathbf{n}}^{(\mathbf{K})}\|$ for some case study and we will show its effects in the CG algorithm.

IV. RF ERROR ANALYSIS

Here we are interested to analyze the error introduced on the matrix-vector operation at step 5. of **Algorithm 1**, when the Gaussian RF is used instead of the discrete Gaussian convolution. As previously explained, in terms of matrices, this is equivalent to change the matrix operator, then **Algorithm 2** can be rewritten as shown in **Algorithm 3**.

Now we are able to give the main result of this paper: indeed the following theorem furnishes an upper bound for the error

Algorithm 3 $(I + \mathbf{F_n^{(K)}}\mathbf{\Psi}\mathbf{F_n^{(K)}})\widetilde{\rho}_k$ Algorithm

1: $\widetilde{z}_1 = \mathbf{F_n^{(K)}}\widetilde{\rho}_k$;
2: $\widetilde{z}_2 = \mathbf{\Psi}\widetilde{z}_1$;
3: $\widetilde{z}_3 = \mathbf{F_n^{(K)}}\widetilde{z}_2$;
4: $\widetilde{\mathbf{q}}_k = \widetilde{\rho}_k + \widetilde{z}_3$;

$\mathbf{q}_k - \widetilde{\mathbf{q}}_k$, made at each single iteration k of the CG (**Algorithm 1**). This bound involves the operator norms

$$\|\mathbf{F_n^{(K)}}\|_\infty, \quad \|\mathbf{\Psi}\|_\infty, \quad \|\mathbf{V}\|_\infty,$$

the distance $\|\mathbf{V} - \mathbf{F_n^{(K)}}\|_\infty$ and the error $\rho_k - \widetilde{\rho}_k$ accumulated on ρ_k at previous iterations.

Theorem 4.1: Let be ρ_k, $\widetilde{\rho}_k$, $\mathbf{q}_k$, $\widetilde{\mathbf{q}}_k$ as in **Algorithm 2** and **Algorithm 3**. Let be $\|\cdot\| = \|\cdot\|_\infty$ and let denote by

$$e_k = \rho_k - \widetilde{\rho}_k$$

the difference between values ρ_k and $\widetilde{\rho}_k$. Then it holds

$$\|\mathbf{q}_k - \widetilde{\mathbf{q}}_k\| \leq (1 + \|\mathbf{V}\|\cdot\|\mathbf{\Psi}\|\cdot\|\mathbf{V}\|)\cdot\|e_k\| + \\ +\|\mathbf{F_n^{(K)}} - \mathbf{V}\|\cdot\|\mathbf{\Psi}\|\cdot\big(\|\mathbf{V}\| + \|\mathbf{F_n^{(K)}}\|\big)\cdot\|\widetilde{\rho}_k\|. \quad (22)$$

Proof: A direct proof follows by using the values z_i and $\widetilde{z}_i$ introduced in Algorithm 2 and in Algorithm 3. It holds:

$$\|z_1 - \widetilde{z}_1\| = \|\mathbf{V}\rho_k - \mathbf{F_n^{(K)}}\widetilde{\rho}_k\| = \|\mathbf{V}\rho_k - \mathbf{V}\widetilde{\rho}_k + \mathbf{V}\widetilde{\rho}_k - \mathbf{F_n^{(K)}}\widetilde{\rho}_k\| \leq \\ \leq \|\mathbf{V}\rho_k - \mathbf{V}\widetilde{\rho}_k\| + \|\mathbf{V}\widetilde{\rho}_k - \mathbf{F_n^{(K)}}\widetilde{\rho}_k\| \leq \\ \leq \|\mathbf{V}\|\cdot\|e_k\| + \|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\widetilde{\rho}_k\|.$$

Then, for the difference $z_2 - \widetilde{z}_2$, we get the bound

$$\|z_2 - \widetilde{z}_2\| = \|\mathbf{\Psi}z_1 - \mathbf{\Psi}\widetilde{z}_1\| \leq \|\mathbf{\Psi}\|\cdot\|z_1 - \widetilde{z}_1\| \leq \\ \leq \|\mathbf{\Psi}\|\cdot\|\mathbf{V}\|\cdot\|e_k\| + \|\mathbf{\Psi}\|\cdot\|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\widetilde{\rho}_k\|.$$

Hence, for the difference $z_3 - \widetilde{z}_3$, we obtain

$$\|z_3 - \widetilde{z}_3\| = \|\mathbf{V}z_2 - \mathbf{F_n^{(K)}}\widetilde{z}_2\| = \|\mathbf{V}z_2 - \mathbf{V}\widetilde{z}_2 + \mathbf{V}\widetilde{z}_2 - \mathbf{F_n^{(K)}}\widetilde{z}_2\| \leq \\ \leq \|\mathbf{V}\|\cdot\|z_2 - \widetilde{z}_2\| + \|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\widetilde{z}_2\| \leq \\ \leq \|\mathbf{V}\|\cdot\|z_2 - \widetilde{z}_2\| + \|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\mathbf{\Psi}\|\cdot\|\mathbf{F_n^{(K)}}\|\cdot\|\widetilde{\rho}_k\| \leq \\ \leq \|\mathbf{V}\|\cdot\|\mathbf{\Psi}\|\cdot\|\mathbf{V}\|\cdot\|e_k\| + \|\mathbf{V}\|\cdot\|\mathbf{\Psi}\|\cdot\|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\widetilde{\rho}_k\| \\ +\|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\mathbf{\Psi}\|\cdot\|\mathbf{F_n^{(K)}}\|\cdot\|\widetilde{\rho}_k\| = \\ \|\mathbf{V}\|\cdot\|\mathbf{\Psi}\|\cdot\|\mathbf{V}\|\cdot\|e_k\| + \|\mathbf{V} - \mathbf{F_n^{(K)}}\|\cdot\|\mathbf{\Psi}\|\big(\|\mathbf{V}\| + \|\mathbf{F_n^{(K)}}\|\big)\cdot\|\widetilde{\rho}_k\|$$

In the second-last inequality we used the fact that

$$\|\widetilde{z}_2\| = \|\mathbf{\Psi}\widetilde{z}_1\| = \|\mathbf{\Psi}\mathbf{F_n^{(K)}}\widetilde{\rho}_k\| \leq \|\mathbf{\Psi}\|\cdot\|\mathbf{F_n^{(K)}}\|\cdot\|\widetilde{\rho}_k\|.$$

Finally, observing that

$$\|\mathbf{q}_k - \widetilde{\mathbf{q}}_k\| = \|\rho_k + z_3 - (\widetilde{\rho}_k + \widetilde{z}_3)\| \leq \\ \leq \|\rho_k - \widetilde{\rho}_k\| + \|z_3 - \widetilde{z}_3\| = \|e_k\| + \|z_3 - \widetilde{z}_3\|,$$

and taking the upper bound of $\|z_3 - \widetilde{z}_3\|$, the thesis is proved. ◇

Previous theorem shows that, at each iteration of the CG algorithm, the error bound on the computed value $\mathbf{q_k}$ at step 5., is characterized by two main terms: the first term can be considered as the contribution of the standard forward error analysis and it is not significant, if $\|e_k\|$ is small; the second term highlights the effect of the introduction of the RF. More in detail, at each iteration step, the computed value $\mathbf{q_k}$ is biased by a quantity proportional to three factors:

- the distance between the original operator (the Gaussian operator $\mathbf{V}$) and its approximation (the operator $\mathbf{F_n^{(K)}}$);
- the norm of $\mathbf{\Psi}$;
- the sum of the operator norms $\|\mathbf{F_n^{(K)}}\|$ and $\|\mathbf{V}\|$.

Table 1: Operator norms

σ	$\|\|\mathbf{F_1^{(1)}}\|\|_\infty$	$\|\|\mathbf{F_3^{(1)}}\|\|_\infty$
5	0.9920	0.9897
20	0.9012	0.8537
50	0.9489	0.8950

As shown in (21) the norm $\mathbf{\Psi}$ is bounded. Besides, the norm of $\mathbf{V}$ is always less or equal to one (because it comes from the discretization of the of the continuous Gaussian convolution). The norm of $\mathbf{F_n^{(K)}}$ is bounded by one too. This fact can be seen by observing the Table 1, where we consider several tests by varying data distributions in the homogeneous case ($\sigma_i = \sigma$), for 1st-RF and 3rd-RF. Starting from these considerations, the error estimate of *Theorem 4.1* can be specialized as:

$$\|\mathbf{q}_k - \widetilde{\mathbf{q}}_k\| \leq (1 + \|\mathbf{\Psi}\|)\,\|e_k\| + 2\|\mathbf{F_n^{(K)}} - \mathbf{V}\|\cdot\|\mathbf{\Psi}\|\|\widetilde{\rho}_k\|. \quad (23)$$

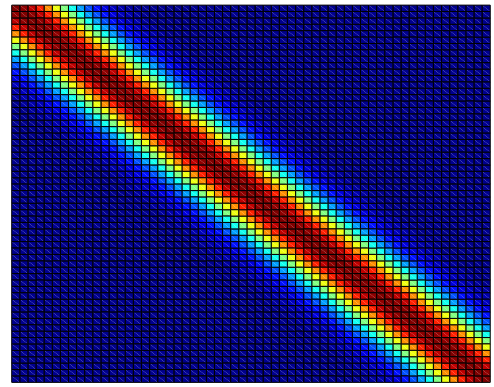
V. Experimental Results

In this section we report some experiments to confirm the discussed theoretical results. In the first part, we deal with the approximations of the discrete operator $\mathbf{V}$ with the first order and of the third order $\mathbf{F_1^{(K)}}$ and $\mathbf{F_3^{(1)}}$ respectively. In the last subsection, we analyze the improving in the performance and in the accuracy terms of the third order RF applied to the case study.

A. 1st-RF and 3rd-RF operators

In the following experiments, we construct the operators $\mathbf{V}$, $\mathbf{F_1^{(1)}}$, $\mathbf{F_1^{(50)}}$ and $\mathbf{F_3^{(1)}}$ in the case of $m = 601$ samples of a random vector $\mathbf{s^0}$. We assume that $\mathbf{s^0}$ comes from a uniform grid with homogeneous condition $\sigma_i = \sigma = 15$. In Figure 1, it is highlighted that the involved discrete operators have different structures. In particular, a first qualitative remark is that the operator $\mathbf{F_1^{(1)}}$ is a poor approximation of $\mathbf{V}$. Conversely, the operator $\mathbf{F_1^{(50)}}$ (Figure 2 on the top) is very close to $\mathbf{V}$ but, as for $\mathbf{F_1^{(1)}}$, there are significant differences with $\mathbf{V}$ in the bottom left and in the top right corners. These dissimilarities in the edges, by a numerical point of view, give some kind of artifacts in the computed convolutions, that determine a vector $\mathbf{s}$ with components, in the initial and final positions, that decay to zero.

Figure 2 bottom shows that the operator $\mathbf{F_3^{(1)}}$ is closer then $\mathbf{F_1^{(1)}}$ and $\mathbf{F_1^{(50)}}$ to the discrete convolution $\mathbf{V}$. In particular, this recursive filter is able to reproduce $\mathbf{V}$ more accurately in the bottom left corner, but unfortunately it does not give good



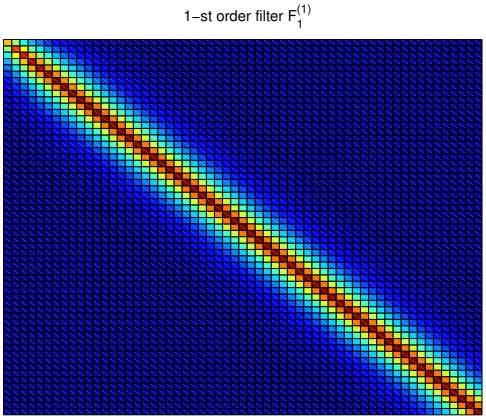
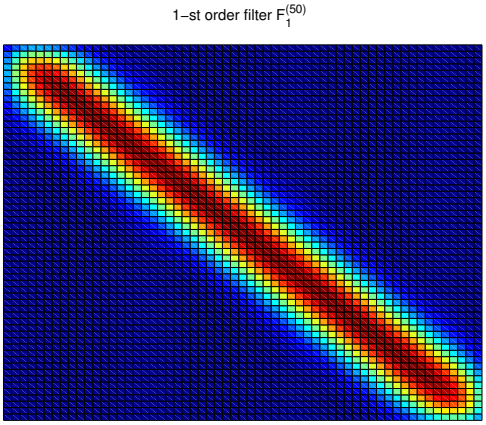


Fig. 1. **Top**. Discrete Gaussian convolution operator $\mathbf{V}$. **Bottom**. 1-st order recursive filter operator F_1



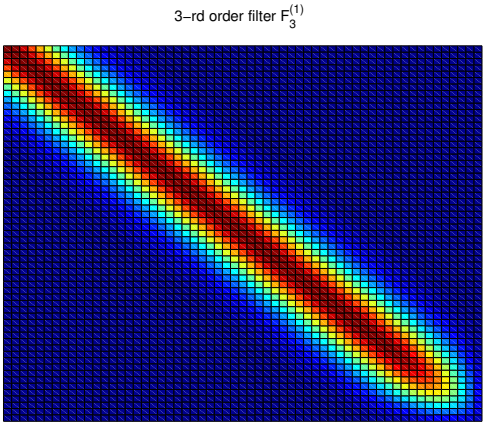


Fig. 2. **Top**. 1-st order recursive filter operator $F_1^{(50)}$ with 50 iterations. **Bottom**. 3-rd order recursive filter operator $F_3^{(1)}$

results on top right corner. In Table 2, for random distributions with homogeneous condition ($\sigma_i = \sigma$), we underline the edge effects by measuring the norms between the discrete convolution $\mathbf{V}$ and the RF filters. Although the $||\mathbf{F_n^{(K)}} - \mathbf{V}||_\infty$ ideally goes to zero as k goes to $+\infty$, this does not happen in practice as observed below.

Table 2: Distance metrics

σ	$\|\|\mathbf{F_1^{(1)}} - \mathbf{V}\|\|_\infty$	$\|\|\mathbf{F_1^{(50)}} - \mathbf{V}\|\|_\infty$	$\|\|\mathbf{F_3^{(1)}} - \mathbf{V}\|\|_\infty$
5	0.2977	0.3800	0.5346
10	0.3895	0.4397	0.5890
25	0.4533	0.4758	0.6221
50	0.4686	0.4809	0.6125

In order to bring out these considerations, we show the application of $\mathbf{V}$, $\mathbf{F_1^{(K)}}$ and $\mathbf{F_3^{(1)}}$ to a periodic signal $\mathbf{s^0}$. We choose $m = 252$ samples of the cos function in $[-2\pi, 2\pi]$ and we perform simulations by using the 1-st RF with $1, 5$ and 50 iterations and 3-rd RF with one iteration. In Figure 3 it is shown the computed Gaussian convolution and the poor approximation of $\mathbf{V}s^0$ on the right side of the test interval, due to the edge effects. A nice result is that our $\mathbf{F_3^{(1)}}$ convolution operator gives better results on the left side of the domain.

Finally, we give some considerations about the accuracy of the studied Gaussian RF schemes, when they are applied to the Dirac rectangular impulse

$$s^0 = (0, \ldots, 0, 1, 0, \ldots).$$

We choose a one-dimensional grid of $m = 301$ points, a constant correlation radius $R = 120, km$, a constant grid space $\Delta x = 6\, km$ and $\sigma = R/\Delta x = 20$. In the numerical experiments to avoid the edge effects, we only consider $\bar{m} = 221$ central values of s^K, i.e.

$$\bar{s}^K = (s^K_{2\sigma}, s^K_{2\sigma+1}, \ldots, s^K_{m-2\sigma-1}, s^K_{m-2\sigma}).$$

Similarly, in Table 3 we measure the operator distances we use $||\bar{\mathbf{F}}_\mathbf{1}^{(1)} - \bar{\mathbf{V}}||_\infty$ and $||\bar{\mathbf{F}}_\mathbf{3}^{(1)} - \bar{\mathbf{V}}||_\infty$, where $\bar{V}$, $\bar{F}_1^{(1)}$ and $\bar{F}_3^{(1)}$ indicate the submatrices obtained, neglecting first and last $2\sigma - 1$ rows and columns.

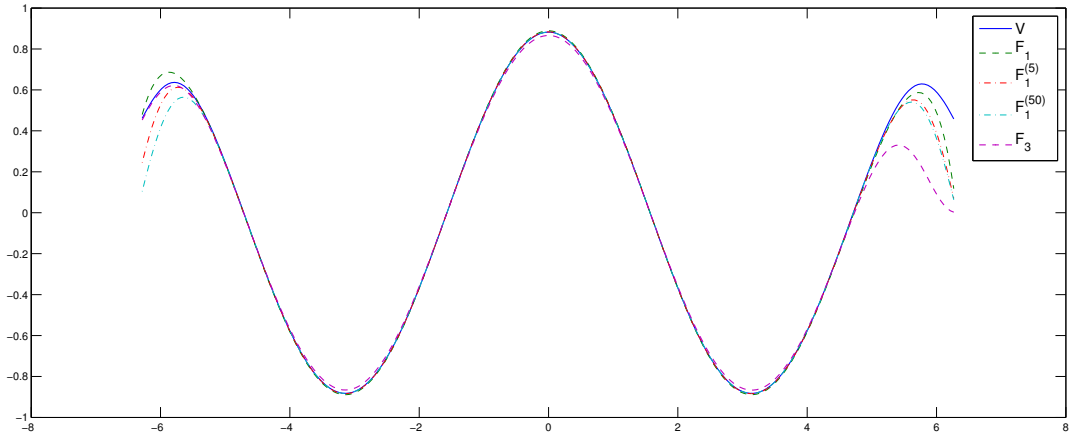


Fig. 3. Discrete convolution $\mathbf{V}$ and Gaussian recursive filtering $\mathbf{F_1^{(K)}}$ with $1, 5, 50$ iterations and $\mathbf{F_3^{(1)}}$ applied to $n = 252$ samples of the periodic function $s^0 = \cos(x)$ in $[-2\pi, 2\pi]$.

Table 3: Convergence history

K	$\|\|\bar{\mathbf{F}}_\mathbf{1}^{(\mathbf{K})} - \bar{\mathbf{V}}\|\|_\infty$	$\|\|\bar{\mathbf{F}}_\mathbf{3}^{(\mathbf{K})} - \bar{\mathbf{V}}\|\|_\infty$
1	0.211	**0.0424**
2	0.13	–
5	0.078	–
50	0.048	–
100	0.0429	–
500	0.0414	–

These case studies show that , neglecting the edge effects, the 3-rd RF filter is more accurate the the 1st-RF order with few iterations. This fact is evident by observing the results in Figure 4 and the operator norms in Table 3. Finally, we remark that the 1-st order RF has to use 100 iteration in order to obtain the same accuracy of the 3-rd order RF. This is a very interesting numerical feature of the third order filter.

B. A case study: Ocean Var

The theoretical considerations of the previous sections are useful to understand the accuracy improvement in the real experiments on Ocean Var. The preconditioned CG is a numerical kernel intensively used in the model minimizations. Implementing a more accurate convolution operators gives benefits on the convergence of GC and on the overall data assimilation scheme [11] . Here we report experimental results of the 3rd-RF in a Global Ocean implementation of OceanVar that follows [22], [12]. These results are extensively discussed in the report [11]. In real scenarios [4], [10], scientific libraries and high performance computing environments are needed. The case study simulations were carried-out on an IBM cluster using 64 processors. The model resolution was about $1/4$ degree and the horizontal grid was tripolar, as described in [19]. This configuration of the model was used at CMCC for global ocean physical reanalyses applications (see [13]). The model has 50 vertical depth levels. The three-dimensional

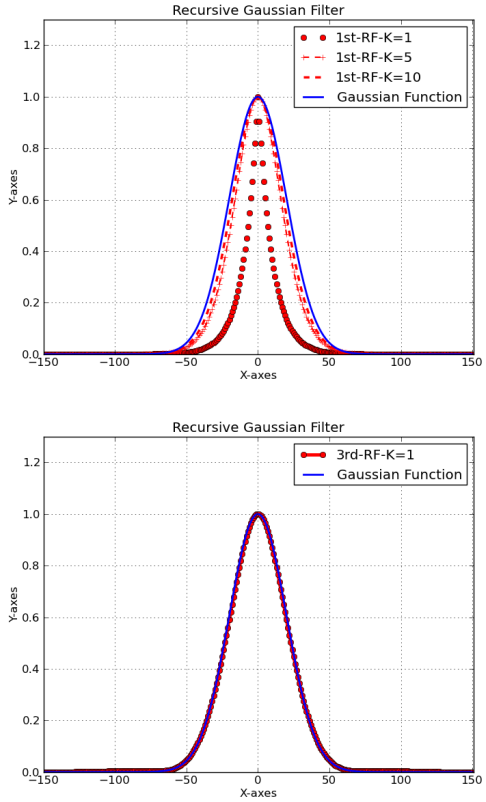


Fig. 4. **Top**. The discrete Gaussian convolution $\mathbf{V}s^0$ (blue) and $\mathbf{F_1^{(K)}}s^0$ for $K = 1, 5, 10$ (red). **Bottom** The discrete Gaussian convolution $\mathbf{V}s^0$ (blue) and $\mathbf{F_3^{(1)}}s^0$ (red).

model grid consists of 736141000 grid-points. The comparison between the 1st-RF and 3rd-RF was carried out for a realistic case study, where all in-situ observations of temperature and

salinity from Expendable bathythermographs (XBTs), Conductivity, Temperature, Depth (CTDs) Sensors, Argo floats and Tropical mooring arrays were assimilated. The observational profiles are collected, quality-checked and distributed by [3]. The global application of the recursive filter accounts for spatially varying and season-dependent correlation length-scales (CLSs). Correlation length-scale were calculated by applying the approximation given in [2] to a dataset of monthly anomalies with respect to the monthly climatology, with inter-annual trends removed.

The obtained performances of a 3Dvar application that uses the 1st-RF with 1, 5 and 10 iterations and the 3rd-RF are shown in Figure 5 with a zoom in the same area of Western Pacific Area as in Figure 5, for the temperature at 100 m of depth. The Figure also displays the differences between the 3rd-RF and the 1st-RF with either 1 or 10 iterations. The patterns of the increments are closely similar, although increments for the case of 1st-RF (K=1) are generally sharper in the case of both short (e.g. off Japan) or long (e.g. off Indonesian region) CLSs. The panels of the differences reveal also that the differences between 3rd-RF and the 1st-RF (K=10) are very small, suggesting once again that the same accuracy of the 3rd-RF can be achieved only with a large number of iterations for the first order recursive filter. Finally, in [12] was also observed that the 3rd-RF compared to the 1st-RF (K=5) and the 1st-RF (K=10) reduces the wall clock time of the software respectively of about 27% and 48%.

VI. Conclusions

Recursive Filters (RFs) are a well known way to approximate the Gaussian convolution and are intensively applied in the meteorology, in the oceanography and in forecast models. In this paper, we deal with the oceanographic 3D-Var scheme OceanVar. The computational kernel of the OceanVar software is a linear system solved by means of the Conjugate Gradient (GC) method. The iteration matrix is related to an error covariance matrix, with a Gaussian correlation structure. In other words, at each iteration, a Gaussian convolution is required. Generally, this convolution is approximated by a first order RF. In this work, we introduced a 3rd-RF filter and we investigated about the main sources of error due to the use of 1st-RF and 3rd-RF operators. Moreover, we studied how these errors influence the CG algorithm and we showed that the third order operator is more accurate than the first order one. Finally, theoretical issues were confirmed by some numerical experiments and by the reported results in the case study of the OceanVar software.

References

[1] M. Abramowitz, I. Stegun - *Handbook of Mathematical Functions*. Dover, New York, 1965.

[2] M. Belo Pereira, L. Berre - *The use of an ensemble approach to study the background-error covariances in a global NWP model*. MOn. Wea. Rev. 134, pp. 2466-2489, 2006.

[3] C. Cabanes, A. Grouazel, K. von Schuckmann, M. Hamon, V. Turpin, C. Coatanoan, F. Paris, S. Guinehut, C. Bppne, N. Ferry, C. de Boyer Montgut, T. Carval, G. Reverding, S. Puoliquen, P.Y. L. Traon - *The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements*. Ocean Sci 9, pp. 1-18, 2013.

[4] S. Cuomo, A. Galletti, G. Giunta and A. Starace-*Surface reconstruction from scattered point via RBF interpolation on GPU* , Federated Conference on Computer Science and Information Systems (FedCSIS), 2013, pp. 433-440.

[5] G. Dahlquist and A. Bjorck - *Numerical Methods*. Prentice Hall, 573 pp. 1974.

[6] J. Derber, A. Rosati - *A global oceanic data assimilation system*. Journal of Phys. Oceanogr. 19, pp. 1333-1347, 1989.

[7] L. D' Amore, R. Arcucci, L. Marcellino, A. Murli- *HPC computation issues of the incremental 3D variational data assimilation scheme in OceanVarsoftware*. Journal of Numerical Analysis, Industrial and Applied Mathematics, 7(3-4), pp 91-105, 2013.

[8] R. Deriche - *Separable recursive filtering for efficient multi-scale edge detection*. Proc. Int. Workshop Machine Vision Machine Intelligence, Tokyo, Japan, pp 18-23, 1987

[9] S. Dobricic, N. Pinardi - *An oceanographic three-dimensional variational data assimilation scheme*. Ocean Modeling 22, pp 89-105, 2008.

[10] R. Farina, S. Cuomo, P. De Michele, F. Piccialli-*A Smart GPU Implementation of an Elliptic Kernel for an Ocean Global Circulation Model*, APPLIED MATHEMATICAL SCIENCES, 7 (61-64), 2013 pp.3007-3021.

[11] R. Farina, S. Dobricic, S. Cuomo-*Some numerical enhancements in a data assimilation scheme*, AIP Conference Proceedings 1558, 2013, doi: 10.1063/1.4826017.

[12] R. Farina, S. Dobricic, A. Storto, S. Masina, S. Cuomo-*A Revised Scheme to Compute Horizontal Covariances in an Oceanographic 3D-VAR Assimilation System*, CoRR, abs/1404.5756, 2014, http://arxiv.org/abs/1404.5756

[13] N. Ferry, B. Barnier, G. Garric, K. Haines, S. Masina, L. Parent, A. Storto, M. Valdivieso, S. Guinehut, S. Mulet - *NEMO: the modeling engine of global ocean reanalysis*. Mercator Ocean Quaterly Newsletter 46, pp 60-66, 2012.

[14] S. Haben, A. Lawless, N. Nichols - *Conditioning of the 3DVar data assimilation problem*. University of Reading, Dept. of Mathematics, Math Report Series 3, 2009;

[15] S. Haben, A. Lawless, N. Nicholos - *Conditioning and preconditioning of the variational data assimilation problem*. Computers and Fluids 46, pp 252-256, 2011.

[16] L. Haglund - *Adaptive multidimensional filtering*. Linkping University, Sweden, 1992.

[17] A.C. Lorenc - *Iterative analysis using covariance functions and filters*. Quartely Journal of the Royal Meteorological Society 1-118, pp 569-591, 1992.

[18] A.C. Lorenc - *Development of an operational variational assimilation scheme*. Journal of the Meteorological Society of Japan 75, pp 339-346, 1997.

[19] G. Madec, M. Imbard - *A global ocean mesh to overcome the north pole singularity*. Clim. Dynamic 12, pp 381-388, 1996.

[20] R.J. Purser, W.-S. Wu, D.F. Parish, N.M. Roberts - *Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: spatially inhomogeneous and anisotropic covariances*. Monthly Weather Review 131, pp 1524-1535, 2003.

[21] C. Hayden, R. Purser - *Recursive filter objective analysis of meteorological field: applications to NESDIS operational processing*. Journal of Applied Meteorology 34, pp 3-15, 1995.

[22] A. Storto, S. Dobricic, S. Masina, P. D. Pietro - *Assimilating along-track altimetric observations through local hydrostatic adjustments in a global ocean reanalysis system*. Mon. Wea. Rev. 139, pp 738-754, 2011.

[23] L.V. Vliet, I. Young, P. Verbeek - *Recursive Gaussian derivative filters*. International Conference Recognition, pp 509-514, 1998.

[24] L.J. van Vliet, P.W. Verbeek - *Estimators for orientation and anisotropy in digitized images*. Proc. ASCI'95, Heijen (Netherlands), pp 442-450, 1995.

[25] A. T. Weaver, P. Courtier - *Correlation modelling on the sphere using a generalized diffusion equation*. Quarterly Journal of the Royal Meteorological Society 127, pp 1815-1846, 2001.

[26] A. Witkin - *Scale-space filtering*. Proc. Internat. Joint Conf. on Artificial Intelligence, Karlsruhe, germany, pp 1019-1021, 1983.

[27] I.T. Young, L.J. van Vliet - *Recursive implementation of the Gaussian filter*. Signal Processing 44, pp 139-151, 1995.

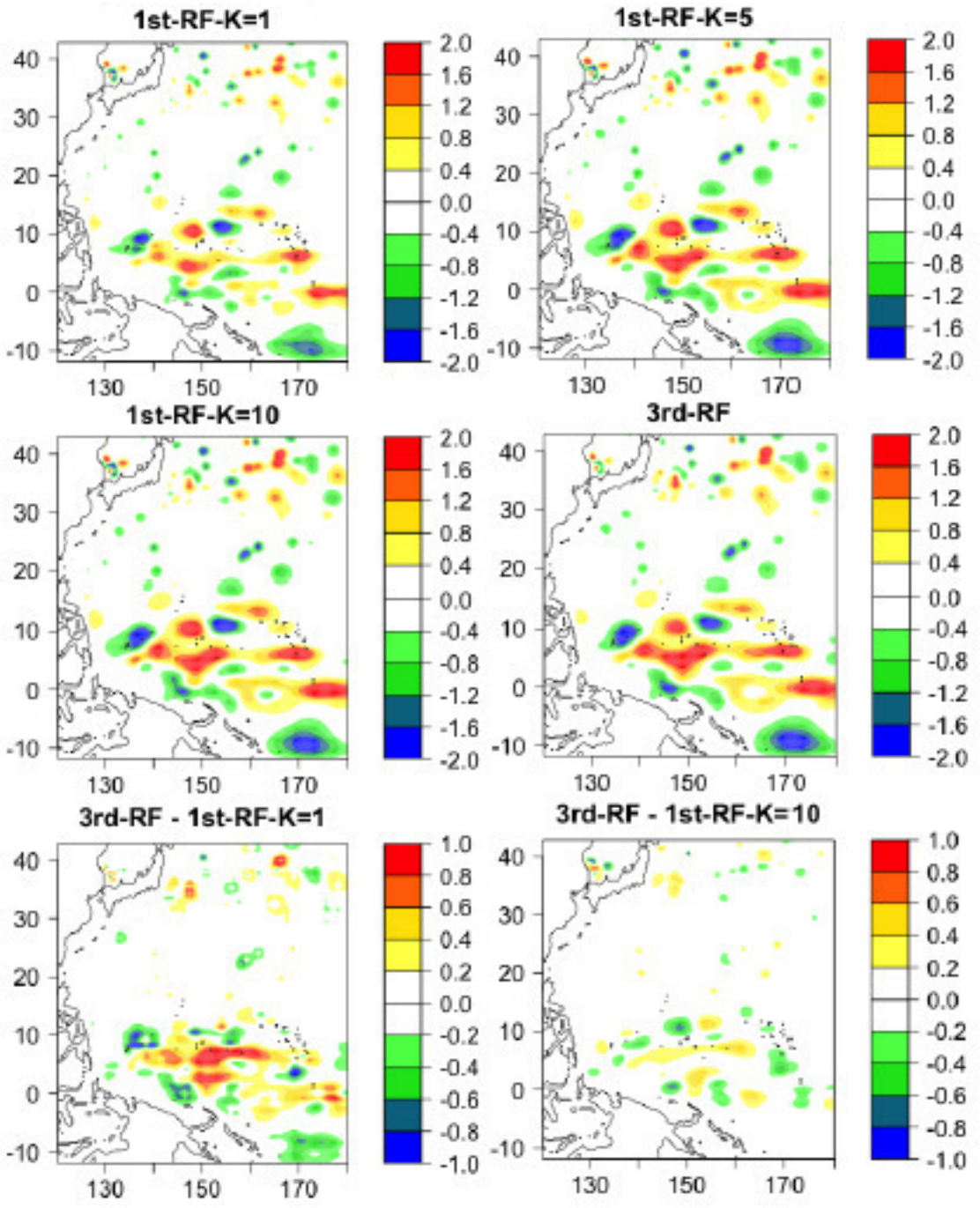


Fig. 5. Analysis increments of temperature at 100 m of depth for the Western Pacific for different configurations of the recursive filter (first two rows of panels). Differences of 100 m temperature analysis increments between 3rd-RF and 1st-RF (K=1) and between 3rd-RF and 1st-RF (K=10) (bottom panels).