European Society of Computational Methods in Sciences and Engineering (ESCMSE)

Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM)
vol. 12, no. 1-2, 2018, pp. 9-28
ISSN 1790–8140

Toward a preconditioned scalable 3DVAR for assimilating Sea Surface Temperature collected into the Caspian Sea

R. Arcucci[ab1], L.Carracciuolo[c] and R.Toumi[d]

a: University of Naples Federico II, Naples, Italy
b: Euro Mediterranean Center on Climate Change, Italy
c: National Research Council, Naples, Italy
d: Imperial College London, London, United Kingdom

Received 1 February, 2017; accepted in revised form 03 April, 2018

Abstract: Data Assimilation (DA) is an uncertainty quantification technique used to incorporate observed data into a prediction model in order to improve numerical forecasted results. As a crucial point into DA models is the ill conditioning of the covariance matrices involved, it is mandatory to introduce, in a DA software, preconditioning methods. Here we present first results obtained introducing two different preconditioning methods in a DA software we are developing (we named S3DVAR) which implements a Scalable Three Dimensional Variational Data Assimilation model for assimilating sea surface temperature (SST) values collected into the Caspian Sea by using the Regional Ocean Modeling System (ROMS) with observations provided by the Group of High resolution sea surface temperature (GHRSST). We present the algorithmic strategies we employ and the numerical issues on data collected in two of the months which present the most significant variability in water temperature: August and March.

Keywords: Data Assimilation, oceanographic data, Sea Surface Temperature, Caspian sea, ROMS

Mathematics Subject Classification: 65Y05, 65J22, 68W10, 68U20

PACS: 02.70.-c

1 Introduction and Motivation

The Caspian Sea has an elongated geometry (1000 km in length and $200-300$ km in width), where the Northern, Middle and Southern Caspian Basins constitute the main geographic divisions [10]. The Sea Surface Temperature (SST) variabilities in the Caspian Sea have different characteristics in the different regions. In the Southern Caspian, the SST reaches a high of $25-29°C$ in the summer months and has a low of $7-10°C$ in the winter. The Northern Caspian experiences a more drastic change in SST throughout the year, with a high of $25-26°C$ in the summer and a

[1]Corresponding author. E-mail: rossella.arcucci@unina.it

below freezing point in the winter [30].

Improvement in Caspian sea temperatures prediction is a crucial point for different climate phenomena simulation. An example is the study on the sea-ice coverage [48] or the prediction of the cyclonicity in winter and anticyclonicity in spring and summer as the water temperature influences the closed atmosphere [43]. This variability may be of interest in the long-term as it may act as an early indicator of large-scale climate change, as well as being an area of interest to industries and vulnerable species.

The current approach in ocean modelling (which includes sea temperature predictions) consists in simulating explicitly only the largest-scale phenomena, while taking into account the smaller-scale ones by means of "physical parametrisations". Due to the inability to resolve the full spectrum of physical mechanisms involved as well as the fundamentally stochastic nature of the turbulent processes in the ocean, all ocean models introduce uncertainty through the selection of scales and parameters that are somewhat inaccurate. Additionally, any computational methodology contributes to uncertainty due to discretization, finite precision and the consequent accumulation and amplification of round-off errors. Taking into account these uncertainties is essential for the acceptance of any numerical simulation.

The Data Assimilation (DA) is an uncertainty quantification technique used to incorporate observed data into a prediction model in order to improve numerical forecasted results [33]. There are many DA methods which are been mostly custom-developed on the ocean model with which they are combined and, today, there are a lot of DA algorithms. Two main methods gained acceptance as powerful methods for data assimilation in the last decennium: the variational approach and the Kalman Filter. The variational approach [1, 6, 39] is based on the minimization of a functional which estimate the discrepancy between numerical results and measures. The Kalman Filter [32] is a recursive filtering instead. Both methods assume that the two sources of information, forecast and observations, have errors that are adequately described by error covariance matrices. The computational kernel of the Kaman Filter is the solution of normal equations. For the variational approach instead, it is the solution of a linear system [44, 33]. Caused by the background error covariance matrices this system is strongly ill conditioned [25, 44].

Due to the scale of the forecasting area used to describe oceans and seas, DA is a "large size problem" which mandates the development of DA software in a High Performance Computing (HPC) environment. HPC gives the opportunity to take full advantage of emerging architectures that can improve performances through the design and implementation of innovative approaches (i.e., [9, 38, 37, 12, 36, 42]) for accessing the computational power that can be used to tackle explicitly the growing complexity of the ocean circulation model [16, 17]. A simulation that yields high-fidelity results is of little use if it is too expensive to run or if it cannot be scaled up to the resolutions required to describe the real-world phenomena of interest. Scalability is mandatory to reduce computational cost as well. Some studies about the minimisation of the total cost by the owners of large scale computing systems, without affecting negatively the quality of service for the users, are been provided in [7].

A suitable DA model must be identified which takes into account both the users/applications requirements and all the mathematical, numerical and algorithmical related issues. Following a problem-to-solve approach, we face the following issues about the

1. *physical and mathematical assumptions concerning the* definition/localization of both forecasting and observed *data*;

2. *algorithmic strategies concerning the definition* of the covariance matrices as well as their preconditioning;

3. *scalability expectation* on computing environment in which the software is implemented.

The article is organized as follows. In section 2, the contribution of the present work with respect to related works is discussed. Section 3 provides mathematical settings and preliminary definitions. Section 4 describes the Preconditioned Scalable 3DVAR (S3DVAR) Computational Kernel while in section 5 we discuss results concerning the accuracy and the efficiency of the S3DVAR results on a cluster of CPUs. In section 6 conclusions are summarized.

2 Related works and contribution of the present work

Compared to other semi-enclosed and enclosed seas of the world, little is known of the Caspian Sea variability in terms of circulation, sea level and air-sea interaction [30]. DA software able to assimilate Sea Surface Temperature (SST) values in the Caspian Sea by keeping application requirements (especially about the size and the resolution of the real domain) does not already exist. In [34] the authors employ a DA model based on the simplified Kalman filter for adjust the variances of the prediction errors by assimilating climatic temperature into the primitive-equation model of water circulation. They present the results of an analysis of the seasonal variability of current fields in the Caspian Sea. In [24], instead, geostrophic velocities calculated from satellite altimetry and SST data were used together with model derived mean dynamic topography to document and try to better understand the seasonal and interannual variations of the Caspian Sea surface circulation. Both the approaches in [24] and [34] employ simplified models or reduced-order approaches [2]. The employment of these simplified models and reduced-order approaches alleviate the computational cost as these methods make the running less expensive and the parameters must still be selected a priori, nevertheless, a consequence is that important informations are missed [11].

The study and analysis of SST data collected into the Caspian Sea are object of interest of operative Centers. The ECMWF (European Centre for Medium-Range Weather Forecasts) provides analyses of SST as interpolations to the model grid of daily global datasets provided by The Metoffice in UK, with backup from OSTIA [47]. However, they provide just analysis of the data, while, at the moment, a scalable software for assimilating SST collected into the Caspian Sea is not available.

In this work, to better adapt the DA model and its implementation into a software with the physical phenomena object of our studies, we follow the problem-to-solve approach introduced into the previous section and we face the following issues:

1. *physical and mathematical assumptions* concerning the forecasting and observed data:
   The forecasting data which represent SST values into the Caspian Sea are produced by using the Regional Ocean Modeling System[3] (ROMS). The observations are satellite data provided by the Group of High resolution sea surface temperature[4] (GHRSST). We employ data collected in two of the months which present the most significant variability in water temperature: August and March [31]. The SST variabilities in the Caspian Sea have different characteristics in the different regions [30]. Caused their diversities, sometimes the studies

[2]The terms *"order reduction"* are used [8] to identify approaches able to lower the computational complexity of simulation problems. By a reduction of the model's associated state space dimension or degrees of freedom, an approximation to the original model is computed. This reduced-order model can then be evaluated with lower accuracy but in significantly less time.
[3]ROMS, Web page: www.myroms.org.
[4]GHRSST, Web page: www.ghrsst.org.

focus on the Northern Caspian, Middle Capian or Southern Caspian separately. This peculiarity suggests that a math DA model able to opportunely assimilate data on different parts of the domain indipendently could be recommended.

In the present work we employ the DA model described in [14, 15] based on a domain decomposition approach which splits the DA problem (let us say, the global problem) into several DA problems which reproduce the DA "global" problem at smaller dimensions (let us say, the local problems). About observed data, we use a spatial distribution named "model distribution" [46] which consists in assigning observed data to their related geographical regions.

2. *the algorithmic strategies*: The Data Assimilation is an ill posed inverse problem. Since a crucial point into DA models is the ill conditioning of the covariance matrices involved, it is mandatory to introduce, in a DA algorithm, preconditioning methods. The inherent ill conditioning of covariance matrices was investigated in the literature in different applications [21, 35]. In DA applications the behaviorof the condition number with respect to sampling distance, number of data points, domain size, for Gaussian-type covariances has been studied in [13, 25, 44]. Some of the relevant DA operative software [1, 6, 20] adopt the Empirical Orthogonal Functions (EOFs) method in order to reduce the ill conditioning and remove the statistically less significant modes which could add noise to the data assimilation estimate. EOFs implement a TSVD method. In order to improve the conditioning, only the Empirical Orthogonal Functions (EOFs) of the first largest eigenvalues of the error covariance matrix are considered. The EOFs (introduced by Edward Lorenz [40]) are the eigenvectors of the error covariance matrix, its condition number is reduced as well. Even if the employment methods as the TSVD, which strongly reduce the dimension, alleviate the computational cost as they make the running less expensive, nevertheless, a consequence is that important informations are missed [11]. This issue introduces a severe drawback to the reliability of the EOFs truncation, hence to the usability of the operative software in different scenarios [27].

   In the present work we employ the Tikhonov regularization which reveals to be more appropriate than the truncation of the EOFs as proved in [5] in which these methods are been applied to the Mediterranean sea data as well. In [5] the regularization parameter is computed by an algorithm based on a Regularization and Perturbation error estimates with respect to a reference solution provided by the EOFs truncation. Here we provide an estimation of the regularization parameter which is independent from any reference solution. We face experimentally the problem concerning the selection of an optimal regularization parameter picked to minimize both:

   – condition number of the DA problem after the preconditioning;
   – a relative Preconditioning Error defined to provide an estimate of how much the preconditioned problem differs from the starting problem.

   We evaluate the order of the error magnitude into the solution of the DA problem which reveals to be smaller by using the Tikhonov regularization with respect to the truncation of the EOFs.

3. *scalability expectation* on the computing environment in which the software is implemented: concerning the design of the algorithm to adapt to the evolutions of the node architectures, we focus on the important feature of the algorithm to be scalable [22]. Here scalability refers

to the capability of the algorithm to exploit performance of emerging computing architectures in order to minimise the time to solution for a given problem with a fixed dimension (strong scaling) (see [19, 18] as examples of works using both different and equivalent approaches).

In the present work, we show that although the Tikhonov regularization method results more expensive that the trucated EOFs in terms of time complexity, it is more efficient in terms of scalability:

- we validate theoretical results based on the evaluation of the Scale-Up factors [3, 4, 2] with experiments on a testbed which is an HPC environment.

Then we can conclude that the Tikhonov regularization method turns out to be more suitable to be used for our implementation on an HPC architecture with respect to the truncation of the EOFs.

3 Preliminaries

In this section we recall some preliminary concepts and definitions that we will use throughout the article [14, 33].

Definition 1 (The Data Assimilation problem) *Let* $\Omega = \{x_j\}_{j=1,\dots,N}$ *be a spatial domain and let* $[0, T_1] = \{t_k\}_{k=0,1,\dots,M}$ *be a time window. Let*

$$u_k^{\mathcal{M}} \equiv u(t_k) \in \Re^N \tag{1}$$

be a vector denoting the state of a sea system (it is often called background). At time t_k *it is* $u(t_k) = \mathcal{M}(u(t_{k-1}))$ *with* $\mathcal{M} : \Re^N \mapsto \Re^N$ *evolutive model often called forecasting model.*
At each time step t_k*, let be*

$$v_k = \mathcal{H}_k(u_k) \in \Re^p \tag{2}$$

the vector of observations where $\mathcal{H}_k : \Re^N \mapsto \Re^p$ *is a non-linear interpolation operator collecting the observations at time* t_k*.*
The aim of DA problem is to find an optimal tradeoff between the current estimate of the system state (the background) in (1) and the available observations v_k *in (2).*

Remark 3.1 *Let be* $p = N$*, that is the observations* v_k *in (2) are collected in the same space where the background* $u_k^{\mathcal{M}}$ *in (1) is defined for each time* t_k*. Then, the interpolation operator* $\mathcal{H}_k$ *is the identical operator and we denote it with* $\mathcal{I}_k$*.*

Definition 2 (3D Variational (3DVAR) Data Assimilation) *For a fixed time* $t_k = t_0$*, the 3DVAR computational model is a non-linear least square problem:*

$$u_0^{DA} = argmin_{u_0} J(u_0)$$

with J *(which is called cost-function) such that:*

$$J(u_0) = \|u_0 - u_0^{\mathcal{M}}\|_B^2 + \|\mathcal{H}(u_0) - v\|_R^2 \tag{3}$$

where R *and* B *are the covariance matrices whose elements provide the estimate of the errors on* v_k *and on* $u_0^{\mathcal{M}}$ *respectively.*

Definition 3 (Domain Decomposition) *The set of overlapping sub-domains*

$$DD(\Omega) = \{\Omega_i\}_{i=1,\ldots,N_{sub}} \tag{4}$$

is a decomposition of the domain $\Omega \subset \Re^N$*, if* $\Omega_i \subset \Re^{r_i}$*,* $r_i \leq N$ *and for* $i = 1, \ldots, N_{sub}$*, it is such that*

$$\cup_{i=1}^{N_{sub}} \Omega_i = \Omega \tag{5}$$

with

$$\Omega_i \neq \emptyset$$

and

$$\Omega_i \cap \Omega_j = \Omega_{ij} \neq \emptyset$$

Definition 4 (Domain Decomposition based 3DVAR (DD-3DVAR) Data Assimilation) *Let* $DD(\Omega) = \{\Omega_i\}_{i=1,\ldots,N_{sub}}$ *be an overlapping decomposition of the physical domain* Ω *as defined in (4).*
For a fixed time $t_k = t_0$*, according to this decomposition, the DD-3DVAR computational model is a system of* N_{sub} *non-linear least square problems:*

$$u_{0_i}^{DA} = argmin_{u_{0_i}} J_i(u_{0_i})$$

and J_i *such that:*

$$J_i(u_{0_i}) = \|u_{0_i} - u_{0_i}^{\mathcal{M}}\|_{B_i}^2 + \|\mathcal{H}_i(u_{0_i}) - v_i\|_{R_i}^2 + \|u_{0_{ij}} - u_{0_{ji}}\|_{B_{ij}}^2 \tag{6}$$

where

- u_{0_i} *and* v_{0_i} *are the same vectors* u_0 *and* v_0 *in (1) and (2) defined on the subdomain* Ω_i*;*
- $u_{0_{ij}}$ *and* $u_{0_{ji}}$ *are the vectors* u_{0_i} *and* u_{0_j} *on* Ω_{ij} *respectively;*
- R_i *and* B_i *are the covariance matrices whose elements provide the estimate of the errors on* v_{0_i} *and on* $u_{0_i}^{\mathcal{M}}$ *respectively;*
- B_{ij} *is the background error covariance matrix defined on* Ω_{ij}*.*

Then

$$u_0^{DA} = \sum_{i=1}^{N_{sub}} \tilde{u}_{0_i}^{DA} \quad \text{where} \quad \tilde{u}_{0_i}^{DA} = \begin{cases} u_{0_i}^{DA} & on \quad \Omega_i \\ 0 & on \quad \Omega - \Omega_i \end{cases} \tag{7}$$

Definition 5 (Singular Value Decomposition) *Let* $A \in \Re^{N\times M}$ *where* $N \geq M$ *and let*

$$A = U\Sigma W^T \tag{8}$$

be the singular value decomposition (SVD) of A *where* $U \in \Re^{N\times N}$ *and* $W \in \Re^{M\times M}$ *are orthogonal (or orthonormal) matrices and*

$$\Sigma = diag(\sigma_j)_{j=1,\ldots,N}$$

where singular values σ_j *appear in decreasing order:*

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N > 0 \quad .$$

If A is a matrix of an over-determined linear system then the discrete problem is ill posed, it is needed to filter out the contribution to the solution corresponding to the smallest singular values [28, 29]. In this case, it might make sense to look at the matrix numerical rank [23] and Singular Value Decomposition (SVD) enables us to deal with this concept. Filtering can be sharp (by recurring to the Truncated Singular Value Decomposition) or smooth (by recurring to the Tikhonov Regularization Matrix) as given in the following definitions:

Definition 6 (Truncated Singular Value Decomposition) *Let $A = U\Sigma W^T$ be the SVD of A as in (8). Let $\Phi_{trnc} \in \Re^{N\times N}$ be a matrix such that*

$$\Phi_{trnc} = diag(\underbrace{1,1,1,\ldots 1}_{trnc},0\ldots,0) \quad , \tag{9}$$

with $1 \leq trnc \leq N$. Then the matrix

$$A^{trnc} := U\Phi_{trnc}\Sigma W^T, \tag{10}$$

is the truncated SVD (TSVD) matrix for S.

Definition 7 (Tikhonov Regularization Matrix) *Let $A = U\Sigma W^T$ be the SVD of A as in (8). Let $\Phi_{Tikh(\lambda)} \in \Re^{N\times N}$ be a matrix such that*

$$\Phi_{Tikh(\lambda)} = diag\left(\frac{\sigma_j^2}{\sigma_j^2+\lambda^2}\right)_{j=1,\ldots,q} \quad , \tag{11}$$

with $\sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1}$. Then, the matrix

$$A^{Tikh(\lambda)} := U\Phi_{Tikh(\lambda)}W^T, \tag{12}$$

is the Tikhonov regularization matrix for S.

4 The Preconditioned Scalable 3DVAR Computational Kernel

Hereafter we provide a synthetic formalization of the data assimilation model we have implemented in Algorithm 1 and Algorithm 2 underlying the corrispondence between the algorithms steps and the mathematical-numerical issues we have faced.

The most popular software, developed in the operative centers, implement the so called incremental formulation of a 3DVAR DA model [1, 6, 20]. Here we consider the scalable version of the incremental DD-3DVAR cost function in (6) which is defined on a decomposition of the domain:

$$J_i(w_i) = \frac{1}{2}w_i^T w_i+\lambda_i\frac{1}{2}(H_iV_iw_i-d_i)^T R_i^{-1}(H_iV_iw_i-d_i)+\mu_i\frac{1}{2}(V_{ij}w_i^+-V_{ij}w_i^-)^T(V_{ij}w_i^+-V_{ij}w_i^-) \tag{13}$$

where

- H_i (Step 2 of Algorthm 1) is the matrix obtained by the first order approximation of the Jacobian of $\mathcal{H}_i$:
$$\mathcal{H}_i(u) = \mathcal{H}_i(u+\delta u) + H_i\ \delta u,$$
- $d_i = [v_i - \mathcal{H}_i(u_i^{\mathcal{M}})]$ (Step 3 of Algorthm 1) is the *misfit*,
- $w_i = V_i^T\delta x_i$, with V_i such that $B_i = V_iV_i^T$ (Steps 6 and 8 of Algorthm 1),

- V_{ij} is such that $B_{ij} = V_{ij}V_{ij}^T$,
- $w_i^+ = w_i$ on Ω_{ij} and $w_j^- = w_j$ on Ω_{ij}.
- R_i is such that

$$R_i = \epsilon_o I_i \tag{14}$$

Steps 9-14 of Algorthm 1 computes the minimum of the cost function J_i in (13) by using the L-BFGS method [45].
The convergence rate of L-BFGS depends on the conditioning of the numerical problem, i.e. it depends on the condition number of the preconditioned Hessian of the cost function (13) [26] which is:

$$D_i = I_i + (H_iV_i)^T R_i^{-1} H_iV_i. \tag{15}$$

Let be $H_i = I_i$, that is under the hypothesis of Remark 3.1, and let $\mu(\cdot)$ denotes the condition number [28], from (14) and for the properties of the condition number, it descends immediately:

$$\mu(D_i) \simeq 1 + \frac{1}{\epsilon_o^2}\mu(V_i)^2 \ . \tag{16}$$

The accuracy and efficiency with which the minimization problem (13) can be solved is determined by the condition number of the error covariance matrix V_i in (15) and (16) [26, 14]. As V_i is ill conditioned, preconditioning methods must be used for improving its conditioning [44]. Then, the matrix V_i (see Step 7 of the Algoritm 1) is computed by Algorithm 2 which implements two preconditioning approaches: the truncation of the Empirical Orthogonal Functions (EOFs) method which consists of a TSVD of the matrix (see (10)) and the Tikhonov regularization method (see (12)). In section 5 we use some parameters to evaluate accuracy and efficiency for some case studies.

Algorithm 1 *the S3DVAR algorithm on each subdomain* Ω_i

1: Input: v_i and $u_{0_i}^{\mathcal{M}}$
2: Define H_i ▷ interpolation operator
3: Compute $d_i \leftarrow v_i - H_i u_{0_i}^{\mathcal{M}}$ ▷ compute the misfit
4: Define R_i ▷ covariance matrix of the observed data v_i
5: Define the initial value of δu_i^{DA}
6: Compute the covariance matrix V_i by a temporal sequence of hystorical data $\{u_{k_i}^{\mathcal{M}}\}_{k=0,\ldots,M}$
7: Compute $\hat{V}_i = PECM(PrecondType, ind, N, M, V_i)$ ▷ See Algorithm 2 for details
8: Compute $w_i \leftarrow \hat{V}_i^T \delta u_i^{DA}$
9: **repeat** ▷ start of the L-BFGS steps
10: Send and Receive the boundary conditions from the adjacent domains
11: Compute $J_i \leftarrow J_i(w_i)$
12: Compute $gradJ_i \leftarrow \nabla J_i(w_i)$
13: Compute new values for w_i
14: **until** (Convergence on w_i is obtained) ▷ end of the L-BFGS steps
15: Compute $u_{0i}^{DA} \leftarrow u_{0_i}^{\mathcal{M}} + \hat{V}_i w_i$

end

In [1, 6, 20] the conditioning of V_i is reduced by truncating the EOFs of V_i, i.e. a matrix V_i^{trnc} is computed by using the TSVD of V_i. Here we introduce in Algorithm 2 the use of the Tikhonov regularization matrix $V_i^{Tikh(\lambda)}$ which reveals to be more appropriate than truncation of EOFs [5].

Algorithm 2 *the Preconditioning Error Covariance Matrix (PECM) algorithm*

1: Input: $PrecondType$, N, M, V_i
2: Compute $V_i = U\Sigma W^T$ ▷ compute the SVD of the matrix V_i
3: **if** $(PrecondType = EOFS)$ **then**
4:     Compute $trnc$ ▷ compute the truncation parameter in (9)
5:     Compute $\hat{V} = TSVD(V_i, trnc)$ ▷ Truncated SVD regularized matrix $\hat{V} = U\Phi_{trnc}\Sigma W^T$
6: **else** ▷ $PrecondType = Tikhonov$
7:     Compute λ_{opt} ▷ compute the regularization parameter in (11)
8:     Compute $\hat{V} = Tikhonov(V_i, \lambda_{opt})$ ▷ Tikhonov regularized matrix $\hat{V} = U\Phi_{Tikh}W^T$
9: **end if**

end

The solutions computed on the subdomains Ω_i by Algorithm 1 and Algorithm 2 are then collected to provide the global solution as in (7). In order to guarantee continuity of the solution along the overlapping regions we force the solution to satisfy the condition:

$$u_0^{DA}/\Omega_{ij} = mean(u_{0i}^{DA}/\Omega_{ij}, u_{0j}^{DA}/\Omega_{ij}), \quad \forall\Omega_{ij} = \Omega_i \cap \Omega_j \neq 0.$$

where $mean(\cdot,\cdot)$ denotes the mean value.
In the next section we face the choice of the regularization parameter λ (Step 7 in Algorithm 2) and the truncation parameter $trnc$ (Step 4 of Algorithm 2) and we provide results of the DA minimization problem (13) obtained by considering both preconditioning methods.

5 Implementation Details and Permormance Analysis

Here we focus on the main computational issues we faced by implementing Algorithm 1 and Algorithm 2 and we present a performance analysis both in terms of accuracy and efficiency of the results obtained by introducing the Tikhonov regularization method compared to the truncated EOFs.

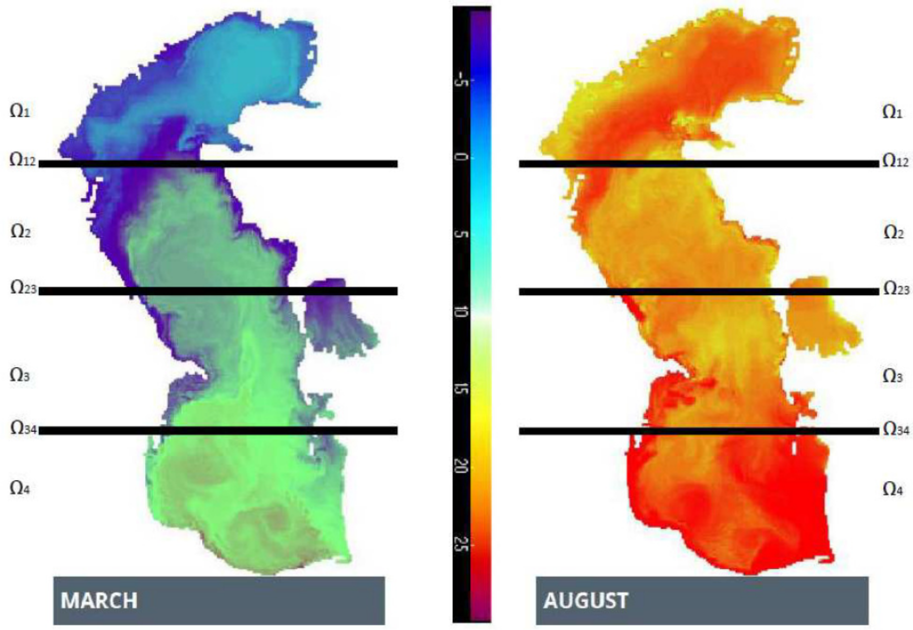
The background data (defined in (1)) we consider are provided by the software ROMS. The satellite observations (defined in (2)) provided by the GHRSST give us information about the SST every day of the selected months at 12:00am according with the data provided by ROMS.

Background and observed data are defined on the same spatial domain, which means that hypothesis of Remark 3.1 are satisfied, and the discretization grid has dimension

$$N = N_1 \times N_2 = 780 \times 560 \ . \tag{17}$$

Then the problem size is $\mathcal{O}(10^6)$.

We employ the DD-3DVAR model in (13). Due the geometry of the domain, we decided to introduce a domain decomposition, as defined in (4), along the coordinates N_1. In Figure 1 is shown an example on horizontal decomposition obtained for $N_{sub} = 4$. Then we emply Algorithm 1 and



(a) (b)

Figure 1: Values of Sea Surface Temperature collected into the Caspian Sea on (a) March 2008 and (b) August 2008 and an example of decomposition obtained along the coordinate N_1 for a number of subdomain $N_{sub} = 4$.

Algorithm 2 on each subdomain Ω_i of the decomposition in parallel.

The architecture we use for developing is a distributed memory architecture made of 8 DELL M600 blades connected by a 10 Gigabit Ethernet technology. Each blade consists of 2 Intel Xeon@2.33GHz quadcore processors sharing the same local 16 GB RAM memory for a total of 16 processors. We implemented the algorithm in the MATLAB®computational environment using some of its native procedures (i.e., to compute SVD and TSVD of matrices) and external toolboxes such as the Parallel Computing Toolbox and the MATLAB interface for L-BFGS-B routine. We observe that the MATLAB Parallel Computing toolbox is able to exploit different kind of computing architectures: multicore nodes, cluster of nodes and heterogeneous computing systems.

Experiments are provided on data collected in two peculiar months: August 2008 and March 2008 [30] and the chosen starting point for assimilating data is been fixed as the first of August and the first of March respectively.

5.1 Setting Up of the regularization and truncation parameters

We face experimentally the problem concerning the selection of an optimal regularization parameter picked to minimize both:

- condition number of V_i after the preconditioning, i.e. condition number of $V_i^{Tikh(\lambda)}$;

- a relative Preconditioning Error defined to provide an estimate of how much the preconditioned problem differs from the starting problem.

We have computed matrix V_i in Step 6 of Algorithm 1 by considering two temporal sequence of data collected in August 2008 and March 2008. Then we have applied the Tikhonov regularization method in Step 8 of Algorithm 2 which has provided $V_i^{Tikh(\lambda)}$ with values of λ such that:

$$0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 49.35$$

for data collected in August and

$$0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 62.89$$

for data collected in March, and we have computed the condition number of $V_i^{Tikh(\lambda)}$ as function of λ.

The trends of the computed condition number in Figure 2 confirm our expectation. Infact, as the condition number $\mu(V_i^{Tikh(\lambda)})$ is such that [29]:

$$\mu(V_i^{Tikh(\lambda)}) \simeq \frac{\sigma_1}{2\sqrt{\lambda}} \ , \tag{18}$$

it is

$$\lim_{\lambda \to 0} \mu(V_i^{Tikh(\lambda)}) = +\infty, \quad \lim_{\lambda \to +\infty} \mu(V_i^{Tikh(\lambda)}) = 0, \tag{19}$$

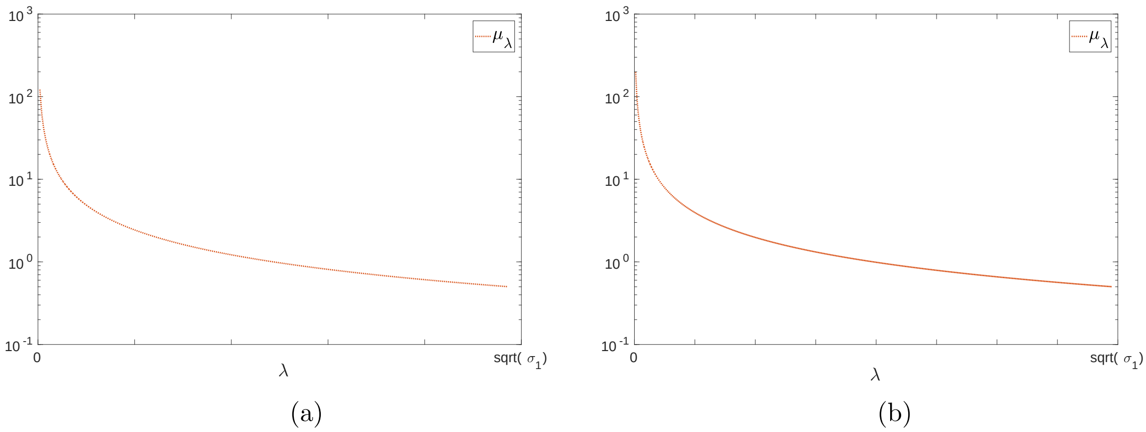
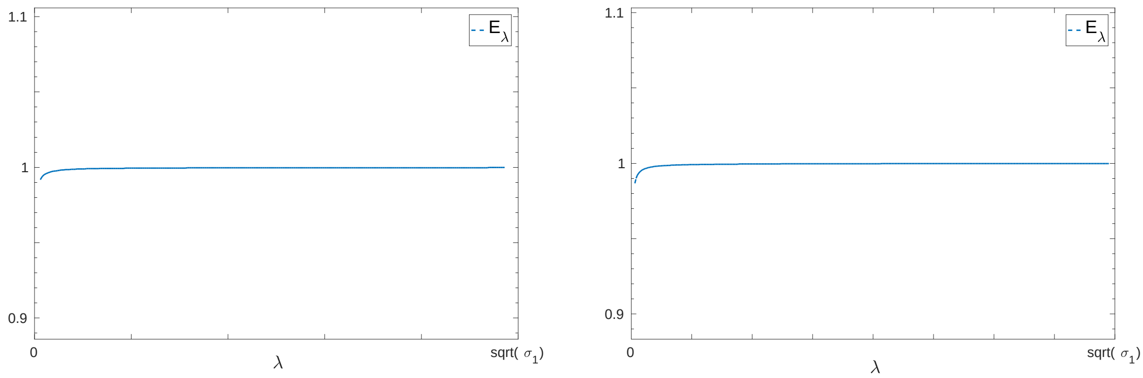


Figure 2: condition numbers of the matrix $V_i^{Tikh(\lambda)}$ where (a) $0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 49.35$ for data collected in August and (b) $0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 62.89$ for data collected in March.

In Figure 3, instead, we have evaluated the relative Preconditioning Error defined as:

$$E_\lambda = \frac{\|\Sigma - \Phi_{Tikh(\lambda)}\|_\infty}{\|\Sigma\|_\infty} \ . \tag{20}$$

which provides an estimate of how much the preconditioned problem differs from the starting problem.



(a) (b)

Figure 3: values of the relative error E_λ where (a) $0 = \sqrt{\sigma_N} \le \lambda \le \sqrt{\sigma_1} = 49.35$ for data collected in August and (b) $0 = \sqrt{\sigma_N} \le \lambda \le \sqrt{\sigma_1} = 62.89$ for data collected in March.

As confirmed by results in Figure 3, it is

$$\lim_{\lambda \to 0} E_\lambda = \frac{\|\Sigma - I\|_\infty}{\|\Sigma\|_\infty} \simeq 1 - \frac{1}{\sigma_1}, \quad \lim_{\lambda \to +\infty} E_\lambda = \frac{\|\Sigma\|_\infty}{\|\Sigma\|_\infty} = 1 \tag{21}$$

As λ is subject to the constraints [28] $\sqrt{\sigma_N} \le \lambda \le \sqrt{\sigma_1}$, we have that, from (19), the smallest value of the condition number is obtained for $\lambda \simeq \sqrt{\sigma_1}$. From (21), however, the smallest error is obtained for $\lambda \simeq 0 = \sqrt{\sigma_N}$. Then, we observe that the optimal value $\lambda = \lambda_{opt}$ should be such that:

$$\lambda_{opt} \simeq mean(\sqrt{\sigma_N}, \sqrt{\sigma_1}). \tag{22}$$

Figure 4 confirms this observation and in Table 1 are reported the values computed as intersection of the curves described by $\mu(V_i^{Tikh(\lambda)})$ and E_λ for data collected in August and March.

The truncation parameter, instead, is been chosen by evaluating the numerical rank of V_i [5, 29, 28] such that

$$\sigma_{trnc} >> trnc >> \sigma_{trnc+1}$$

then by studying the spectrum of the matrix V_i computed in Step 6 of Algorithm 1. In our case study, it is $trnc = 20$ both for August and March;

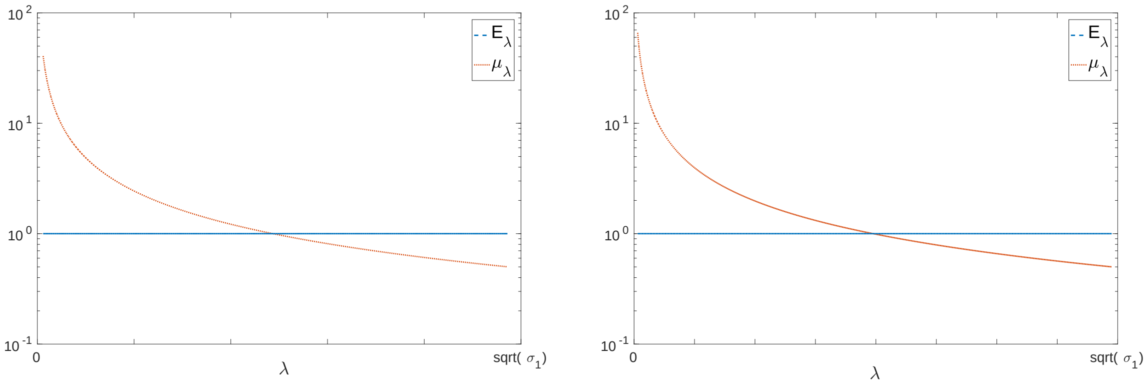
5.2 Validation of the results

Let p denotes the number of processors involved, in the following we assume

$$N_{sub} \leftrightarrow p \ , \tag{23}$$

i.e. each subdomain is assimilated on a processor.

Validation is carried out by performing the following steps:



(a) (b)

Figure 4: values of (a) $\lambda_{opt} = 34.899$ and (b) $\lambda_{opt} = 44.463$ computed as intersection of the curves described by $\mu(V_i^{Tikh(\lambda)})$ and E_λ where (a) $0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 49.35$ for data collected in August and (b) $0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 62.89$ for data collected in March.

1. Analysis of the results in terms of *accuracy* based on the *errors* evaluation:

   We evaluate the order of magnitude of the errors into the solution computed by using both the Tikhonov regularization and the EOFs

   $$err^{Tikh(\lambda_{opt})} = \|u_0^{DA(Tikh(\lambda_{opt}))} - v_C\|_\infty \ , \quad err^{EOFs} = \|u_0^{DA(EOFs)} - v_C\|_\infty \qquad (24)$$

   where v_C is the control variable provided by the Space and Atmospheric Physics Group at Imperial College London for these hystorical data collected into the Caspan Sea. We show the results obtained as function of the number N_{sub} of subdomains which constitute the domain decomposition.

   Figure 5 shows the values of the errors $err^{Tikh(\lambda_{opt})}$ and err^{EOFs} as function of $p = N_{sub}$ for data collected in (a) August 2008 and in (b) March 2008. Details about these values are also provided in Table 2.
   Results carried out from the assimilation of the data collected in March 2008 present a small increase of the error into the DA solution. It is because in March the presence of ice [30, 48] imply a reduction into the accuracy of the data, then (as also stated in [27]) a consequence is that the reliability of the EOFs truncation, hence its usability into operative software, is disadvantageous. Accuracy into the solution provided by considering the Tikhonov regularization method is instead maintained.
   For both sets of results it holds that:

   $$\frac{err^{Tikh(\lambda_{opt})}}{err^{EOFs}} \simeq \mathcal{O}(10^{-1})$$

   which means that the Tikhonov regularization method provides a solution with an error of one order of magnitude smaller then the truncation of the EOFs.

2. Analysis of the results in terms of *efficiency*:

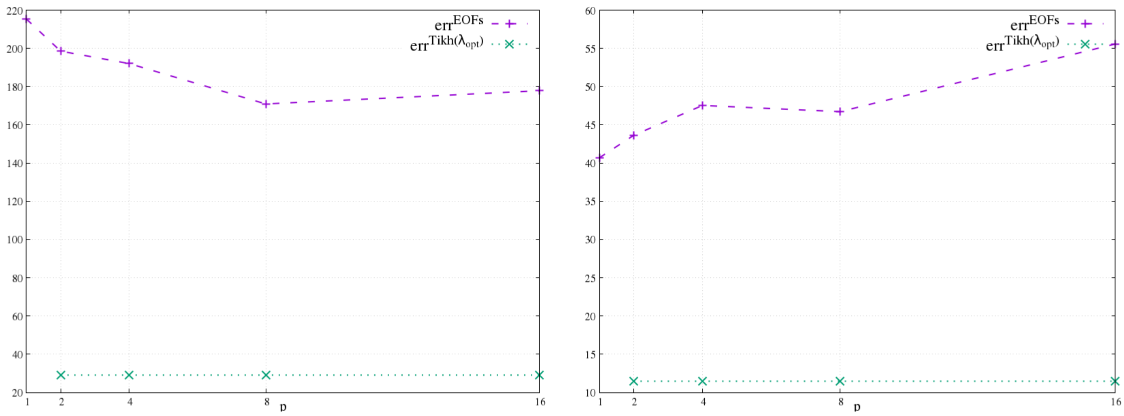
	$\lambda_{opt}-\Delta\lambda$					λ_{opt}					$\lambda_{opt}+\Delta\lambda$
λ	14.764	20.445	24.859	28.601	31.906	34.899	37.656	40.224	42.638	44.922	47.095
E_λ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
μ_λ	5.5860	2.9133	1.9705	1.4887	1.1962	0.9998	0.8588	0.7526	0.6698	0.6034	0.5490

(a)

	$\lambda_{opt}-\Delta\lambda$					λ_{opt}					$\lambda_{opt}+\Delta\lambda$
λ	31.257	34.307	37.108	39.711	42.154	44.463	46.658	48.754	50.764	52.697	54.561
E_λ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
μ_λ	2.0240	1.6801	1.4360	1.2539	1.1128	1.0002	0.9083	0.8319	0.7673	0.7121	0.6642

(b)

Table 1: values of (a) $\lambda_{opt} = 34.899$ and (b) $\lambda_{opt} = 44.463$ computed as intersection of the curves described by $\mu(V_i^{Tikh(\lambda)})$ and E_λ where (a) $0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 49.35$ for data collected in August and (b) $0 = \sqrt{\sigma_N} \leq \lambda \leq \sqrt{\sigma_1} = 62.89$ for data collected in March.



(a) (b)

Figure 5: Values of errors $err^{Tikh(\lambda_{opt})}$ and err^{EOFs} as function of $p = N_{sub}$ for data collected in (a) August 2008 and in (b) March 2008.

We evaluate the performance of the software in terms of execution times and scalability. We evaluate the Scale Up factor [3, 14] which is (by assuming (23)) defined as:

$$S_{p_1,p_2}^{Method} = \frac{\mathcal{T}(\mathcal{A}(p_1, Method))}{(p_2/p_1)\mathcal{T}(\mathcal{A}(p_2, Method))} \tag{25}$$

where

- $Method$ denotes one of the methods: Tikhonov regularization or truncation of EOFs implemented in Algorithm 2;
- $\mathcal{A}(p, Method)$ denotes the Algorithm obtained from Algorithm 1 on p processor and Algorithm 2 with $PrecondType = Method$,

p	$err^{Tikh(\lambda_{opt})}$	err^{EOFs}
1	-	2.153218920796193458500056e+02
2	2.922676086425781250000000e+01	1.985537308080609477656253e+02
4	2.922676086425781250000000e+01	1.920829574240959800590645e+02
8	2.922676086425781250000000e+01	1.709302357650879002903821e+02
16	2.922676086425781250000000e+01	1.778700303366430546248012e+02

(a)

p	$err^{Tikh(\lambda_{opt})}$	err^{EOFs}
1	-	4.069147530943650536983114e+01
2	1.148108196258544921875000e+01	4.363353211828277267159137e+01
4	1.148108196258544921875000e+01	4.754367594091701221259427e+01
8	1.148108196258544921875000e+01	4.672725439802213998063962e+01
16	1.148108196258544921875000e+01	5.556904118757078947510308e+01

(b)

Table 2: Values of errors $err^{Tikh(\lambda_{opt})}$ and err^{EOFs} as function of $p = N_{sub}$ for data collected in (a) August 2008 and in (b) March 2008.

- $\mathcal{T}(\mathcal{A}(p_1, Method))$ denotes the time complexity of the Algorithm $\mathcal{A}(p, Method)$ by using p_1 processor and $\mathcal{T}(\mathcal{A}(p_2, Method))$ denotes the time complexity of the Algorithm $\mathcal{A}(p, Method)$ by using p_2 processor ($p_2 \le p_1$).

Concerning the estimate of the theoretical Scale Up Factor values as defined in [14] we can affirm that

$$S_{p_1,p_2}^{Tikh(\lambda_{opt})} > S_{p_1,p_2}^{EOFS}. \tag{26}$$

Infact, as

$$\mathcal{T}(\mathcal{A}(p_1, Tikh(\lambda))) \simeq \mathcal{O}(N^3) \tag{27}$$

and

$$\mathcal{T}(\mathcal{A}(p_1, EOFs)) \simeq \mathcal{O}(trnc \cdot N^2). \tag{28}$$

Let be $N = N_1 \times N_2$ as in (17). By considering (25), and by considering that we decompose just along N_1, it holds that for:

- $Method = Tikh(\lambda)$, then by using (27) in (25), it is:

$$S_{p_1,p_2}^{Tikh(\lambda_{opt})} \simeq \frac{\frac{N_1^3}{p_1^3}\ N_2^3}{(p_2/p_1)\ \frac{N_1^3}{p_2^3}\ N_2^3} \simeq \left(\frac{p_2}{p_1}\right)^2 . \tag{29}$$

- $Method = EOFs$, then by using (28) in (25), it is:

$$S_{p_1,p_2}^{EOFS} \simeq \frac{trnc\ \frac{N_1^2}{p_1^2}\ N_2^2}{(p_2/p_1)\ trnc\ \frac{N_1^2}{p_2^2}\ N_2^2} \simeq \frac{p_2}{p_1} ; \tag{30}$$

Then from (29) and (30), the (26) follows.

The result provided in (26) suggests that the Tikhonov regularization method better exploit the resurces provided by an HPC computing environment for solving the minimization DA problem in (13). It is been confirmed by the experimental results we carried out. In fact, the Tikhonov regularization method results more expensive that the trucated EOFs in terms of execution times (as shown in Figure 6 (a) and (b)). However, values of measured Scale Up, as in Figure 6 (c) and (d), show as it is more efficient in terms of scalability.

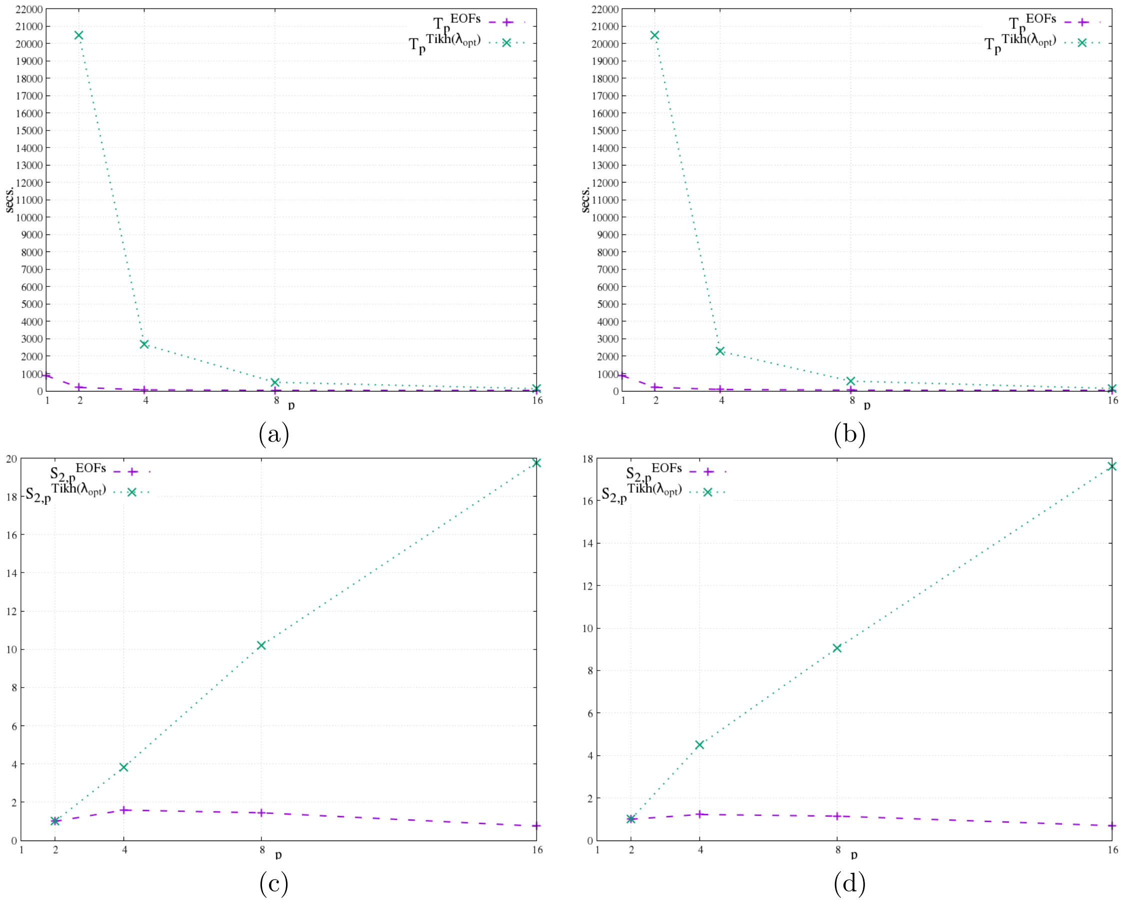


Figure 6: Execution times as function of the number of subdomains for data collected in (a) August and (b) March 2008 and values of the Measured Scale Up $S_{p_1,p_2}^{Tikh(\lambda_{opt})}$ and S_{p_1,p_2}^{EOFs} as function of the number of subdomains for data collected in (c) August and (d) March 2008

Then we may conclude that the employment of the Tikhonov regularization method is more suitable for the data assimilation algorithm we are developing for the data collected into the Caspian sea when it is implemented on an HPC architecture such as a cluster of CPUs.

6 Conclusions

We have presented first results obtained introducing two different preconditioning methods (namely the Tikhonov regularization method and the truncated EOFs) in a DA software we are developing (we named S3DVAR) which implements a Scalable Three Dimensional Variational Data Assimilation model for assimilating sea surface temperature (SST) values collected into the Caspian Sea by using the Regional Ocean Modeling System (ROMS) with observations provided by the Group of High resolution sea surface temperature (GHRSST). We have presented the algorithmic strategies we have employed and the numerical issues on data collected in two of the months which present the most significant variability in water temperature: August and March. We have evaluated the performance obtained both in terms of accuracy and efficiency. Results we carried out show how the Tikhonov regularization method is more accurate in terms of mean error. Also, although the Tikhonov regularization method results more expensive that the trucated EOFs in terms of execution time, we proved that it is more efficient in terms of scalability on HPC architectures. Then we can conclude that the Tikhonov regularization method is more suitable for the data assimilation algorithm we are developing for the data collected into the Caspian sea.

Acknowledgment

This work was developed within the research activity of the H2020-MSCA-RISE-2015 NASDAC Project N. 691184. It has been realized thanks to the use of the S.Co.P.E. computing infrastructure at the University of Naples Federico II.

References

[1] E. Andersson et al., *The ECMWF implementation of three dimensional variational assimilation (3DVar). Part III: Experimental results*, Quarterly Journal Royal Met. Society. Vol. 124, No. 550, pp. 1831-1860, 1998.

[2] R. Arcucci, L. D'Amore, L. Carracciuolo, *On the Problem Decomposition of Scalable 4D-Var Data Assimilation Models*, HPCS-IEEE, 978-1-4673-7812-3, pages "589–594", 2015.

[3] R. Arcucci, L. D'Amore, L. Carracciuolo, G. Scotti, G. Laccetti, *A Decomposition of the Tikhonov Regularization Functional Oriented to Exploit Hybrid Multilevel Parallelism*, Int J Parallel Prog, vol. 45, Issue 5, pp 12141235, 2017.

[4] R. Arcucci, L. D'Amore, S. Celestino, G. Laccetti, A. Murli, *A scalable numerical algorithm for solving Tikhonov regularization problems*, Lecture Notes in Computer Science, vol. 9574, pages "45–54", 2016.

[5] R. Arcucci, L. D'Amore, J. Pistoia, R. Toumi and A. Murli, *On the variational data assimilation problem solving and sensitivity analysis*, Journal of Computational Physics, vol. 335, pages 311-326 2017.

[6] D. M. Baker, W. Huang, Y.R. Guo, J. Bourgeois, Q.N. Xiao, *Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results*, Mon. Wea. Rev., 132, pp. 897-914, 2004.

[7] G. B. Barone, V. Boccia, D. Bottalico, R. Campagna, L. Carracciuolo, G. Laccetti, M. Lapegna, *An Approach to Forecast Queue Time in Adaptive Scheduling: How to Mediate System Efficiency and Users Satisfaction*, Int J Parallel Prog, vol. 45, pp 1164-1193, 2017

[8] P. Benner, H. Fassbender, *Model Order Reduction: Techniques and Tools*, Encyclopedia of Systems and Control, Springer, doi:10.1007/978-1-4471-5102-9_142-1, 2014.

[9] V. Boccia, L. Carracciuolo, G. Laccetti, M. Lapegna, V. Mele, *HADAB: enabling fault tolerance in parallel applications in distributed environments*, chapter of the book Parallel Processing and Applied Mathematics PPAM 2011, LNCS n. 7203, Springer, 2012, pag. 700-709

[10] N. Bruneau, R.Toumi, *A fully-coupled atmosphere-ocean-wave model of the Caspian Sea*, Ocean Modelling vol 107, pp 97–111, 2016.

[11] D.G. Cacuci, I. M. Navon, M. Ionescu-Bujor, *Computational Methods for Data Evaluation and Assimilation*, CRC Press 2013.

[12] L. Carracciuolo, L. D'Amore, V. Mele, *Toward a fully parallel Multigrid in Time algorithm in PETSc environment: a case study in ocean models*, in proceedings of International Conference on High Performance Computing & Simulation (HPCS) 2015, Amsterdam, 2015, pp. 595-598.

[13] S. Cuomo, A. Galletti, R. Farina, L. Marcellino, *An error estimate of Gaussian Recursive Filter in 3Dvar problem*, Preprints of the Federated Conference on Computer Science and Information Systems, ACSIS, Vol. 2, pp: 593601, 2014

[14] L. D'Amore, R. Arcucci, L. Carracciuolo, A. Murli, *A Scalabale Variational Data Assimilation*, Journal of Scientific Computing, vol. 61, pages "239–257", 2014.

[15] L. D'Amore and R. Arcucci and L. Carracciuolo and A. Murli, DD-OceanVar: a Domain Decomposition fully parallel Data Assimilation software in Mediterranean Sea, Procedia Computer Science 18, pages "1235–1244", 2013.

[16] L. D'Amore, R. Arcucci, L. Marcellino, A. Murli, *HPC computation issues of the incremental 3D variational data assimilation scheme in OceanVar software*, Journal of Numerical Analysis, Industrial and Applied Mathematics, vol.7, pages "91–105", 2013.

[17] L. D'Amore and R. Arcucci and L. Marcellino and A. Murli, A parallel three-dimensional variational data assimilation scheme, AIP Conference Proceedings 1389, pages "1829–1831", 2011.

[18] L. D'Amore, V. Mele, G. Laccetti, A. Murli, *Mathematical Approach to the Performance Evaluation of Matrix-matrix Multiply Algorithm on a Two Level Parallel Architecture*, chapter of the book Parallel Processing and Applid Mathematics PPAM 2015, Part II, LNCS 9574, Springer, 2016, pp. 25-34.

[19] L. D'Amore, V. Mele, A. Murli, *Performance analysis of the Taylor expansion coefficients computation as implemented by the software package TADIFF*, Journal of Numerical Analysis, Industrial and Applied Mathematics, Vol. 8; p. 1-10.

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

[21] M.P. Ekstrom, *A spectral characterization of the ill conditioning in numerical deconvolution*, IEEE Trans. Audio Electroacoust. AU-21(4) (1973) 344348.

[22] G.C. Fox, R.D. Williams, P.C. Messina , *P.C.: Parallel Computing Works!*, Morgan Kaufmann Publishers Inc., Los Altos, CA, 1994

[23] G.H. Golub, C.F. Van Loan, Matrix Computations, John Hopkins University Press, 1996.

[24] M. Gunduz, *Caspian Sea surface circulation variability inferred from satellite altimeter and sea surface temperature*, Journal of Geophysical Research: Oceans 119, pages 1420-1430, 2014.

[25] S. A. Haben, *Conditioning and Preconditioning of the Minimisation Problem in Variational Data Assimilation*, PhD. thesis, The University of Reading, 2011.

[26] S. A. Haben, A. S. Lawless, N . K . Nichols, *Conditioning of incremental variational data assimilation, with application to the Met Office system*, Tellus 63A, 782792, 2011.

[27] A. Hannachi, *A Primer for EOF Analysis of Climate Data*, Department of Meteorology, University of Reading Reading RG6 6BB, U.K., 2004.

[28] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems, numerical aspects of linear inversion*, SIAM, 1998.

[29] P. C. Hansen, J. G. Nagy, D. P. OLeary, *Deblurring Images: Matrices, Spectra, and Filtering*, SIAM 2006.

[30] R. Ibrayev, E. Ozsoy, C. Schrum, H. Sur, *Seasonal variability of the caspian sea three-dimensional circulation, sea level and air-sea interaction*, Ocean Science Discussions 6, pages "1913–1970", 2009.

[31] S. Jamshidia, N. B. Abu Bakar, *Seasonal Variations in Temperature, Salinity and Density in the Southern Coastal Waters of the Caspian Sea*, Oceanology Vol. 52 No. 3, pages 380-396, Pleiades Publishing Inc., 2012.

[32] R.E. Kalman, *A new approach to linear filtering and prediction problems*, Trans. ASME J. Basic Eng.82(Series D), 3545, 1960.

[33] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge, MA 2003.

[34] V. V. Knysh, R. A. Ibrayev, G. K. Korotaev, N. V. Inyushina, *Seasonal Variability of Climatic Currents in the Caspian Sea Reconstructed by Assimilation of Climatic Temperature and Salinity into the Model of Water Circulation*, Atmospheric and Oceanic Physics, Vol. 44, No. 2, pages 236-249, 2008.

[35] A. B. Kostinski, A.C. Koivunen, *On the condition number of Gaussiansample covariance matrices*, IEEE Trans. Geosci. Remote Sens. 38 329332, 2000.

[36] G. Laccetti, M. Lapegna, V. Mele, *A Loosely Coordinated Model for Heap-Based Priority Queues in Multicore Environments*, International Journal of Parallel Programming, Springer, 44(4), 901-921

[37] G. Laccetti, M. Lapegna, V. Mele, D. Romano, *A study on adaptive algorithms for numerical quadrature on hybrid GPU and multicore based systems*, chapter of the book Parallel Processing and Applid Mathematics PPAM 2013, LNCS n. 8384, Springer, pag. 704-713, 2014.

[38] G. Laccetti, M. Lapegna, V. Mele, D. Romano, A. Murli, *A double adaptive algorithm for multidimensional integration on multicore based HPC systems*, International Journal Of Parallel Programming, Springer, vol. 40, pp. 397-409

[39] F.X. Le Dimet, O. Talagrand, *Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects*, Tellus 38A, 97110, 1986.

[40] E. N. Lorenz , *Empirical orthogonal functions and statistical weather prediction*, Sci.Rep. No. 1, Statistical Forecasting Project, M.I.T., Cambridge, MA, 1956.

[41] B.A. Marx, R.W.E. Potthast, *On Instabilities in Data Assimilation Algorithms*, University of Reading, Department of Mathematics and Statistics, Preprint MPS-2012-06, 2012.

[42] R. Montella, G. Coviello, G. Giunta, G. Laccetti, F. Isaila, J.G. Blas, *A general-purpose virtualization service for HPC on cloud computing: An application to GPUs*, chapter of the book Parallel Processing and Applied Mathematics, PPAM 2011, LNCS vol. 7203, part 1, Springer, 2012, pages 740-749.

[43] J. F. Nicholls, R. Toumi, *On the lake effects of the Caspian Sea*, Quarterly Journal of the Royal Meteorological Society, Soc. 140, pages " 1399–1408", 2014.

[44] N. Nichols, *Mathematical Concepts in Data Assimilation*, W. Lahoz et al. (eds), Data Assimilation, Springer, 2010.

[45] J. Nocedal, R.H. Byrd, P. Lu, C. Zhu, *L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization*, ACM Transactions on Mathematical Software, Vol. 23, No. 4, pages "550–560", 1997.

[46] J. Rantakokko, *Strategies for parallel variational data assimilation*, Parallel Computing 23, Elsevier, pp. 2017-2039, 1997.

[47] J. D. Stark, C. J. Donlon, M. J. Martin, M. E. McCulloch, *Ostia : An operational, high resolution, real time, global sea surface temperature analysis system*, Oceans 07 IEEE Aberdeen, conference proceedings. Marine challenges: coastline to deep sea. Aberdeen, Scotland. IEEE., 2007.

[48] H. Tamura-Wicks, R. Toumi, W. Paul Budgell, *Sensitivity of Caspian sea-ice to air temperature*, Quarterly Journal of the Royal Meteorological Society, Soc. 141, pages "3088–3096", 2015.