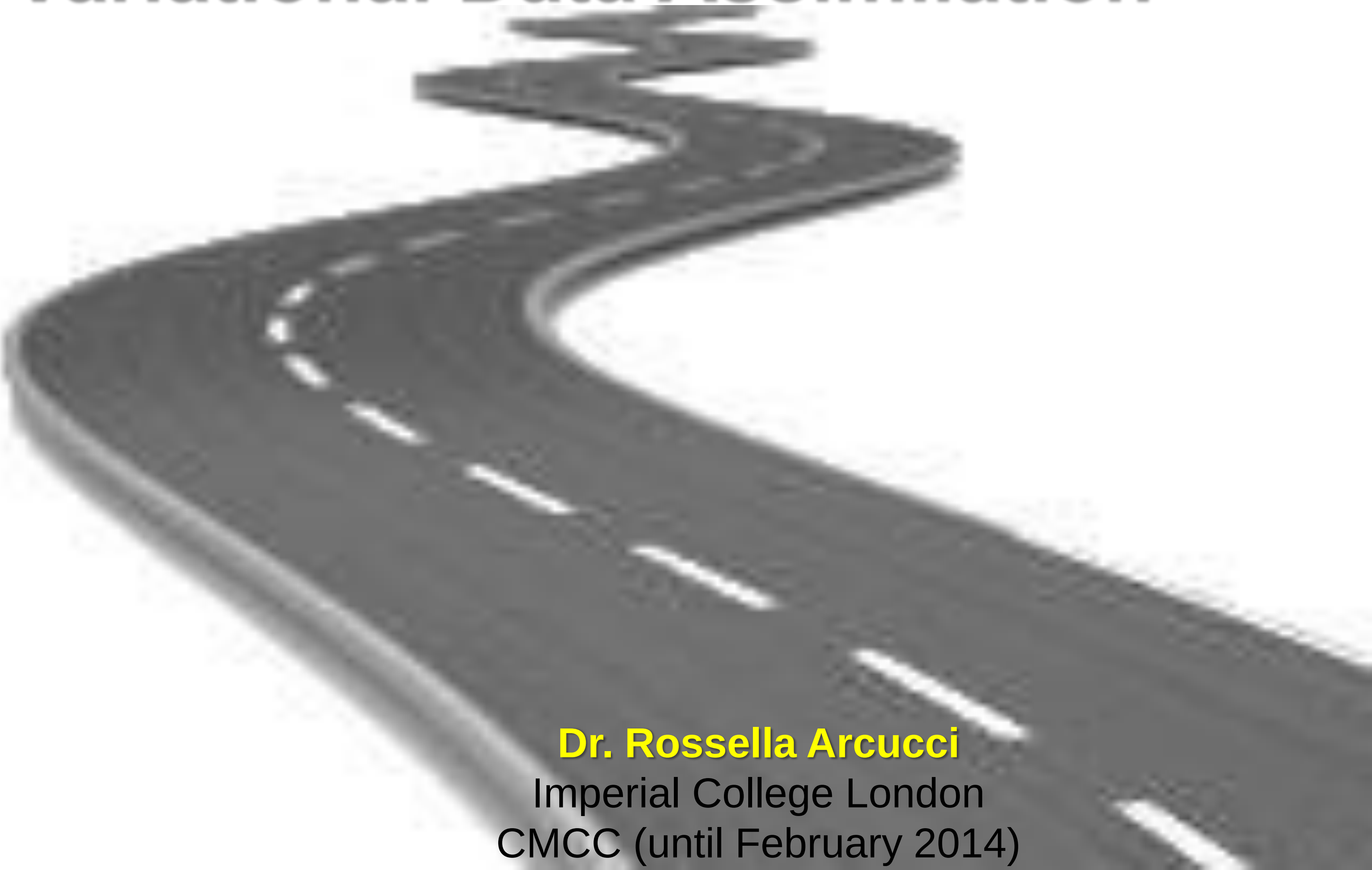
A Scalable Approach for Variational Data Assimilation



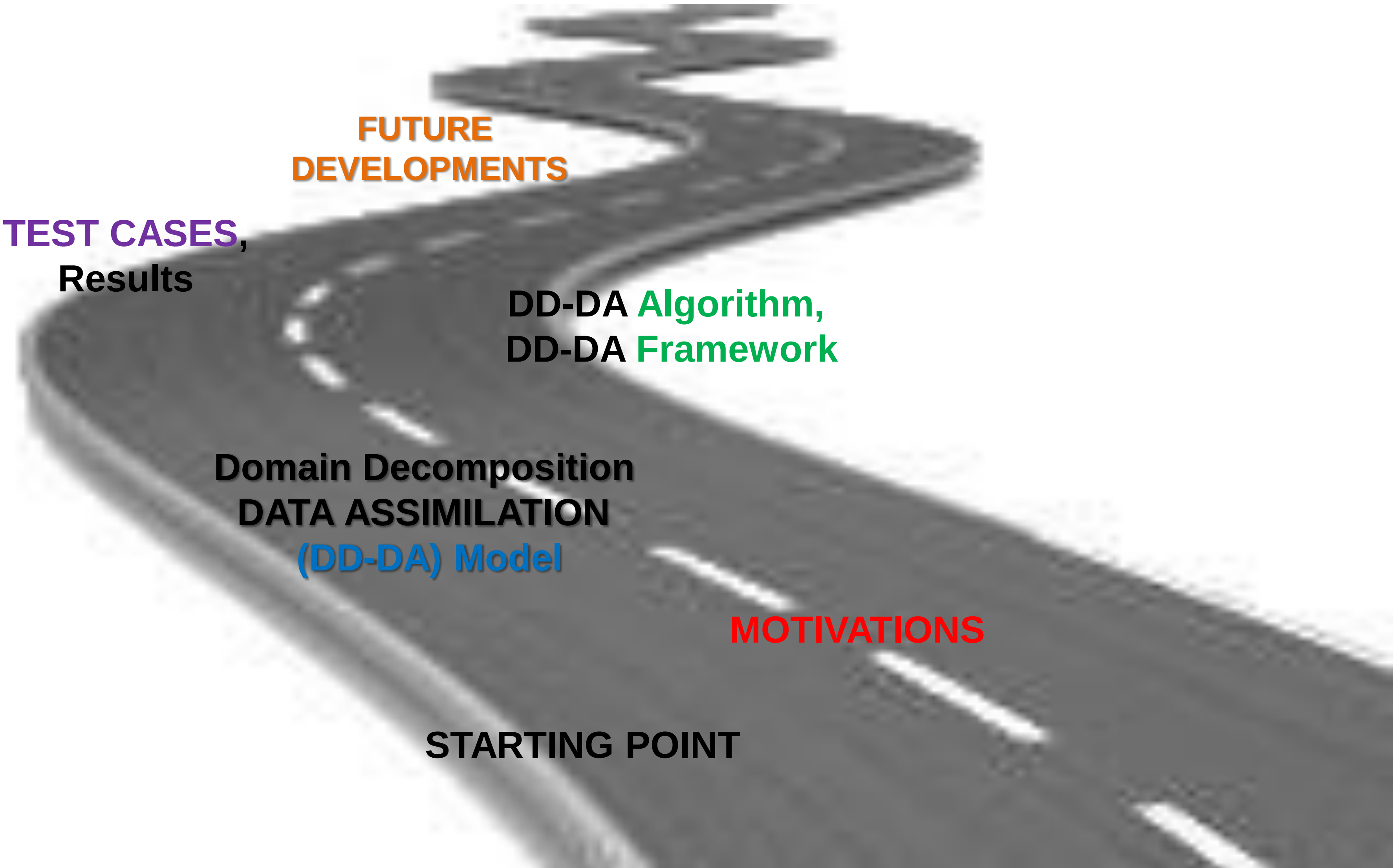
Dr. Rossella Arcucci
Imperial College London
CMCC (until February 2014)

Collaborators

This research results from a collaboration between the **H**igh **E**nd **C**omputing line of **CMCC**-SCO division (a)**,** the **University of Naples “Federico II”** (b), the **SPACI** consortium (c) and the **CNR** (d) in Italy.

Prof. Luisa D’Amore (a) ((b) Head of Numerical Computing group)
Prof. Almerico Murli (a) (c)
Dr Luisa Carracciuolo (d)

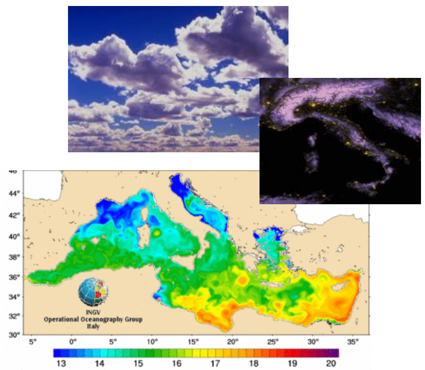
Contents


FUTURE DEVELOPMENTS
TEST CASES, Results
DD-DA Algorithm,
DD-DA Framework
Domain Decomposition
DATA ASSIMILATION
(DD-DA) Model
MOTIVATIONS
STARTING POINT

STARTING POINT: Data Assimilation

data
(initial condition)

u_0



observed data

y_0

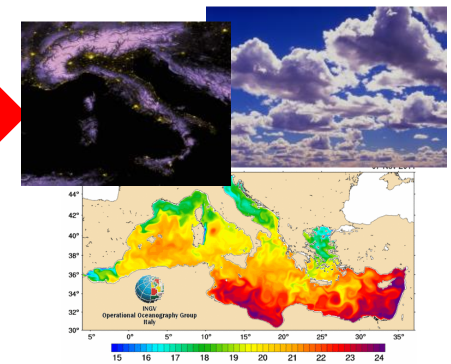
D.A. $\underline{u_0}$

M(P):

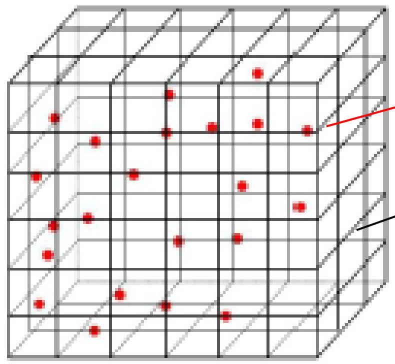
$$\begin{cases} L(u(t))=f \\ u(t_0)=\underline{u_0} \end{cases}$$

$u(t),\ t>t_0$

forecast



STARTING POINT: An ill posed inverse problem



$\mathbf{y}$ = p-dimensional observations vector

$\mathbf{x}_M$ = n-dimensional vector of forecasted value (n > p)

$\mathbf{x}$ = n-dimensional unknown vector (n > p)

y=*H*(x), *H* non linear interpolating function

linearizing...

$$H(x) = H(z) + \mathbf{H}(x - z)$$

H = (p × n)-dimensional matrix, with rank(H) = p

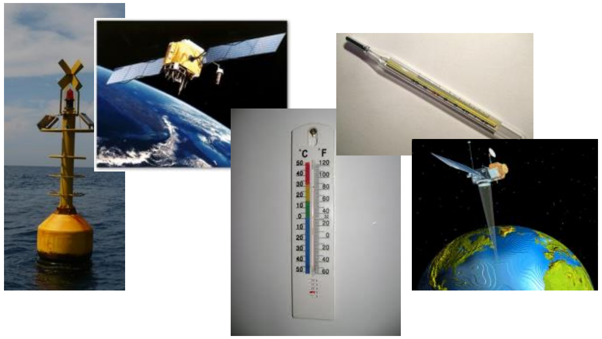
H is the matrix obtained by the first order approximation of the Jacobian of *H*

$y = \mathbf{H}x + R$

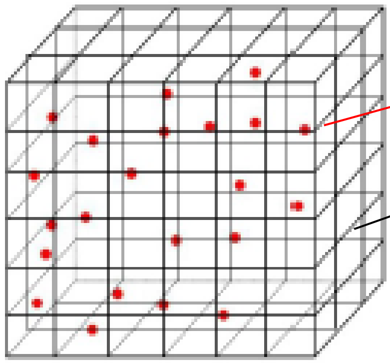
Inverse problem

Ill posed

R is a covariance matrix



STARTING POINT: An ill posed inverse problem



y = p-dimensional observations vector

$\mathbf{x}_M$ = n-dimensional vector of forecasted value (n > p)

x = n-dimensional unknown vector (n > p)

y=*H*(x), *H* non linear interpolating function

linearizing... $$H(x) = H(z) + \mathbf{H}(x - z)$$

H = (p × n)-dimensional matrix, with rank(H) = p

H is the matrix obtained by the first order approximation of the Jacobian of *H*

$y = \mathbf{H}x + R$ — Inverse problem

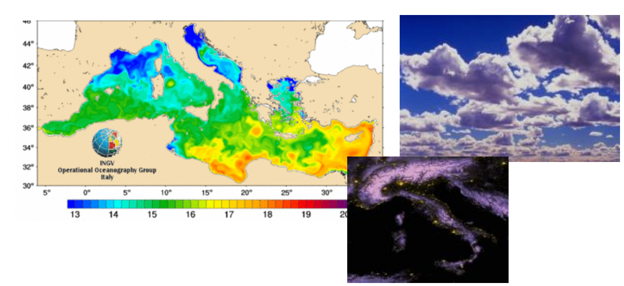
+

$x_M = x + B$ — constraints

R is a covariance matrix

B is a covariance matrix

Regularization… Tikhonov



STARTING POINT: An ill posed inverse problem

we reformulated tha Data Assimilation problem as an ill posed inverse problem and we solved it as a Least Squares problem plus constraints

Tikhonov $$min\{\|Ax - b\|_2^2 + \lambda^2 \|L(x - x_0)\|_2^2\}$$

$\lambda = 1$, $L = \mathbf{I}$

$$\min\left\{\|\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{R}}^2 + \|\boldsymbol{x} - \boldsymbol{x}_M\|_{\boldsymbol{B}}^2\right\}$$

$$\min_{x(t)} J[x(t)]$$

$$J(\mathrm{x}) = (\mathrm{x} - \mathrm{x_M})^T \mathrm{B}(\mathrm{x} - \mathrm{x_M}) + (\mathrm{y} - \mathrm{Hx})^T \mathrm{R}(\mathrm{y} - \mathrm{Hx})$$

- **HPC computation issues of the incremental 3D variational data assimilation scheme in OceanVar software -** L.D'Amore, R.Arcucci, L.Marcellino, A.Murli - Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM) vol. 7, no. 3-4, 2012, pp. 91-105 ISSN 1790–8140.
- **On A Parallel Three-dimensional Variational Data Assimilation Scheme-** - L.D'Amore, R.Arcucci, L.Marcellino, A.Murli - German Symposium on Data Assimilation 2011 (28-30 Sept 2011 - DWD - Offenbach, Germany)
- **A Parallel Three-dimensional Variational Data Assimilation Scheme -** L.D'Amore, R.Arcucci, L.Marcellino, A.Murli - Numerical Analysis and Applied Mathematics, AIP C.P. 1389, 1829-1831 (19-25 Settembre 2011, International Conference of Numerical Analysis and Applied Mathematics 2011, Halkidiki, Grecia 2011) - ISBN: 978-0-7354-0956-9

STARTING POINT: An ill posed inverse problem

NOT considering the temporal variable

$$\min\left\{\|\boldsymbol{H}\boldsymbol{x}-\boldsymbol{y}\|_R^2+\|\boldsymbol{x}-\boldsymbol{x}_M\|_B^2\right\}$$

Variational approach

$$J[x(t)]=\|\boldsymbol{H}\boldsymbol{x}-\boldsymbol{y}\|_R^2+\|\boldsymbol{x}-\boldsymbol{x}_M\|_B^2$$

$$\min_{x(t)} J[x(t)]$$

3DVariational

$$J(\mathrm{x}_i)=(\mathrm{x}_i-\mathrm{x}_{\mathrm{M}i})^T\mathrm{B}(\mathrm{x}_i-\mathrm{x}_{\mathrm{M}i})+(\mathrm{y}_i-\mathrm{H}\mathrm{x}_i)^T\mathrm{R}(\mathrm{y}_i-\mathrm{H}\mathrm{x}_i)$$

Normal equations

$$S=\|\mathrm{H}\mathrm{x}-\mathrm{y}\|_{\mathrm{R}}^2+\|\mathrm{x}-\mathrm{x}_{\mathrm{M}}\|_{\mathrm{B}}^2$$

$$\frac{\partial S}{\partial \mathrm{x}}=0$$

Kalman Filter, Optimal Interpolation

$$\mathrm{x}_i=\mathrm{x}_{\mathrm{M}i}+\mathrm{K}_i(\mathrm{y}_i-\mathrm{H}\mathrm{x}_i)$$

$$\mathrm{K}_i=\mathrm{R}_i\mathrm{H}^T(\mathrm{H}^T\mathrm{R}_i\mathrm{H}+\mathrm{B}_i)^{-1}$$

considering the temporal variable

$$min\{\sum_{j=0}^{n}\|\mathrm{H}\mathrm{x}_\mathrm{j}-\mathrm{y}_\mathrm{j}\|_{\mathrm{R}}^2+\|\mathrm{x}-\mathrm{x}_M\|_{\mathrm{B}}^2\}$$

Variational approach

$$J[x(t)]=\sum_{j=0}^{n}\|\mathrm{H}\mathrm{x}_\mathrm{j}-\mathrm{y}_\mathrm{j}\|_{\mathrm{R}}^2+\|\mathrm{x}-\mathrm{x}_M\|_{\mathrm{B}}^2$$

$$\min_{x(t)} J[x(t)]$$

4DVariational

$$J(\mathrm{x}_i)=(\mathrm{x}_i-\mathrm{x}_{\mathrm{M}i})^T\mathrm{B}(\mathrm{x}_i-\mathrm{x}_{\mathrm{M}i})+\sum_{j=0}^{n}(\mathrm{y}_i^j-\mathrm{H}^j\mathrm{x}_i^j)^T\mathrm{R}^\mathrm{j}(\mathrm{y}_i^j-\mathrm{H}^j\mathrm{x}_i^j)$$

Normal equations

$$S=\sum_{j=0}^{n}\|\mathrm{H}\mathrm{x}_\mathrm{j}-\mathrm{y}_\mathrm{j}\|_{\mathrm{R}}^2+\|\mathrm{x}-\mathrm{x}_M\|_{\mathrm{B}}^2$$

$$\frac{\partial S}{\partial \mathrm{x}}=0$$

Ensemble Kalman Filter

$$\mathrm{x}_i=\mathrm{x}_{\mathrm{M}i}+\sum_{j=0}^{n}\mathrm{K}_i^j(\mathrm{y}_i^j-\mathrm{H}^j\mathrm{x}_i^j)$$

$$\mathrm{K}_i^j=\mathrm{R}_i^j\mathrm{H}^{\mathrm{j}'}(\mathrm{H}^{\mathrm{j}'}\mathrm{R}_i^j\mathrm{H}^\mathrm{j}+\mathrm{B}_i^j)^{-1}$$

MOTIVATIONS: OceanVAR

Centro Euro-Mediterraneo sui Cambiamenti Climatici

OceanVAR model, a 3DVariational model used at CMCC*

(*) developed by Srdjan Dobricic and Nadia Pinardi (2008);

$$J(x) = \frac{1}{2}(x - x_M)^T B^{-1}(x - x_M) + \frac{1}{2}(H[x] - y)^T R^{-1}(H[x] - y)$$

State variable $x_k = [T, S, \eta, u, v]^*$

- Three-dimensional Temperature field
- Three-dimensional Salinity field
- Two-dimensional free surface elevation
- Horizontal velocity components

Linearizing… $\delta x = x - x_M$ $\quad \mathbf{v} = \mathbf{V}^{+}\delta \mathbf{x},$

The matrix *V is*: $B = VV^T$ Cholesky factorization

$$J(x) = \frac{1}{2}v^T v + \frac{1}{2}(HVv - d)^T R^{-1}(HVv - d)$$ preconditioned problem

This function is minimized using the L-BFGS** method (a quasi-Newton method).

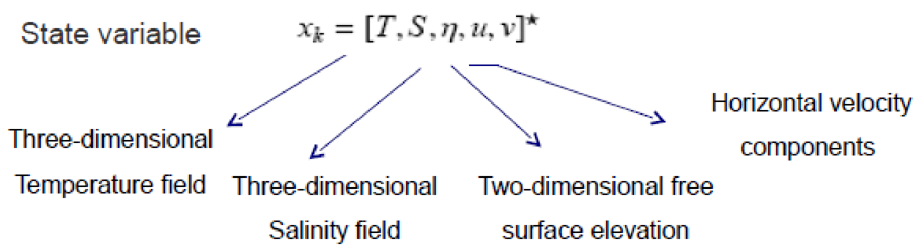
(**) J. Nocedal R.H. Byrd, P. Lu and C. Zhu,
L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization, ACM Transactions on Mathematical Software, Vol. 23, No. 4, December 1997, Pages 550-560.

MOTIVATIONS: OceanVAR

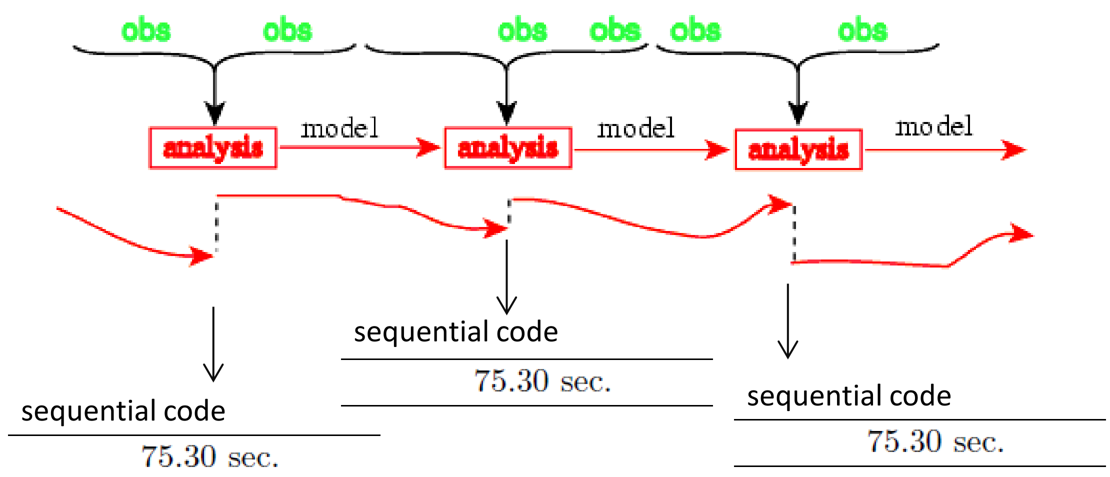
OceanVAR model, a 3DVariational model used at CMCC*

(*) developed by Srdjan Dobricic and Nadia Pinardi (2008);

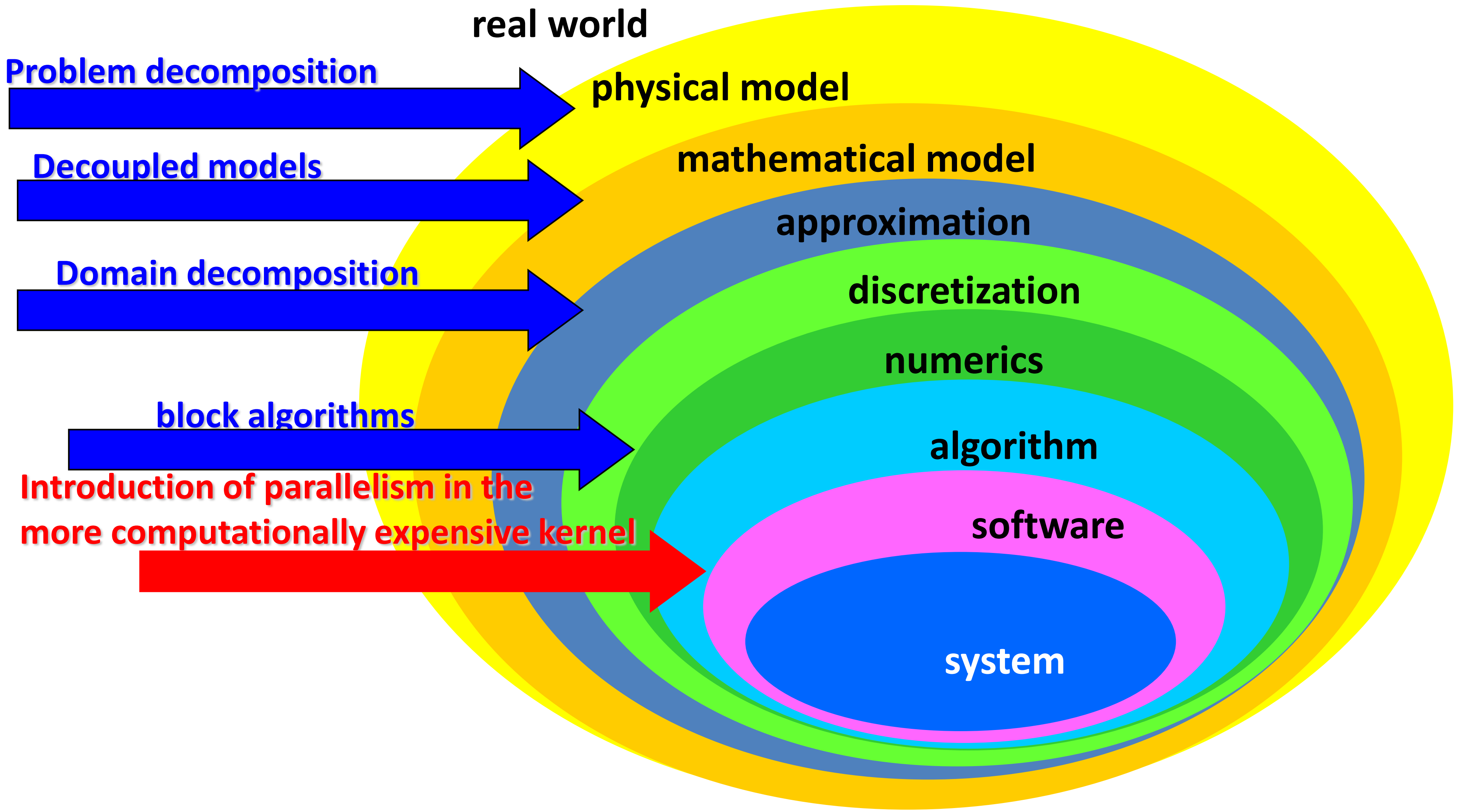
$$J(x) = \frac{1}{2}(x - x_M)^T B^{-1}(x - x_M) + \frac{1}{2}(H[x] - y)^T R^{-1}(H[x] - y)$$



The code needs to be parallelized to reduce the execution time



MOTIVATIONS: Introduction of parallelism



MOTIVATIONS: Optimization and parallelization

Multigrain (Multilevel) Parallelization

LEVEL 1

Fine grained Parallelism

(on Multicores)

+

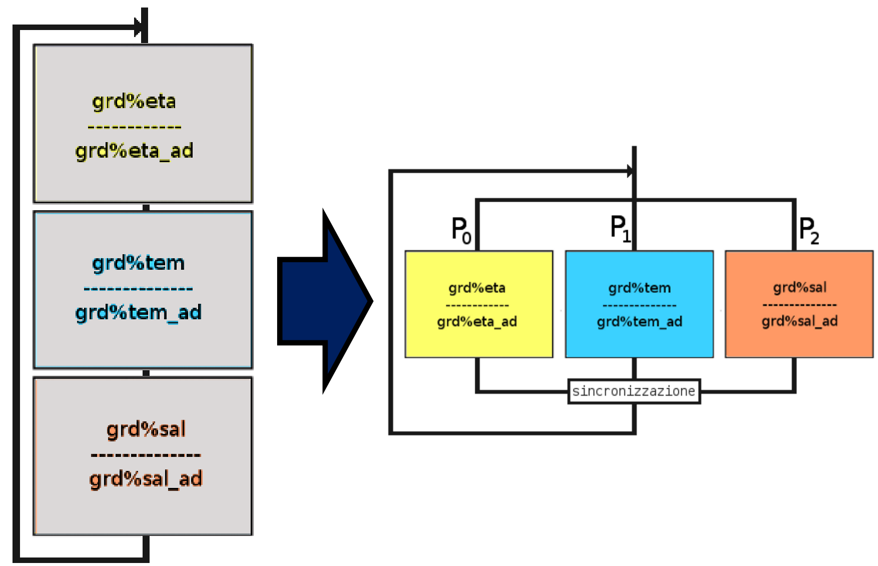
LEVEL 2

Coarse grained Parallelism

(on Multiprocessors)

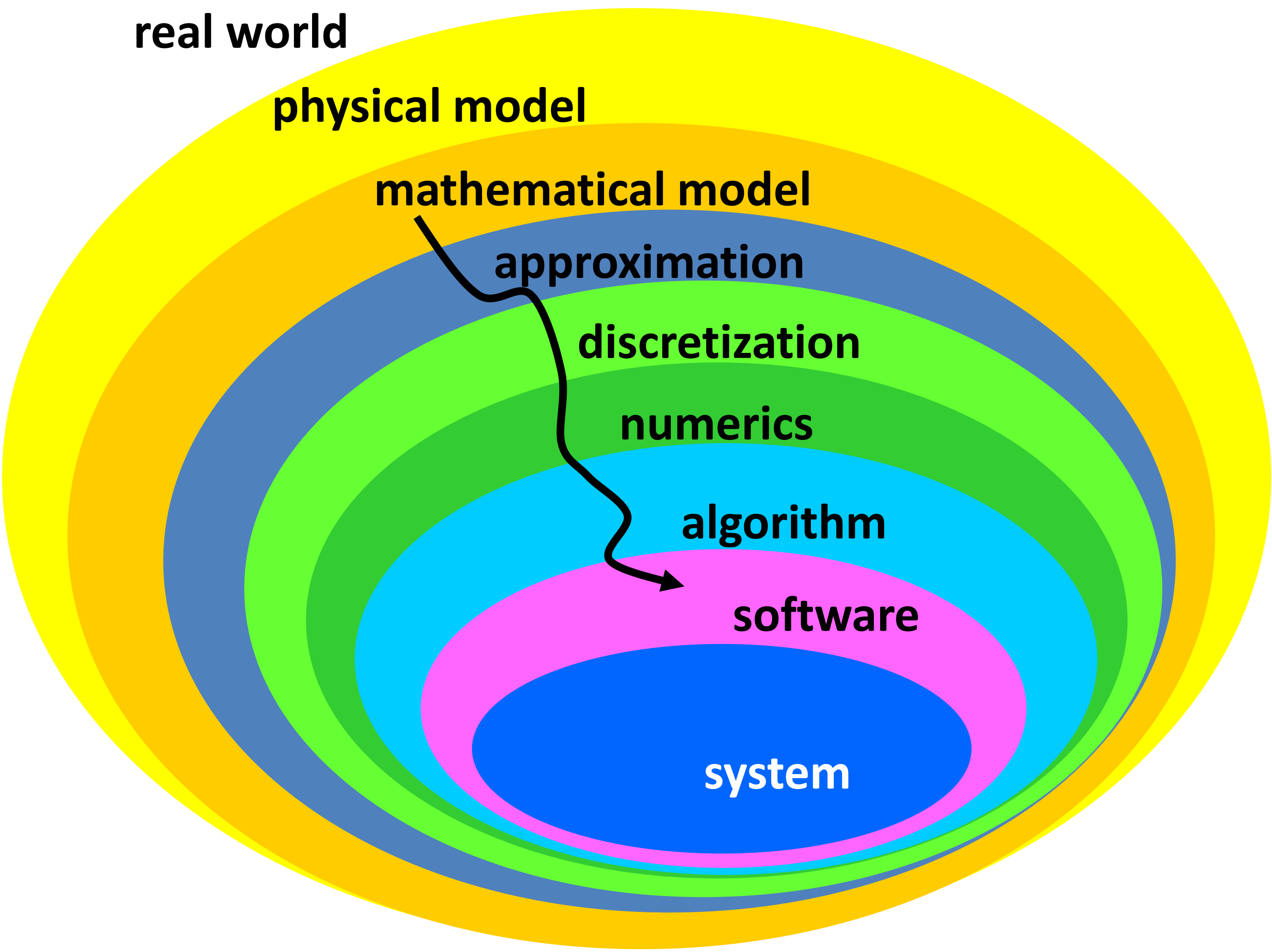
Introduction of the libraries **Blas** and **Lapack**

- Matrix-**Vector Operations**
- **Linear Algebra Operations**



MOTIVATIONS: Introduction of parallelism (2)

«Adapting old programs to fit new machines usually means adapting new machine to behave like old ones.»



MOTIVATION!

OceanVAR model, a 3DVariational model used at CMCC*

(*) devel… by Srdjan Dobricic and N…

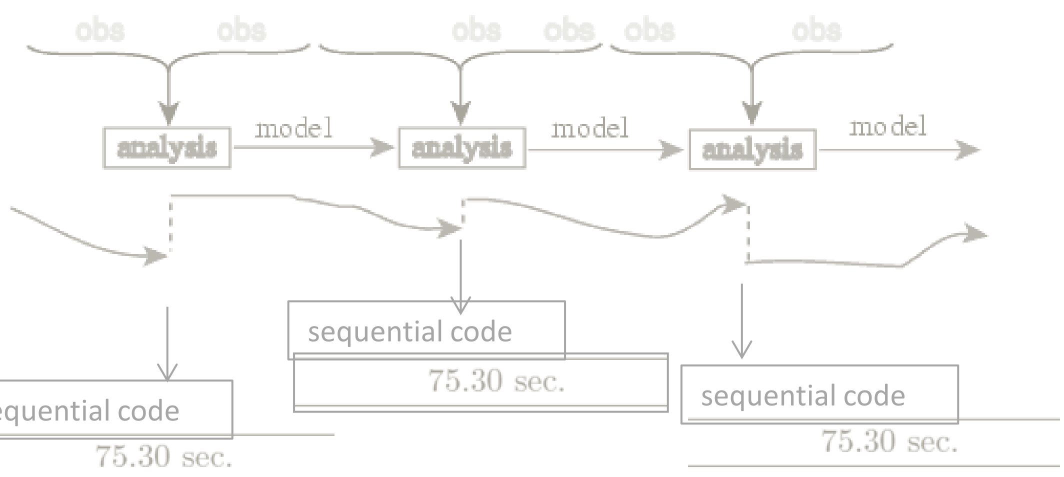
$$J(x) = \frac{1}{2}(x - x_M)^T B^{-1}(x - x_M) + \frac{1}{2}(H[x] - y)^T R^{-1}(H[x] - y)$$

we needed to change the model…

State variable $x_k = [T, S, \eta, u, v]^*$

- Three-dimensional Temperature field
- Three-dimensional Salinity field
- Two-dimensional free surface elevation
- Horizontal velocity components

The code needs to be parallelized to reduce the execution time



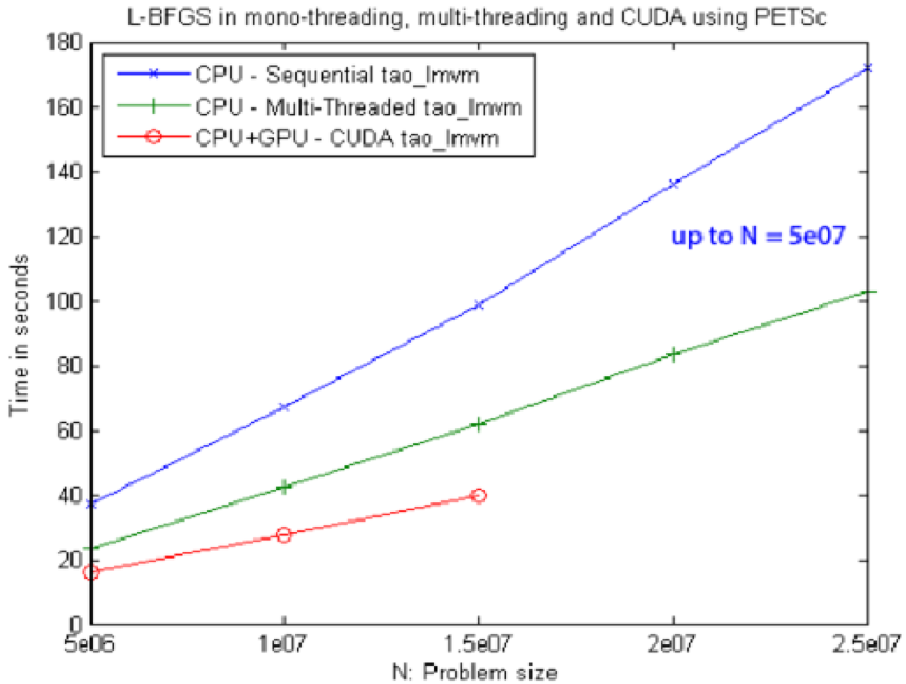
MOTIVATIONS: HPC resources

Cluster of CPU and GPU accelerators

Some tests:

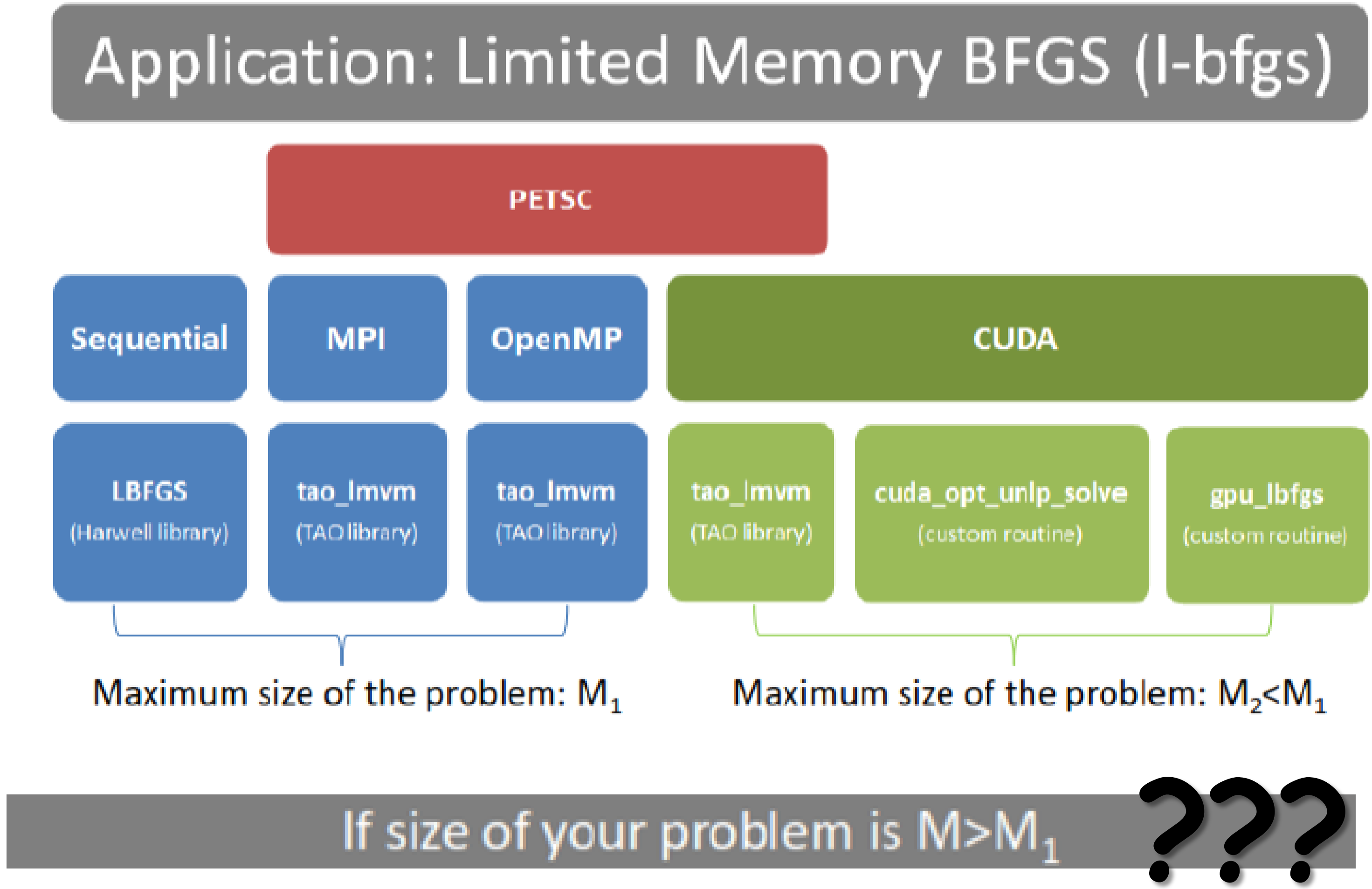
We produced a CUDA version of l-bfgs routine and we also tested the TAO version of the l-bfgs routines on cluster of CPU and GPU

On **GPU** there is a **LIMIT** on the **PROBLEM SIZE**

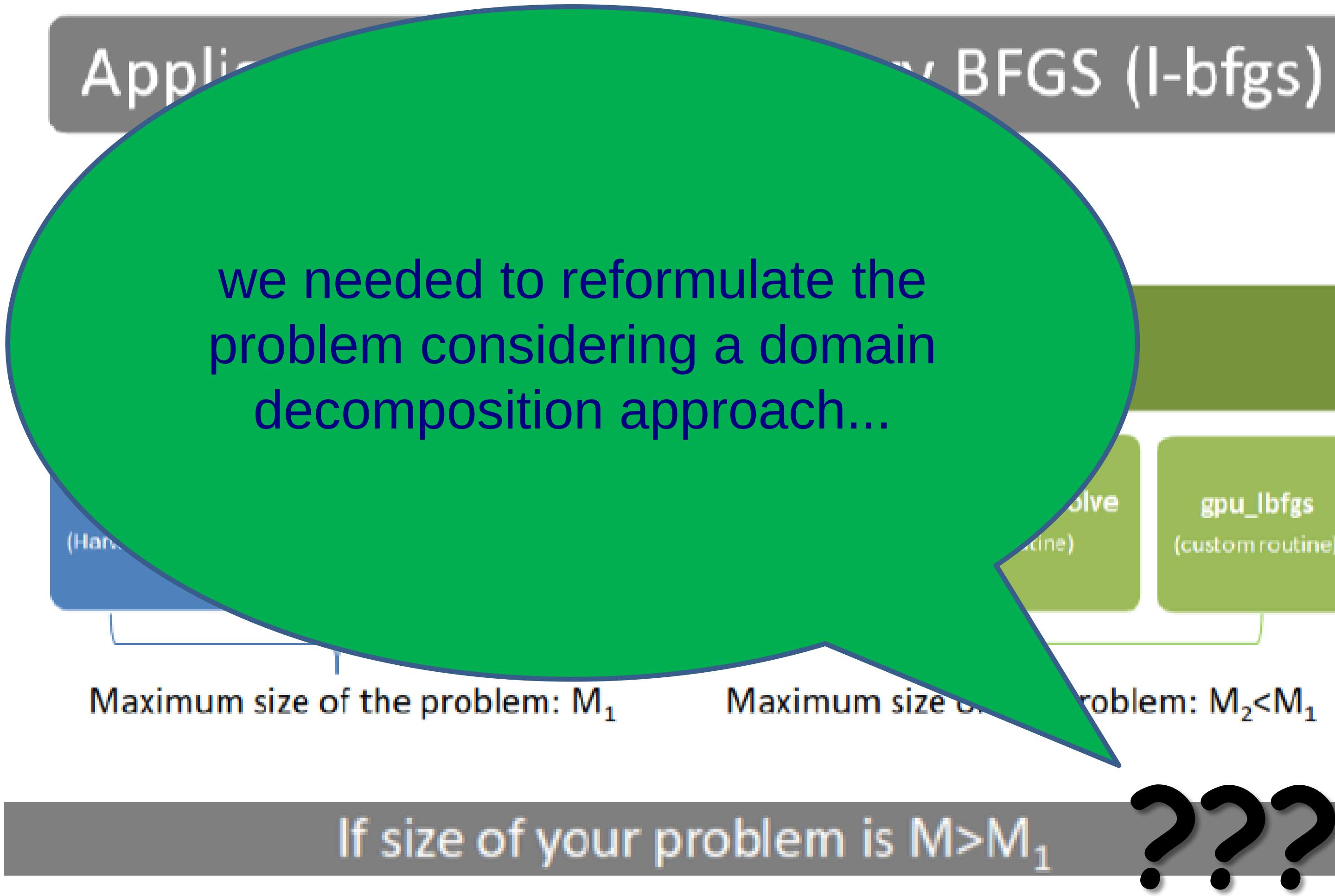


- L.D'Amore, R.Arcucci, G.Scotti,V.Mele, A.Murli– Technical documentation LBFGS for GPUCUDA, Reference Manual and User’s Guide – WN CMCC – Feb 2013 - RP0167
- A Feasibility analysis of a domain decomposition-based approach for solving Variational Data Assimilation problems- - L.D'Amore, R.Arcucci, L.Carracciuolo, A.Murli - Summer School/Creative Workshop: Data Assimilation & Inverse Problems From Weather Forecasting to Neuroscience - July (22-26) 2013 - University of Reading, UK
- Data Assimilation achievements on HPC systems: experiments on OceanVar in the Mediterranean Sea - L.D'Amore, R.Arcucci, L.Carracciuolo A.Murli - Annual Meeting 2013 CMCC June (3-4) 2013 – Marina di Ugento (LE)- Italy

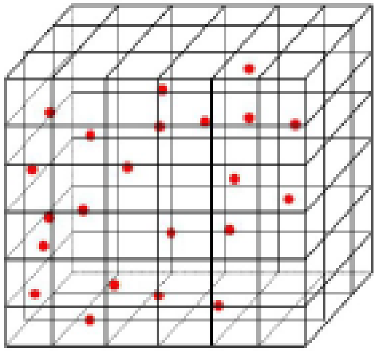
MOTIVATIONS: Some tests.



MOTIVATIONS: Some tests.



DD-DA Model

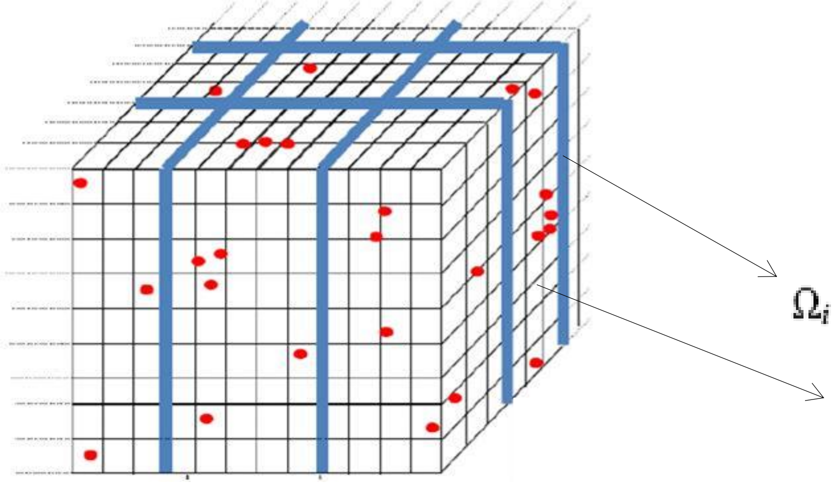


Let

$$J(x_k) = \|\mathcal{H}_k(x_k) - y_k\|_{\mathbf{R}}^2 + \|x_k - x_{M_k}\|_{\mathbf{B}}^2$$

be the function defined on the domain Ω

Let us consider the following overlapping decomposition of the physical domain



$$\Omega = \bigcup_{i=1}^{N} \Omega_i$$

$\Omega_i \cap \Omega_j = \Omega_{ij} \neq 0$ if Ω_i and Ω_j are adjacent.

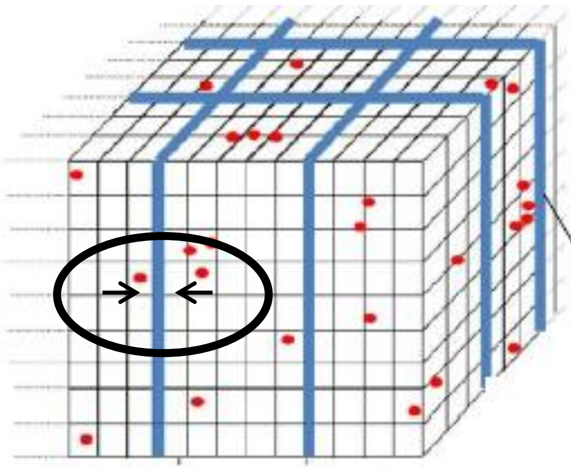
$$\tilde{x}_{DA_i} = \begin{cases} x_{DA_i} & on \quad \Omega_i \\ 0 & on \quad \Omega - \Omega_i \end{cases}$$

be the vector extension of x_{DA_i} on Ω, we define the *DD-analysis* as:

$$x_{DA}^{DD} = \sum_{i=1}^{n} \tilde{x}_{DA_i}$$

DD-DA Model

We reformulated the problem considering a domain decomposition approach...

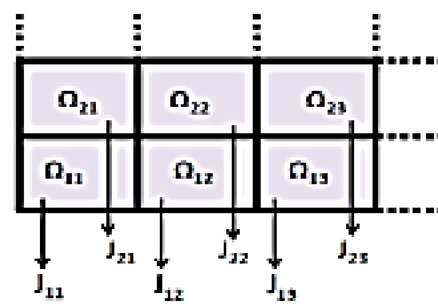


$$\Omega = \bigcup_{i=1}^{N} \Omega_i$$

$$\Omega_i \cap \Omega_j = \Omega_{ij} \neq 0 \text{ if } \Omega_i \text{ and } \Omega_j \text{ are adjacent.}$$

We changed the mathematical model and we imposed to the solution a "continuity" on the overlapping region.

$$J_i(x_k) = J(x_k) + \boxed{\|\mathbf{x}_{k_i}/\Omega_{ij} - \mathbf{x}_{k_j}/\Omega_{ij}\|^2_{B_{ij}}}$$



J_i with $i = (1,1), (1,2), (2,1), \ldots$

DD-DA Model

Domain Decomposition $\Omega = \bigcup_{i=1,p} \Omega_i$

Def: *Extension function*

$$EO_i : J_{\Omega_i} \mapsto J_{\Omega_i}^{EO_i}$$

$$EO_i[J_{\Omega_i}] = \begin{cases} J_{\Omega_i} & x \in \Omega_i \\ 0 & elsewhere \end{cases}$$

Theorem

Global solution

$$\mathbf{u}^{DA} = argmin_{\mathbf{u} \in \Re^{NP}} J(\mathbf{u}) = argmin_u \left\{ \|\mathbf{H}\mathbf{u} - \mathbf{v}\|_{\mathbf{R}}^2 + \lambda \|\mathbf{u} - \mathbf{u}^b\|_{\mathbf{B}}^2 \right\}$$

Local solution

$$\mathbf{u}_i^{DA} = argmin_{\mathbf{u}} J_i(\mathbf{u})$$

$$\tilde{\mathbf{u}}^{DA} \overbrace{=}^{DEF} \sum_{i=1,p} (\mathbf{u}_i^{EO_i})^{DA} \Longrightarrow \tilde{\mathbf{u}}^{DA} = \mathbf{u}^{DA}$$

A Scalable Approach for Variational Data Assimilation, Arcucci R., D'Amore L., Carracciuolo L., Murli A , **Jurnal of Scientific Computing,** DOI 10.1007/s10915-014-9824-2

DD-DA Model: preconditioning

$$J_i(x_k) = \|\mathcal{H}_{k_i}(\mathbf{x}_{k_i}) - \mathbf{y}_{k_i}\|^2_{\mathbf{R}_i} + \|\mathbf{x}_{k_i} - \mathbf{x}_{M_{k_i}}\|^2_{\mathbf{B}_i} + \|\mathbf{x}_{k_i}/\Omega_{ij} - \mathbf{x}_{k_j}/\Omega_{ij}\|^2_{\mathbf{B}_{ij}}$$

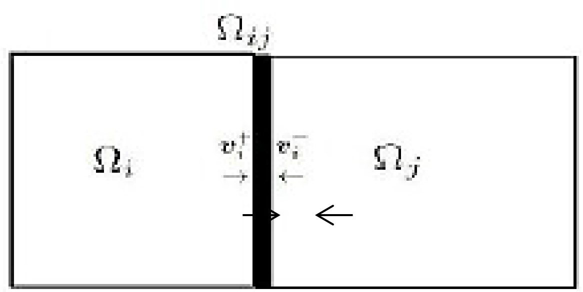
posed $\delta x = x - x_M$ $\quad\quad$ $\mathbf{v} = \mathbf{V}^{+}\delta\mathbf{x},$

To compute the matrix V we apply the TSVD (Truncated SVD) to the matrix B. So, we get

$$B = UdU^T = Ud^{\frac{1}{2}}d^{\frac{1}{2}}U^T = (Ud^{\frac{1}{2}})(Ud^{\frac{1}{2}})^T$$

and by posing $V = Ud^{\frac{1}{2}}$ we get V such that

$$B = VV^T.$$



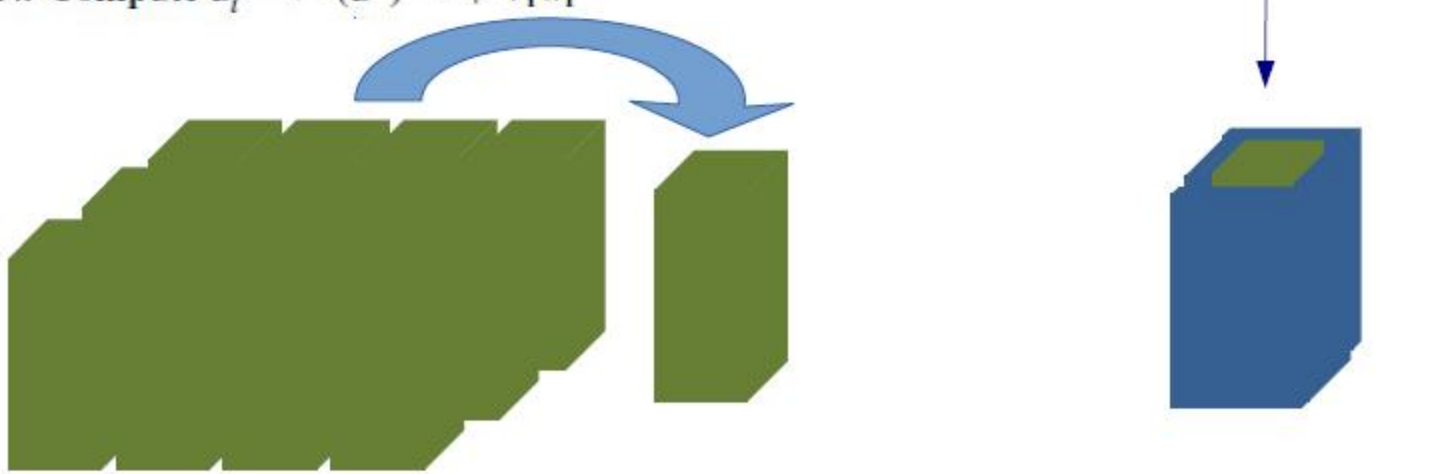
$$J_i(\mathbf{v}_i) = \frac{1}{2}\mathbf{v}_i^T\mathbf{v}_i + \frac{1}{2}(H_iV_i\mathbf{v}_i - d_i)^T R_i^{-1}(H_iV_i\mathbf{v}_i - d_i) + \frac{1}{2}(V_{ij}\mathbf{v}_i^+ - V_{ij}\mathbf{v}_i^-)^T(V_{ij}\mathbf{v}_i^+ - V_{ij}\mathbf{v}_i^-)$$

DD-DA Algorithm

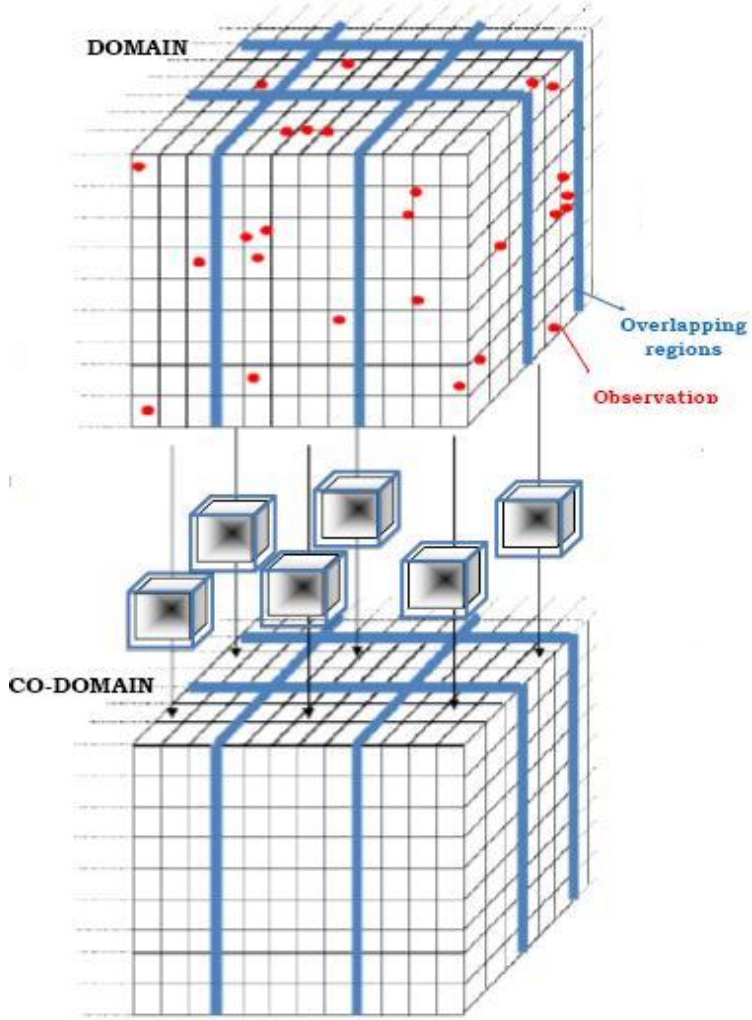
Algorithm 1 $\mathcal{A}_{M3D}(r_i)$: *Modified 3D-Var algorithm on* $\Omega_i, i = 1, \ldots, p$

Input: $\mathbf{v}_i$ and $(\mathbf{u^b})^{RO_i}$
Define $\mathbf{H_i}$
Compute $\mathbf{d_i} \leftarrow \mathbf{v}_i - \mathbf{H_i}(\mathbf{u^b})^{RO_i}$
Define $\mathbf{R_i}$ and $\mathbf{B_i}$
Compute the matrix $\mathbf{V_i}$ from $\mathbf{B_i}$
Define the initial value of $\mathbf{u}^{DA_i}$
Compute $\mathbf{w}_i \leftarrow \mathbf{V_i}^T \mathbf{u}^{DA_i}$
repeat
  Send and Receive the boundary conditions from the adjacent domains
  Compute $\mathbf{J_i} \leftarrow \mathbf{J_i}(\mathbf{w_i})$
  Compute $\mathbf{grad J_i} \leftarrow \nabla \mathbf{J_i}(\mathbf{w_i})$
  Compute new values for $\mathbf{w}_i$ by the L-BFGS steps
until (Convergence on $\mathbf{w_i}$ is obtained)
Compute $\mathbf{u}_i^{DA} \leftarrow (\mathbf{u^b})^{RO_i} + \mathbf{V_i w_i}$

Few modifications



DD-DA Framework



we developed a framework to implement the Data Assimilation Method based on domain decomposition

we tested the framework with a benckmark based on shallow water equations

- **A Scalable Approach for Variational Data Assimilation,** Arcucci R., D'Amore L., Carracciuolo L., Murli A , Jurnal of Scientific Computing
- ***DD-OceanVar: a Domain Decomposition fully parallel Data Assimilation software for the Mediterranean Forecasting System*****,** D'Amore,L., Arcucci,R., Carracciuolo L., Murli A., ICCS 2013, Procedia computer Science, 2013
- ***A Domain Decomposition-Based Parallel Software for Data Assimilation in the Mediterranean*****,** D'Amore L., Arcucci, R., Carracciuolo L., Murli, 2013 SIAM Conference on Mathematical and Computational Issues in the Geosciences June 17-20, Padua, Italy

DD-DA Test Case

Centro Euro-Mediterraneo sui Cambiamenti Climatici

In order to validate the proposed DD approach we tested the model with a benckmark.
We analysed the results related to the **quality of the numerical** results
and in terms of **reduction in computation time**

accuracy

n	$\|h_{true} - h_{M_k}\|_\infty$	$\|u_{true} - u_{M_k}\|_\infty$	$\|v_{true} - v_{M_k}\|_\infty$
64	4.999763e-03	4.999357e-03	4.999357e-03
128	4.999910e-03	4.999468e-03	4.999505e-03

$$x_{true_k} = [h_{true}, u_{true}, v_{true}]^\star$$

$$x_{M_k} = [h_{M_k}, u_{M_k}, v_{M_k}]^\star$$

	n	$nproc = p \times q$	$\|h_{true} - h_{Comp}\|_\infty$	$\|u_{true} - u_{Comp}\|_\infty$	$\|v_{true} - v_{Comp}\|_\infty$
	64	$nproc = 1, p = 1, q = 1$	5.778788e-03	5.778319e-03	5.778319e-03
O(10^6)		$nproc = 2, p = 2, q = 1$	5.455278e-03	5.454835e-03	5.454488e-03
		$nproc = 4, p = 2, q = 2$	5.239452e-03	5.243863e-03	5.244196e-03
		$nproc = 8, p = 4, q = 2$	5.120223e-03	5.127610e-03	5.127936e-03
	128	$nproc = 1, p = 1, q = 1$	–	–	–
O(10^7)		$nproc = 2, p = 2, q = 1$	6.442963e-03	6.442394e-03	6.442441e-03
		$nproc = 4, p = 2, q = 2$	5.850961e-03	5.850444e-03	5.850183e-03
		$nproc = 8, p = 4, q = 2$	5.484419e-03	5.472367e-03	5.478658e-03

DD-DA Test Case

scale-up factor

$$T^{DD}_{nproc} = O\left(p \times q\left(n_z \times \frac{n_x}{p} \times \frac{n_y}{q}\right)^3\right)$$

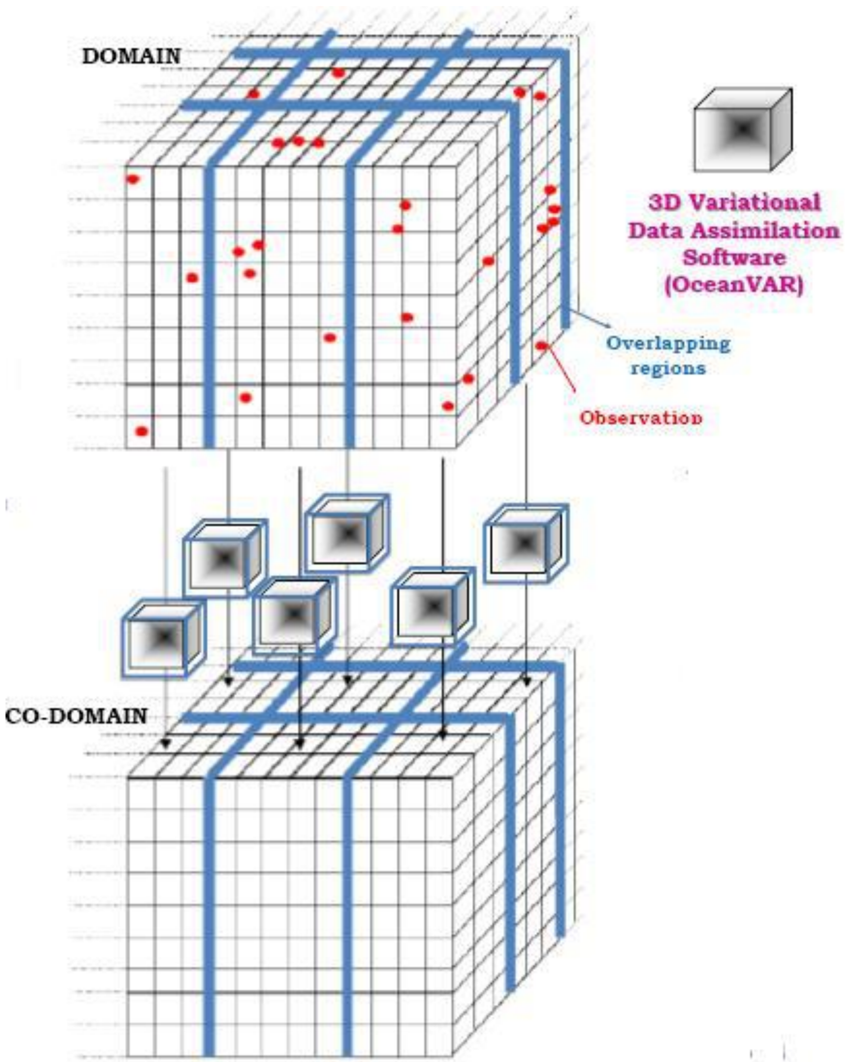
$$T^{TSVD}_{loc}(N_{loc}) = O\left(\left(n_z \times \frac{n_x}{p} \times \frac{n_y}{q}\right)^3\right)$$

$$S^{DD}_{nproc,nproc_1} = \frac{T^{DD}_{nproc_1}}{T^{DD}_{nproc}} = \frac{nproc^2}{nproc_1^2}$$

			real	theoretical
n	$nproc$	$T^{nproc}(NP)$	$S^{DD}_{nproc,nproc_1}$	$S^{DD}_{nproc,nproc_1}$
64	8	2.0545e+02	1.0	1
	16	3.1658e+01	3.25	4
	32	5.0012e+00	10.27	16
	64	1.0979e+00	23.39	64
n	$nproc$	$T^{nproc}(NP)$	$S^{DD}_{nproc,nproc_1}$	$S^{DD}_{nproc,nproc_1}$
128	8	–	–	–
	16	3.9091e+03	1.0	1
	32	4.9976e+02	3.91	4
	64	6.8960e+01	14.17	16

The code scale in agreement with the theoretical scale-up factor

Work in progress.



NEMO

Possible Future developments

Using the DD-DA framework for a global domain as Mediterranean sea

Developing a DD-4DVar model

…………………………………..

Any Questions