Parallel Computing: Technology Trends
I. Foster et al. (Eds.)

© 2020 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/APC200040

A Domain Decomposition Reduced Order Model with Data Assimilation (DD-RODA)

Rossella Arcucci [a], César Quilodrán Casas [a], Dunhui Xiao [b,c,a], Laetitia Mottet [c], Fangxin Fang [c,a], Pin Wu [d], Christopher Pain [c,a] and Yi-Ke Guo [a]

[a] *Data Science Institute, Department of Computing, Imperial College London, UK*
[b] *ZCCE, College of Engineering, Swansea University, UK*
[c] *Department of Earth Science & Engineering, Imperial College London, UK*
[d] *School of Computer Science and Engineering, Shanghai University, Shanghai, China*

Abstract. We present a Domain Decomposition Reduced Order Data Assimilation (DD-RODA) model which combines Non-Intrusive Reduced Order Modelling (NIROM) method with a Data Assimilation (DA) model. The NIROM is defined on a partition of the domain in sub-domains with overlapping regions and the DA is defined on a partition of the domain in sub-domains without overlapping regions. This choice allows to avoid communications among the processes during the Data Assimilation phase. However, during the balance phase, the model exploits the domain decomposition implemented in DD-NIROM which balances the results among the processes exploiting overlapping regions. The model is applied to the pollutant dispersion within an urban environment. Simulations are performed using the open-source, finite-element, fluid dynamics model Fluidity.

Keywords. Numerical simulations, Reduced Order Models, Data Assimilation, Domain Decomposition

1. Introduction

It is estimated that by 2050, around four-million deaths per year will be attributable to outdoor air pollution (twice the current mortality rate) [1]. This mandates the development of techniques that can be used for emergency response, real-time operational prediction and management. Numerical simulations are extensively used as a predictive tool to better understand complex air flows and pollution transport on the scale of individual buildings, city blocks and entire cities [2]. Fast-running Non-Intrusive Reduced Order Model (NIROM) for predicting the turbulent air flows has been proved to be an efficient method to provide numerical forecasting results [3]. However, due to the reduced space on which the model operates, the solution includes uncertainties that are somewhat ambiguous [3]. Additionally, any computational methodology contributes to uncertainty due to 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 main question is how to incorporate data (e.g. from physical

measurements) in models in a suitable way, in order to improve model predictions and quantify prediction uncertainty.

Here, the focus is on the prediction of nonlinear dynamical systems: the classical application example being weather forecasting. In this paper, we combine a Domain Decomposition NIROM (DD-NIROM) method [4] with Data Assimilation (DA). DA is an uncertainty quantification technique used to incorporate observed data into a prediction model in order to improve numerical forecasted results. The DD-RODA (Domain Decomposition Reduced Order Data Assimilation) model we propose in this paper achieves both efficiency and accuracy by including Variational DA (VarDA) into DD-NIROM.

The DD-NIROM can be constructed by a combination of proper orthogonal decomposition (POD) and machine learning methods or interpolation methods. The key idea of the DD-NIROM is that it constructs a set of hypersurfaces representing the reduced system (including linear and non-linear processes). The novelty of the NIROM and DD-NIROM, presented in [3], lies in how they are generated, i.e. how the hypersurfaces are calculated using a machine learning method. The model we introduce in this paper combines the state of the art of domain decomposition reduced order models with an efficient variational DA model defined on an optimal reduced space [5,6,7,8] and on a decomposition of the domain in sub-domains named (DD-DA). Even if the DD-DA method we employ is efficient, it lacks of efficiency in the pre-processing phase which mainly consists in evaluating and computing the background error covariance matrices. Modelling and specification of the covariance matrix of background error constitute important components of any data assimilation system [9]. The main attributes of the background error covariance matrix are: to spread out the information from the observations; to provide statistically consistent increments at the neighboring grid points; and to ensure that observations of one model variable produce dynamically consistent increments in the other model variables. The use of DD-NIROM for the pre-processing phase of the DD-DA process can improve the efficiency of the whole prediction-correction cycle with a consequent improvement of the operational prediction model fidelity.

In summary, in this paper we combine a Domain Decomposition Non-intrusive Reduced Order Modelling method [4] with Domain Decomposition Data Assimilation [8,7] in a Domain Decomposition Reduced Order Data Assimilation (DD-RODA) model in order to achieve both accuracy and efficiency in our simulations. An important advantage of the DD-RODA approach is that once the DD-NIROM model is obtained, there is no need to refer to the full model while performing DD-DA. With this approach, in fact, we improve

- the accuracy of the DD-NIROM model by introducing information from observed data using the variational DD-DA process.
- the efficiency of the DD-DA process in the pre-processing phase: we use the DD-NIROM results to train our background error covariance matrices resulting in a strong reduction of the overall execution time.

We demonstrate the accuracy and the scalability of our approach. A mathematical formulation of the model is provided.

The model is tested on the pollutant dispersion within an urban environment. Simulations are performed using the open-source, finite-element, fluid dynamics software Fluidity (http://fluidityproject.github.io/). The details of the equations solved and their

implementations can be found in [10,11,12]. In this paper, the state variable consists of values of pollution concentration. However, the algorithm and numerical methods proposed in this work can be applied to other physical problem involving other equations and/or state variables.

2. Reduced Order Assimilation model

Let $\mathbf{u}$ be a state variable and let f represent a full physical system:

$$\dot{\mathbf{u}} = f(\mathbf{u}, t) \tag{1}$$

where t denotes the time.
Let Ω be the discrete spatial domain and let $\mathscr{P}(\Omega) = \{\Omega_j\}_{j=1,\dots,s}$ be a partition of Ω in s sub-domains. Let $\mathbf{u}_j$ be the restriction of the state variable $\mathbf{u}$ on the sub-domain Ω_j.
In this work, a Domain Decomposition Non-intrusive Reduced Order Modelling (DD-NIROM) method is used to enhance the computational efficiency. The reduced order model projects the sub-domains of the full physical system with a big dimensional size onto a reduced space sub-domains with a much smaller dimensional size, therefore it is faster to solve.
Let n denotes a time level, the DD-NIROM uses a Proper Orthogonal Decomposition (POD) method and Gaussian Process Regression (GPR) method to approximate the solutions of equation (1).

In DD-NIROM based on the POD method, any variables $\mathbf{u}_j^n$ (for example, the velocity or tracers) at time level n can be expressed by the expansion,

$$\mathbf{u}_j^n = \overline{\mathbf{u}}_j + \sum_{i=1}^{M} \alpha_{ji}^n \phi_{ji}, \tag{2}$$

where α_{ji}^n ($i \in \{1,2,\dots,M\}$) denotes the POD coefficients of the POD basis functions at the time level n. ϕ_{ji} are the POD basis functions. M is the number of POD basis functions ($M << N$) which can represent most (99% for example) of energy within the chosen solution snapshots. $\overline{\mathbf{u}}_j$ represents the mean of the snapshots.

Let n be a fixed time level and let $\mathbf{u}_j^n$ be a state variable expressed by DD-NIROM as described in equation (2). Let $\mathbf{e}_j^n = \mathbf{u}_j - \mathbf{u}_j^n$ be the error introduced by replacing the full physical system in (1) by the NIROM model (2). We introduce a DD-Reduce Order Assimilation process by which the DD-NIROM model in (2) is combined with a DD-Data Assimilation method in order to improve the accuracy of the solution $\mathbf{u}_j^n$ (i.e. reduce $\mathbf{e}_j^n$) introducing information by observation of the state variable $\mathbf{u}_j$.

Let $\mathbf{v}_j^n$ be an observation of the state variable at time n, the aim of DD-Reduced Order Data Assimilation (DD-RODA) problem is to find an optimal trade-off between the prediction made based on the DD-NIROM system state $\mathbf{u}_j^n$ (background) defined in (2) and the available observation $\mathbf{v}_j^n$. For a fixed time step n, given $\mathbf{u}_j^n$, $\mathbf{v}_j^n$ and a mapping

$$H_j : \mathbf{u}_j^n \mapsto \mathbf{v}_j^n, \tag{3}$$

the DD-RODA process consists in finding $\mathbf{u}_j^{DD-RODA}$ as inverse solution of

$$\mathbf{v}_j^n = H_j\left(\mathbf{u}^{DD-RODA}\right), \tag{4}$$

subject to the constraint that $\mathbf{u}_j^{DD-RODA} = \mathbf{u}_j^n$, i.e.:

$$\mathbf{u}_j^{DD-RODA} = \overline{\mathbf{u}}_j + \sum_{i=1}^{M} \alpha_{ji}^n \phi_{ji}. \tag{5}$$

where ϕ_{ji} denotes the POD basis functions. Since H_j is typically rank deficient, equation (4) is an ill-posed inverse problem [13,14]. The Tikhonov formulation [15,16] leads to an unconstrained least squares problem, where the term in (5) provided by DD-NIROM ensures the existence of a unique solution in (4). The DD-RODA process can then be described as follows:

$$\mathbf{u}_j^{DD-RODA} = argmin_{\mathbf{u}_j}\left\{\|\mathbf{u}_j - \mathbf{u}_j^n\|^2_{B_j^{-1}} + \|\mathbf{v}_j^n - H\left(\mathbf{u}_j\right)\|^2_{R_j{}^{-1}}\right\} \tag{6}$$

where R_j and B_j are the observation and model error covariance matrices respectively, defined on each subdomain Ω_j, $j = 1,\ldots,s$:

$$R_j := \sigma_0^2 I_j, \tag{7}$$

with $0 \leq \sigma_0^2 \leq 1$ and I_j the identity matrix,

$$B_j = V_j V_j^T \tag{8}$$

where V_j is the deviance matrix [6]. If equation (6) is linearised around the background state [17], we have:

$$\mathbf{u}_j = \mathbf{u}_j^n + \delta\mathbf{u}_j \tag{9}$$

where $\delta\mathbf{u}_j = \mathbf{u}_j - \mathbf{u}_j^n$ denotes the increments. The DD-RODA problem is formulated by the following form:

$$\delta\mathbf{u}_j^{DD-RODA} = argmin_{\delta\mathbf{u}_j} J_j(\delta\mathbf{u}_j) \tag{10}$$

where

$$J_j(\delta\mathbf{u}_j) = \frac{1}{2}\delta\mathbf{u}_j^T\mathbf{B}_j^{-1}\delta\mathbf{u}_j + \frac{1}{2}(\mathbf{H}_j\delta\mathbf{u}_j - \mathbf{d}_j^{DD-NIROM})^T\mathbf{R}_j^{-1}(\mathbf{H}_j\delta\mathbf{u}_j - \mathbf{d}_j^{DD-NIROM}) \tag{11}$$

and

$$\mathbf{d}_j^{DD-NIROM} = [\mathbf{v}_j^n - H_j\left(\mathbf{u}_j^n\right)] \tag{12}$$

is the misfit between the observation and the solution computed by DD-NIROM (see Algorithm 1) and

$$H_j(\mathbf{u}_j) \simeq H_j\left(\mathbf{u}_j^n\right) + \mathbf{H}_j \delta \mathbf{u}_j \tag{13}$$

denotes the linearised observational and model operators evaluated at $\mathbf{u}_j = \mathbf{u}_j^n$ where $\mathbf{H}_j$ is the Hessian of H_j. In equation (10), the minimisation problem is defined on the field of increments [18]. In order to avoid the inversion of $\mathbf{B}_j$, as $\mathbf{B}_j = \mathbf{V}_j \mathbf{V}_j^T$, the minimisation can be computed with respect to a new variable [17] $\mathbf{w}_j = \mathbf{V}_j^+ \delta \mathbf{u}_j$ and $\mathbf{V}_j^+$ denotes the generalised inverse of $\mathbf{V}_j$:

$$\mathbf{w}_j^{DD-RODA} = argmin_{\mathbf{w}_j} J_j(\mathbf{w}_j) \tag{14}$$

where

$$J_j(\mathbf{w}_j) = \frac{1}{2}\mathbf{w}_j^T \mathbf{w}_j + \frac{1}{2}(\mathbf{H}\mathbf{V}_j \mathbf{w}_j - \mathbf{d}_j^{DD-NIROM})^T \mathbf{R}_j^{-1} (\mathbf{H}\mathbf{V}_j \mathbf{w}_j - \mathbf{d}_j^{DD-NIROM}) \tag{15}$$

In the next section, DD-RODA is applied to improve the pollutant dispersion prediction within an urban environment. Simulations are performed using the open-source, finite-element, fluid dynamics model Fluidity.

3. Numerical example

The capability of DD-RODA has been estimated using an urban environment located in London South Bank University (LSBU) area (London, UK) shown in Figure 1. The computational domain has a size of $[0,2041] \times [0,2288] \times [0,250]$ (metres). This work uses the 3D non-hydrostatic Navier-Stokes equations as the full physical system,

$$\nabla \cdot \mathbf{u} = 0, \tag{16}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nabla \cdot \boldsymbol{\tau}, \tag{17}$$

where $\mathbf{u} \equiv (u,v,w)^T$ is the velocity, $p = \tilde{p}/\rho_0$ is the normalised pressure ($\tilde{p}$ being the pressure and ρ_0 the constant reference density) and $\boldsymbol{\tau}$ denotes the stress tensor.

Simulations were carried out for the study area using Fluidity, an open-source, finite-element, fluid dynamics model [12]. The dispersion of pollutant is described by the classic advection-diffusion equation with the pollutant concentration treated as a passive scalar. A source term was added to the advection-diffusion equation to mimic a constant release of pollutant generated by traffic in a busy intersection for example. The location of the point source is depicted by the red sphere in Figure 1(a). The time step was adaptive based on the Courant (CFL) number defined by the user, and the Crank-Nicholson scheme was used for the time discretization [11,10]. The mesh is shown in Figure 1(c). The outlet boundary condition was defined by a zero-pressure (no-stress) condition; perfect slip boundary conditions were applied at the top and on the sides of the domain and no-slip boundary conditions were applied on all building facades and the bottom surface of the domain. A synthetic incoming-eddy method was used at the inlet [20] to mimic the behaviour of the boundary layer. The mean velocity profile was prescribed as in equation (18):

Algorithm 1 DD-RODA algorithm on each sub-domain Ω_j, $j = 1, ..., s$

▷ The following are known: $\{f_i\}_{i=1}^m$, $\{\boldsymbol{\phi}_i\}_{i=1}^m$, H_j and R_j

$\boldsymbol{\alpha}_j^0 = \boldsymbol{\alpha}_j(t_0)$ ▷ Initialisation of POD coefficients

for $n = 1$ *to* $\mathcal{N}_t$ **do**
  $t = t_0 + n\Delta t$ ▷ Current time

  ▷ Step (a): calculate the POD coefficients, $\boldsymbol{\alpha}_j^n$, at the current time step:

  for $i = 1$ *to* m **do**
    $\alpha_{ji}^n = f_{ji}(\alpha_{j1}^{n-1}, \alpha_{j2}^{n-1}, \ldots, \alpha_{jm}^{n-1})$
  endfor

  ▷ Step (b): obtain the solution $\mathbf{u}_j^n$ in the full space at the current time, t, by projecting α_{ji}^n onto the full space using equation (2):

  $\mathbf{u}_j^n = \mathbf{0}$
  for $i = 1$ *to* m **do**
    $\mathbf{u}_j^n = \mathbf{u}_j^n + \alpha_{ji}^n \boldsymbol{\phi}_{ji}$
  endfor

  ▷ Step (c): compute the optimal background error covariance matrix:

  $\mathbf{V}_j^n = \mathbf{u}_j^n - \bar{\mathbf{u}}_j$
  $\mathbf{V}_j = \{\mathbf{V}_j, \mathbf{V}_j^n\}$

endfor

$\mathbf{V}_{j\tau} = TSVD(\mathbf{V}, \tau)$. ▷ Truncated SVD regularised matrix [6,19]

▷ Step (d): solve the reduced order assimilation process (6):

$\mathbf{d}_j^{DD-NIROM} \leftarrow \mathbf{v}_j^n - H_j \mathbf{u}_j^n$ ▷ Compute the misfit

$\mathbf{G}_j \leftarrow \mathbf{H}_j \mathbf{V}_{j\tau} \mathbf{w}_j - \mathbf{d}_j^{DD-NIROM}$

$\mathbf{w}_j^{DD-RODA} = argmin_{\mathbf{w}_j} \left\{ \frac{1}{2} \mathbf{w}_j^T \mathbf{w}_j + \frac{1}{2} \mathbf{G}_j^T \mathbf{R}_j^{-1} \mathbf{G}_j \right\}$ ▷ Compute the minimum

$\delta \mathbf{u}_j^{DD-RODA} \leftarrow \mathbf{V}_{j\tau} \mathbf{w}_j$ ▷ From the reduced to physical space

$\mathbf{u}_j^{DD-RODA} \leftarrow \mathbf{u}_j^n + \delta \mathbf{u}_j^{DD-RODA}$

$$(u, v, w) = \left(0.97561 \ln\left(\frac{z}{0.01}\right), 0, 0\right) \tag{18}$$

where z denotes the height (in m). The inlet length-scale $\boldsymbol{L}$ and Reynolds stresses $\boldsymbol{Re}$ are prescribed constant and equal to 100 m and 0.8 respectively, for the diagonal components, and zero elsewhere. Zero velocity is prescribed on the bottom and on the wall boundaries. Zero stress conditions is set to be $p = 0$ at the outlet boundary and a perfect slip condition is specified on the vertical lateral boundaries. Experiments have been implemented and tested on 3 high performance nodes equipped with bi-Xeon E5-2650 v3 CPU and 250GB of RAM with Python 3.5

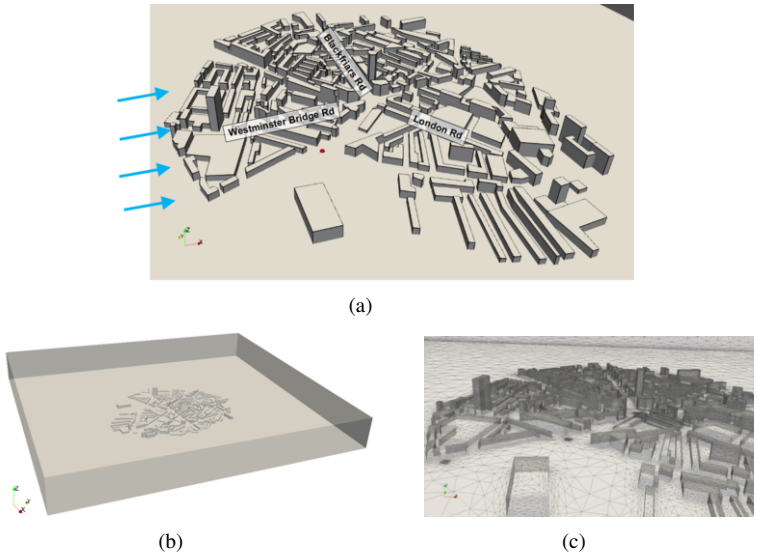


Figure 1. (a) London South Bank University (LSBU) test case area. The red sphere denotes the location of the pollution source and the blue arrows denote the wind direction. (b) 3D computational domain and (c) surface mesh of the test site.

The accuracy of the DD-RODA results is evaluated by the mean squared error on each sub-domain:

$$MSE(u_j) = \frac{\|u_j - u_j^C\|_{L^2}}{\|u_j^C\|_{L^2}} \tag{19}$$

computed with respect to a control variable u_j^C provided by observed data, for $j = 1, \ldots, s$ and s denotes the number of sub-domains. Figure 2 shows the values of $MSE(u^{DD-NIROM})$ and $MSE(u^{DD-RODA})$ for a decomposition made of $s = 16$ sub-domains running on $p = 16$ processors. We can observe that the error decreases for each sub-domain. We observe a bigger gain in terms of accuracy reduction in sub-domains where the pollution concentration is more diffused. Fox example, the sub-domain number 11 presents a bigger gain as shown in Figure 2, this sub-domain is the central sub-domain in Figure 4 (orange colour).

We evaluated the execution time needed to compute the solution of the DD-RODA model using Algorithm 1. Let T_s denote the execution time of Algorithm 1 for a domain decomposition made of s sub-domains. We assume that $p = s$, where p denotes the number of processors and we pose:

$$T_s = max\{T_{s_i}\}_{i=1,\ldots,s} \tag{20}$$

where T_{s_i} denotes the execution time for each processor on each sub-domain. The total execution time is shown in Figure 3(a). There is a clear decreasing trend in the total execution time with the increase of number of processors. Figure 3(b) shows the values of execution time of DD-NIROM and Fluidity on $p = 4, 16, 32$ processors for a decomposition of $s = 4, 16, 32$ sub-domains. The gain in terms of execution time provided by using

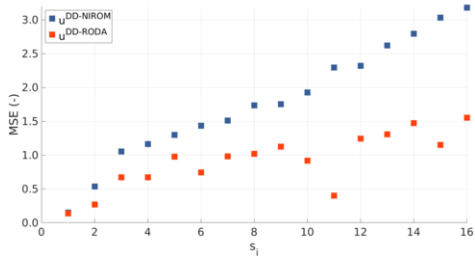
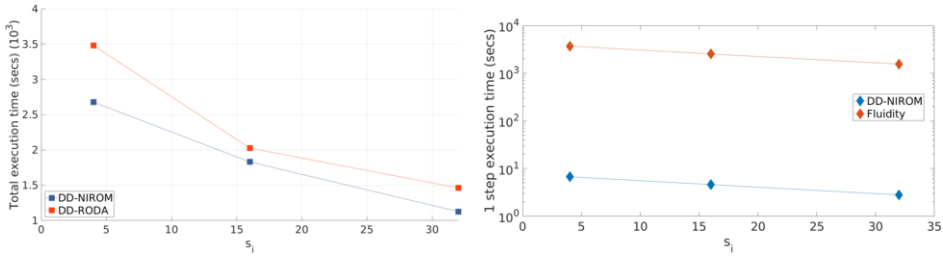


Figure 2. Values of $MSE(u^{DD-NIROM})$ and $MSE(u^{DD-RODA})$ for a decomposition made of $s = 16$ sub-domains running on $p = 16$ processors

DD-NIROM instead of Fluidity strongly impact on the efficiency of the pre-processing to the Data Assimilation phase (Step (c) of Algorithm 1) for computing the covariance matrices V_j for each time step n.



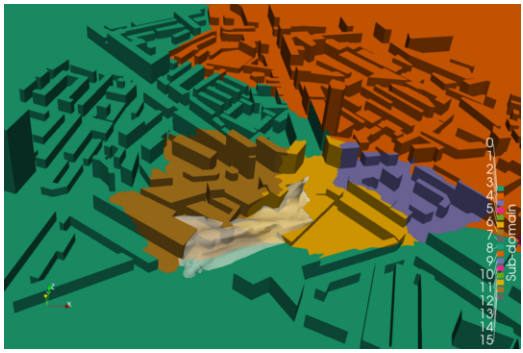
(a) (b)

Figure 3. (a) Values of execution times of DD-NIROM and DD-RODA for a number of sub-domains $s = 4, 16, 32$ running on $p = 4, 16, 32$ processors (plot in linear scale) (b) Values of execution time for running 1 time step of DD-NIROM and Fluidity on $p = 4, 16, 32$ processors (plot in log scale).

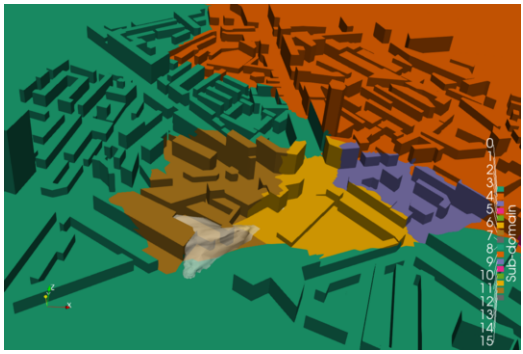
Figure 4 shows the impact of DD-RODA on the iso-surface of the pollutant concentration for $5.10^{-1} kg/m^3$ computed in parallel with $p = 16$ processors and generated by a point source. Figure 4(a) shows the results predicted by DD-NIROM, i.e. $u^{DD-NIROM}$, while Figure 4(b) shows the observed data, i.e. v. Values v are assimilated in parallel by DD-RODA to correct the forecasting data $u^{DD-NIROM}$. The assimilated data after the DD-RODA process, i.e. $u^{DD-RODA}$, are then obtained (Figure 4(c)).

4. Conclusions

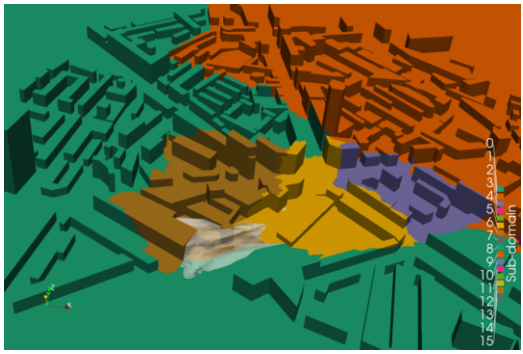
In this paper, we have presented a Domain Decomposition Reduced Order Data Assimilation (DD-RODA) model which is a fusion of the Non-Intrusive Reduced Order Modelling method with a 3D Data Assimilation both defined on a decomposition of the domain in sub-domains. We proved that our approach improves both accuracy of the DD-NIROM model and efficiency of the DA process. The accuracy of the DD-NIROM model is improved by introducing information from observed data exploiting the 3D variational DA process. The efficiency of the DA process is mainly improved in the pre-processing phase. In fact, we used the DD-NIROM results to train our background error covariance



(a) $u^{DD-NIROM}$: predicted pollutant iso-surface by DD-NIROM.



(b) v: observed pollutant iso-surface.



(c) $u^{DD-RODA}$: assimilated pollutant iso-surface by DD-RODA.

Figure 4. Iso-surface, in white, of the pollutant concentration for $5.10^{-1} kg/m^3$ computed in parallel with $p = 16$ and generated by a point source.

matrices and we have shown that this implies a strong reduction of the overall execution time. The efficiency and the accuracy of our model were discussed and tested using a 3D case of air flows and pollution transport in an urban environment. The algorithms and the method proposed are, however, enough generic and can be used easily for other physical problems.

Acknowledgments

This work is supported by the EPSRC Grand Challenge grant "Managing Air for Green Inner Cities" (MAGIC) EP/N010221/1 and by the EPSRC Centre for Mathematics of Precision Healthcare EP/N0145291/1.

References

[1] Jos Lelieveld, John S Evans, Mohammed Fnais, Despina Giannadaki, and Andrea Pozzer. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525(7569):367, 2015.

[2] Bert Blocken. Computational fluid dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. *Building and Environment*, 91:219–245, 2015.

[3] D Xiao, CE Heaney, L Mottet, F Fang, W Lin, IM Navon, Y Guo, OK Matar, AG Robins, and CC Pain. A reduced order model for turbulent flows in the urban environment using machine learning. *Building and Environment*, 148:323–337, 2019.

[4] D Xiao, CE Heaney, F Fang, L Mottet, R Hu, DA Bistrian, E Aristodemou, IM Navon, and CC Pain. A domain decomposition non-intrusive reduced order model for turbulent flows. *Computers & Fluids*, 2019.

[5] R. Arcucci, C. Pain, and Y. Guo. Effective variational data assimilation in air-pollution prediction. *Big Data Mining and Analytics*, 1(4):297 – 307, 2018.

[6] R. Arcucci, L. Mottet, C. Pain, and Y. Guo. Optimal reduced space for variational data assimilation. *Journal of Computational Physics*, 2018.

[7] R. Arcucci, L. DAmore, L. Carracciuolo, G. Scotti, and G. Laccetti. A decomposition of the tikhonov regularization functional oriented to exploit hybrid multilevel parallelism. *International Journal of Parallel Programming*, 45(5):1214–1235, 2017.

[8] R. Arcucci, L. Carracciuolo, and R. Toumi. Toward a preconditioned scalable 3dvar for assimilating sea surface temperature collected into the caspian sea. *Journal of Numerical Analysis, Industrial and Applied Mathematics*, 12(1-2):9–28, 2018.

[9] D.G. Cacuci, I. M. Navon, and M. Ionescu-Bujor. Computational methods for data evaluation and assimilation. *CRC Press*, 2013.

[10] R. Ford, C. C. Pain, M. D. Piggott, A. J. H. Goddard, C. R. E. de Oliveira, and A. P. Umpleby. A nonhydrostatic finite-element model for three-dimensional stratified oceanic flows. part i: Model formulation. *Monthly Weather Review*, 132:2816–2831, 2004.

[11] E. Aristodemou, T. Bentham, C. Pain, and A. Robins. A comparison of mesh-adaptive les with wind tunnel data for flow past buildings: Mean flows and velocity fluctuations. *Atmospheric Environment*, 43:6238–6253, 2009.

[12] Imperial College London AMCG. Fluidity manual v4.1.12. 4 2015.

[13] H. K. Engl, M. Hanke, and A. Neubauer. Regularization of inverse problems. *Kluwer*, 1996.

[14] N. Nichols. Mathematical concepts in data assimilation. *W. Lahoz, et al. (Eds.), Data Assimilation, Springer*, 2010.

[15] P.C. Hansen. Rank deficient and discrete ill-posed problems. *SIAM, Philadelphia*, 1998.

[16] Y. Wang, I. M. Navon, X. Wang, and Y. Cheng. 2D Burgers equation with large Reynolds number using POD/DEIM and calibration. *International Journal for Numerical Methods in Fluids*, 82(12):909–931, 2016.

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

[18] JP. Courtier. A strategy for operational implementation of 4d-var, using an incremental approach. *Q J R Meteorol Soc*, 120(519):1367–1387, 1994.

[19] 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*, pages 311–326, 2017.

[20] D. Pavlidis, G.J. Gorman, J.L.M.A. Gomes, C. Pain, and H. ApSimon. Synthetic-Eddy Method for Urban Atmospheric Flow Modelling. *Boundary-Layer Meteorology*, 136:285–299, 2010.