

Preconditioning Weak-Constraint 4D-Var: A Parallelisable Implementation in Firedrake

Jemima M. Tabcart, David Ham, Josh Hope-Collins

Abstract

Data assimilation refers to a class of methods which seek to find the most likely state of a dynamical system by combining information from a (numerical) model of the system of interest with measurements of the system [2]. The most mature application of data assimilation is to numerical weather prediction, where large dimensional problems (10^9 dimensional states and 10^7 measurements) need to be solved in a very short amount of time. Algorithms also need to be highly parallelisable in order to exploit high performance computing resources available at meteorological centres.

In variational data assimilation methods, a non-linear least squares problem is solved via a Gauss-Newton approach [5]. One computationally expensive component of this implementation consists of approximately solving large linear systems. Preconditioners can help to speed up the convergence of iterative methods, but it can be challenging to design effective and efficient preconditioning methods. This is particularly true for the weak-constraint 4D-Var problem, which accounts for the fact that the numerical model itself is imperfect. Relaxing the assumption of a perfect model increases the size of the state space, but introduces the possibility of using parallel-in-time [3] methods, compared to the strong constraint method where model evaluations must be performed in serial.

For a fixed time window $[t_0, t_N]$, $\mathbf{x}_i^t \in \mathbb{R}^s$ denotes the true state of a dynamical system of interest at time t_i , with observations $\mathbf{y}_i \in \mathbb{R}^{p_i}$ made at times t_i . Prior information obtained from a numerical model, $\mathbf{x}_b \in \mathbb{R}^s$, is then combined with the observation information to find $\mathbf{x}_i \in \mathbb{R}^s$, the most likely state of the system at time t_i . The prior, or background state, is valid at initial time t_0 and can be written as an approximation to the true state via $\mathbf{x}_b = \mathbf{x}_0^t + \epsilon^b$. We assume that the background errors $\epsilon^b \sim \mathcal{N}(0, B)$. We define a, possibly non-linear, observation operator $\mathcal{H}_i : \mathbb{R}^s \rightarrow \mathbb{R}^{p_i}$ which maps from state variable space to observation space at time t_i . Observations are written as $\mathbf{y}_i = \mathcal{H}_i[\mathbf{x}_i^t] + \epsilon_i \in \mathbb{R}^{p_i}$, for $i = 0, 1, \dots, N$, where the observation error $\epsilon_i \sim \mathcal{N}(0, R_i)$ for $R_i \in \mathbb{R}^{p_i \times p_i}$.

This weak constraint 4D-Var problem then leads to a non-linear objective function of the form:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T B^{-1}(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} \sum_{i=0}^N (\mathbf{y}_i - \mathcal{H}_i[\mathbf{x}_i])^T R_i^{-1} (\mathbf{y}_i - \mathcal{H}_i[\mathbf{x}_i]) \\ + \frac{1}{2} \sum_{i=1}^N (\mathbf{x}_i - \mathcal{M}_i(\mathbf{x}_{i-1}))^T Q_i^{-1} (\mathbf{x}_i - \mathcal{M}_i(\mathbf{x}_{i-1})),$$

Within each outer loop, the inner loop minimises a quadratic objective function to find $\delta \mathbf{x}^{(l)} \in \mathbb{R}^{s(N+1)}$, where $\delta \mathbf{x}^{(l)} = \mathbf{x}^{(l+1)} - \mathbf{x}^{(l)}$. Writing $\delta \mathbf{x} = (\delta \mathbf{x}_0^T, \delta \mathbf{x}_1^T, \dots, \delta \mathbf{x}_N^T)^T$, the full non-linear observation operator \mathcal{H}_i (similarly the model operator \mathcal{M}_i) is linearised about the current best guess $\mathbf{x}_i^{(l)}$ to obtain the linearised operator $H_i^{(l)}$ (respectively $M_i^{(l)}$). The updated initial guess $\delta \mathbf{x}_0^{(l)}$ is propagated forward between observation times by $M_i^{(l)}$ to obtain $\delta \mathbf{x}_{i+1}^{(l)} = M_i^{(l)} \delta \mathbf{x}_i^{(l)}$.

The aim of the inner loop is to solve the symmetric positive definite system given by

$$\mathbf{S} \delta \mathbf{x} = \mathbf{L}^T \mathbf{D}^{-1} \mathbf{b} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{d}, \quad \mathbf{S} = \mathbf{L}^T \mathbf{D}^{-1} \mathbf{L} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad (1)$$

where

$$\mathbf{D} = \begin{pmatrix} B & & \\ & Q_1 & \\ & & \ddots \\ & & & Q_N \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} -M_1 & I & & \\ & \ddots & \ddots & \\ & & -M_N & I \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} R_0 & & \\ & R_1 & \\ & & \ddots \\ & & & R_N \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} H_0 & & \\ & H_1 & \\ & & \ddots \\ & & & H_N \end{pmatrix},$$

with M_i, H_i being the linearisations of \mathcal{M}_i and \mathcal{H}_i about the current solution.

However, the primal formulation, (1), has limited potential for acceleration via preconditioning approaches. In particular, it is difficult to exploit the inherent parallelism in the forward problem when designing preconditioners. Recent work has focused on a reformulation of the linearised objective function as a saddle point system [6] which takes the form

$$\begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{L} \\ \mathbf{0} & \mathbf{R} & \mathbf{H} \\ \mathbf{L}^\top & \mathbf{H}^\top & \mathbf{0} \end{pmatrix} \begin{pmatrix} \delta\eta \\ \delta\nu \\ \delta\mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{d} \\ \mathbf{0} \end{pmatrix}. \quad (2)$$

A number of approaches to precondition the model term, \mathbf{L} , have been proposed (see e.g. [6, 7, 12, 11]), many of which impose parallel-in-time structure on $\hat{\mathbf{L}}^{-1} \approx \mathbf{L}^{-1}$. However, testing and comparing these new preconditioners in realistic frameworks is not straightforward. Variational data assimilation requires access to linearised model and adjoint operators. The time cost of implementing these for a new problem means that researchers often test their new approaches on a limited number of toy problems, such as the Lorenz 96 problem [9] or the shallow water equations. If available (and accessible to the researchers), the next step is a full scale implementation within an operational code, meaning that even in the best case there is a gap in test problems. Data assimilation methods are also increasingly being applied to other dynamical systems, and the properties of the usual toy models that make them appropriate for weather applications may no longer be desirable or relevant for other applications.

In this project we integrate both weak- and strong-constraint 4D-Var algorithms within the Firedrake [8] and PETSc [1] frameworks, for both the primal (1) and saddle point (2) problems. Firedrake is an automated system for the solution of partial differential equations using the finite element method. In particular, this means that tangent linear and adjoint operators are available automatically, significantly reducing the implementational burden of applying variational data assimilation methods to new models. For the user, the cost of setting up the 4D-Var system is not substantially higher than the cost of a single run of the forward model \mathcal{M}_i and application of the observation operators \mathcal{H}_i . This implementation is also done in parallel.

In this talk I will present the integration of variational data assimilation problems within Firedrake, and demonstrate how this ‘plug-and-play’ approach allows users to focus on the numerical linear algebra aspects of their problem rather than the implementation of test problems. In particular I will present a theoretical and practical comparison of existing and new preconditioners for the model term, \mathbf{L} , including the block Toeplitz approaches of [6, 11], block diagonal approximations as studied in [12], and block (α) -circulant preconditioners as used in the all-at-once setting (see e.g [10, 4]).

References

- [1] S. Balay, S. Abhyankar, M. F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E.

- Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, J. Faibussowitsch, W. D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M. G. Knepley, F. Kong, S. Kruger, D. A. May, L. Curfman McInnes, R. Tran Mills, L. Mitchell, T. Munson, J. E. Roman, K. Rupp, P. Sanan, J. Sarich, B. F. Smith, H. Suh, S. Zampini, H. Zhang, H. Zhang and J. Zhang, *PETSc/TAO Users Manual*, Argonne National Laboratory, ANL-21/39 - Revision 3.22, = 10.2172/2205494, 2024.
- [2] A. Carrassi, M. Bocquet, Marc, L. Bertino, and G. Evensen. *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*. Wiley Interdisciplinary Reviews: Climate Change, 9 (5): e535, 2018.
- [3] M. J. Gander, *50 Years of Time Parallel Time Integration* Multiple Shooting and Time Domain Decomposition Methods: MuS-TDD, Heidelberg, May 6-8, 2013. Cham: Springer International Publishing, 69-113, 2015.
- [4] M. J. Gander, and S. L. Wu. *Convergence analysis of a periodic-like waveform relaxation method for initial-value problems via the diagonalization technique*. Numerische Mathematik, 143, 489-527, 2019.
- [5] S. Gratton, A. S. Lawless, and N. K. Nichols. *Approximate Gauss-Newton Methods for Nonlinear Least Squares Problems*. SIAM Journal on Optimization, 18(1):106–132, 2007.
- [6] M. Fisher and S. Gürol. *Parallelization in the time dimension of four- dimensional variational data assimilation*. Quarterly Journal of the Royal Meteorological Society, 143(703):1136–1147, 2017.
- [7] M. A. Freitag and D. L. H. Green. *A low-rank approach to the solution of weak constraint variational data assimilation problems*. Journal of Computational Physics, 357:263–281, 2018.
- [8] D. A. Ham, P. H. J. Kelly, L. Mitchell, C. J. Cotter, R. C. Kirby, K. Sagiya, N. Bouziani, S. Vorderwuelbecke, T. J. Gregory, J. Betteridge, D. R. Shapero, R. W. Nixon-Hill, C. J. Ward, P. E. Farrell, P. D. Brubeck, I. Marsden, T. H. Gibson, M. Homolya, T. Sun, A. T. T. McRae, F. Luporini, A. Gregory, M. Lange, S. W. Funke, F. Rathgeber, G.-T. Bercea and G. R. Markall. *Firedrake User Manual*. First edition, doi:10.25561/104839, 2023.
- [9] E. N. Lorenz, *Predictability: a problem partly solved*, in Seminar on Predictability, 4–8 September 1995, ECMWF, Reading, 1995.
- [10] E. McDonald, J. Pestana, and A. J. Wathen. *Preconditioning and Iterative Solution of All-at-Once Systems for Evolutionary Partial Differential Equations*. SIAM J. Sci. Comput., 40, 2018.
- [11] D. Palitta and J. M. Tabart. *Stein-based preconditioners for weak-constraint 4D-Var*. Journal of Computational Physics, 482:112068, 2023.
- [12] J. M. Tabart and J. W. Pearson. *Saddle point preconditioners for weak-constraint 4D-Var*. ETNA - Electronic Transactions on Numerical Analysis, 60:197–220, 2024.