Department of Meteorology and Data Assimilation Research Centre (DARC)

University of Reading

A weak-constraint 4DEnsembleVar

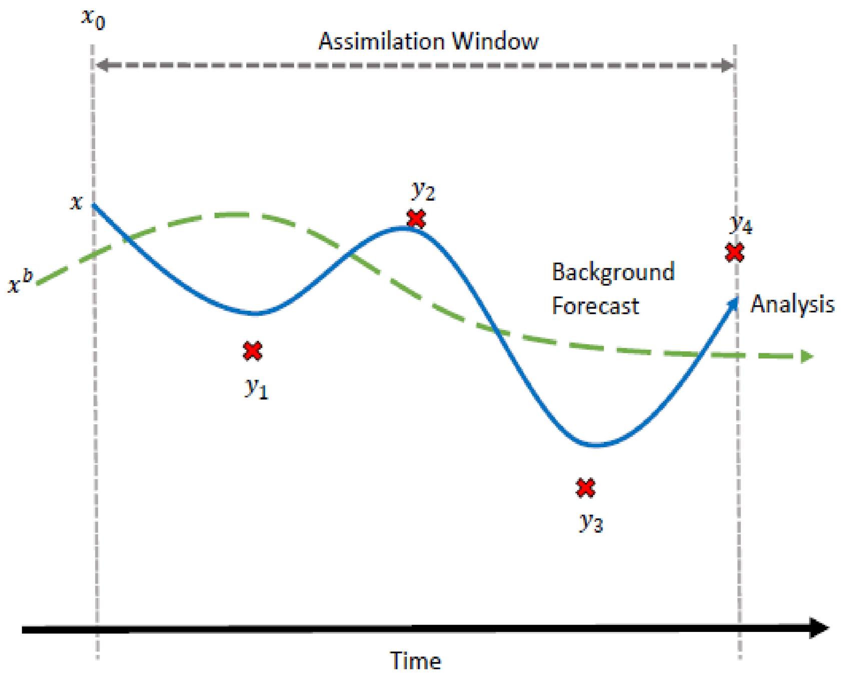
Javier Amezcua (j.amezcuaespinosa@reading.ac.uk) | Michael Goodliff | Peter Jan van Leeuwen

Summary

We formulate a 4DEnsembleVar for the weak-constraint data assimilation problem. Namely, we take into consideration (and make use of) the presence of model error. This formulation partially alleviates the problem of statically localised 4D cross-time covariances, which affects strong-constraint 4DEnsembleVar.

From SC4DVar to SC4DEnVar

4DVar is a minimisation problem over an assimilation window. In the **strong-constraint** case, it reduces to an **initial value problem**.



Model evolution: $\mathbf{x}^t = m^{0\to t}(\mathbf{x}^0)$

Observations: $\mathbf{y}^{o,t} = h(\mathbf{x}^t) + \boldsymbol{\eta}^t$

SC4DVar is solved by minimising the **cost function**:

$$J(\mathbf{x}^0) = \frac{1}{2}(\mathbf{x}^0 - \mathbf{x}^{b,0})^{\mathbf{T}}\mathbf{B}^{-1}(\mathbf{x}^0 - \mathbf{x}^{b,0}) + \frac{1}{2}\sum_{t=1}^{\tau}\rho^t(\mathbf{y}^{o,t} - h(m^{0\to t}(\mathbf{x}^0)))^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{y}^{o,t} - h(m^{0\to t}(\mathbf{x}^0)))$$

which can be solved (for a convex function) by **equating** the **gradient to zero**:

$$\nabla J(\mathbf{x}^0) = \mathbf{B}^{-1}(\mathbf{x}^0 - \mathbf{x}^{b,0}) - \sum_{t=1}^{\tau}\rho^t(\mathbf{M}^{0\to t})^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}\mathrm{R}^{-1}(\mathbf{y}^{o,t} - h(m^{0\to t}(\mathbf{x}^0)))$$

Indicator function: 1 if there are observations, 0 otherwise

Tangent linear observation model (**adjoint** after transpose)

Tangent linear evolution model (**adjoint** after transpose)

The **adjoint models** have the role of **communicating observational impacts at time *t* back to time *0***. **(SC)4DEnsembleVar** uses **sample cross-time covariances** to do this job.

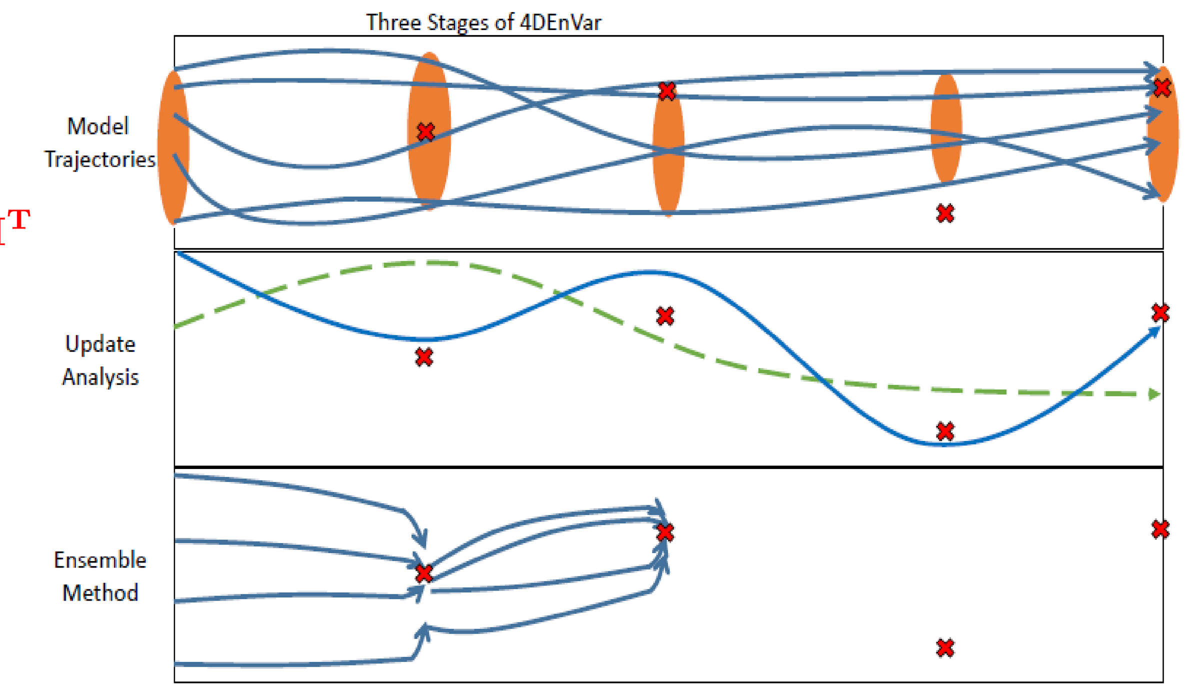
$(\mathbf{M}^{0\to t})^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}$

$(\mathbf{P}^0)^{-1}\mathbf{P}^0(\mathbf{M}^{0\to t})^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}$

$(\mathbf{P}^0)^{-1}\hat{\mathbf{X}}^0\left(\hat{\mathbf{X}}^0\right)^{\mathbf{T}}(\mathbf{M}^{0\to t})^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}$

$(\mathbf{P}^0)^{-1}\hat{\mathbf{X}}^0\left(\hat{\mathbf{X}}^t\right)^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}$

$(\mathbf{P}^0)^{-1}\hat{\mathbf{X}}^0\left(\hat{\mathbf{Y}}^t\right)^{\mathbf{T}}$



Sample estimators are often noisy and need to be **localised**.

$\hat{\mathbf{X}}^0\left(\hat{\mathbf{Y}}^t\right)^{\mathbf{T}} \to \left(\hat{\mathbf{X}}^0\left(\hat{\mathbf{Y}}^t\right)^{\mathbf{T}}\right)\circ\mathbf{L}_{\mathbf{x}=0,\mathbf{y}=t}$ Localization **needs to evolve** with the flow

What is the structure of $\mathbf{L}_{\mathbf{x}=0,\mathbf{y}=t}$? Not an easy answer since it is **flow dependent**. But we **often use static** $\mathbf{L}_{\mathbf{x}=0,\mathbf{y}=0}$ since it is easier.

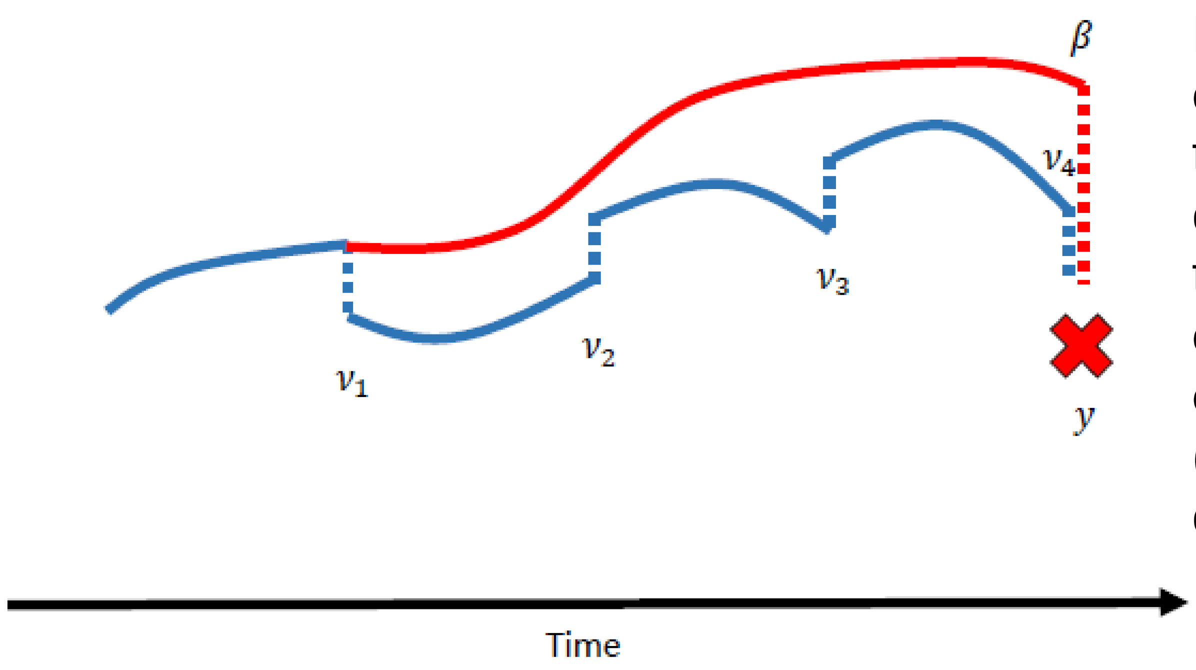
Making use of the model error

Models are **seldom perfec**t, since they contain unresolved processes, missing physics, numerical integration errors, etc. A 1-model-step **model error** can be represented as:

Lag-1 model error $\mathbf{x}^t = m^{(t-1)\to t}(\mathbf{x}^{t-1}) + \nu^t$ $\quad\boldsymbol{\nu} \sim N(\mathbf{0},\mathbf{Q})$

In the **weak-constrain**t 4DVar there is not a unique way to define the control variables. For simplicity, we choose the '**model bias correction**' formulation of weak-constraint 4DVar (Tremolet, 2006).

Effective model error (statistics not trivial) $\mathbf{x}^t = m^{0\to t}(\mathbf{x}^0) + \beta^t$



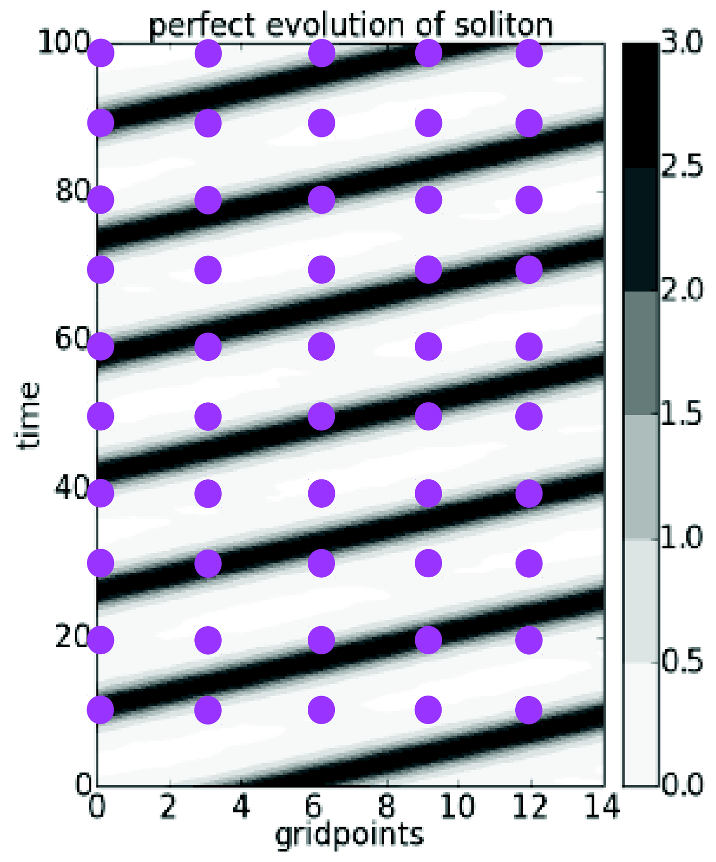
Instead of making **corrections** every time step (ideal yet costly), one finds **the effective correction at the observational time** (simpler and fewer control variables).

An example in the KdV model

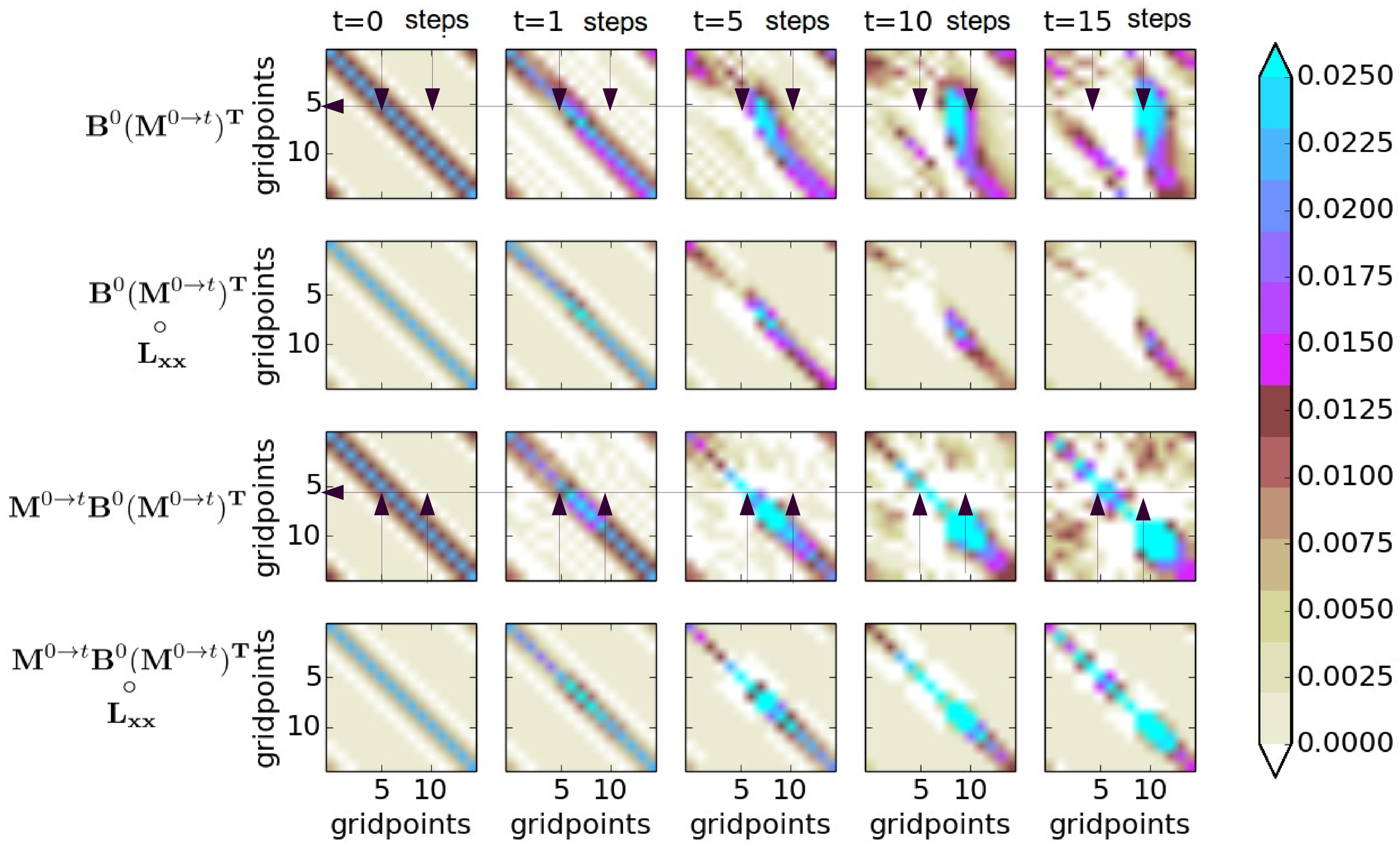
The **Korteweg-de Vries equation** is a nonlinear partial differential equation that describes the **evolution of a soliton in** 1 dimension**.**

We discretise this equation over space (periodic domain) and time and use it as our **model**.

We compute explicitly the **tangent linear model**, and perform a **data assimilation experiment** with observations as indicated in the figure.

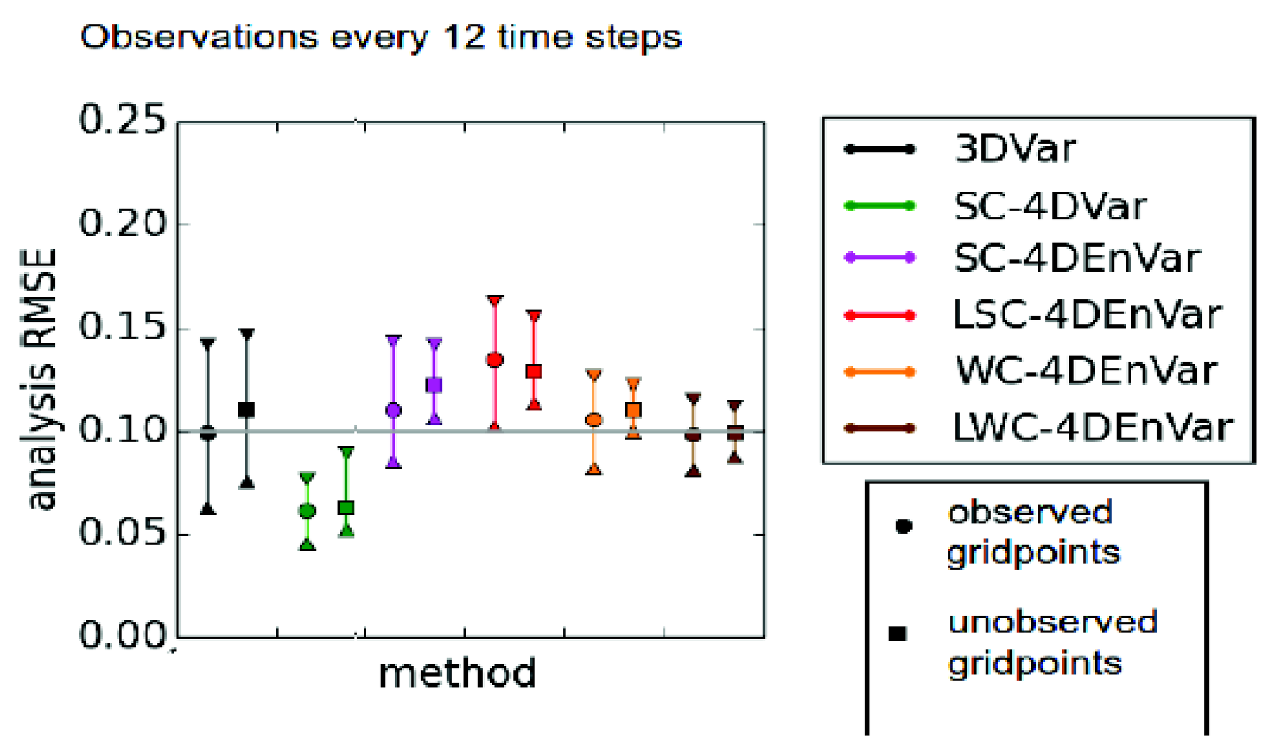


The following figure illustrates the **structure of cross covariances between time *0* and time t**, as well as **covariances at time t**. A sample estimator would be quite noisy and require localisation, so we provide the **localised versions** as well.



As time progresses, **most of the information in the cross-time covariances** lie in the **region killed by localisation**. The **covariances at time *t***, however, are more diagonal dominant, so that **information remains even after localisation**. The **WC formulation allows for increments to be applied at these times.**

Not surprisingly, the **performance of SC4DEnVar is not improved by localisation**, as shown in the following fgure. The **performance of WC4DEnVar** is **better**.



References

1. J. Amezcua, M. Goodliff and P. J. van Leewen. A weak-constraint 4DEnsembleVar. Part I: theory and simple model experiments. *Tellus A*. In review.
2. M. Goodliff, J. Amezcua and P. J. van Leeuwen. A weak-constraint 4DEnsembleVar. Part II: large model experiments. *Tellus A*. In review.
3. Y. Tremolet, 2006. Accounting for an imperfect model in 4D-Var. *Q. J. R. Meteorol. Soc.*, **132**, 2483–2504.

Acknowledgements

- The authors kindly acknowledge the funding of NCEO and NERC via the DIAMET project.