Recent developments of the BUMP software

Benjamin Ménétrier - IRIT, Toulouse (JCSDA funding)

18th JCSDA Technical Review Meeting and Science Workshop

June 8, 2021

Outline

BUMP overview

Training process

Affordable training

Conclusions

Outline

BUMP overview

Training process

Affordable training

Conclusions

What is B?

In the variational framework, the background error covariance matrix B describes the uncertainty of the background x^b:

$$\mathcal{J}\left(\mathrm{x}\right) = \boxed{\underbrace{\left(\mathrm{x}-\mathrm{x}^b\right)^{\mathrm{T}} \mathrm{B}^{-1} \left(\mathrm{x}-\mathrm{x}^b\right)}_{\mathcal{J}_b(\mathrm{x})}} + \underbrace{\left(\mathrm{y}^o - \mathcal{H}\left(\mathrm{x}\right)\right)^{\mathrm{T}} \mathrm{R}^{-1} \left(\mathrm{y}^o - \mathcal{H}\left(\mathrm{x}\right)\right)}_{\mathcal{J}_o(\mathrm{x})}$$

In the equivalent BLUE framework, B is the last operator applied to define the analysis increment:

$$\mathrm{x}^a = \mathrm{x}^b + \mathrm{B}\mathrm{H}^{\mathrm{T}} \left(\mathrm{H}\mathrm{B}\mathrm{H}^{\mathrm{T}} + \mathrm{R}\right)^{-1} \left(\mathrm{y}^o - \mathcal{H}\left(\mathrm{x}^b\right)\right)$$

B contains complex multivariate 3D structures (potentially 4D for the 4DEnVar algorithm).

Hybrid background error modeling

Hybrid background error covariance matrix:

$$\mathrm{B} = \underbrace{\beta^{e2}\, \mathrm{L} \circ \widetilde{\mathrm{B}}}_{\text{“ensemble” term}} + \underbrace{\beta^{s2}\, \mathrm{B}^s}_{\text{“static” term}}$$

where:

- $\widetilde{\mathrm{B}}$ is sampled from an ensemble of forecasts
- L is the localization matrix
- $\circ$ denotes a Schur product (element-by-element)
- β^e and β^s are scalar coefficients
- B^s is a “static” covariance matrix, usually modeled as:

$$\mathrm{B}^s = \mathrm{K}\boldsymbol{\Sigma}\mathrm{C}\boldsymbol{\Sigma}\mathrm{K}^{\mathrm{T}}$$

where:

- K is a balance operator
- $\boldsymbol{\Sigma}$ is the diagonal matrix of standard-deviations
- C is a correlation operator

BUMP overview

- **BUMP** stands for “Background error on an Unstructured Mesh Package”.
- **BUMP** is one of the background error covariance libraries of **SABER** (System Agnostic Background Error Representation), a component of the **JEDI** project.
- **BUMP** works with any model grid and is able to take complex boundaries into account (important for ocean or land models).
- Written in modern Fortran 90 ($\sim$ 25.000 lines), the code can be easily called through Fortran and C++ interfaces.
- Interfaces with **BUMP** are implemented for most of JEDI models, via OOPS.

BUMP functionalities

BUMP provides tools based on ensembles of forecasts to estimate:

- Horizontal and vertical localization length-scales of L and hybrid weights β^e and β^s, estimated locally using the theory of Ménétrier and Auligne (2015) [HDIAG].
- Vertical statistical regressions of the balance operator K estimated locally [VBAL].
- Standard-deviations Σ objectively filtered with the method of Ménétrier *et al.* (2015a,b) [VAR].
- Horizontal and vertical correlation length-scales of C estimated locally and simultaneously [HDIAG], or alternatively Local Correlation Tensors [LCT] for anisotropic functions.

A grid smoother is required to apply L and C in a variational framework: **BUMP** implements the **NICAS** method (Normalized Interpolated Convolution from an Adaptive Subgrid) [NICAS].

Outline

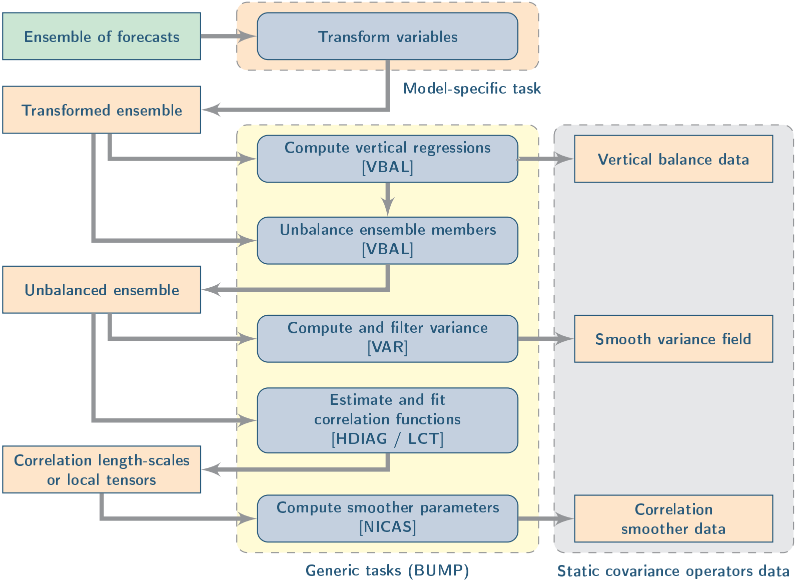
BUMP overview

Training process

Affordable training

Conclusions

Static covariance training



Ensemble of forecasts

Transform variables

Model-specific task

Transformed ensemble

Compute vertical regressions [VBAL]

Vertical balance data

Unbalance ensemble members [VBAL]

Unbalanced ensemble

Compute and filter variance [VAR]

Smooth variance field

Estimate and fit correlation functions [HDIAG / LCT]

Correlation length-scales or local tensors

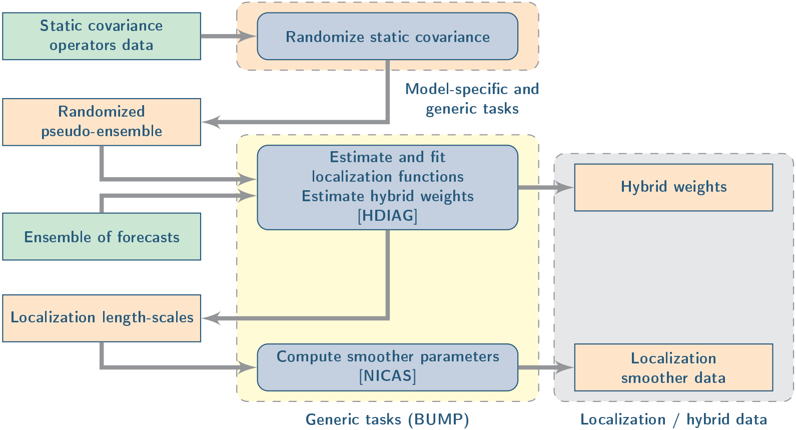
Compute smoother parameters [NICAS]

Correlation smoother data

Generic tasks (BUMP)

Static covariance operators data

Localization and hybrid weights training


Static covariance operators data
Randomize static covariance
Model-specific and generic tasks
Randomized pseudo-ensemble
Estimate and fit localization functions
Estimate hybrid weights
[HDIAG]
Hybrid weights
Ensemble of forecasts
Localization length-scales
Compute smoother parameters
[NICAS]
Localization smoother data
Generic tasks (BUMP)
Localization / hybrid data

Outline

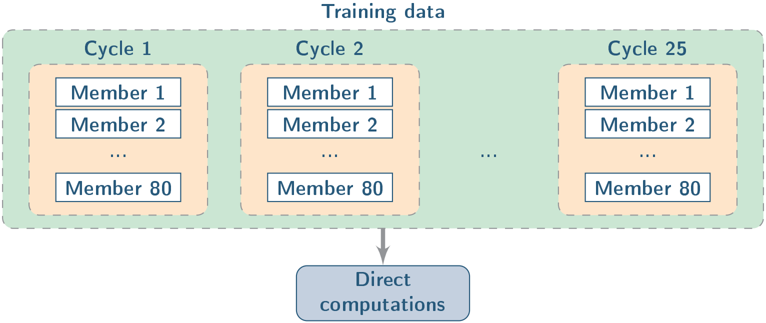
BUMP overview

Training process

Affordable training

Conclusions

Critical memory issue

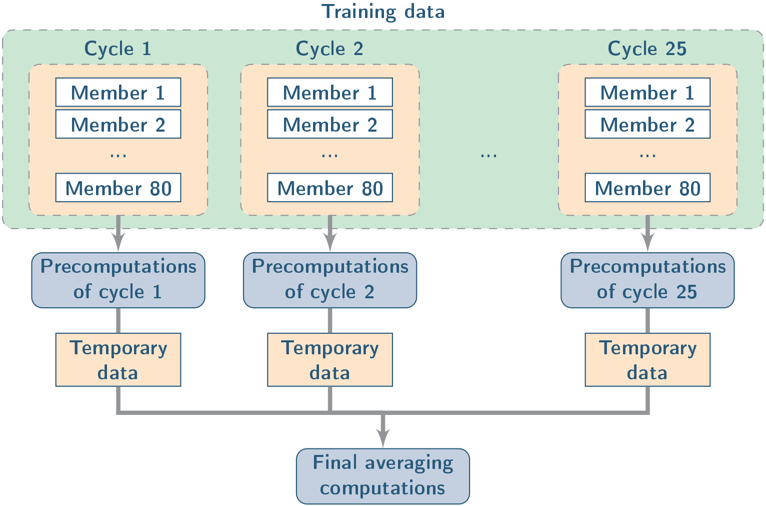


For FV3 at resolution C384:

- 1 member: $\sim$ 11 Gb
- 1 cycle of 80 members: $\sim$ 880 Gb
- 25 cycles of 80 members: $\sim$ 22 Tb

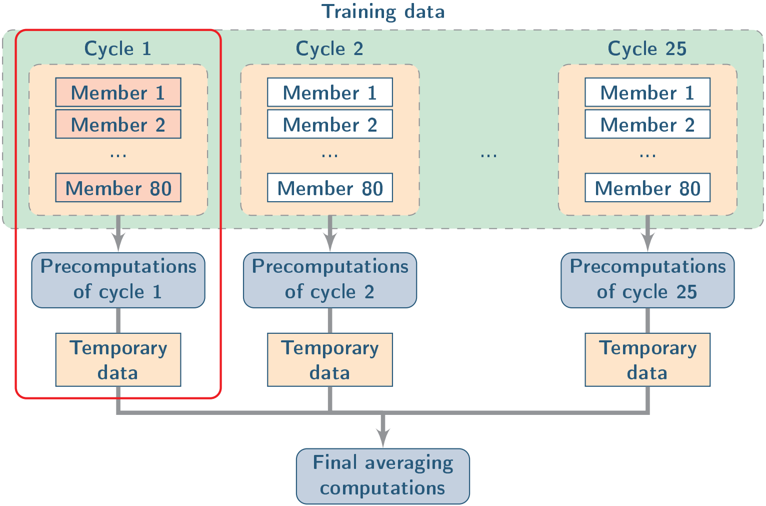
Difficult to load simulatneously on disk, not to say in memory...

BUMP solution: split the work



Easy to do for variance [VAR] and correlation [HDIAG / LCT], required theoretical developments for vertical regressions [VBAL].

BUMP solution: load members sequentially



Members are loaded sequentially to reduce the memory footprint.
No slow down, no precision loss.

Outline

BUMP overview

Training process

Affordable training

Conclusions

Conclusions

In the recent months, **BUMP** has been upgraded to improve the **B** components training:

- work split in cycle-based precomputations, with a final averaging step,
- reduced memory footprint, reading ensemble members sequentially.

Tests are underway on Orion for FV3-JEDI at full resolution (C384). Once the new branch is validated, it will be merged into `develop` and available for all users. A working example of a full static B training workflow will be provided.

Thanks for your attention!