ISDA-Online

January 8, 2021 / 3 – 5 pm UTC
(16 – 18 pm CET / 8 – 10 am MST / 12 pm – 2 am JST)

“Data Assimilation Methodology”

Presentations:

Feature Data Assimilation - Theory, Algorithms and Examples
Roland Potthast and Christian Welzbacher

Space-Time Multigrid for the Maximum Likelihood Ensemble Filter Method
Xinfeng Gao, Milija Zupanski, Jacob B. Schroder, Robert D. Falgout

A method for representing spatially correlated observation errors for wind data
Oliver Guillet, Yann Michel, Mathilde Moureaux

Efficient nonlinear data assimilation using synchronization in a particle filter
Flavia Pinheiro, Peter Jan van Leeuwen

Understanding the differences between EnVar and LETKF solvers in an operational NWP setting
Anna Shlyaeva, Jeff Whitaker

Feature Data Assimilation - Theory, Algorithms and Examples

Roland Potthast and Christian Welzbacher

Deutscher Wetterdienst, Offenbach, Germany

The classical setup of atmospheric data assimilation employs direct (in-situ) or indirect remote sensing measurements (e.g. satellite radiances) within a variational data assimilation framework to find the best possible state of some dynamical system as initial condition for forecasting. Modern ensemble data assimilation systems exploit the covariance information provided by an ensemble of forcasted states within their analysis step. Hybrid approaches such as EnVAR combine the advantages of both worlds. When classical continuous variables are measured directly or indirectly, the classical approach has been proven to provide high-quality assimilation and forecasting results.

With the growth of temporally and spatially high-resolution measurement data such as 3D-Volume RADAR as well as hyper-spectral satellite radiances and further temporally high-resolved remote sensing techniques, both variational as well as ensemble-based approaches are significantly challenged by the strong non-linearity of atmospheric processes linked to cloud formation and precipitation processes. The complex processes, which take place for example in thunderstorms, do not allow to fully fit the full dynamical behaviour of a process to observed phenomena based on measurements.

The goal of feature or object data assimilation is to move from the assimilation of snapshots of some process recorded by classical direct or indirect measurements to the assimilation of properties or features of a whole process. Often, the process leads to the formation of objects such as clouds or thunderstorms, which have a typical behaviour with a life cycle which consists of birth, growth, stability and decay.

Here, we describe a proper mathematical framework for the description and assimilation of features of phenomena and objects within an ensemble data assimilation systems. Starting from a generic Bayesian approach we describe the natural derivation of feature assimilation methods. We then discuss the design and properties of feature or object forward operators and their use within an ensemble Kalman filter (LETKF) or particle filter (LAPF/LMCPF) based assimilation system. Examples will be shown for the popular Lorenz 63 & 96 benchmark systems as well as for the convective scale ICON model, which is in preparation for operational use of the COSMO consortium with about 40 weather services and, in particular, by Deutscher Wetterdienst from Q1/2021.

Space-Time Multigrid for the Maximum Likelihood Ensemble Filter Method

Xinfeng Gao*, Milija Zupanski+, Jacob B. Schroder†, Robert D. Falgout‡

*Colorado State University, CFD and Propulsion Laboratory, Fort Collins, Colorado, USA

+Cooperative Institute for Research in the Atmosphere, Fort Collins, Colorado, USA

†University of New Mexico, Department of Mathematics and Statistics, Albuquerque, New Mexico, USA

‡Lawrence Livermore National Laboratory, Center for Applied Scientific Computing, Livermore, California, USA

In this work, we propose a new method to improve the computational efficiency of minimizing a 4D-variational-ensemble optimal control system. The new method employs multigrid in space and time to aim for model reduction by successively restricting the full system to lower resolutions for more efficient assimilation, while incorporating the covariance information during the multigrid process to ensure the essen- tial dynamical properties are retained and predicted with sufficient accuracy. This work is motivated by applying data assimilation to high dimensional nonlinear engineering applications such as turbulent combustion. While the meteorological community already employs spatial multigrid methods with data assimilation, applying multi-grid reduction in time (MGRIT) to data assimilation is new. Using MGRIT, parallel in time can provide further speedups and effectively make use of future architecture.

The 4D-variational-ensemble assimilation method is based on the maximum likelihood ensemble filter (MLEF). The cost function is derived based on a Gaussian probability density function framework and Bayes theorem. While the original MLEF is an ensemble-based sequential data assimilation method, we herein introduce the assimilation of data over a time window, a summation of multiple time terms instead of a single time term in the original MLEF, which adds the ingredient of strong constraint 4D-variational assimilation. The core concept is achieved by solving the nonlinear system using the full approximation scheme of the space-time multigrid method. The proposed space-time multigrid is a natural fit for solving the data assimilation optimal control system, because the optimization constrained by the nonlinear dynamical model can be solved on a sequence of low-resolution space-time operators with computational efficiency. This process can be simply depicted by J^m, J^c, J^{cst}, which denote cost functions at various resolutions, in Fig. 1. The technical novelty of the parallel-in-space-time optimization algorithm will be presented and discussed in detail at the symposium.

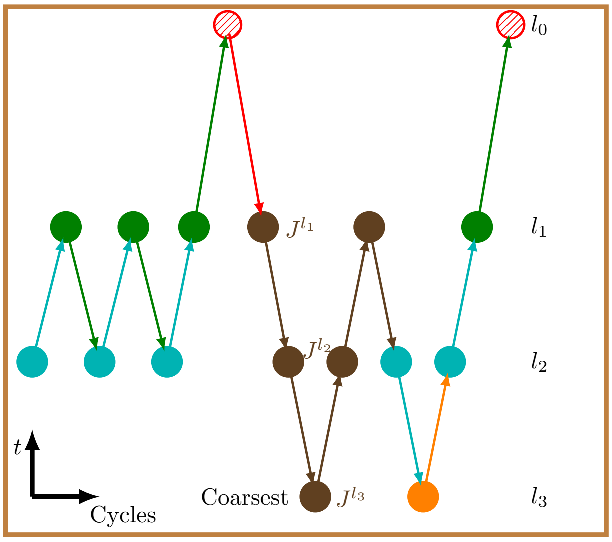

Figure 1: An illustration of MG cycles in time. Levels 0,1,2,3 indicate the meshes from the finest to the coarsest in both space and time.

A method for representing spatially correlated observation errors for wind data

Oliver Guillet, Yann Michel, Mathilde Moureaux

Météo-France, Toulouse, France

The objective of this research project is to drastically improve the spatial density of the observations exploited in the numerical prediction systems ARPEGE and AROME. This requires accurately representing the observation error correlations. To this end, we use a technique coming from the field of oceanography, based on the solution of a diffusion equation, which we decide to apply on unstructured meshes.

A first study dealing with scalar data makes use of the finite element method. It provides a way to represent horizontal error correlations for scalar data, such as brightness temperature from satellite data. Experimental validation was achieved using measurements from the infrared imager MSG/SEVIRI, which are assimilated both in ARPEGE and AROME.

We propose to extend this method to the representation of wind error correlations. These are vectorial data, meaning that every observation location is associated with two values, zonal and meridional. We focus specifically on scatterometer measurements, that are available every 25km but only assimilated every 100km.

First, the wind field is decomposed into one divergent component and and rotational component. Then, the scalar correlation operator is applied to each component. Finally, the wind field is reconstructed while maintaining the symmetry and the positivity of the correlation operator. Experiments show agreement with the analytical results.

In the future, all types of observations will be considered, whether or not they are conventional, and we will extend the method to the three dimensional case to address the specific case of radars for instance.

Efficient nonlinear data assimilation using synchronization in a particle filter

Flavia Pinheiro[1], Peter Jan van Leeuwen[1,2]

[1] University of Reading, Department of Meteorology, Reading, United Kingdom.
[2] Colorado State University, Department of Atmospheric Science, Fort Collins, Colorado, USA.

Current data assimilation methods still face problems in strongly nonlinear cases. A promising solution is a particle filter, which provides a representation of the state probability density function (pdf) by a discrete set of particles. To allow a particle filter to work in high- dimensional systems, the proposal density freedom is explored. We used a proposal density from synchronisation theory, in which one tries to synchronise the model with the true evolution of a system using one-way coupling, via the observations. This is done by adding an extra term to the model equations which will control the growth of instabilities transversal to the synchronisation manifold. In this work, an efficient ensemble-based synchronisation scheme is used as a proposal density in the implicit equal-weights particle filter, which avoids filter degeneracy by construction. Tests using the Lorenz96 model for a 1,000-dimensional system show successful results, where particles efficiently follow the truth, both for observed and unobserved variables. These first tests show that the new method is comparable to, and slightly outperforms, a well-tuned Local Ensemble Transform Kalman Filter. We also look at another variant of synchronisation, in which observations back in time are also included. The advantage is that the synchronisation has more time to influence the particle trajectories, leading to better filter performance. This Synchronisation Particle Filter is a promising solution for high-dimensional nonlinear problems in the geosciences, such as numerical weather prediction.

Understanding the differences between EnVar and LETKF solvers in an operational NWP setting

Jeff Whitaker[1] and Anna Shlyaeva[2]

[1] NOAA Earth System Research Lab, Boulder, Colorado, USA
[2] Joint Center for Satellite Data Assimilation, Boulder, Colorado, USA.

In the next upgrade, the NOAA operational global hybrid ensemble-variational data assimilation system will implement an LETKF solver with model-space vertical localization to update ensemble perturbations. Previous work has shown that the inferior performance of the serial EnSRF and LETKF solvers compared to the (non-hybrid) EnVar solver in the NOAA system was primarily due to the of observation-space localization in the EnKF when assimilating radiances. Now that model-space localization has been implemented, the LETKF solver seems to perform slightly better than EnVar. In this talk, a hierarchy of simpler models is utilized to understand the reason for this. The results show that observation-error (R) localization (used in the LETKF) outperforms covariance (B) horizontal localization (used in EnVar) under certain conditions. In particular, when the horizontal scale of the Kalman Gain is narrower than the horizontal scale of the background-error covariance, R-localization performs better since it acts directly to localize the gain matrix. This tends to occur in the simple models studied when there are dense and/or accurate obs (the 'strong assimilation' limit) and the background-error covariance has heavy tails (the covariance is closer to exponential than it is to Gaussian). Both of these conditions are present in the operational NWP setting.