Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-18 (2016) 188–192

Bayesian Reduced-Resolution Data Assimilation

Daniel Hodyss*, Sarah King**

*Naval Research Laboratory, Monterey, CA 93943
USA (Tel: 831-656-4860; e-mail: daniel.hodyss@nrlmry.navy.mil).
**Naval Research Laboratory, Monterey, CA 93943 USA (e-mail: sarah.king@nrlmry.navy.mil)

Abstract: The numerical simulation of geophysical problems invariably leads to using a mesh that is coarser than what is required to resolve all of the important physical processes being described by the set of governing partial differential equations. This coarse mesh will therefore miss important physical phenomena that the observational instruments used for data assimilation will see. The performance of a data assimilation algorithm can be improved by accounting for these missing physical processes. We briefly review recent work describing how to properly use Bayes' rule when the model is attempting to predict a truncated version of a much higher resolution state-vector and the observations that are being assimilated are observing the elements of this high-resolution state-vector. Then, we go on to describe a practical ensemble (Monte Carlo) data assimilation system that makes use of this theory in a simple problem which has the property that data assimilation at low-resolution works very poorly unless the aforementioned theory is properly accounted for.

Keywords: Bayesian, Data Assimilation, Ensemble Methods, Dimension Reduction

1. INTRODUCTION

In geophysical applications we are rarely if ever able to simulate the problem at hand at a resolution for which all important scales of motion are fully resolved. Almost universally we must truncate the continuous variables of interest to a discrete set and then concatenate those variables into a state-vector that does not fully describe the problem. Typically, we model this state-vector with a discretized partial differential equation (PDE) that coarsely models the entire physical system. The result of this coarsening of the simulation of the system is that the numerical model does not simulate the actual variables of interest but simulates some (unknown) function of the variables of interest. For example, a coarse spatial mesh used to solve the typical hyperbolic PDEs of geophysical fluid dynamics delivers a solution that is smoother than a fine spatial mesh, and therefore the solution for each element of the coarse mesh model is some function of many elements of the fine mesh model. This has particular consequences on the data assimilation (the process by which prior simulations are combined with observations to produce a state estimate) that has not been accounted for previously in a rigorous Bayesian framework. Observations of the actual physical system observe state variables on the high-resolution mesh that are not actually simulated by our coarse mesh forecast model, at least not directly. The questions to be examined in this manuscript include: what does Bayes' rule mean in this context? What is the best way to make use of these kinds of observations? And, what data assimilation system should be used in this situation? We will review some recent research towards a solution to this problem and illustrate its application in an example problem.

2. THEORY AND GAUSSIAN EXAMPLE

In this section we review the general theory for reduced-resolution data assimilation presented in Hodyss and Nichols (2015). Our review will only cover the aspects necessary to build a data assimilation system that uses a reduced state-vector. Here, we will build a Gaussian covariance model and from it deduce the correct state-estimation procedure for the case where the observations are viewing a state with a higher dimension than the available forecast model is capable of simulating.

2.1 Problem Setup

A simple way to construct a Gaussian problem that is amenable to analysis is through the use of a discrete Fourier series representation. To this end we assert a Gaussian covariance model for the high-resolution states of the form

$$\mathbf{x}_H = \overline{\mathbf{x}}_H + \mathbf{Z}\boldsymbol{\eta} \tag{2.1}$$

where $\overline{\mathbf{x}}_H$ is an N-vector, $\mathbf{Z}$ is the square-root of the true covariance matrix,

$$\mathbf{P}_H = \mathbf{Z}\mathbf{Z}^T \tag{2.2}$$

and $\boldsymbol{\eta}$ is an N-vector of random numbers drawn from $N(\mathbf{0},\mathbf{I})$. We construct (2.2) using a sinusoidal basis in which the columns of $\mathbf{E}_H$ ($N \times N$) contain the sinusoids such that

$$\mathbf{P}_H = \mathbf{E}_H \boldsymbol{\Gamma} \mathbf{E}_H^T \tag{2.3}$$

10.1016/j.ifacol.2016.10.161

$\boldsymbol{\Gamma}$ is a diagonal matrix whose i^{th} element of the diagonal determines the weight given to a particular basis function.

We connect the high-resolution (*N*-vector) states to the low-resolution (*M*-vector) states

$$\mathbf{x}_L = \mathbf{S}\mathbf{x}_H \tag{2.4}$$

through a "smoother" S ($M \times N$) that operates as:

$$\mathbf{S} = \mathbf{E}_L \begin{bmatrix} \mathbf{D}^{1/2}\mathbf{T} & \mathbf{0} \end{bmatrix} \mathbf{E}_H^T \tag{2.5}$$

where $\mathbf{D}$ ($M \times M$) is a diagonal matrix, $\mathbf{E}_L$ ($M \times M$) is the low-resolution basis whose columns are also the sinusoids, $\mathbf{T}$ ($M \times M$) is a diagonal matrix with the value $\sqrt{M/N}$ along the diagonal. As alluded to in the introduction, the equation (2.4) is meant to describe the relationship between the field variables (e.g. winds, temperature, etc.) described by our coarse spatial mesh model and the actual fine spatial mesh reality from which the observations are taken.

If we assume that the columns of $\mathbf{E}_L$ are simply the subsampled columns of $\mathbf{E}_H$ then the interpretation of (2.5) becomes straightforward. The matrix $\mathbf{D}$ represents the climatological "model error" on the resolved scales and would be equal to the identity matrix if the forecast model's climate at the resolved scales was identical to the true model's climate at those same scales. The matrix implied by the bracket in (2.5) performs a truncation of the high-resolution basis to the *M*-dimensional subspace while the matrix $\mathbf{T}$ assures that the Fourier coefficients calculated from the high-resolution basis are reweighted consistently with respect to the low-resolution basis.

Equation (2.4) allows for the creation of the low-resolution states from the high-resolution states in (2.1). This implies that the low-resolution error covariance matrix may be written as

$$\mathbf{P}_L = \mathbf{S}\mathbf{P}_H\mathbf{S}^T = \mathbf{E}_L \begin{bmatrix} \mathbf{D}^{1/2}\mathbf{T} & \mathbf{0} \end{bmatrix} \boldsymbol{\Gamma} \begin{bmatrix} \mathbf{D}^{1/2}\mathbf{T} & \mathbf{0} \end{bmatrix}^T \mathbf{E}_L^T . \tag{2.6}$$

Because $\boldsymbol{\Gamma}$ are the true, high-resolution eigenvalues, equation (2.6) shows that the forecast (low-resolution) covariance matrix would be correct up to its *M* eigenvalues if the climatological model error $\mathbf{D}$ could be removed. The data assimilation method that we describe next will remove this climatological model error from the state estimate by accounting for the error from the truncated state-space.

2.2 Data Assimilation

It is shown in Hodyss and Nichols (2015) that the best linear unbiased estimate for the problem setup here is

$$\overline{\mathbf{x}}_L^a = \overline{\mathbf{x}}_L + \mathbf{G}\left[\mathbf{v}_L - \langle \mathbf{v}_L \rangle\right] \tag{2.7}$$

where

$$\mathbf{G} = \left[\mathbf{P}_L\mathbf{H}_L^T + \mathbf{P}_{LH}\right]\left[\mathbf{H}_L\mathbf{P}_L\mathbf{H}_L^T + \overline{\mathbf{R}}_L^*\right]^{-1}, \tag{2.8}$$

$$\mathbf{P}_{LH} = \mathbf{S}\mathbf{P}_H\left(\mathbf{H}_H - \mathbf{H}_L\mathbf{S}\right)^T , \tag{2.9}$$

$$\overline{\mathbf{R}}_L^* = \mathbf{R}_{ins} + \mathbf{H}_H\mathbf{P}_H\mathbf{H}_H^T - \mathbf{H}_L\mathbf{P}_L\mathbf{H}_L^T , \tag{2.10}$$

$$\mathbf{v}_L = \mathbf{y} - \mathbf{H}_L\overline{\mathbf{x}}_L , \tag{2.11}$$

$$\langle \mathbf{v}_L \rangle = \mathbf{H}_H\overline{\mathbf{x}}_H - \mathbf{H}_L\overline{\mathbf{x}}_L . \tag{2.12}$$

In (2.8) through (2.12), $\mathbf{y}$ is a *p*-vector of observations, $\mathbf{H}_H$ is a $p \times N$ observation operator, $\overline{\mathbf{x}}_H$ is the high-resolution prior mean, $\overline{\mathbf{x}}_L$ is low-resolution prior mean, and $\mathbf{R}_{ins}$ is the instrument observation covariance matrix whose diagonal contains the instrument error variances.

3. METHODS AND APPLICATION

In this section we will first describe a standard way by which one invokes a truncation of the state vector and then move on to an application of the theory of section 2. We then describe an example problem where the truncation of the state vector leads to serious issues with the data assimilation. Application of the two methods described in this section will reveal the impact of not accounting for the truncated state space in the data assimilation.

3.1 A Contemporary Approach

We will approach the problem here assuming ensemble methods (Monte Carlo) are available. We assume however that for computational reasons we cannot perform data assimilation at the resolution of the true high-resolution (*N*-vector) state. We will however assume that we can run an ensemble at this resolution, but then must perform our data assimilation at a reduced resolution of length *M*. Using the low-resolution ensemble we may make the following state-estimate at low-resolution

$$\overline{\mathbf{x}}_L^c = \overline{\mathbf{x}}_L + \mathbf{G}_c\left[\mathbf{v}_L - \langle \mathbf{v}_L \rangle\right] \tag{3.1}$$

where

$$\mathbf{G}_c = \mathbf{P}_L\mathbf{H}_L^T\left[\mathbf{H}_L\mathbf{P}_L\mathbf{H}_L^T + \overline{\mathbf{R}}_c\right]^{-1} \tag{3.2}$$

$$\overline{\mathbf{R}}_c = \mathbf{R}_{ins} + \mathbf{R}_c \tag{3.3}$$

$$\mathbf{v}_L = \mathbf{y} - \mathbf{H}_L\overline{\mathbf{x}}_L \tag{3.4}$$

and $\mathbf{H}_L$ is a $p \times M$ observation operator. The low-resolution prior covariance matrix, $\mathbf{P}_L$, and the prior mean, $\overline{\mathbf{x}}_L$, are both calculated from the low-resolution ensemble using sample statistics. Here, the matrix $\mathbf{R}_c$ is simply a diagonal matrix and along that diagonal the elements are tuned to produce the