

## Deterministic predictability of the most probable state and reformulation of variational data assimilation

TSUYUKI, Tadashi<sup>1\*</sup>

<sup>1</sup>Meteorological Research Institute

Four-dimensional variational data assimilation (4DVar) and ensemble Kalman filter (EnKF) have been widely used for data assimilation in meteorology and oceanography. Since prior probability density functions (PDFs) used in 4DVar and EnKF are usually assumed to be Gaussian, those two methods may not work well for a strongly nonlinear system. Thus data assimilation with nonlinear systems or non-Gaussian PDFs is a challenge in geophysics. Theoretical study of the two methods for nonlinear systems may be expected to provide insight for further advancement of data assimilation. The present study addresses this issue by reformulating variational data assimilation.

A necessary condition for deterministic predictability may be that the forecast state starting from the most probable state at an initial time remains close to the most probable state at a forecast time. It is found from the Liouville equation that if the trace of the Jacobian matrix of a deterministic nonlinear system does not depend on the state variables (hereafter referred to as the trace condition), the mode of a PDF of the state variables evolves according to the governing equations of the system. A condition for the forecast state to be close to the mode of the PDF is derived under an assumption of small prediction error for general deterministic nonlinear systems. This condition depends on the sensitivity of the trace to an initial condition, the size of initial condition error, and the length of forecast lead time.

Since the dynamical cores of numerical models of the atmosphere and the ocean are based on fluid dynamics, it is interesting to examine whether the governing equations of fluid dynamics satisfy the trace condition. The trace of the Jacobian matrix is calculated for finite-dimensional analogs of several Eulerian equations of ideal fluids. It is readily found that the trace condition generally holds for unbounded fluids under periodic boundary conditions. It is shown that the trace condition also holds for the quasigeostrophic equations with rigid boundaries, the Boussinesq approximation with rigid boundaries, and the shallow water equations on a sphere, by expanding the state variables in eigenfunctions of the Laplacian operator or the curl operator. The shallow water equations in a channel and compressible fluid with top and bottom boundaries do not satisfy the trace condition due to divergence at the boundaries.

A new formulation of variational data assimilation is presented for deterministic nonlinear systems that satisfy the trace condition. Though the cost function in the new formulation takes the same form as the conventional one, it makes clear an advantage of 4DVar over EnKF. If the trace condition holds, the forecast state starting from the mode of the posterior PDF at the last analysis time is the mode of the prior PDF. In the new formulation, the logarithm of the prior PDF is expanded around the forecast state, and covariance globalization is introduced to take into account the global distribution of the prior PDF that may be non-Gaussian. A feasible method for the covariance globalization may be to replace the local covariance matrix at the mode with a forecast error covariance matrix taken from EnKF. It is proved that a non-Gaussian prior PDF that evolves according to the Liouville equation is implicitly used for assimilating observational data in 4DVar. Results from an assimilation experiment with a toy model suggest that 4DVar thus formulated outperforms EnKF if the global minimum of the cost function is found.

Finally, it is pointed out that enough observational data are necessary for variational data assimilation to work well for a deterministic nonlinear system that does not satisfy the trace condition. Otherwise, EnKF and ensemble prediction may be a better choice for data assimilation and prediction of the system.

Keywords: variational data assimilation, deterministic predictability, non-Gaussianity