A Unified Framework for Four-Dimensional Ensemble-Variational Hybrid Data Assimilation: Relationships among Ensemble Variational Algorithms with Full and Approximate Ensemble Covariance Localization

Chengsi Liu[1] and Ming Xue[1,2]
[1]Center for Analysis and Prediction of Storms and [2]School of Meteorology
University of Oklahoma, Norman Oklahoma 73072, U.S.A
cliu@ou.edu

Ensemble variational data assimilation (DA) algorithms that can incorporate the time dimension (four-dimensional or 4D) and combine static and ensemble-derived background error covariances (BECs) (hybrid) are formulated in general forms based on the extended control variable (ECV) and the observation-space-projection approaches. The properties and relationships of these algorithms and their approximated formations are discussed. The main algorithms discussed include: (1) the standard ensemble 4DVar (En4DVar) algorithm incorporating ensemble-derived BEC through the ECV approach, (2) the En4DVar neglecting the time-propagation of the extended control variable (En4DVar-NPC), (3) the 4D ensemble-variational algorithm based on observation space projection (4DEnVar), and (4) the 4DEnVar with no propagation of localization (4DEnVar-NPL). Without the static BEC term, none of the algorithms requires the adjoint model (AJM) except for En4DVar. Costly applications of the tangent linear model (TLM) to localized ensemble perturbations can be avoided by making the NPC and NPL approximations. It is proven that En4DVar and 4DEnVar are mathematically equivalent, while En4DVar-NPC and 4DEnVar-NPL are mathematically equivalent. Such equivalences are also demonstrated by single-observation assimilation experiments with a 1-D linear advection model. The effects of the non-flow-following localization approximations are also examined through the experiments.

All of the above algorithms can include the static BEC term to establish a hybrid formulation. When the static term is included, all algorithms will require TLM and AJM. The ‘first guess at appropriate time (FGAT)’ approximation is proposed to avoid the requirement for TLM and AJM. Computational costs of the algorithms are also discussed.