Geophysical Research Abstracts
Vol. 14, EGU2012-9892, 2012
EGU General Assembly 2012

Why perturbing observations in data assimilation is incorrect.

P..J. van Leeuwen

University of Reading, DARC, Department of Meteorology, Reading, United Kingdom (p.j.vanleeuwen@reading.ac.uk, +441183788905)

Several popular Gaussian ensemble data-assimilation schemes are available to the geosciences community that seem to perturb observations. Examples are the 'Perturbed observations Ensemble Kalman Filter', 'Randomised Maximum Likelihood', and recently 'Perturbed observation ensemble incremental 4DVar' methods. It is easy to show that these schemes generate posterior ensembles that do have the correct posterior covariance in the limit of an infinite number of ensemble members. Some argue that observations should always be perturbed in data assimilation, because one has to perturb 'anything' when using ensemble methods.

However, no fundamental justification for perturbing observations is given other than by doing so one generates an ensemble with the correct covariance statistics. Furthermore, the notion of perturbing observations is inconsistent with Bayes Theorem, which states that the posterior probability density function is conditioned on the specific set of observations present, not perturbed ones. Also, square-root ensemble Kalman filters generate similarly consistent posterior ensembles without perturbing observations. In this presentation this inconsistency will be solved.

All schemes generate ensembles from the posterior pdf, using an ensemble from the prior (this is done implicitly in 4DVar implementations). In square-root Ensemble Kalman filters the symmetric square root of the posterior covariance is used to find a transformation equation from prior to posterior ensemble. It turns out that 'perturbed observation' schemes also use a square root of the posterior, but an asymmetric square root. The square root is of size n x (n+m) in which n is the dimension of the state vector, and m is the number of observations (also the 4DVar scheme can be written in this form). So, the 'perturbed observation' ensemble methods are in fact asymmetric square root implementations, and the interpretation of perturbing observations is misleading, if not incorrect.

These results allow us to unify all Gaussian ensemble data-assimilation schemes, including incremental 4DVar, in one single framework. Furthermore, while there is one unique positive definite symmetric square root of the posterior covariance, there is an infinite number of asymmetric positive definite square roots. This allows us to generate a whole new family of Gaussian ensemble data-assimilation schemes, and a symmetric square-root ensemble incremental 4DVar scheme.