Hybrid EnVar

Because the error covariance estimated from limited sized ensemble can be unreliable, and in particular is rank deficient, one possible way is to linearly combine static background error covariance **B** typically used in 3DVar and ensemble derived covariance **P**, forming the so-called hybrid covariance.

Hamill and Snyder (2000) first suggested this approach, where the forecast error covariance **B** in 3DVar is replaced by a linear combination of the (static) 3D-Var covariance $\mathbf{B}_{3DVAR}$

$$\mathbf{P}_l^{f\,(hybrid)} = (1-\alpha)\mathbf{P}_l^f + \alpha\mathbf{B}_{3DVAR}.$$

α is a tunable parameter that varies from 0, corresponding to using 100% ensemble covariances, to 1 corresponding to 100% static **B**. pure 3D-Var.

For a small enough problem, one can simply define **P** according to the above equation, and use it within a 3DVar cost function. As discussed earlier, for practical NWP problems, this **P** is usually too large to define and use explicitly. More computational efficient procedure is needed.

Hamill and Snyder (2000) tested the hybrid En3DVar with a low-resolution quasi-geostrophic model and simulated data in a perfect model setting. By running the hybrid analysis system multiple times with perturbed observations, the system is able to provide an ensemble of analyses. It was found that the analysis performs the best when BEC is estimated almost fully from the ensemble, especially when the ensemble size was large (100 in their case). When the ensemble is smaller, the system benefits from a lesser weighting given to the ensemble-based covariances.

Lorenc (2003) proposed an elegant, alternative hybrid formulation, in which the control variables of the regular variational cost function are augmented by extended control variables (hereafter, ECV), which are preconditioned upon the square root of ensemble covariance. The ECV formulation involves adding an additional term to the variational cost function for the ECVs which has a similar form as the original background term, and is therefore relatively easy to implement based on an existing variational DA framework. Wang et al. (2007)

proved that the ECV formulation is mathematically equivalent to that of Hamill and Snyder (2000).

The potential for the hybrid system to perform better than a pure EnKF when the ensemble size is relatively small makes it attractive for operational implementation where computational constraint is often a significant issue. A variational framework used by the hybrid scheme also makes it easier to include additional equation constraints in the cost function.

Furthermore, for observations whose forward operators are non-local, such as those of satellite radiance data, the state-space-based covariance localization used in the hybrid formulation is potentially advantageous (Campbell et al. 2010).

As suggested by Lorenc (2003), Buehner et al. (2010a, b), both (traditional) 3DVar and 4DVar can be formulated to use the ensemble covariance with the extended control variable method, and such ensemble-variational formulations are called 3DEnVar and 4DEnVar, respectively, or EnVar in general.

Based on Lorenc (2003) and Wang et al (2007), the analysis increment $\delta\mathbf{x}$ is a sum of two terms, defined as

$$\delta\mathbf{x} = \delta\mathbf{x}_1 + \sum_{k=1}^{K} (\mathbf{a}_k \circ \mathbf{x}_k^{'}) \quad , \qquad (1)$$

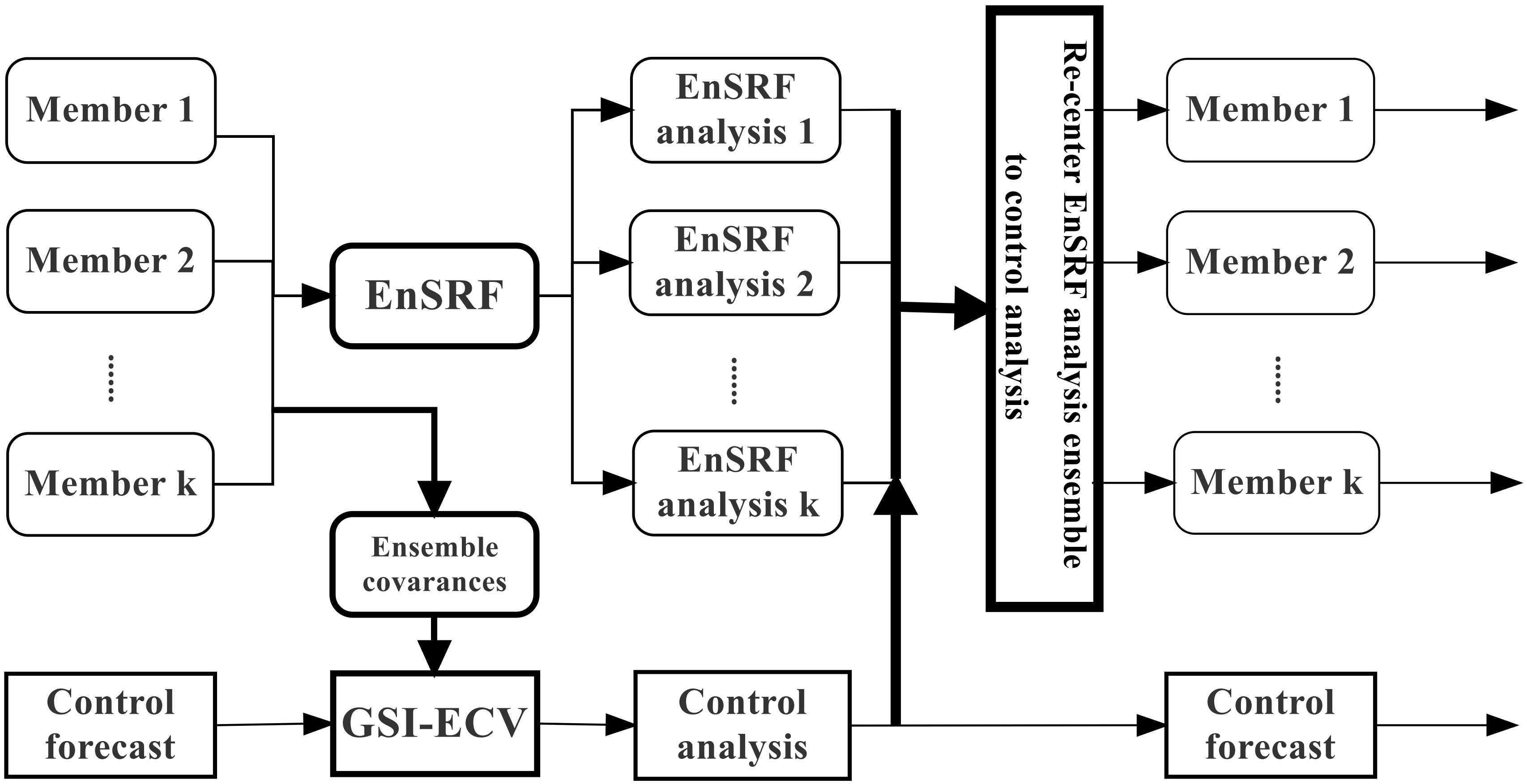
where $\delta\mathbf{x}_1$ is the analysis increment associated with static BEC $\mathbf{B}$ and the second term on the right hand side is the increment associated with the ensemble covariance. $\mathbf{x}_k^{'}$ is the k^{th} ensemble background perturbation normalized by $\sqrt{K-1}$, where K is ensemble size. Vectors $\mathbf{a}_k\,(k = 1,\cdots,K)$ in the second term are the extended control variables. Analysis increment $\delta\mathbf{x}$ is obtained by minimizing the following cost function:

$$\begin{aligned} J(\delta\mathbf{x}_1,\mathbf{a}) &= \beta_1 J_b + \beta_2 J_e + J_o \\ &= \frac{1}{2}\beta_1 \delta\mathbf{x}_1^T \mathbf{B}^{-1}\delta\mathbf{x}_1 + \frac{1}{2}\beta_2 \mathbf{a}^T \mathbf{A}^{-1}\mathbf{a} + \frac{1}{2}[\mathbf{y}_o - H(\mathbf{x}_b + \delta\mathbf{x})]^T \mathbf{R}^{-1}[\mathbf{y}_o - H(\mathbf{x}_b + \delta\mathbf{x})], \end{aligned} \qquad (2)$$

which gives the solutions of partial increment $\delta\mathbf{x}_1$ and ECV $\mathbf{a}$. Vector $\mathbf{a}$ is formed by concatenating K vectors $\mathbf{a}_k$. Compared to a traditional 3DVar cost function, a weighted sum of J_b and J_o is replaced by the sum of weighted J_b and J_e terms and J_o, where J_b is the traditional background term associated with static covariance $\mathbf{B}$, J_o is the observation term as in traditional 3DVar. J_e is the additional term associated with flow-dependent covariance for the ECV. Weighting factors β_1 and β_2 are placed in front of J_b and J_e terms, respectively, and they are constrained by

$$\frac{1}{\beta_1} + \frac{1}{\beta_2} = 1, \tag{3}$$

to conserve the total variances in the hybrid implementation.



Flowchart of a full EnSRF-En3DVar hybrid data assimilation cycle, with one-way or two-way coupling between the EnSRF (upper portion) and En3DVar hybrid control analysis (lower portion denoted En3DVar). The thick upward pointing arrow indicates the feedback of the En3DVar hybrid analysis to the EnSRF in the two-way coupling procedure, when the En3DVar hybrid control analysis is used to replace the ensemble mean of the EnSRF analyses.

The NCEP GFS system currently employs a hybrid 4DEnVar DA system, coupled with an EnKF system using the EnSRF algorithm. The operational RAP and NAM systems also employ a hybrid algorithm, borrowing ensemble perturbations from the global EnKF system. Significant forecast improvement was achieved with GFS when moving from the original 3DVar algorithm to the EnVar algorithm.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon. Wea. Rev.*, **138,** 1550-1566.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.*, **138,** 1567-1586.

Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138,** 282-290.

Hamill, T. M. and C. Snyder, 2000: A hybrid ensemble Kalman filter - 3D variational analysis scheme. *Mon. Wea. Rev.*, **128,** 2905-2919.

Lorenc, A., 2003: The potential of the ensemble Kalman filter for NWP - a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129,** 3183-3204.

Wang, X., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble/VAR hybrid analysis schemes. *Mon. Wea. Rev.*, **135,** 222–227.