

1           **Hybrid Data Assimilation without Ensemble Filtering**

2                   RICARDO TODLING\* AND AMAL EL AKKRAOUI<sup>†</sup>

*Global Modeling and Assimilation Office, NASA/GSFC, Greenbelt, Maryland*

---

\* *Corresponding author address:* Dr. Ricardo Todling, Global Modeling and Assimilation Office,  
NASA/GSFC, Code 610.1, Greenbelt, MD 20771.

E-mail: ricardo.todling@nasa.gov

<sup>†</sup> *Additional Affiliation:* Science Systems and Applications, Inc., Lanham, MD 20706.

## ABSTRACT

4 The Global Modeling and Assimilation Office is preparing to upgrade its three-dimensional  
5 variational system to a hybrid approach in which the ensemble is generated using a square-  
6 root ensemble Kalman filter (EnKF) and the variational problem is solved using the Grid-  
7 point Statistical Interpolation system. As in most EnKF applications, we found it necessary  
8 to employ a combination of multiplicative and additive inflations, to compensate for sampling  
9 and modeling errors, respectively and, to maintain the small-member ensemble solution  
10 close to the variational solution, we also found it necessary to re-center the members of  
11 the ensemble about the variational analysis. During tuning of the filter we have found re-  
12 centering and additive inflation to play a considerably larger role than expected, particularly  
13 in a dual-resolution context when the variational analysis is ran at larger resolution than the  
14 ensemble. This led us to consider a hybrid strategy in which the members of the ensemble  
15 are generated by simply converting the variational analysis to the resolution of the ensemble  
16 and applying additive inflation, thus bypassing the EnKF. Comparisons of this, so-called,  
17 filter-free hybrid procedure with an EnKF-based hybrid procedure and a control non-hybrid,  
18 traditional, scheme show both hybrid strategies to provide equality significant improvement  
19 over the control; more interestingly, the filter-free procedure was found to give qualitatively  
20 similar results to the EnKF-based procedure.

# 1. Introduction

It is now generally accepted that a practical feasible way to introduce flow dependence in the background error covariances needed for either sequential or variational data assimilation procedures is to rely on an ensemble of short-range forecasts. Multiple works have now shown (Whitaker et al. 2008, Buehner et al. 2010, and Clayton et al. 2012) that combining the time-varying background error covariance derived from an ensemble of forecasts with the typical, stationary, climatological background error covariance leads to non-trivial improvements to the resulting, so-called, hybrid data assimilation system (Lorenc 2003). Most operational weather centers use three- or four-dimensional variational (3D/4DVar) techniques and have implemented hybrid approaches in these contexts. With the variational component capable of accepting hybrid formulations of its underlying background error covariance, what remains to be specified is a methodology to generate the required ensemble of forecasts. Presently, the Global Modeling and Assimilation Office, follows the National Centers for Environmental Predictions, and uses the square-root-based ensemble Kalman filter (EnKF; Whitaker et al. 2008) for this purpose. The small number of ensemble members used in practice requires care to render adequate spread from the ensemble of forecasts to represent forecast errors. It is thus necessary to fiddle with the ensemble of analyses and: (i) apply multiplicative inflation to compensate for sampling errors; (ii) apply additive inflation to represent model uncertainties; and (iii) re-center the ensemble of analyses around the, hybrid, variational analysis to prevent possible divergence between the two assimilation systems.

During the process of implementation and testing of the EnKF to provide initial conditions for the ensemble of forecasts for a hybrid strategy to be adopted for the Goddard Earth

43 Observing System (GEOS) atmospheric data assimilation system (ADAS), we have found  
44 steps (ii) and (iii) above to play a significant role in determining the behavior of the ensemble  
45 of forecasts. This is particularly noticeable when the ensemble and the (hybrid) variational  
46 analyses are produced at different resolutions in a, so-called, dual resolution approach. That  
47 re-centering and additive inflation are such key components of the hybrid strategy is illus-  
48 trated in Fig. 1, where the incremental contribution to the 500 hPa temperature field is  
49 shown for an arbitrarily selected member of the ensemble, at an arbitrarily selected time,  
50 after the EnKF has cycled beyond a spin up period. The panels in the figure correspond  
51 to increments at various stages in the ensemble analysis procedure: directly from the EnKF  
52 (top left), when only multiplicative inflation has been applied; when the EnKF increment  
53 is re-centered around the (hybrid) variational (higher resolution) analysis (top right); when  
54 applying additive inflation to the EnKF increment (bottom left); and when multiplicative  
55 inflation, additive inflation, and re-centering have been applied to form the total increment  
56 (bottom right). Re-centering is clearly a larger contributor to the total increment. Still,  
57 the main features in the increment obtained from the EnKF assimilation of observations are  
58 visibly identified after re-centering and additive inflation have taken place. At first, these  
59 results might suggest the EnKF to be poorly tuned, however, as we will show later, this is  
60 far from being the case. One key factor is that the EnKF analyses are at coarser resolution  
61 than the (hybrid) variational analysis used for re-centering; when the ensemble is at full  
62 resolution, the contribution from re-centering is much lesser (not shown).

63 The crucial role played by steps (ii) and (iii) prompted us to investigate what would  
64 happen if we bypassed the EnKF step altogether. This led us to the, so-called, filter-free

65 ensemble scheme when ensemble analyses are generated by simply adding perturbations to  
66 the central, hybrid, variational analysis – that is, steps (ii) and (iii) are what constitute  
67 the ensemble analysis strategy. The additive perturbations used in this procedure corre-  
68 spond to samples of the scaled, 48-minus-24-hour forecast differences, similar to those used  
69 to generate the climatological background error covariance of the traditional assimilation  
70 approach; these are also the perturbations used when the EnKF is exercised. The remaining  
71 of this manuscript presents a comparison of results obtained from dual-resolution hybrid  
72 3DVar procedures when either the EnKF or the filter-free approach is used for the ensemble  
73 analysis generation.

## 74 2. Brief overview and the filter-free strategy

75 The basic idea of hybrid variational data assimilation is to use an ensemble of background  
76 fields to introduce instantaneous, flow-dependent, features to the traditionally non-evolving  
77 (static) background error covariance. In 3DVar this can be done by augmenting the control  
78 vector with an extra set of variables, usually referred to as alpha-control variables. The cost  
79 function of a hybrid incremental 3DVar system can be written as

$$J(\delta\mathbf{z}) = \frac{1}{2}\delta\mathbf{z}^T [\beta_s\mathbf{B}_s + \beta_e\mathbf{T}^T(\mathbf{B}_e \circ \mathbf{L})\mathbf{T}]^{-1} \delta\mathbf{z} + \frac{1}{2}(\mathbf{d} - \mathbf{H}\delta\mathbf{z})^T \mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\delta\mathbf{z}), \quad (1)$$

80 where the control variable  $\delta\mathbf{z}$  is a combined contribution from the  $n$ -vector solution  $\delta\mathbf{x}$  of the  
81 standard variational problem and a component that comes from an  $M$ -member ensemble,  
82 that is,

$$\delta\mathbf{z} = \beta_s\delta\mathbf{x} + \beta_e\mathbf{T}^T \sum_{m=1}^M \boldsymbol{\alpha}_m \circ \Delta\mathbf{w}_m^e. \quad (2)$$

83 Here, the symbol  $\circ$  stands for the Hadamard-Schur (element-wise) product of two vectors,  
 84  $\boldsymbol{\alpha}_m$  is the  $m$ -th control vector related to the  $m$ -th ensemble member, and, using the symbol  
 85  $\Delta$  to denote deviation from the mean,  $\Delta\mathbf{w}_m^e = (\mathbf{w}_m^b - \bar{\mathbf{w}})/\sqrt{M-1}$  is the  $m$ -th ensemble  
 86 perturbation created from the  $m$ -th member background  $n_w$ -vector state  $\mathbf{w}_m^b$ , with respect to  
 87 the ensemble mean  $\bar{\mathbf{w}}^b$ . The formulation allows for the ensemble members to be of different  
 88 (usually lower) resolution, than the primary  $n$ -vector control  $\delta\mathbf{x}$ , with the operator  $\mathbf{T}^T$  being  
 89 responsible for resolution conversion. In (1), the matrices  $\mathbf{B}_s$  and  $\mathbf{B}_e$  stand for the static and  
 90 ensemble background error covariances, respectively; the matrix  $\mathbf{L}$  stands for a correlation  
 91 matrix responsible for localization of the ensemble; the last term is the usual observation-fit  
 92 term involving the observation error covariance matrix  $\mathbf{R}$ , and the observation residual  $p$ -  
 93 vector  $\mathbf{d} = \mathbf{y} - \mathbf{h}(\mathbf{x}^g)$  created from differencing the observation  $p$ -vector  $\mathbf{y}$  with the projection  
 94 of the first-guess state-vector  $\mathbf{x}^g$  onto observation space by the observation operator  $\mathbf{h}$ , whose  
 95 linearization is represented by the matrix  $\mathbf{H}$ . The parameters  $\beta_s$  and  $\beta_e$  specify the interplay  
 96 between the static and the ensemble background error covariances, respectively. The problem  
 97 is reset to its traditional 3DVar configuration, with solution  $\delta\mathbf{x}$ , when  $\beta_s = 1$  and  $\beta_e = 0$ .  
 98 Details of the hybrid variational problem can be found in Hamill and Snyder (2000), Lorenc  
 99 (2003) and Wang et al. (2007).

100 The *first* hybrid implementation studied in the present work relies on the ensemble square-  
 101 root Kalman filter formulation of Whitaker and Hamill (2002). Each 6-hours the ensemble

102 analysis updates the ensemble mean and its members through the sequence

$$\bar{\mathbf{w}}^a = \bar{\mathbf{w}}^b + \sum_{j=1}^p \mathbf{k}_j [y_j - h_j(\bar{\mathbf{w}}^b)] \quad (3a)$$

$$\Delta \mathbf{w}_m^a = \Delta \mathbf{w}_m^b - \sum_{j=1}^p \mathbf{k}_j \gamma_j \delta h_{m;j}, \quad (3b)$$

103 where  $y_j$  is the  $j$ -th observation,  $\delta h_{m;j}$  is the  $j$ -th element of the incremental factor  $\delta \mathbf{h}_m \equiv$

104  $\mathbf{H} \Delta \mathbf{w}_m \approx \mathbf{h}(\mathbf{w}_m^b) - \mathbf{h}(\bar{\mathbf{w}}^b)$  resulting from the fact that observations are not perturbed in this

105 formulation, and the  $n_w$ -vector  $\mathbf{k}_j$  is the  $j$ -th column of the gain matrix,  $\mathbf{K}$ , and is given by

$$\mathbf{k}_j = \frac{1}{M-1} \sum_{m=1}^M \Delta \mathbf{w}_m^{j-1} \delta h_{m;j} / \sigma_j^2 \quad (4a)$$

$$\Delta \mathbf{w}_m^j = \Delta \mathbf{w}_m^{j-1} - \mathbf{k}_j \gamma_j \delta h_{m;j} \quad (4b)$$

106 for  $j = 1, 2, \dots, p$ ,  $\Delta \mathbf{w}_m^0 \equiv \Delta \mathbf{w}_m^b$ , and scalar coefficients  $\sigma_j^2$  and  $\gamma_j$  given by

$$\sigma_j^2 \equiv \frac{1}{M-1} \sum_{m=1}^M (\delta h_{m;j})^2 + (\sigma_j^o)^2, \quad (5)$$

$$\gamma_j \equiv 1 / \left[ \sqrt{M-1} (1 + \sigma_j^o / \sigma_j) \right], \quad (6)$$

107 Here only the diagonal elements  $(\sigma^o)_j^2 \equiv (\mathbf{R})_{jj}$  of the observation error covariance are referred

108 to, given that observation errors are assumed to be uncorrelated thus allowing observations

109 to be processed serially (e.g., Houtekamer and Mitchell 2001); the algorithm above is a direct

110 application of the expressions in Appendix II.E of Bierman (1977) for when the square-root

111 of the background error covariance is made up of column vectors  $\Delta \mathbf{w}_m^b$ , for  $m = 1, 2, \dots, M$ .

112 After all  $p$  observations are processed,  $\Delta \mathbf{w}_m^p = \Delta \mathbf{w}_m^a$ , which is obtained by a backward recur-

113 sion of (4b) from  $j = p$  to  $j = 1$  to obtain (3b). Just as when solving the variational hybrid

114 problem, localization is also needed and used in the square-root Kalman filter formulation

115 of Whitaker and Hamill (2002), though it is left out of the equations above for the sake of

116 notational simplicity.

117 The final ensemble of analyses, ultimately used to serve as initial conditions for the  
 118 ensemble of forecasts, are typically re-centered around the variational analysis and inflated  
 119 by scaled perturbations  $\epsilon_m$ . That is, the  $m$ -th member final analysis is given by

$$\mathbf{w}_m^a := \mathbf{w}_m^a - \bar{\mathbf{w}}^a + \mathbf{T}\mathbf{x}^a + \mu\epsilon_m, \quad (7)$$

120 where the parameter  $\mu$  specifies the magnitude of the additive perturbation, and ideally, the  
 121 operator  $\mathbf{T}$  converting the high-resolution variational analysis onto the  $n_w$ -dimensional space  
 122 of the ensemble satisfies the relation  $\mathbf{T}\mathbf{T}^T = \mathbf{I}_{n_w}$ , though presently in our implementation  
 123 this is not the case. Note that, in the application to GEOS ADAS, the operator  $\mathbf{T}$  involves  
 124 remapping of the central analysis to the topography of each member. Re-centering prevents  
 125 the ensemble from steering far from the (hybrid) variational analyses, and additive inflation  
 126 is one way of boosting error growth (e.g., Mitchell et al. 2002, Houtekamer et al. 2005, and  
 127 Hamill and Whitaker 2005).

128 The *second* hybrid strategy examined in the present work relies on the “filter-free” pro-  
 129 cedure, constructed by simply replacing expression (7) with

$$\mathbf{w}_m^a = \mathbf{T}\mathbf{x}^a + \alpha\epsilon_m, \quad (8)$$

130 completely removing the EnKF component from the cycle. By construction, the mean en-  
 131 semble analysis equals the variational (hybrid) analysis, aside from differences in resolution.  
 132 Notice that both strategies (7) and (8) employ the same additive perturbation  $\epsilon_m$ , which in  
 133 practice means pooling from the same database on 48-minus-24-hour forecast NMC-method-  
 134 like differences.



### 3. GEOS ADAS 3DVar Ensemble Hybrid

In GOES ADAS the variational problem of minimizing (1) is solved using the Grid-point Statistical Interpolation (GSI; Kleist et al. 2009a) analysis and the preconditioning formulation of (Derber and Rosati 1989). The static background error covariance matrix is implemented as a series of recursive filters producing nearly Gaussian and isotropic correlation functions following Wu et al. (2002), and tuned from GEOS forecasts (Wei Gu contribution in Rienecker et al. 2008); the hybrid background error covariance matrix uses an ensemble of GEOS background fields in a hybrid-capable GSI (David F. Parrish, personal communication). Satellite radiances are processed using the Community Radiative Transfer Model (CRTM; Kleespies et al. 2004) and the online variational bias-correction procedure of Derber and Wu (1998). A normal-mode-based balance constraint term following Kleist et al. (2009b) is applied to the static increment as well as to the ensemble part of the increment whenever the hybrid analysis is used.

The ensemble hybrid-capable GEOS ADAS relies on the GEOS global atmospheric general circulation model (AGCM), developed at NASA/Goddard. The GEOS AGCM is built under the infrastructure of the Earth System Modeling Framework (ESMF; Collins et al. 2005) and couples a cubed-sphere hydrodynamics (Putman and Lin 2007) with various physics packages including a modified version of the Relaxed Arakawa-Schubert convective parameterization scheme of Moorthi and Suarez (1992), the catchment-based hydrological model of Koster et al. (2000), the multi-layer snow model of Stieglitz et al. (2001), and the radiative transfer model of Chou and Suarez (1999), which uses interactive climatological aerosols from the Goddard Global Ozone Chemistry Aerosol Radiation and Transport

157 (GOCART; Collarco et al. 2010) package.

158 In GEOS ADAS, assimilation is performed using the incremental analysis update (IAU)  
159 procedure of Bloom et al. (1996). A schematic representation of standard IAU appears in the  
160 top panel of Fig. 2. Considering for example the availability of observations around 00 UTC  
161 and of three-hourly AGCM background fields, the GSI analysis (purple boxes) produces an  
162 increment that is converted into a tendency and used to force a 6-hour (corrector) model  
163 integration (red triangles); this is followed by a 6-hour (predictor) integration period when  
164 the model is then set to run free from the analysis forcing as to produce backgrounds (green,  
165 upside-down, triangles) for the next assimilation cycle; the prediction period can be extended  
166 beyond 6-hours to complete, say, a 5-day forecast (horizontal orange-dashed lines). The cycle  
167 of running GSI and AGCM takes place whether GEOS ADAS is performing its traditional  
168 3DVar procedure or its hybrid extension. The only difference between these two options is  
169 that in the latter case, an ensemble of background fields is required for GSI to internally  
170 augment its background error covariance information, through (1). Hereafter, this cycle  
171 will be referred to as the *central* ADAS. It usually operates at a higher resolution than the  
172 ensemble ADAS (see below).

173 Generation of the ensemble of background fields to make up the ensemble background  
174 error covariance  $\mathbf{B}_e$  involves AGCM integrations similar to those of the central ADAS, but  
175 generally carried at lower resolution. In turn, the ensemble of backgrounds requires an  
176 ensemble of “initial conditions” (analyses) to be available. At least three options exist within  
177 GEOS ADAS to generate an ensemble of analyses. The standard option follows Whitaker et  
178 al. (2008), as described earlier, and relies on the ensemble Kalman filter (EnKF) software of

179 J. S. Whitaker, from NOAA/ESRL. This is the same software presently used in the NCEP  
180 operational global data assimilation system. Alternatively, one can generate an ensemble  
181 of GSI analyses, but this is considerably more computationally demanding than using the  
182 EnKF since it involves a complete variational analysis for each member of the ensemble. And  
183 lastly, an option to exercise the filter-free ensemble analysis is also available. Regardless of the  
184 ensemble of analyses scheme, once analyses are available, a corresponding set of background  
185 fields is generated through IAU-based AGCM integrations, similar to those of the central  
186 ADAS. The IAU-based ensemble procedure is illustrated in the bottom panel of Fig. 2.  
187 Availability of observations and an ensemble of backgrounds triggers one of the ensemble  
188 analysis options (EnAna; right-placed, purple boxes), including re-centering and additive  
189 inflation, generating an ensemble of analyses which are then turned into an ensemble of  
190 tendencies used to initialize the ensemble of AGCM integrations — forced during the first  
191 6 hours (light-red triangles), and unforced during the 6-hour background prediction period  
192 (light-green, upside-down triangles).

193 There is a subtle difference to note related to how the GEOS ADAS IAU-based ensemble  
194 evolves its members when the EnKF is used versus when the filter-free strategy is used  
195 instead. With the EnKF, each member permanently cycles its corresponding set of initial  
196 conditions needed by the GEOS AGCM each cycle. With the filter-free strategy, the initial  
197 conditions for the ensemble of AGCM integrations are generated by simply converting the  
198 (high-resolution) initial conditions from the central (hybrid) cycle to the configuration of  
199 the ensemble; namely, at each cycle, all members start from the exact same set of initial  
200 conditions; the only thing making these integrations distinct is the corresponding IAU forcing

201 term used by each member, each derived from the ensemble analysis equation (8).

## 202 **4. Evaluation of hybrid strategies in GEOS ADAS**

203 In what follows, we present a discussion of results obtained for experiments from single  
204 analysis as well as fully cycled ADAS. Regular, non-hybrid, 3DVar results are compared with  
205 results from hybrid 3DVar analyses produced at 0.5-degree resolution on 72 vertical levels  
206 and relying on a 32-member, 1-degree, 72-level ensemble generated by either the EnKF or  
207 the filter-free procedure described above.

### 208 *a. Non-cycling hybrid analysis*

209 When an ensemble of backgrounds is used in a hybrid GSI analysis, one of the first  
210 things we examined was how the analysis increment changed with respect to its non-hybrid  
211 counterpart. Figure 3 provides an illustration for the change in analysis increment, measured  
212 in total energy units, for an analysis calculated at a single synoptic time using: (i) regular  
213 3DVar, with only the static background error covariance (left); (ii) 3DVar with a background  
214 error covariance matrix that is fully determined by the 32-member ensemble (center); and (iii)  
215 3DVar hybrid, when 50% of background error covariance matrix comes from the ensemble  
216 and the remaining 50% comes from its regular static background error covariance matrix  
217 (right). The ensemble-only case (center) shows considerably more activity in the tropics  
218 than when compared with the static-only case (left); the resulting hybrid (right) increment  
219 shows slight, but noticeable, energy increase in the mid-tropospheric and low-stratospheric

220 levels — a little less energy seems to be present along the Southern tropospheric jet in the  
221 ensemble (center) when compared to the static case (left), with the resulting hybrid retaining  
222 the energy in this region (right).

223 Another aspect of relevance when introducing upgrading to hybrid analyses relates to  
224 how balance gets affected. In its 3DVar configuration, GSI has the capability of applying a  
225 tangent linear normal mode constraint (TLNMC) to the increment (see Kleist et al. 2009b).  
226 The constraint can be applied to either part of the increment (essentially to either of the  
227 two terms in eq. 2, or both; see Kleist 2012). Figure 4 shows two illustrations of the result  
228 of balancing the increment in various configurations of GSI. The panel on the left shows the  
229 total cost function during the iterations of the GSI minimization when using: traditional  
230 3DVar without TLNMC (black curve); traditional 3DVar with TLNMC (red curve); hybrid  
231 3DVar with TLNMC applied only to the static part of increment (green); and hybrid 3DVar  
232 when TLNMC is applied to the full increment. The behavior is typical of when adding  
233 constraints to the analysis, that is, with balance, the cost settles a little higher than when  
234 no constraint is applied. The hybrid minimization tends to reduce the cost when compared  
235 to the static-balanced configuration; particularly noticeable in the first outer minimization  
236 (first 100 iterations; compare green and blue curves with red curve, respectively). This  
237 is indication that the hybrid minimization recovers the fit to the observations somewhat  
238 deteriorated when the constraint is added to traditional 3DVar.

239 The real measure of improved balance is displayed in the right panel of Fig. 4 where  
240 the spectra of the vertically integrated mass-wind divergence increment is shown for the  
241 same four configurations. The color scheme is preserved, and the curves show clearly that

242 TLNMC brings in considerable improvement in balance when applied to traditional 3DVar  
243 (compare black and red curves). It is also clear from the figure that applying TLNMC only  
244 to the static part of the increment when hybrid 3DVar is used is rather troublesome (green  
245 curve). This is natural since nothing guarantees the ensemble contribution to the increment,  
246 through its background error covariance matrix  $\mathbf{B}_e$ , to be balanced in any way; TLNMC  
247 must be applied to the full increment (blue curve) for balance to be acceptable in the hybrid  
248 configuration. However, this latter case is not completely perfect since some power in the  
249 spectrum still remains for large wave numbers which would best be reduced. As pointed  
250 out by Kleist (2012; see Figure 4.2 on page 108, in that work), this is a consequence of the  
251 dual-resolution aspect of the hybrid analysis and some aliasing of the winds. It is possible to  
252 use scale-dependent weights to reduce some of the aliasing issue (see Kleist 2012, Fig. 4.4, in  
253 that work), but this is part of future work. At present, the default in GEOS hybrid ADAS  
254 is to apply TLNMC to the full increment.

255 The remaining illustrations in this section summarize results and comparisons from three  
256 experiments covering the month of April 2012. The abbreviations and brief explanation of  
257 each experiment follows:

- 258 • Control (CTL): traditional 3DVar, similar to what is used by GMAO Operations,  
259 though experiments here are at, coarser, 0.5-degree resolution.
- 260 • Hybrid (HY5): Dual-resolution hybrid ADAS using 50% static and 50% ensemble  
261 background error covariance contributions, with an ensemble of analyses generated by  
262 the EnKF.
- 263 • Hybrid (HYA): similar to HY5, but using the filter-free procedure, that is, at each

264 cycle, an ensemble of analyses is generated by adding scaled NMC-like perturbations  
265 to the hybrid (central) variational analysis.

266 Evaluation of results of these experiments examine familiar diagnostics: observation-minus-  
267 analysis (OMA), observation-minus-background (OMB), and observation-minus-forecast (OMF)  
268 residual statistics, monthly mean comparison with corresponding means from other numer-  
269 ical weather prediction (NWP) centers, and forecast skills scores. Additionally, ensemble-  
270 related diagnostics have also been examined to evaluate the performance of the ensemble  
271 itself. These included monthly-mean of the ensemble mean analyses and/or backgrounds,  
272 OMA, OMB and OMF residual statistics for the mean and ensemble members, and also time  
273 evolution of ensemble spread. Rank histograms (of say, OMB residuals) have been looked  
274 but we have found them to be rather difficult to interpret given the uncertainties associated  
275 with the observations (see Hamill 2001), therefore we refrain from discussing them here.

276 *b. About the ensemble itself*

277 We have seen in Fig. 1 how much re-centering and additive inflation participate to modify  
278 the analysis increments calculated by the EnKF. In addition to what was said earlier, we  
279 should point out that we have found re-centering and additive inflation to be necessary within  
280 the context of the small-size ensemble GEOS hybrid ADAS. Without re-centering the EnKF  
281 analyses were found to diverge from the central hybrid analysis; without additive inflation the  
282 ensemble was found to collapse rather quickly. Furthermore, finding the scaling parameter  
283  $\alpha$  multiplying the additive inflation term requires careful tuning. We have found a value of  
284 0.25 to be rather reasonable when the EnKF is used. This is considerably lower than value

285 of 0.40 presently used in the NCEP hybrid 3DVar (Daryl Kleist, pers. comm.). However,  
286 when using the filter-free approach, the value of 0.40 was found to be more adequate.

287 In a cycling situation, the interplay between re-centering and inflation must lead to  
288 reasonable forecast spread. Figure 5 illustrates the time evolution of the global (largely tro-  
289 pospheric) spread of a 32-member ensemble for typical experiments performed with GEOS  
290 hybrid ADAS. The panel on the left uses the EnKF for its ensemble analysis and shows how  
291 the initial spread (blue curve) changes as the members evolve within the 9-hour background  
292 period (green, red, and black for the 3-, 6- and 9-hour backgrounds, respectively). The  
293 resulting hybrid ADAS performs rather well (see below), even when there is not much error  
294 growth within the 9-hour background period — note the green, red and black curves are very  
295 close to each other; however, the growth of error is consistent within the same period, with  
296 the smallest error seen in the 3-hr background and the largest in the 9-hour background.  
297 The panel on the right shows similar forecast spread for various times within the background  
298 period, but now when the filter-free approach is used to generate the ensemble of analyses.  
299 The initial spread is zero by construction (blue curve); the overall error growth is smaller  
300 than when the EnKF is used, and the error growth for within the 6-hour background period  
301 is now considerably larger. However, as we will see shortly, even with this difference in fore-  
302 cast spread within the 6-hour background period, the end result between the two ensemble  
303 generation procedures is very similar to the corresponding hybrid ADAS performing rather  
304 closely.



305 *c. Evaluation with respect to observations*

306 Figure 6 shows vertical profiles of *monthly averaged* zonal wind (top) and temperature  
307 (bottom) radiosonde OMB residuals over three regions of the globe, namely, Northern Hemi-  
308 sphere (NH; left), tropics (center), and Southern Hemisphere (SH; right). Two hybrid ex-  
309 periments, one using the EnKF (HY5, red) and another using the filter-free scheme (HYA,  
310 green), are compared to the traditional 3DVar control experiment (CTL, blue). The only  
311 noticeable differences are in the tropics and SH for zonal winds, where the hybrid experi-  
312 ments show reduced biases with respect to the control; the EnKF and simplified (filter-free)  
313 scheme are rather comparable to each other. Results for temperature remain rather neutral.  
314 Examination of *standard deviation* of the OMB residuals for both winds and temperature  
315 indicate negligible differences among all three experiments (not shown).

316 It is also possible to examine the impact of observations on the analysis following Todling  
317 (2013). This is an observation-space approach that uses the inverse of the observation  
318 error variances to define a measure for evaluating the contribution of various observing  
319 systems to the cycling assimilation. Fig. 7 displays impact results for the three experiments  
320 under consideration: control (black), EnKF-based hybrid (cyan), and filter-free-based hybrid  
321 (magenta). Regardless of the underlying analysis procedure, all three experiments show  
322 aircraft, radiosondes, and Aqua AIRS as the dominating observing systems in GEOS ADAS.  
323 These observing systems tend to display smaller impact when the cycling analysis is based  
324 on a hybrid approach as compared to traditional 3DVar — the hybrid strategies seem to rely  
325 slightly more on these observing systems than does traditional 3DVar.

326 Figure 8 shows vertical profiles of standard deviations, calculated over the month of April

327 2012, for zonal wind radiosonde OMF residuals of the 24 hour forecasts. Though rather small,  
328 the benefit of using a hybrid assimilation strategy shows in both the tropics and Southern  
329 Hemisphere. Again here, the difference between the EnKF-based system and that using the  
330 filter-free configuration is very small, with some advantage shown for the latter in the SH.

331 *d. Evaluation with respect to independent analysis*

332 We routinely compare monthly mean analyses with those from other NWP centers. Fig-  
333 ure 9 shows the differences of the April 2012 zonally-averaged zonal wind for each experiment  
334 with the corresponding ECMWF operational analysis. Panels in the figure are differences for  
335 the control (CTL, top left), the filter-free hybrid scheme (HYA, top right), and the EnKF-  
336 based hybrid (HY5, bottom left). Compared with the control, both hybrid procedures ob-  
337 tains monthly mean analysis considerably closer to ECMWF’s monthly mean analysis; this is  
338 especially noticeable in the tropics. The bottom-right panel shows the monthly mean of the  
339 ensemble mean EnKF analysis (from HY5) difference with ECMWF operational analysis.  
340 Comparing this result with, say, that in the bottom-left panel, illustrates the behavior and  
341 reliability of the underlying EnKF ensemble analyses, though in the presence of re-centering  
342 it serves mainly as a sanity check to show that inflation averages away.

343 *e. Evaluation with respect to self analysis*

344 Lastly, we show some results when comparing forecasts from each of the three experiments  
345 with their own respective analyses. Figure 10 displays the zonally-averaged wind RMS error

346 of the 24 hour forecast, as a function of pressure, for three regions of interest. Results are  
347 for the three experiments under consideration: control (blue), and the two EnKF (HY5,  
348 red) and filter-free (HYA, green) hybrid strategies. Both hybrid strategies yield the same  
349 improvement in RMS error in the Northern and Southern Hemispheres, but result in some  
350 deterioration in Tropical mid-troposphere, with the filter-free procedure being less damaging  
351 than the EnKF. This behavior is opposite to that seen when examining both the monthly  
352 mean analyses and mean OMB radiosonde residuals, in which hybrid strategies amounted to  
353 improvement over traditional 3DVar. This remains an issue to tackle in future studies with  
354 GEOS Hybrid ADAS.

355 In many ways, successful procedures must amount to improvement in the 500 hPa geopo-  
356 tential height anomaly correlations. Self-analysis evaluation results appear in Fig. 11 for  
357 5-day forecasts in both Northern (top-right) and Southern Hemisphere (top-left). Curves for  
358 the control experiment are in blue, those for the EnKF-based hybrid are in red, and those for  
359 the filter-free strategy are in green. The corresponding statistical significance curves appear  
360 at the bottom panels. The NH scores are pretty much neutral, but those in the SH show  
361 significant benefit from hybrid assimilation (bottom-left shows red and green curves outside  
362 and above significance boxes). Both hybrid strategies bring comparable and non-negligible  
363 improvements up to 5 days in their forecasts. We must stress the word comparable, as we see  
364 the filter-free procedure amounting to rather indistinguishable performance from a system  
365 using the EnKF to generate the ensemble of analyses.

## 5. Closing remarks

In the process of implementing a 3DVar hybrid strategy for the Goddard Earth Observing System (GEOS) atmospheric data assimilation system (ADAS) using the ensemble Kalman filter (EnKF) of Whitaker and Hamill (2002), under a dual resolution approach, we have found re-centering and additive inflation to play a fundamental role in determining the behavior of the ensemble. Examination of some preliminary results led us to consider generating the ensemble by simply adding NMC-method-like perturbations to the central (hybrid) variational analysis at each cycle, thus completely bypassing the EnKF. This so-called filter-free procedure was put to the same evaluation test suite as that used to examine the quality of our EnKF-based 3DVar hybrid implementation. Both schemes are shown to perform rather similarly, bringing statistically significant improvements to GEOS ADAS. Indeed, the improvements to GEOS ADAS due to hybridization are comparable in magnitude to those seen at NCEP when upgrading its 3DVar system to a hybrid strategy, around May 2012. The successful evaluation of the filter-free approach is encouraging since one of its main advantages relates to not having to maintain two considerably different analysis systems, namely, one to perform the EnKF and another to perform the 3DVar hybrid analysis (the Grid-point Statistical Interpolation analysis, in the present case). Though not the main driving motivation for this work, it is also important to stress the computational advantages of the filter-free approach over the EnKF, or any alternative ensemble filter scheme, since the filter-free scheme does not explicitly analyze the members of the ensemble.

At this point, we can only attempt to speculate on the reasons why the EnKF and filter-free procedures perform so similarly. Factors that are likely to contribute to this are the small

388 size of the ensemble, and the dual resolution aspect of the GEOS ADAS implementation.  
389 Future tests are planned to accurately evaluate the role solely due to the resolution interplay.  
390 Further tests are also planned to look at the role played by the size of the ensemble, though  
391 we expect these to be harder to accurately provide conclusive results since they may require  
392 too large an ensemble to possibly afford in real applications such as the ones presented here.

393 *Acknowledgments.*

394 The authors are thankful to David F. Parrish, Darly Kleist, Russ Treadon, and John  
395 Derber, from NOAA/NCEP, for the multiple discussions during the period of implementation  
396 of the hybrid components of the GEOS ADAS, particularly in what refers to GSI- and EnKF-  
397 related settings. The authors are also thankful to Jeffrey S. Whitaker, from NOAA/ESRL,  
398 for a number of discussions throughout the course of implementation of his EnKF software in  
399 our system. Results were obtained using support and computing resources from the NASA  
400 Center for Climate Simulation at Goddard Space Flight Center.

401

## 402 REFERENCES

403 Bierman, G. L., 1977: *Factorization Methods for Discrete Sequential Estimation*. Dover  
404 Publications, 241 pp.

405 Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using  
406 incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.

407 Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercompari-  
408 son of variational data assimilation and the ensemble kalman filter for global deterministic  
409 nwp. part ii: One-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567–  
410 1586.

- 411 Chou, M.-D. and M. J. Suarez, 1999: A shortwave radiation parameterization for atmo-  
412 spheric studies. Tech. rep., NASA. TM 104606, Vol. 15, 40 pp.
- 413 Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2012: Operational implementation of a  
414 hybrid ensemble/4d-var global data assimilation system at the met office. *Q. J. Royal*  
415 *Meteorol. Soc.*, doi:10.1002/qj.2054.
- 416 Collarco, P., A. da Silva, M. Chin, and T. Diehl, 2010: Online simulations of global aerosol  
417 distributions in the nasa geos-4 model and comparisons to satellite and ground-based  
418 aerosol optical depth. *J. Geophys. Res.*, **155**, D14 207.
- 419 Collins, N., et al., 2005: Design and implementation of components in the Earth System  
420 Modeling Framework. *Intl. J. High Perform. Comput. Appl.*, **19(3)**, 341–350.
- 421 Derber, J. C. and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys.*  
422 *Oceanogr.*, **19**, 1333–1347.
- 423 Derber, J. C. and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP  
424 SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299.
- 425 Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon.*  
426 *Wea. Rev.*, **129**, 550–560.
- 427 Hamill, T. M. and C. Snyder, 2000: A hybrid ensemble kalman filter-3d variational analysis  
428 scheme. *Mon. Wea. Rev.*, **128**, 2905–2919.
- 429 Hamill, T. M. and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in

430 ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.*, **133**,  
431 3132–3147.

432 Houtekamer, P. L. and H. L. Mitchell, 2001: A sequential ensemble kalman filter for atmo-  
433 spheric data assimilation. *Mon. Wea. Rev.*, **129**, 123137.

434 Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and  
435 B. Hansen, 2005: Atmospheric data assimilation with an ensemble kalman filter: Results  
436 with real observations. *Mon. Wea. Rev.*, **133**, 604620.

437 Kleespies, T. J., P. van Delst, L. M. McMillin, and J. Derber, 2004: Atmospheric trans-  
438 mittance of an absorbing gas. 6. OPTRAN status report and introduction to the NES-  
439 DIS/NCEP Community Radiative Transfer Model. *Appl. Opt.*, **43**, 3103–3109.

440 Kleist, D. T., 2012: An evaluation of hybrid variational-ensemble data assimilation for  
441 the NCEP GFS. Tech. rep., Univ. of Maryland. Ph.D. Thesis, 163 pp. [Available  
442 at [http://www.emc.ncep.noaa.gov/gmb/wd20dk/docs/phd/DarylKleist\\\_PhDThesis\\\_](http://www.emc.ncep.noaa.gov/gmb/wd20dk/docs/phd/DarylKleist\_PhDThesis\_Revised.pdf)  
443 [\\_Revised.pdf](http://www.emc.ncep.noaa.gov/gmb/wd20dk/docs/phd/DarylKleist\_PhDThesis\_Revised.pdf)].

444 Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009b:  
445 Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137**,  
446 1046–1060.

447 Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009a: Intro-  
448 duction of the GSI into the NCEPs Global Data Assimilation System. *Wea. Forecasting*,  
449 **24**, 1691–1705.



450 Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP–A comparison  
451 with 4D-Var. *Q. J. Royal Meteorol. Soc.*, **129**, 3183–3203.

452 Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble, balance, and model-error  
453 representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.

454 Moorthi, S. and M. Suarez, 1992: Relaxed Arakawa-Schubert: A parameterization of moist  
455 convection for general-circulation models. *Mon. Wea. Rev.*, **120**, 978–1002.

456 Putman, W. M. and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids.  
457 *J. Comp. Phys.*, **227**, 55–78.

458 Rienecker, M. M., et al., 2008: The GEOS-5 Data Assimilation System - Documentation of  
459 Versions 5.0.1, 5.1.0, and 5.2.0. Tech. rep., NASA. TM 104606, Technical Report Series  
460 on Global Modeling and Data Assimilation, Vol. 27, 118 pp.

461 Stieglitz, M., A. Ducharne, R. Koster, and M. Suarez, 2001: The impact of detailed snow  
462 physics on the simulation of snow cover and subsurface thermodynamics at continental  
463 scales. *J. Hydromet.*, **2**, 228–242.

464 Todling, R., 2013: Comparing two approaches for assessing observation impact. *Mon. Wea.*  
465 *Rev.*, **141**, 1484–1505.

466 Wang, X., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently  
467 proposed ensemble/3d-var hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222–227.

468 Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed  
469 observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

- 470 Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the ncep global forecast system. *Mon. Wea. Rev.*, **136**, 463–482.
- 471
- 472 Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spacially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.
- 473

## 474 List of Figures

- 475 1 Illustration of contribution from each step taking place after the EnKF ensemble  
476 of analyses are generated. The panels show 500 hPa temperature: analysis  
477 increment for a given ensemble member (top left); effect of re-centering this  
478 given member about the central GSI analysis (top right); effect of applying  
479 additive inflation to the member analysis with a coefficient of 0.25 (bottom  
480 left); and resulting increment after both re-centering and additive inflation  
481 are applied (bottom right). 30
- 482 2 Schematic of AU as implemented in GEOS hybrid ensemble-variational atmo-  
483 spheric data assimilation system. 31
- 484 3 Zonal mean analysis increment, in total wet energy (J/kg) norm, using a stan-  
485 dard 3DVar (left), a 3DVar when the background error covariances are fully  
486 determined by the ensemble (center), and a hybrid 3DVar when the covari-  
487 ances are a 50% weighted sum of the static- and ensemble-derived background  
488 error covariances (right). 32

- 489 4 The panel on the left shows the total cost function as it changes during the  
490 iterations of the GSI minimization; all cases are calculated for the same synop-  
491 tic time but GSI is configured as follows: static (non-hybrid) 3DVar without  
492 balance constraint (black curve); (non-hybrid) 3DVar with TLNMC balance  
493 constraint (red curve); hybrid 3DVar without balance constraint applied to  
494 hybrid part of increment (green curve); and hybrid 3DVar with balance con-  
495 straint applied to full increment (blue curve). The panel on the right shows  
496 the integrated mass-wind divergence spectra of the analysis increment as a  
497 function of wave number for the same four configurations; color scheme of  
498 curves is as in panel on the left. 33
- 499 5 Global spread of a 32-member ensemble measured in total energy units (J/kg);  
500 when EnKF is used to generate ensemble (top), and when filter-free ensemble  
501 scheme is used instead (bottom). The curves are for: analysis spread before  
502 re-centering and inflation (blue); 3-, 6- and 9-hour backgrounds (green, red,  
503 and black respectively). Totals exclude levels roughly above 10 hPa. 34
- 504 6 Regionally-averaged, monthly mean of radiosonde OMB residuals of zonal  
505 wind (top) and temperature (bottom) for three experiments: control (blue),  
506 EnKF-based hybrid (red), and filter-free hybrid (green), shown for: Northern  
507 Hemisphere (left column), tropics (center column), and Southern Hemisphere  
508 (right column). 35

- 509 7 Observation impact on the analysis for three 3DVar experiments: control,  
510 non-hybrid (black bars); hybrid using EnKF (cyan bars); and hybrid using  
511 simplified, filter-free approach (magenta bars). In addition to the observa-  
512 tion types shown, all experiments use GPS radio occultation, but results are  
513 not shown here due to a little glitch in the output files saving their corre-  
514 sponding information (basically, GPS impacts are of the magnitude of those  
515 of radiosondes, and are comparable among the difference analysis approaches). 36
- 516 8 Similar to Fig. 6, but for standard deviation. Only zonal winds are shown  
517 since temperature have neutral results. 37
- 518 9 April 2012 monthly mean of zonally-averaged zonal wind analysis differences  
519 with ECMWF operational analysis from four different ADAS scenarios: con-  
520 trol, traditional 3DVar (top left); filter-free-based hybrid 3Dvar (top right);  
521 EnKF-based hybrid 3DVar (bottom left); and EnKF ensemble mean (bottom  
522 right). 38
- 523 10 Twenty-four hour forecast RMS error, with respect to self-analysis, of regionally-  
524 averaged zonal winds for the three experiments under consideration: con-  
525 trol (blue), EnKF-based hybrid (red), and filter-free hybrid (green); Northern  
526 Hemisphere (left), tropics (center), and Southern Hemisphere (right). 39

527 11 Anomaly correlation of the 500 hPa height of 5-day forecasts (top) verified  
528 with respect to own analysis, and shown for Northern (left) and Southern  
529 (right) Hemispheres for the three experiments under consideration: the con-  
530 trol (blue), EnKF-based hybrid (red), and filter-free hybrid (green). Signifi-  
531 cance plots appear beneath anomaly correlations with significance boxes color  
532 according to experiment designation; results are statistically significant when  
533 curve appear outside, and above, corresponding box.

40

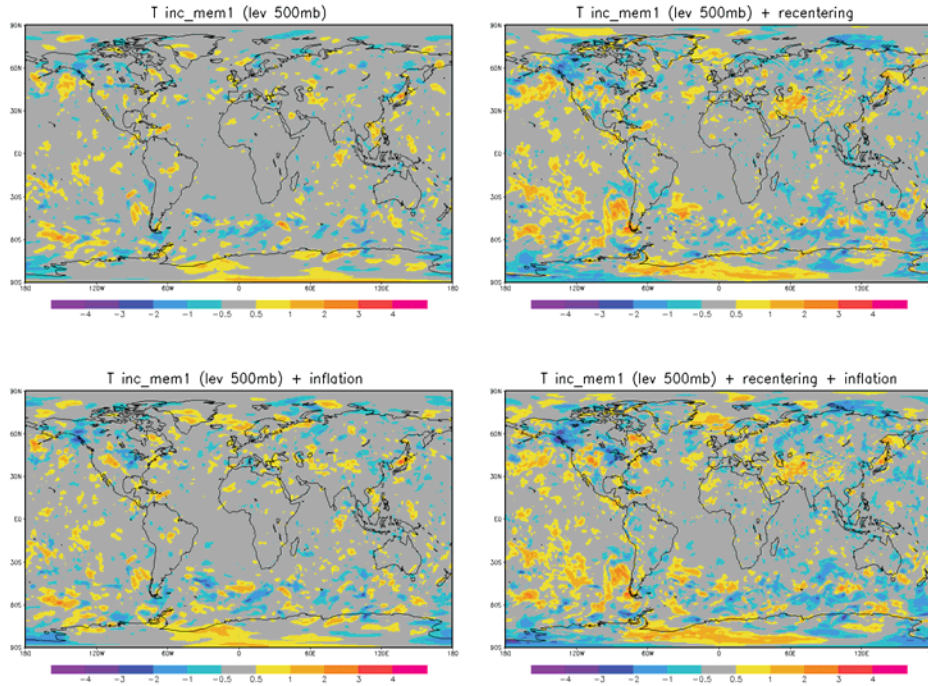


FIG. 1. Illustration of contribution from each step taking place after the EnKF ensemble of analyses are generated. The panels show 500 hPa temperature: analysis increment for a given ensemble member (top left); effect of re-centering this given member about the central GSI analysis (top right); effect of applying additive inflation to the member analysis with a coefficient of 0.25 (bottom left); and resulting increment after both re-centering and additive inflation are applied (bottom right).

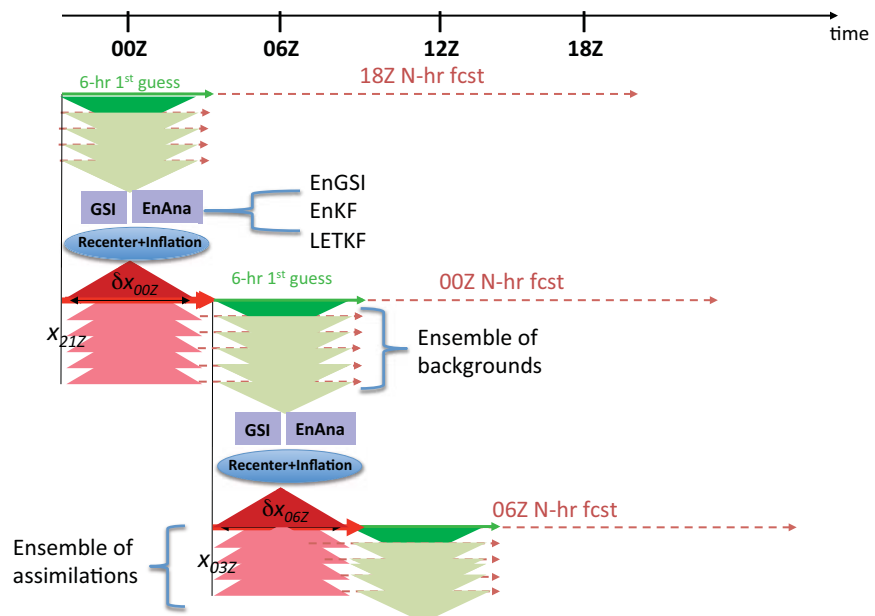


FIG. 2. Schematic of AU as implemented in GEOS hybrid ensemble-variational atmospheric data assimilation system.



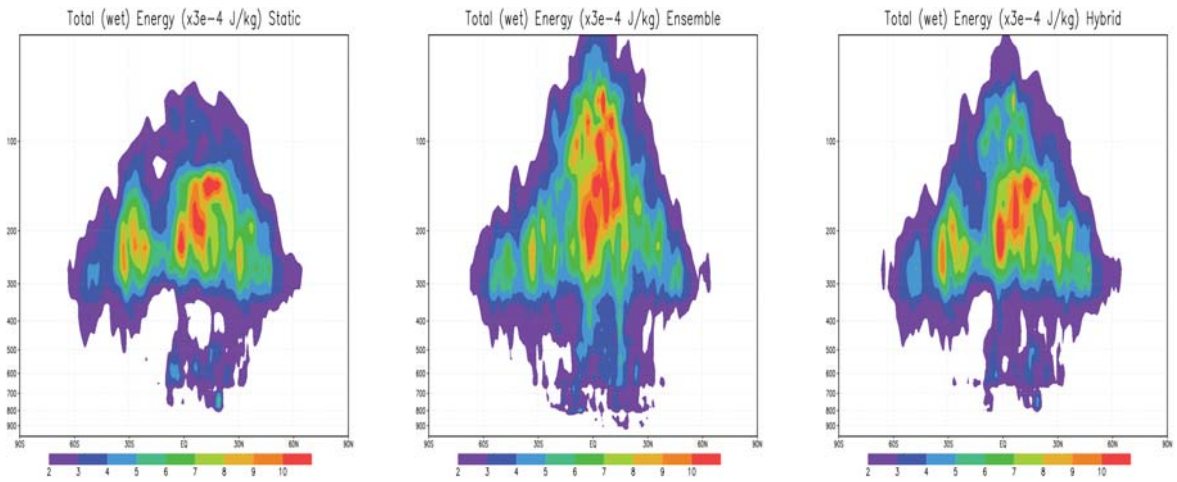


FIG. 3. Zonal mean analysis increment, in total wet energy ( $\text{J/kg}$ ) norm, using a standard 3DVar (left), a 3DVar when the background error covariances are fully determined by the ensemble (center), and a hybrid 3DVar when the covariances are a 50% weighted sum of the static- and ensemble-derived background error covariances (right).

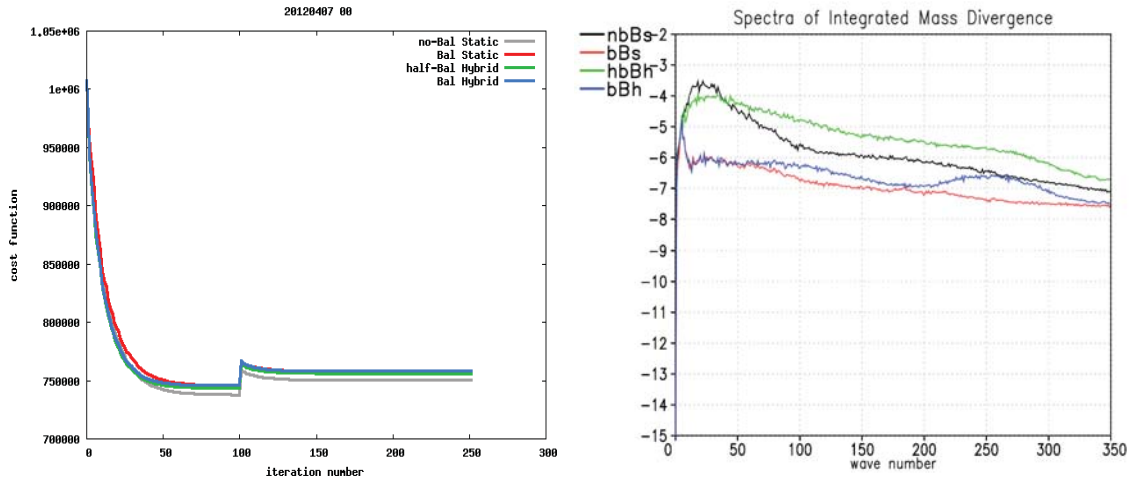


FIG. 4. The panel on the left shows the total cost function as it changes during the iterations of the GSI minimization; all cases are calculated for the same synoptic time but GSI is configured as follows: static (non-hybrid) 3DVar without balance constraint (black curve); (non-hybrid) 3DVar with TLNMC balance constraint (red curve); hybrid 3DVar without balance constraint applied to hybrid part of increment (green curve); and hybrid 3DVar with balance constraint applied to full increment (blue curve). The panel on the right shows the integrated mass-wind divergence spectra of the analysis increment as a function of wave number for the same four configurations; color scheme of curves is as in panel on the left.

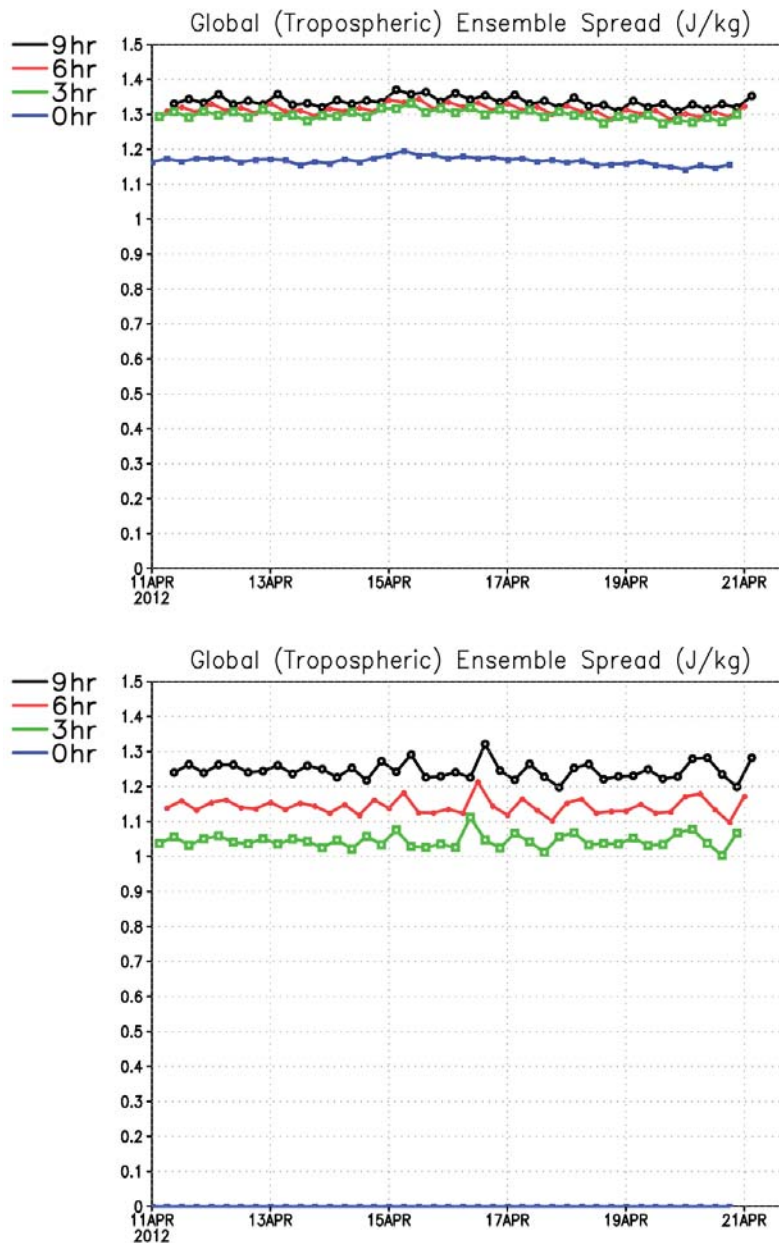


FIG. 5. Global spread of a 32-member ensemble measured in total energy units (J/kg); when EnKF is used to generate ensemble (top), and when filter-free ensemble scheme is used instead (bottom). The curves are for: analysis spread before re-centering and inflation (blue); 3-, 6- and 9-hour backgrounds (green, red, and black respectively). Totals exclude levels roughly above 10 hPa.

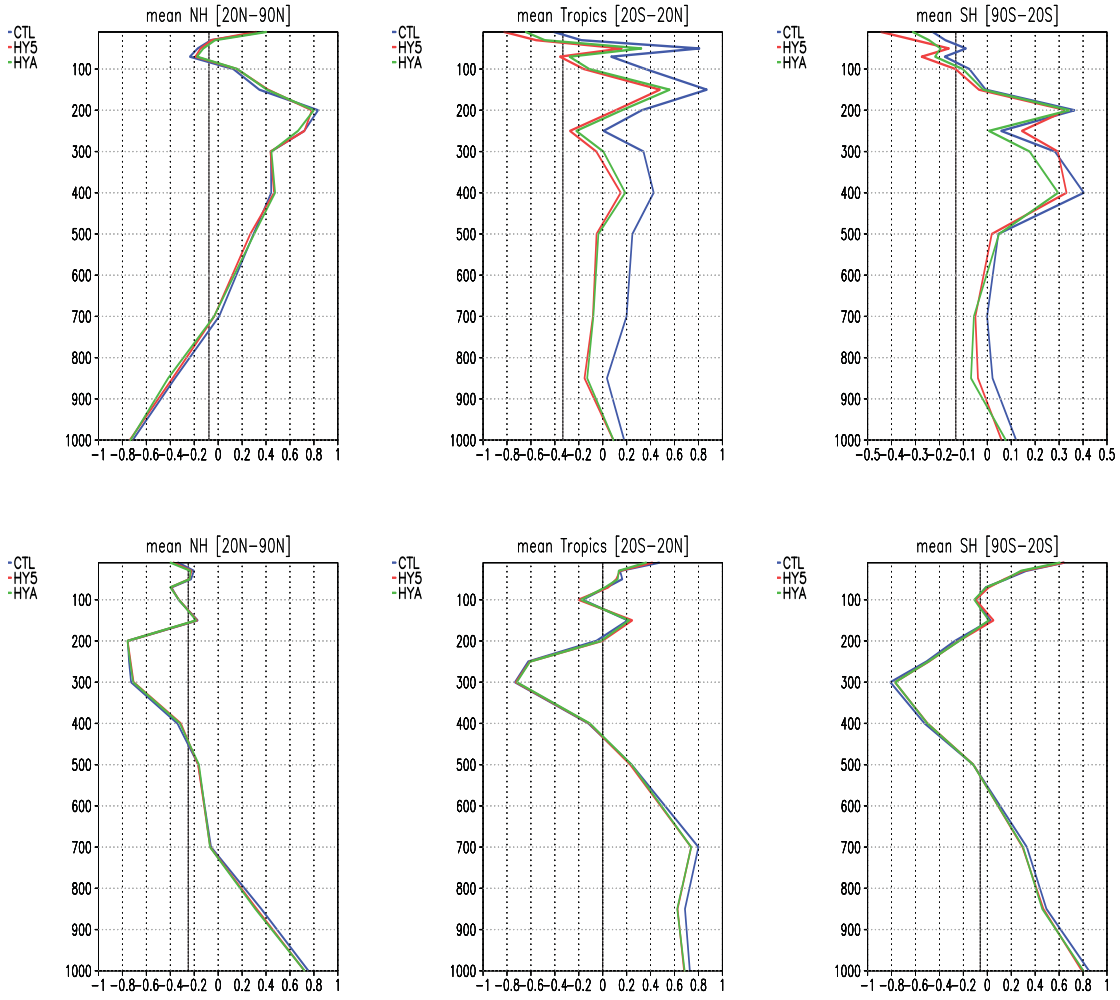


FIG. 6. Regionally-averaged, monthly mean of radiosonde OMB residuals of zonal wind (top) and temperature (bottom) for three experiments: control (blue), EnKF-based hybrid (red), and filter-free hybrid (green), shown for: Northern Hemisphere (left column), tropics (center column), and Southern Hemisphere (right column).

GEOS-5 Summary  
1 Apr 2012-30 Apr 2012  
Global Domain Total Impact

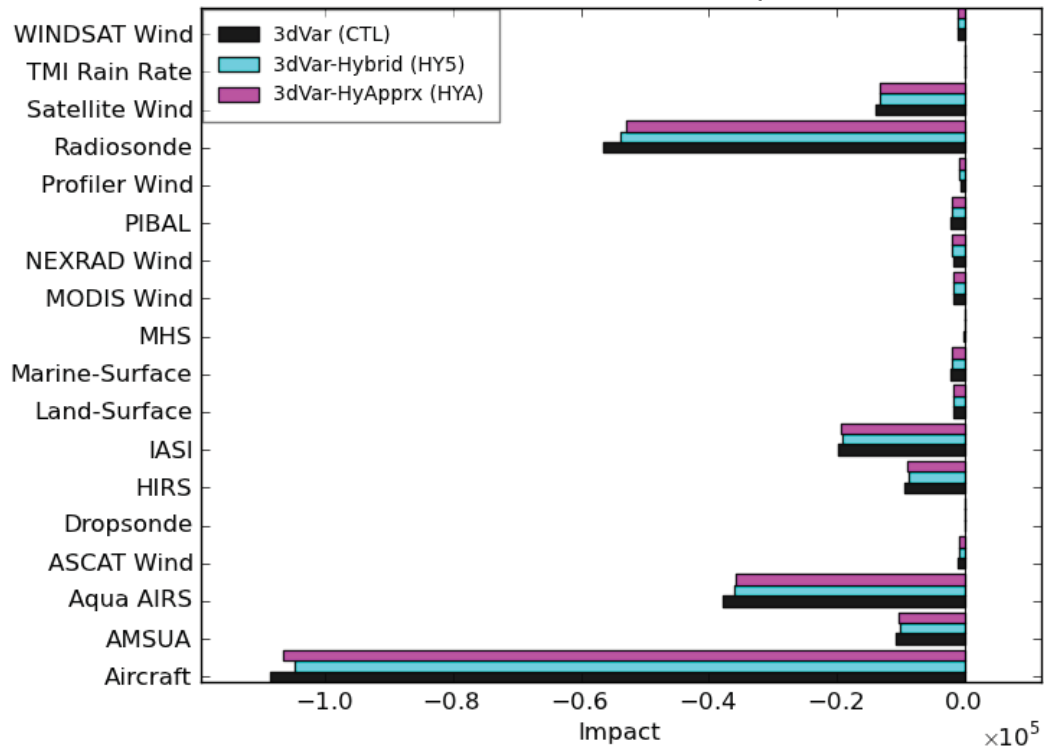


FIG. 7. Observation impact on the analysis for three 3DVar experiments: control, non-hybrid (black bars); hybrid using EnKF (cyan bars); and hybrid using simplified, filter-free approach (magenta bars). In addition to the observation types shown, all experiments use GPS radio occultation, but results are not shown here due to a little glitch in the output files saving their corresponding information (basically, GPS impacts are of the magnitude of those of radiosondes, and are comparable among the difference analysis approaches).

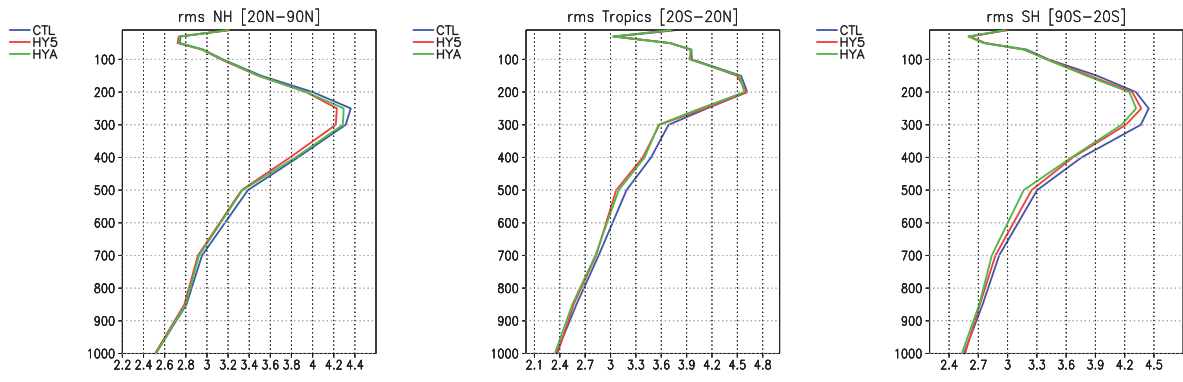


FIG. 8. Similar to Fig. 6, but for standard deviation. Only zonal winds are shown since temperature have neutral results.

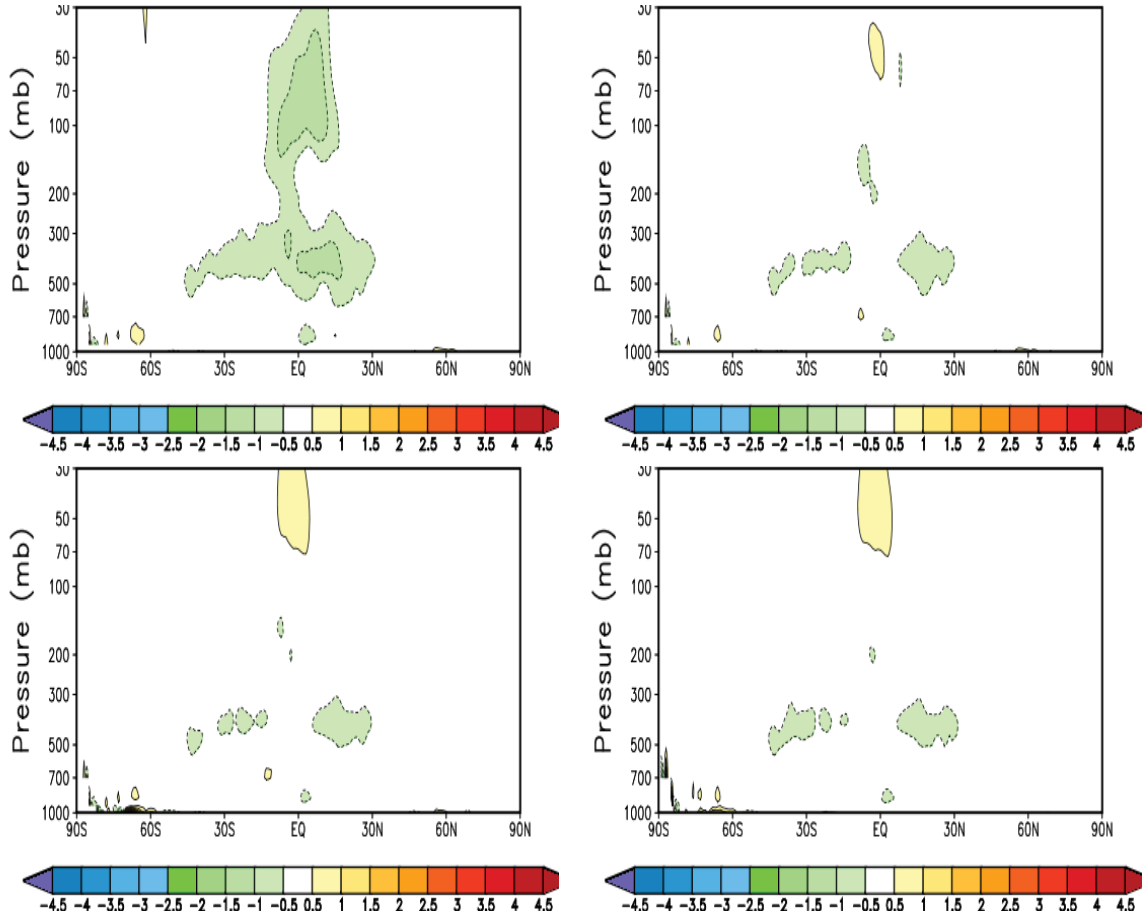


FIG. 9. April 2012 monthly mean of zonally-averaged zonal wind analysis differences with ECMWF operational analysis from four different ADAS scenarios: control, traditional 3DVar (top left); filter-free-based hybrid 3Dvar (top right); EnKF-based hybrid 3DVar (bottom left); and EnKF ensemble mean (bottom right).

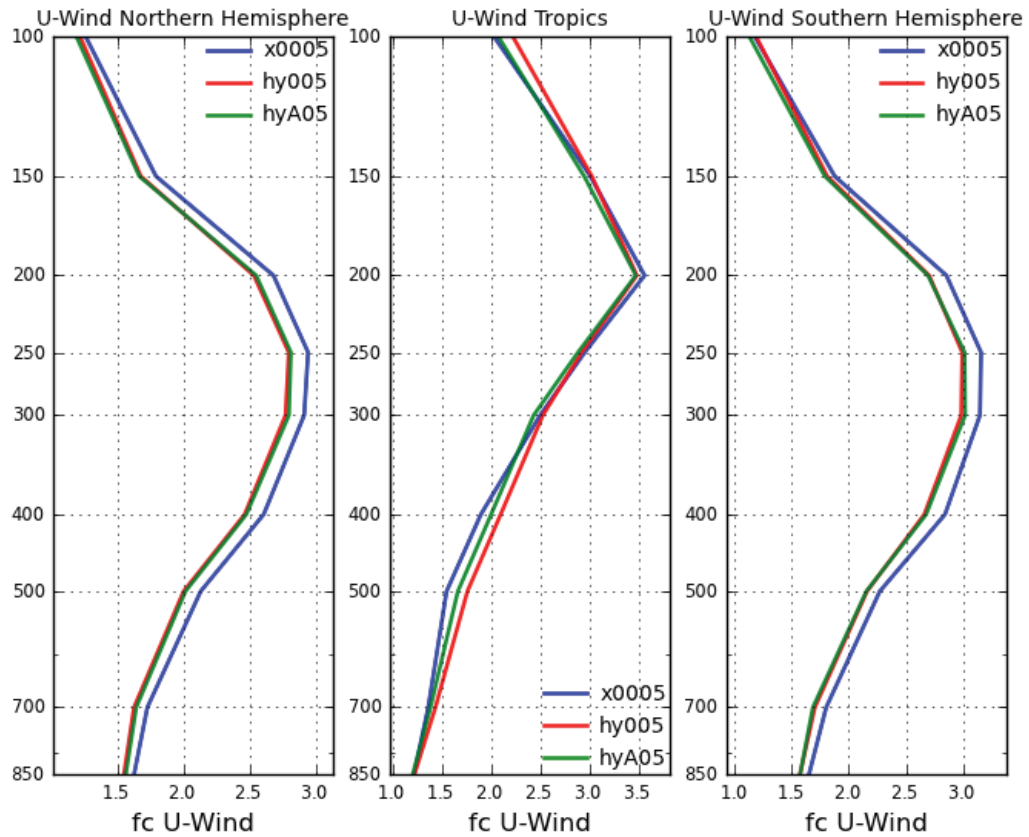


FIG. 10. Twenty-four hour forecast RMS error, with respect to self-analysis, of regionally-averaged zonal winds for the three experiments under consideration: control (blue), EnKF-based hybrid (red), and filter-free hybrid (green); Northern Hemisphere (left), tropics (center), and Southern Hemisphere (right).



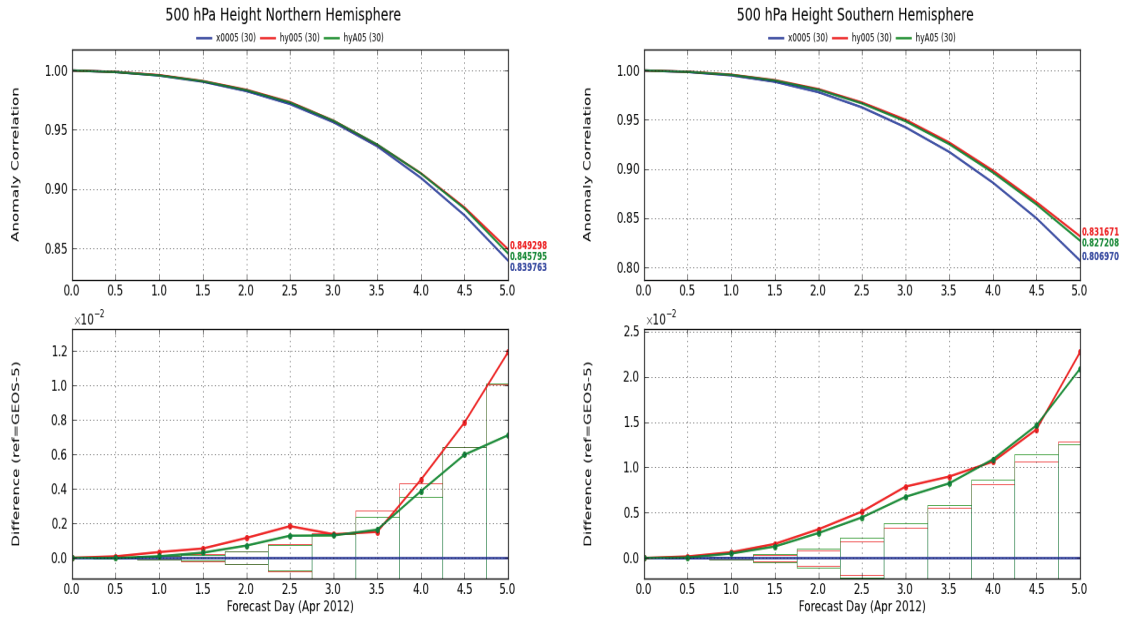


FIG. 11. Anomaly correlation of the 500 hPa height of 5-day forecasts (top) verified with respect to own analysis, and shown for Northern (left) and Southern (right) Hemispheres for the three experiments under consideration: the control (blue), EnKF-based hybrid (red), and filter-free hybrid (green). Significance plots appear beneath anomaly correlations with significance boxes color according to experiment designation; results are statistically significant when curve appear outside, and above, corresponding box.