JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2022MS003036

A Comparison of Hybrid-Gain Versus Hybrid-Covariance Data Assimilation for Global NWP

Jeffrey S. Whitaker[1], Anna Shlyaeva[2], and Stephen G. Penny[3,4]

[1]NOAA Physical Sciences Laboratory, Boulder, CO, USA, [2]Joint Center for Satellite Data Assimilation, University Corporation for Atmospheric Research, Boulder, CO, USA, [3]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, [4]Sofar Ocean Technologies, San Francisco, CA, USA

Key Points:

- A hybrid-gain approach that implicitly combines ensemble Kalman filter (EnKF) and 3DVar Kalman gain matrices by recentering EnKF analyses around an updated mean created by blending 3DVar and EnKF analysis increments is tested and compared to the current NOAA operational hybrid-covariance approach
- The hybrid-gain approach is found to be a simpler and computationally less expensive alternative that performs similarly to the operational hybrid-covariance approach if the incremental normal-model balance constraint in the hybrid-covariance update is disabled

Correspondence to:
J. S. Whitaker,
jeffrey.s.whitaker@noaa.gov

Citation:
Whitaker, J. S., Shlyaeva, A., & Penny, S. G. (2022). A comparison of hybrid-gain versus hybrid-covariance data assimilation for global NWP. *Journal of Advances in Modeling Earth Systems*, *14*, e2022MS003036. https://doi.org/10.1029/2022MS003036

Received 8 FEB 2022
Accepted 30 JUN 2022

Author Contributions:
Methodology: Jeffrey S. Whitaker, Anna Shlyaeva, Stephen G. Penny

Abstract Two methods for incorporating a time-invariant, high-rank covariance estimate in an ensemble-based data assimilation system for global weather prediction are compared. The hybrid-covariance approach linearly combines the static and ensemble-based covariance estimate in a four-dimensional variational solver, whereas the hybrid-gain approach blends analysis increments computed separately using a three-dimensional variational solution and an ensemble Kalman filter solution. Results show that the simpler and less expensive hybrid-gain approach performs similarly if the incremental normal-mode balance constraint applied to the ensemble-part of the hybrid-covariance update is turned off.

Plain Language Summary Accurate forecasts of the earth system rely on accurate estimates of the initial state of the system, which are created by updating a model forecast with the latest observations using a technique called data assimilation. A crucial aspect of the problem is an accurate estimation of the error of the background model forecast being updated. Ensemble-based estimates of background-forecast error covariances are often used, but their accuracy is limited by sample size. To mitigate this, the ensemble-based estimate is often combined with a static, or climatological estimate that is not dependent on the current model state but can be higher dimensional. Rather than blending the two error covariance estimates directly within the update (the hybrid-covariance approach), the updates computed separately using the static and ensemble-based covariances can be combined (the hybrid-gain approach). The simpler and less expensive hybrid-gain approach is shown here to perform similarly to the hybrid-covariance approach, aside from the beneficial impact of extra dynamical balance constraints that are available in the hybrid-covariance approach but have not yet been implemented in the hybrid-gain system.

1. Introduction

Most ensemble-based data assimilation systems for global weather prediction use a blend of a time-invariant (also called static or climatological) background error covariance estimate with a dynamic localized ensemble-based estimate. The static estimate generally compensates for deficiencies in the ensemble-based estimate. For example, the static estimate is typically of higher rank than the ensemble-based estimate, and thus can counter the tendency for filter divergence when the ensemble size is insufficient to represent the subspace of unstable modes that correspond to the positive and neutral Lyapunov exponents of the cycling data assimilation system (Bocquet & Carrassi, 2017; Penny, 2017). This blending can be implemented by linearly combining the background-error covariance estimates before computing the Kalman Gain (the hybrid-covariance approach, first proposed by Hamill and Snyder (2000)) or by combining the static and ensemble-based Kalman gain estimates (the hybrid-gain approach, first proposed by Penny (2014)). In many current operational systems, the hybrid-covariance approach is applied by computing the update using a hybrid four-dimension ensemble variational solver (4DEnVar; see e.g., Buehner et al. (2010)). The ensemble perturbations are typically updated with an ensemble Kalman filter (EnKF) that does not utilize a static background error covariance estimate. The current operational NOAA Global Data Assimilation System (GDAS) uses such an approach (Kleist & Ide, 2015), in which the ensemble mean update produced by the EnKF is replaced by the hybrid 4DEnVar solution, and all of the ensemble members are then recentered around the hybrid-covariance updated ensemble mean analysis. In this study, we compare the hybrid-covariance approach with a hybrid-gain approach that implicitly blends the Kalman gain matrices produced by the EnKF and 3DVar by recentering the ensemble perturbations around a blended ensemble mean analysis. There are two primary motivations for investigating the hybrid-gain approach in the NOAA system. First, since the hybrid-gain approach uses 3DVar instead of the more expensive hybrid 4DEnVar solver, it runs

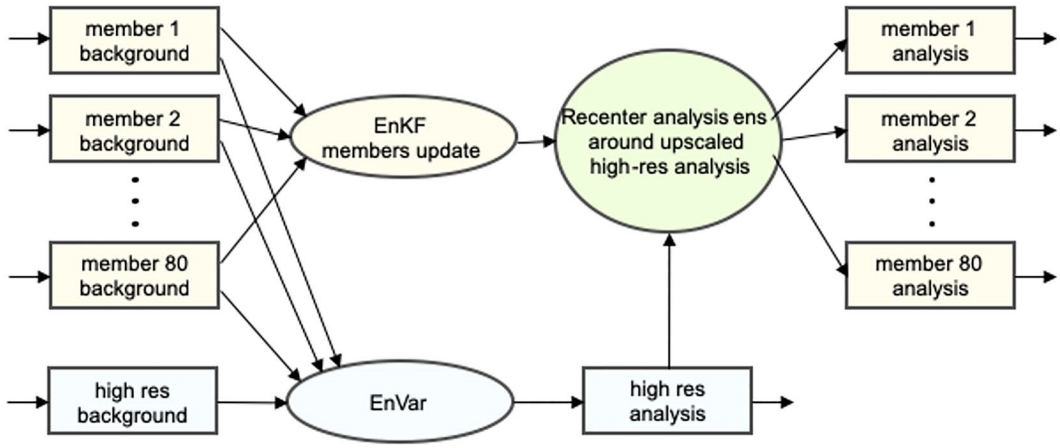

Figure 1. Operational configuration of the NOAA Global Data Assimilation System data assimilation update. For the "deterministic" analysis part hybrid four-dimension ensemble variational solver (4DEnVar) is run using high-resolution background and lower-resolution ensemble perturbations. Ensemble members are updated using the local ensemble transform Kalman filter and recentered around the upscaled high-resolution EnVar analysis before being propagated with ensemble forecast system to produce backgrounds for the next data assimilation window.

faster. Second, since the EnKF solution is still used in the ensemble mean update (unlike the hybrid-covariance approach where it is completely replaced by the hybrid 4DEnVar solution), it is more useful for testing improvements to the EnKF solver.

2. Experimental Design

The current operational hybrid 4DEnVar configuration in the GDAS uses 80 ensemble members of the Global Forecast System (GFS) model and a single higher-resolution deterministic forecast run at twice the horizontal resolution. A hybrid 4DEnVar algorithm is used to update the deterministic forecast, using the lower-resolution ensemble perturbations, and an EnKF algorithm (the gain-form of the local ensemble transform Kalman filter (LGETKF) with model-space vertical localization as described in Lei et al. (2018)) is used to update the ensemble members. The analysis ensemble is fully recentered around the up-scaled hybrid 4DEnVar analysis before being propagated forward to the next analysis time by the forecast model. This process is illustrated schematically in Figure 1. We note that it is not necessary to fully recenter the EnKF analysis ensemble around the hybrid 4DEnVar analysis—Houtekamer et al. (2019) found that a "partial recentering" approach in which only half the EnKF members were recentered performed better in the Canadian operational system.

In this study, we use a lower-resolution ensemble (~50 km horizontal resolution and 64 vertical levels instead of the ~25 km and 127 vertical level resolution used in operations), and there is no high-resolution deterministic forecast. Instead, the hybrid 4DEnVar solver uses the ensemble mean background to compute an analysis at the same resolution as the ensemble, and the LGETKF analyzed ensemble members are recentered around this analysis (Figure 2). Another difference with the operational setup is that the four-dimensional incremental analysis update (Lei & Whitaker, 2016) implemented in NOAA operations for GFS version 16 is not used. For the hybrid-gain configuration of the GDAS, we replace the hybrid 4DEnVar step in Figure 2 with a 3DVar solution. A weighted average of the 3DVar analysis increment and the LGETKF analysis increment is computed and added to the ensemble mean background, and the analysis ensemble is recentered around this new analysis (Figure 3). This is equivalent to blending the 3DVar and LGETKF Kalman Gains, as described in Appendix B of Penny (2014). A similar approach was used to blend 4DVar and LETKF gains by Bonavita et al. (2015). Note that in this approach the part of the ensemble mean analysis increment computed by the LGETKF is retained in the cycled system, unlike in the hybrid-covariance approach where it is completely replaced by the variational analysis.

A key parameter in the hybrid-covariance (hybrid-gain) algorithms is the relative weight given to the static and ensemble-based part of the covariance (gain). Currently in NOAA operations, the static covariance estimate is given a weight of 0.125 and the ensemble-based part of the covariance is given a weight of 0.875 in the hybrid 4DEnVar solution. The weights sum to one—but this is not a requirement. Because of the smoothing inherent in

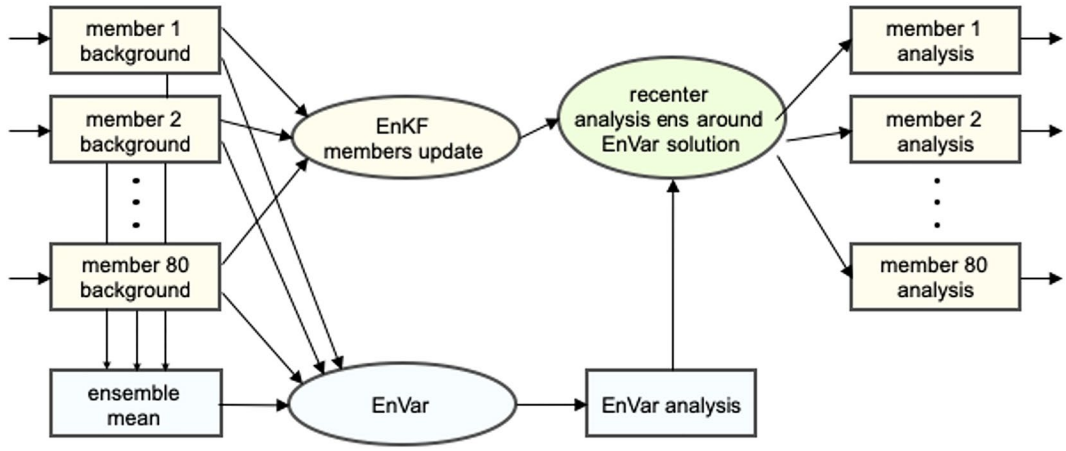

Figure 2. Data assimilation configuration used for the hybrid-covariance experiments in this study. Note the differences between operational setup (Figure 1) and this setup. The hybrid four-dimension ensemble variational solver system is run at the same spatial resolution as the ensemble system, using the ensemble mean background. There is no higher-resolution deterministic forecast.

the estimation of the static covariance, the static part of the increment is primarily representative of larger-scale errors, and the ensemble part of the increment dominates at the smallest scales. This has motivated us to fix the weight given to the ensemble part of covariance (increment) in the hybrid-covariance (hybrid-gain) experiments in this study to 1.0, so as not to reduce the amplitude total increment at small scales. This does not degrade the performance of either the hybrid-covariance or hybrid-gain algorithms as compared to requiring that the weights sum to 1.0 (not shown), and we note that the UK Met Office also uses a nonaffine weighting scheme (Bowler et al., 2017). Using the full ensemble-based increment is also consistent with the interpretation of the role of the static part of the update as a "stabilizer" to prevent filter divergence due to insufficient ensemble size, by preventing the larger scales from drifting away from the true state in the cycling system.

All of the observations that are assimilated operationally at NOAA are also assimilated in the experiments described here, including all-sky microwave AMSU-A and ATMS radiances (Zhu et al., 2019). Cross-channel error correlations for IASI and CRIS infrared sounders are also accounted for using the method described by Bathmann and Collard (2021). The method applies a linear transformation to a space in which the observation errors are uncorrelated, so no changes to the LGETKF algorithm (which assumes a diagonal observation error covariance and uses model-space vertical localization) are needed. In contrast, traditional observation-space

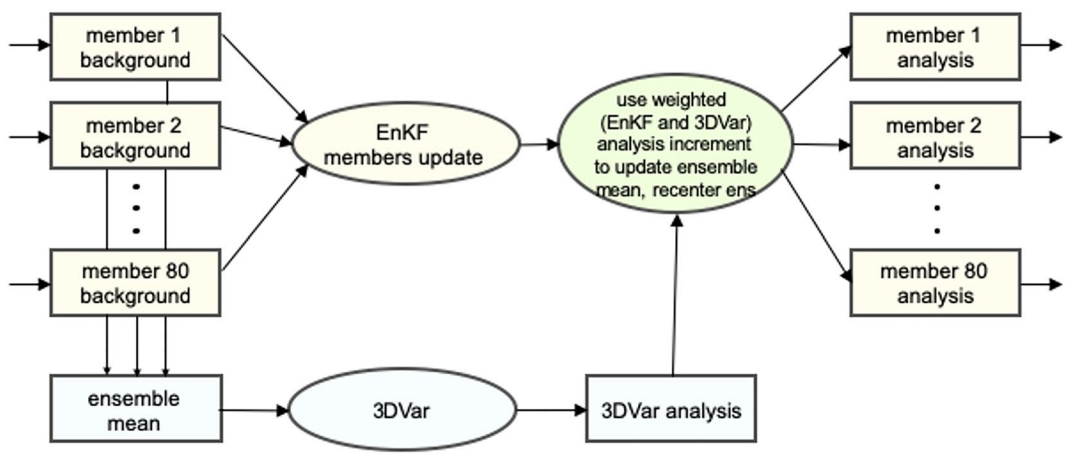

Figure 3. Data assimilation configuration used for the hybrid-gain experiments in this study. The hybrid-gain setup differs from the hybrid-covariance setup in the recentering step, which also blends the ensemble Kalman filter (EnKF) ensemble mean and 3DVar increments to create a new ensemble mean analysis.

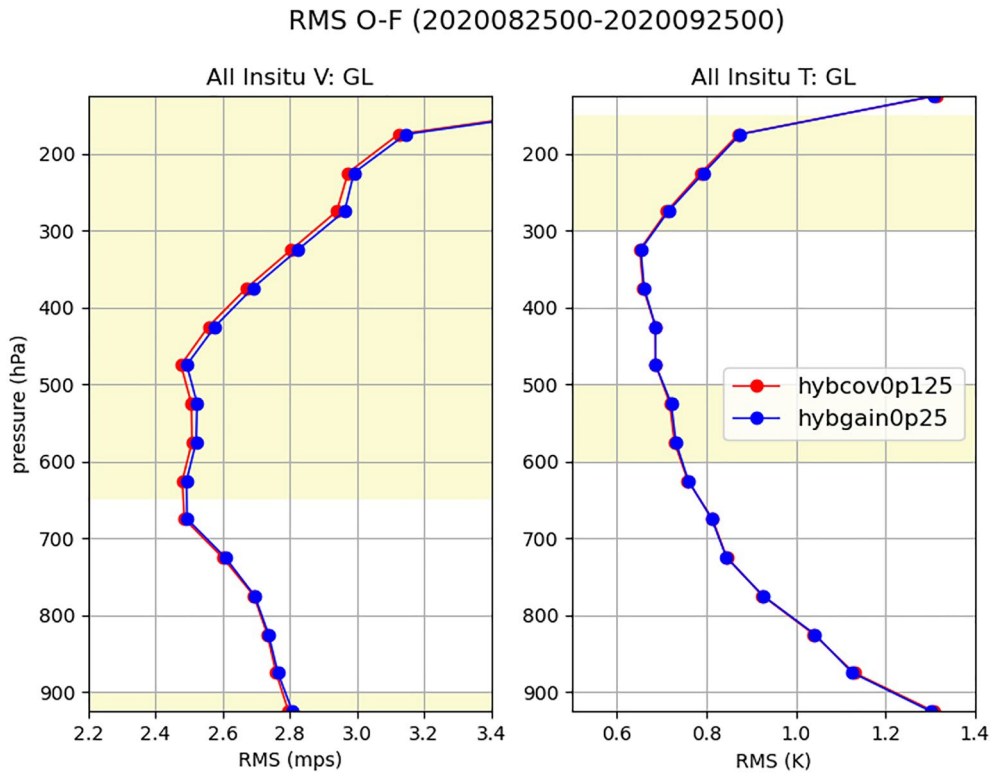

Figure 4. Hybrid-covariance and hybrid-gain innovation statistics for in situ vector wind and temperature observations. Only the two best performing experiments ($\alpha = 0.125$ for hybrid-covariance and = 0.25 for hybrid-gain) are shown. Vertical levels where the differences are deemed statistically significant at the 99% level are shaded yellow. Significance is evaluated using a Student's *t*-test with inflation to account for temporal autocorrelations (Geer, 2016, Appendix A).

vertical localization schemes such as those typically used in the LETKF are problematic since they require observations be assigned a vertical location, which is not well defined in the transformed space in which the observation errors are uncorrelated. Horizontal and vertical localization length scales (the scale at which the localization function goes to zero) are 1,250 km and 14 vertical levels, respectively, for both the LGETKF and 4DEnVar. In the LGETKF algorithm, vertical localization is applied in model space using the modulated ensemble approach described by Lei et al. (2018). As was shown by Lei et al. (2018), this is crucial for obtaining a solution with the LGETKF that matches the 4DEnVar solution when radiances are assimilated. To represent 98% of the variance of the vertical localization matrix, 13 eigenvectors are required, so the total (modulated) ensemble size used in the LETKF is 1,030 for each local volume. Since the number of ensemble members in each local volume implies a limit on the amount of observational information that can be effectively utilized, we cap the number of observations assimilated in each local volume to 10,000 (roughly 10 times the effective ensemble size) and select the subset that is closest (in horizontal distance) to the analysis grid point. Experimentation has shown that including more observations in each local volume does not improve (and in some cases degrades) the result. We also use the linearized observation operator in the LGETKF to calculate ensemble perturbations in observation space (Shlyaeva & Whitaker, 2018), so that the full nonlinear observation operator needs only to be computed for the ensemble mean background. This makes the forward observation operator calculation for the ensemble much more efficient, otherwise the full nonlinear observation operator would have to be computed for each of the 1,030 modulated ensemble members. To represent the model uncertainty component of the ensemble-based covariance estimate, we use relaxation-to-prior spread inflation (Whitaker & Hamill, 2012) with a coefficient of 0.85 and stochastic physics enabled in the forecast model. The stochastic parameterizations (stochastic kinetic energy backscatter, stochastically perturbed physics tendencies, and stochastically perturbed boundary layer humidity) in the GFS are configured as in NOAA operations (Pegion et al., 2016).

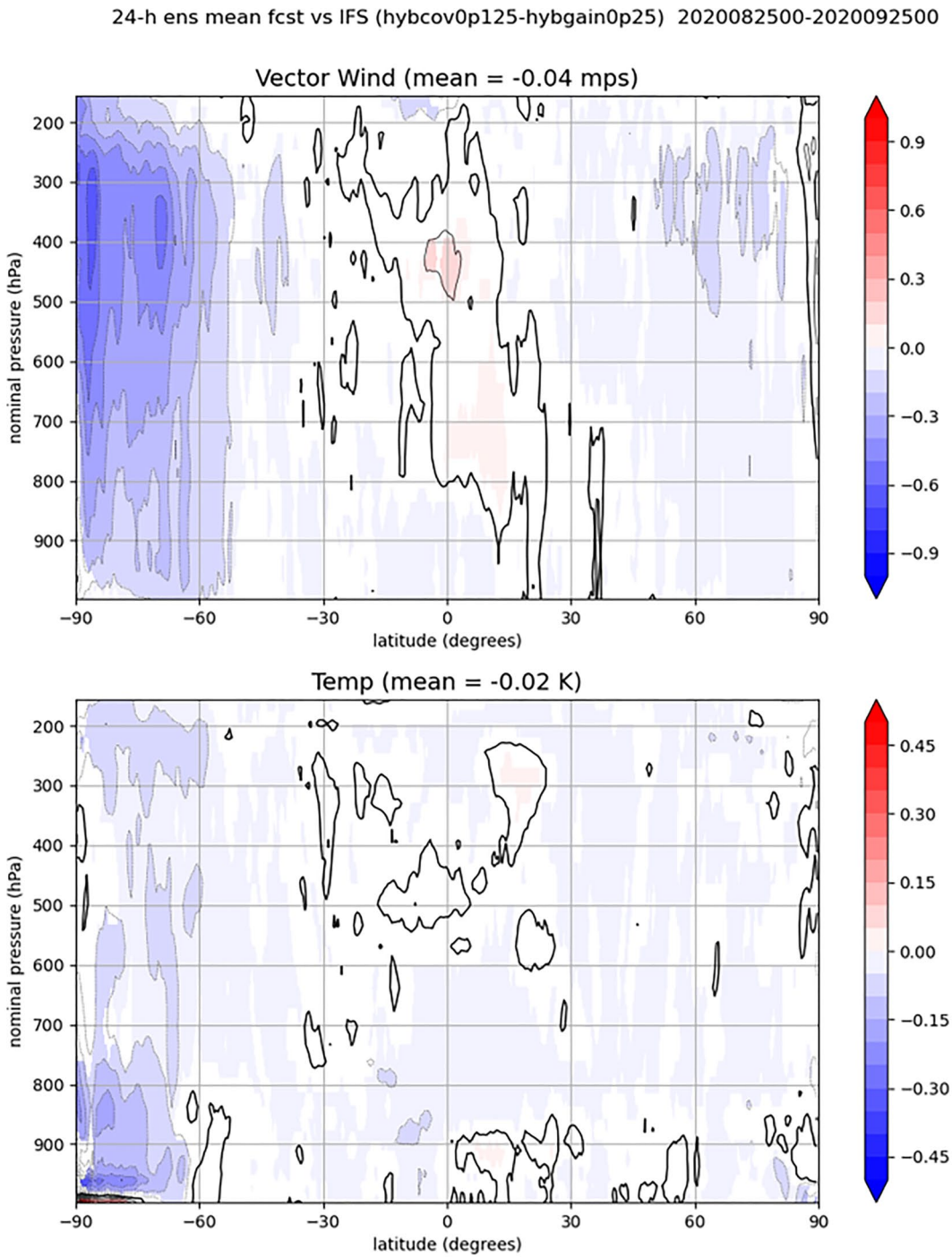

Figure 5. Zonal mean root-mean-squared (RMS) 24-hr ensemble-mean forecast error difference (vector wind and temperature) for the best performing hybrid-covariance (=0.125) and hybrid-gain (=0.25) experiments. Errors are computed relative to operational ECMWF analyses. Shaded areas indicate statistical significance at the 95% level including the effects of serial correlation. The global mean RMS error difference between the surface and the 150 hPa level is shown in the title of each plot. Negative values indicate the hybrid-covariance experiment performs better. Negative contours are dotted, positive contours are thin solid, and the zero contour is thicker solid.

Table 1
Global-Mean Root-Mean-Squared 24-hr Ensemble-Mean Forecast Errors (Vector Wind in m/s Temperature in K) Between the Surface and 150 hPa, Relative to ECMWF Analyses for All of the Experiments

Static B weight	Hybrid-cov with TLNMC	Hybrid-cov without TLNMC	Hybrid-gain
$\alpha = 0.0$	3.070/0.883	3.117/0.892	3.133/0.901
$\alpha = 0.125$	3.015/0.870	3.0578/0.876	3.072/0.889
$\alpha = 0.25$	3.035/0.875	3.071/0.880	3.058/0.886
$\alpha = 0.5$	3.080/0.884	3.102/0.888	3.096/0.894

Note. For $\alpha = 0.0$, the hybrid-covariance experiment is pure ensemble-four-dimension ensemble variational solver, and for the hybrid-gain it is pure gain-form local ensemble transform Kalman filter.

3. Comparison Between the Hybrid-Covariance and Hybrid-Gain Systems

Experiments were initialized with the upscaled operational EnKF ensemble from 2020-08-20:06, and run until 2020-09-25:00. Twice per day at 00 and 12 UTC the background forecast for the first 20 ensemble members was extended out to 24 hr. The performance metrics used here are the root-mean-squared (RMS) fit of the first guess to the in situ wind and temperature observations, and the RMS difference between the ensemble mean 24-hr forecast and the operational ECMWF analyses. Statistics were computed for the period 2020-08-25:00–2020-09-25:00. Experiments were conducted for four different static **B** weighting coefficients (hereafter denoted as α)—0, 0.125, 0.25, and 0.5. In Figures 4 and 5, comparisons of the innovation statistics and 24-hr forecast errors for the best performing hybrid-gain ($\alpha = 0.25$) and hybrid-covariance ($\alpha = 0.125$) experiments are shown. The fact that the minimum error is achieved with a different value of α in the two systems is not surprising, since the weight is applied to the Kalman gain in the hybrid-gain system, and Kalman gain is not a linear function of the background-error covariance. Table 1 shows the global mean temperature and vector wind RMS errors relative to ECMWF analyses for all of the experiments. For $\alpha = 0$, the hybrid-covariance experiment reduces to pure 4DEnVar, and the hybrid-gain experiment reduces to pure LGETKF. In the hybrid-gain $\alpha = 0$ case, the 3DVar solution is only used to update the bias-correction coefficients for radiance and aircraft temperature observations.

The best-performing hybrid-covariance experiment ($\alpha = 0.125$) outperforms the best-performing hybrid-gain experiment ($\alpha = 0.25$) especially for vector wind in the extra-tropics. However, the hybrid-covariance experiment

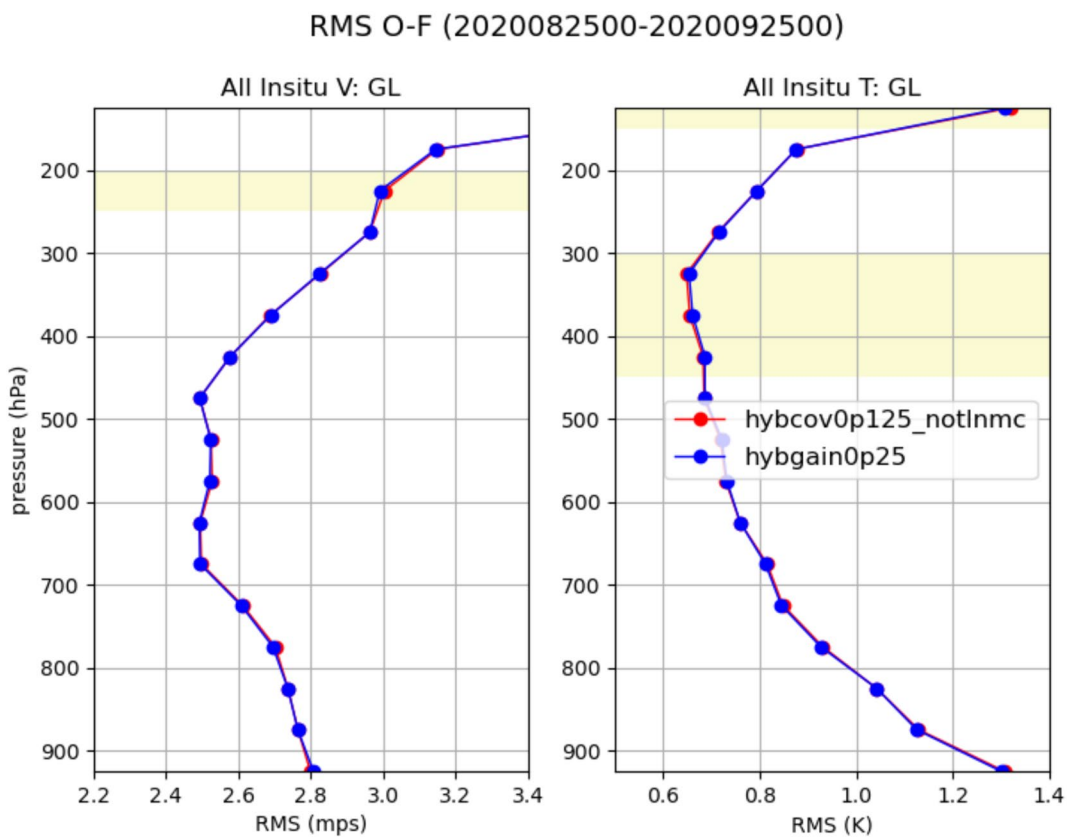

Figure 6. As in Figure 4, but without the tangent-linear normal mode constraint applied to the ensemble-based component of the increment in the hybrid-covariance experiment.

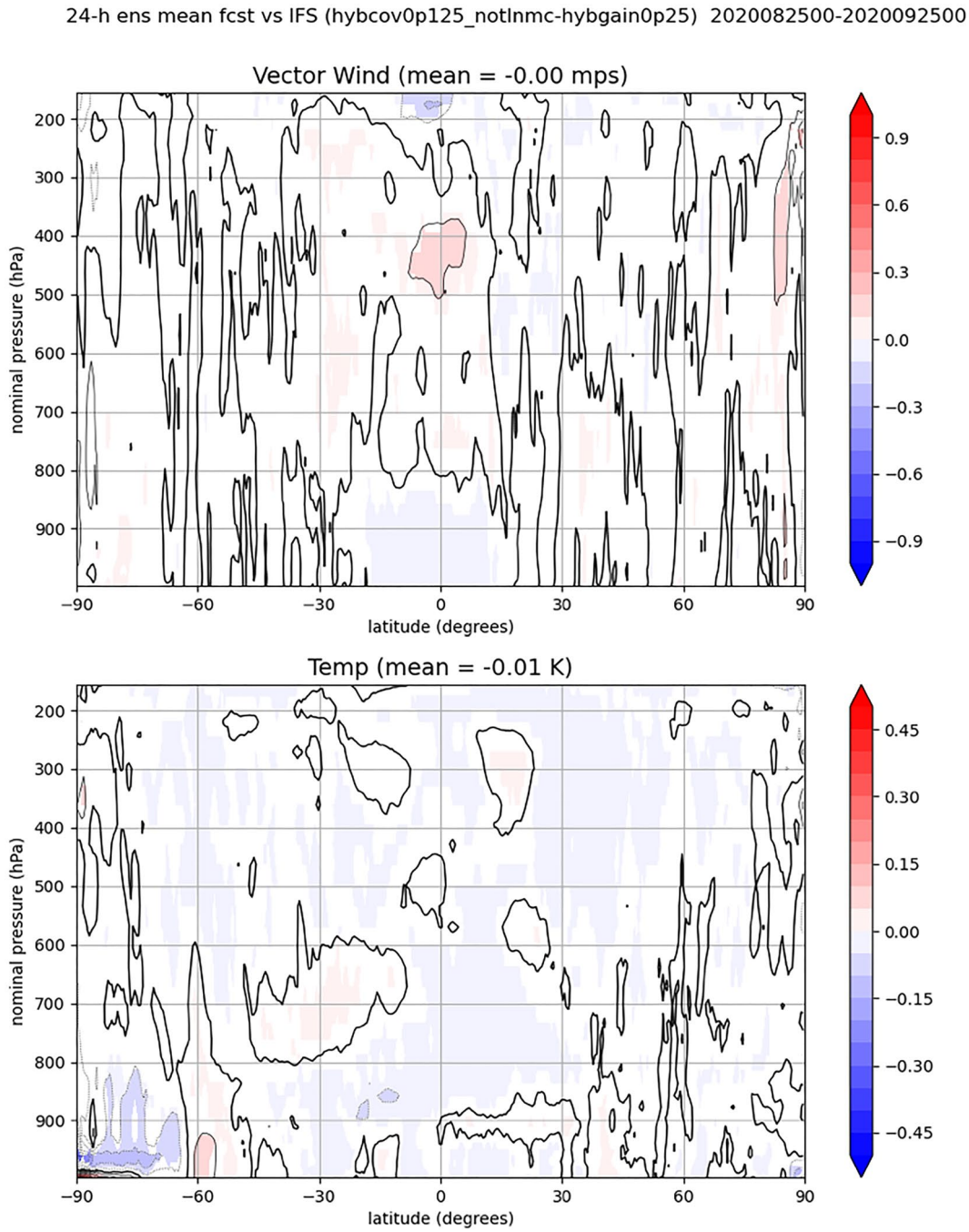

Figure 7. As in Figure 5, but without the tangent-linear normal mode constraint applied to the ensemble-based component of the increment in the hybrid-covariance experiment.

includes a tangent-linear normal mode incremental balance constraint (TLNMC; Kleist et al., 2009) on the ensemble-based part of the analysis increment while the hybrid-gain experiment only includes the TLNMC on the static **B** part of the increment. Figures 4, 5, 6 and 7 compare the hybrid-gain solution with a hybrid-covariance solution that only includes the TLNMC on the static **B** part of the increment. Without the TLNMC applied to the ensemble part of the variational increment in the hybrid-covariance solution, the differences are reduced considerably, and are not statistically significant over much of the domain, especially for vector-wind. We argue that these differences are small enough that they could be explained by differences in the way localization is performed in

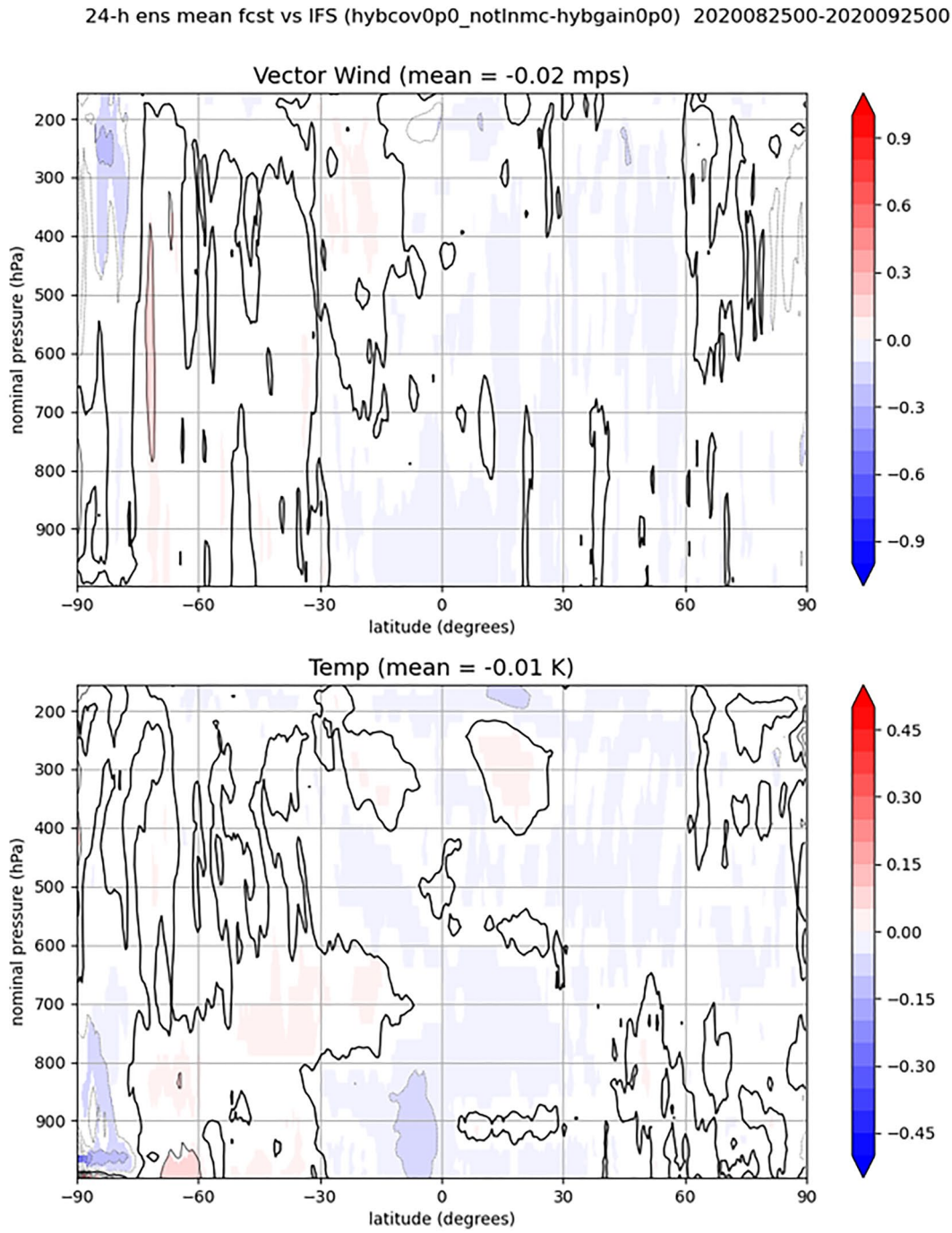

Figure 8. As in Figure 7, but for $\alpha = 0.0$, to isolate the impact of the method used to compute the ensemble-based part of the ensemble mean analysis increment (four-dimension ensemble variational solver vs. gain form local ensemble transform Kalman filter).

the variational and LGETKF systems, rather than the way the static component of the background-error covariance is treated. This is illustrated by Figure 8, which shows the difference in 24-hr ensemble-mean forecast errors for the $\alpha = 0.0$ case, which isolates the impact of the solver used to compute the ensemble-based part of the ensemble-mean analysis increment (4DEnVar vs. LGETKF). The differences shown in Figure 8, although small, are slightly larger than those shown in Figure 7. Since the primary difference between the LGETKF and 4DEnVar solvers is the way horizontal localization is handled (**B**-localization for 4DEnVar as opposed to **R**-localization

in the LGETKF, see e.g., Sakov and Bertino (2011)), this supports the conjecture that the differences in the way localization is treated is at least, if not more, important than the way the static covariances are handled in the hybrid-covariance and hybrid-gain algorithms. We note that Nerger et al. (2012) examined the impact of the localization method in ensemble-based solvers, and found that although the two methods produced similar results, the optimal length scale for **R**-localization is shorter than for **B**-localization.

4. Conclusions

Experiments using a low-resolution version of the NOAA operational GDAS show that a hybrid-gain approach performs similarly to a hybrid-covariance approach, if the tangent-linear incremental balance constraint applied to the ensemble-part of the increment is turned off. The TLNMC allows the hybrid-covariance algorithm to outperform the hybrid-gain algorithm, which does not apply a balance constraint to the ensemble-based part of the increment. This suggests that the hybrid-gain approach could be a viable and computationally efficient alternative to the current operational hybrid-covariance approach, if an incremental balancing procedure were developed for the EnKF update.

The TLNMC algorithm works by performing a minimization in the space of the "slow" normal modes of the primitive equations linearized about a state of rest. In this context, "slow" means the geostrophically balanced part of the spectrum, excluding the fast-propagating gravity waves. The 4DIAU algorithm used in the NOAA operational GDAS also acts to reduce the amplitude of fast motions since it acts as a frequency-selective digital filter (see Section 2 of Lei and Whitaker (2016)). In the experiments described here, we elected not to use 4DIAU to save computation time. This leaves open the possibility that application of the 4DIAU would reduce the gravity wave noise introduced by the ensemble-based part of the analysis increments, thereby reducing the impact of the TLNMC. To test this, we have performed additional EnVar and LGETKF experiments with $\alpha = 0.0$ and 4DIAU (not shown), and found that the apparent benefit of the TLNMC in the EnVar solution remains quite similar to that show in Figure 8. This suggests that the TLNMC is more effective at filtering out the deleterious gravity-wave oscillations introduced by the ensemble part of the analysis increment (caused by localization). Further study is needed to understand why this is the case, and how best to introduce a similar capability in the LGETKF solver.

Data Availability Statement

The operational NOAA GSI data assimilation system (including the variational and EnKF solvers) is publicly available at https://github.com/NOAA-EMC/GSI. The NOAA Unified Forecast System that forms the basis of the NOAA operational GFS is publicly available at https://github.com/ufs-community/ufs-weather-model. The observational data sets used for assimilation experiments described here (excluding some commercial platforms that cannot be redistributed) are available at https://www.ncei.noaa.gov/products/weather-climate-models/global-data-assimilation. The python scripts and netCDF data needed to generate Figures 4–8 are available at https://doi.org/10.5281/zenodo.6506468 (Whitaker, 2022).

Acknowledgments
This research was supported by the NOAA Physical Sciences Laboratory and the NOAA Unified Forecast System Research to Operation (UFS-R2O) Project, which is jointly funded by NOAA's Office of Science and Technology Integration (OSTI) of National Weather Service (NWS) and Weather Program Office (WPO) of the Office of Oceanic and Atmospheric Research (OAR). S. G. Penny also acknowledges support from the Office of Naval Research (ONR) (Grant Nos. N00014-19-1-2522 and N00014-20-1-2580) and the NOAA (Grant No. NA20OAR4600277). The authors thank Laura Slivinski for her comments on an early draft of this manuscript.

References

Bathmann, K., & Collard, A. (2021). Surface-dependent correlated infrared observation errors and quality control in the FV3 framework. *Quarterly Journal of the Royal Meteorological Society*, *147*(734), 408–424. https://doi.org/10.1002/qj.3925

Bocquet, M., & Carrassi, A. (2017). Four-dimensional ensemble variational data assimilation and the unstable subspace. *Tellus A: Dynamic Meteorology and Oceanography*, *69*, 1. https://doi.org/10.1080/16000870.2017.1304504

Bonavita, M., Hamrud, M., & Isaksen, L. (2015). EnKF and hybrid gain ensemble data assimilation. Part II: EnKF and hybrid gain results. *Monthly Weather Review*, *143*(12), 4865–4882. https://doi.org/10.1175/MWR-D-15-0071.1

Bowler, N. E., Clayton, A. M., Jardak, M., Jermey, P. M., Lorenc, A. C., Wlasak, M. A., et al. (2017). The effect of improved ensemble covariances on hybrid variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 785–797. https://doi.org/10.1002/qj.2964

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., & He, B. (2010). Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Monthly Weather Review*, *138*(5), 1550–1566. https://doi.org/10.1175/2009MWR3157.1

Geer, A. J. (2016). Significance of changes in medium-range forecast scores. *Tellus A: Dynamic Meteorology and Oceanography*, *68*, 1. https://doi.org/10.3402/tellusa.v68.30229

Hamill, T. M., & Snyder, C. (2000). A hybrid ensemble Kalman filter–3D variational analysis scheme. *Monthly Weather Review*, *128*(8), 2905–2919. https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2

Houtekamer, P. L., Buehner, M., & De La Chevrotière, M. (2019). Using the hybrid gain algorithm to sample data assimilation uncertainty. *Quarterly Journal of the Royal Meteorological Society*, *145*(Suppl. 1), 35–56. https://doi.org/10.1002/qj.3426

Kleist, D. T., & Ide, K. (2015). An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. *Monthly Weather Review*, *143*(2), 452–470. https://doi.org/10.1175/MWR-D-13-00350.1

Kleist, D. T., Parrish, D. F., Derber, J. C., Treadon, R., Errico, R. M., & Yang, R. (2009). Improving incremental balance in the GSI 3DVAR analysis system. *Monthly Weather Review*, *137*(3), 1046–1060. https://doi.org/10.1175/2008MWR2623.1

Lei, L., & Whitaker, J. S. (2016). A four-dimensional incremental analysis update for the ensemble Kalman filter. *Monthly Weather Review*, *144*(7), 2605–2621. https://doi.org/10.1175/MWR-D-15-0246.1

Lei, L., Whitaker, J. S., & Bishop, C. (2018). Improving assimilation of radiance observations by implementing model space localization in an ensemble Kalman filter. *Journal of Advances in Modeling Earth Systems*, *10*(12), 3221–3232. https://doi.org/10.1029/2018MS001468

Nerger, L., Janjić, T., Schröter, J., & Hiller, W. (2012). A regulated localization scheme for ensemble-based Kalman filters. *Quarterly Journal of the Royal Meteorological Society*, *138*(664), 802–812. https://doi.org/10.1002/qj.945

Pegion, P., Bates, G., Gehne, M., Hamill, T., Kolczynski, W., Whitaker, J., & Zhu, Y. (2016). Stochastic parameterization development in the NOAA/NCEP global forecast system. In *Workshop on model uncertainty*, 47. ECMWF/WWRP. Retrieved from https://www.ecmwf.int/en/elibrary/16741-stochastic-parameterization-development-noaa-ncep-global-forecast-system

Penny, S. G. (2014). The hybrid local ensemble transform Kalman filter. *Monthly Weather Review*, *142*(6), 2139–2149. https://doi.org/10.1175/MWR-D-13-00131.1

Penny, S. G. (2017). Mathematical foundations of hybrid data assimilation from a synchronization perspective. *Chaos*, *27*(12), 126801. https://doi.org/10.1063/1.5001819

Sakov, P., & Bertino, L. (2011). Relation between two common localisation methods for the EnKF. *Computational Geosciences*, *15*(2), 225–237. https://doi.org/10.1007/s10596-010-9202-6

Shlyaeva, A., & Whitaker, J. S. (2018). Using the linearized observation operator to calculate observation space ensemble perturbations in ensemble filters. *Journal of Advances in Modeling Earth Systems*, *10*(7), 1414–1420. https://doi.org/10.1029/2018MS001309

Whitaker, J. S. (2022). *jswhit/hybcov_vs_hybgain_da_paper: version 1.0.0 (v1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.6506468

Whitaker, J. S., & Hamill, T. M. (2012). Evaluating methods to account for system errors in ensemble data assimilation. *Monthly Weather Review*, *140*(9), 3078–3089. https://doi.org/10.1175/MWR-D-11-00276.1

Zhu, Y., Gayno, G., Purser, R. J., Su, X., & Yang, R. (2019). Expansion of the all-sky radiance assimilation to ATMS at NCEP. *Monthly Weather Review*, *147*(7), 2603–2620. https://doi.org/10.1175/MWR-D-18-0228.1