Observation Impact and Information Retention in the Lower Troposphere of the GMAO GEOS Data Assimilation System

Yanqiu Zhu, Ricardo Todling, and Nathan Arnold

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center

Corresponding author: Yanqiu Zhu, Yanqiu.Zhu@nasa.gov

ABSTRACT: In this study, we have assessed the effectiveness of the use of existing observing systems in the lower troposphere in the GEOS hybrid–4DEnVar data assimilation system through a set of observing system experiments. The results show that microwave radiances have a large impact in the Southern Hemisphere and Tropical ocean, but the large influence is mostly observed above 925 hPa and dissipates relatively quickly with longer forecast lead times. Conventional data information holds better in the forecast ranging from the surface to 100 hPa, depending on the field evaluated, in the Northern Hemisphere and lowest model levels in the Tropics. Infrared radiances collectively have much less impact in the lower troposphere. Removing surface observations has small but persistent impact on specific humidity in the upper atmosphere, but small or negligible impact on planetary boundary layer (PBL) height and temperature. The model responses to the incremental analysis update (IAU) forcing are also analyzed. In the IAU assimilation window, the physics responds strongly to the IAU forcing in the lower troposphere, and the changes of physics tendency in the lower troposphere and hydrodynamics tendency in the mid- and upper troposphere are viewed as beneficial to the reduction of state error covariance. In the subsequent forecast, the model tendencies continue to deviate further from the original free forecast with forecast lead times around 300–400 hPa, but physics tendency has showed signs of returning to its original free forecast mechanisms at 1-day forecast in the lower troposphere.

1. Introduction

The planetary boundary layer (PBL) is an important interface between the Earth's surface and the atmosphere, core to the understanding of flux balances across the Earth system components. Its importance is well-established for its applications in weather, climate, and air dispersion, and as such the PBL has been listed as an 'Incubation'-class targeted observable in the 2018 NASEM Earth Science Decadal Survey (National Academies of Sciences and Medicine 2018). Due to the high degree of spatial and temporal heterogeneity of the near surface processes, including diurnal variations and complex interactions between the land/ocean surface and atmosphere, the PBL has been challenging to accurately simulate and observe.

Given the capability of advanced data assimilation systems combining model physics with data from multiple observing systems coherently to provide optimal initial conditions for models, an increasingly large number of observations have been assimilated in the Goddard Earth Observing System (GEOS) global hybrid 4D Ensemble-Variational (4DEnVar) data assimilation system at the Global Modeling and Assimilation Office (GMAO). The GEOS data assimilation system is aimed to constrain the PBL atmospheric thermodynamic structure and to reduce the uncertainties of land surface model and PBL parameterization schemes, but in this study we will focus on the lower troposphere for simplicity and set the stage for follow-on PBL work. While each individual available observing system offers unique advantages in measuring the earth, however, there is no perfect observing system for the lower troposphere. Conventional data such as radiosonde data provide reliable temperature, specific humidity and wind profiles, but they are mainly concentrated over land and lack adequate temporal resolution to capture diurnal variations. Global Navigation Satellite System (GNSS) radio occultation (RO) data has high-vertical but coarse along-ray resolution. Furthermore, they are currently not used or assigned very large observation errors in the low troposphere, where super-refraction is common and biases are large. Satellite radiance data have good global coverage, but introduce other challenges. Microwave observations can penetrate through clouds but with broad weighting function and coarse resolution; hyperspectral infrared observations offer high spectral resolution but are unable to provide information beneath clouds. The various limitations in the existing observations, or in our ability to use them, have made it essential to use these observations coherently in the GEOS global data assimilation system. So far, no known study has been performed to comprehensively assess the impact of the various observing systems on the analysis and forecasts of the lower troposphere.

Several approaches can be used to assess observation impact in data assimilation systems, such as forecast-based sensitivity observation impact (FSOI, Langland and Baker (2004), Gelaro et al. (2010)) and observation-based observation-minus-forecast residuals (Todling (2013). Another commonly used way of studying the impact of different observations in data assimilation systems is to employ the approach of observing system experiments (OSEs). Many previous OSE studies have focused on free atmosphere above the PBL, including the stratosphere. For example, Kelly and Thépaut (2007) and Lord et al. (2016). Recently, Duncan et al. (2021) investigated the impact of microwave sounders on the analysis and model forecast in the ECMWF system; Lawrence et al. (2019) investigated the impact of observations in the polar regions. The present study investigates the impact of observations in the low troposphere using the GEOS atmospheric data assimilation system (ADAS). The goal is to try to identify specific weaknesses of data usages associated with analyzing and predicting thermodynamic structure of the lower troposphere. Additionally, results and discussion are presented to illustrate the model responses to analysis increments. It is well known that rapidly changing physical, rather than hydrodynamical, processes are hard to constrain with data assimilation; basically the model physics tends to forget rather quickly changes induced by the assimilation of observations and falls back into its own mechanisms. The responses of model physics and hydrodynamics tendencies are worth closer examination.

This article is organized as the following: section 2 gives an overview of the GEOS global hybrid–4DEnVar data assimilation system; a brief summary of the observations used in GEOS ADAS is given in section 3 by looking at a summary of its FSOI tool. Data denial experiments results are presented in section 4, model responses to analysis increments and evolutions are discussed in section 5, and conclusions are provided in section 6.

2. Brief overview of the GEOS hybrid–4DEnVar ADAS

The GEOS atmospheric data assimilation system is a hybrid–4DEnVar system that produces estimates of global atmospheric states by analyzing observations within a 6-hour time window. The deterministic hybrid–4DEnVar uses the Gridpoint Statistical Interpolation (GSI) of Kleist et al. (2009) with the preconditioning strategy of El Akkraoui et al. (2013). Its cost function can be written as

$$\begin{aligned} J(\delta \mathbf{x}_k, \mathbf{a}) &= \frac{1}{2}\beta_c \delta \mathbf{x}_c^T \mathbf{B}_c^{-1} \delta \mathbf{x}_c + \frac{1}{2}\beta_e \delta \mathbf{a}^T \mathbf{L}^{-1} \delta \mathbf{a} \\ &+ \sum_{k=1}^{k=K} (\mathbf{H}_k \delta \mathbf{x}_k - \mathbf{d}_k)^T \mathbf{R}^{-1} (\mathbf{H}_k \delta \mathbf{x}_k - \mathbf{d}_k) \\ &+ J_d\,, \end{aligned} \tag{1}$$

where $\delta\mathbf{x}$ is the total analysis increment; $\mathbf{d}$ is the observation-minus-background departure vector at time t_k; $\mathbf{H}$ is the Jacobian of the nonlinear observation operator; $\mathbf{a}$ is the extension to the control vector account for the ensemble contribution; $\mathbf{B}_c$ and $\mathbf{R}$ are the prescribed climatological background and observation error covariances; $\mathbf{L}$ is a localization matrix; β_c and β_e represent weights given to the climatological and ensemble background terms. The last term, J_d, represents additional constraints, for example, a dry mass conservation term (see Takacs et al. 2016). Minimization of the cost function leads to a four dimensional increment, which is given as the sum of the contribution from the climatological term $\delta\mathbf{x}_c$ and a term composed of linear combination of ensemble perturbations $\delta\mathbf{x}^m$ and optimal coefficients $\mathbf{a}^m$, that is,

$$\delta \mathbf{x}_k = \delta \mathbf{x}_c + \sum_{m=1}^{M} \mathbf{a}^m \bullet \delta \mathbf{x}_k^m\,. \tag{2}$$

Here the symbol $\bullet$ stands for the Hadamard-Schur (element–wise) product of two vectors, and the ensemble of forecast perturbations $\delta\mathbf{x}_k^m$ are derived from running the ensemble square-root filter (EnSRF) of Whitaker et al. (2008). In the current GEOS ADAS, the coefficients β_c and β_e change with the vertical analysis levels, equally weighting the two terms up to about 5 hPa, and smoothly transitioning to a purely climatological error term above that (see Todling and El Akkraoui 2018).

The *assimilation* of observations in GEOS ADAS is performed through a 4D incremental analysis update (IAU). Instead of using the 4D incremental solution provided by the minimization of (1) to correct model initial conditions and subsequent model states at given frequency (hourly), the incremental solutions are used to form tendency terms that are applied at each model time step during a so-called corrector interval that lines up with the 6-hour assimilation window. In addition, the present formulation of IAU is a revision of the Bloom et al. (1996) version, following Takacs et al. (2018), that guarantees IAU stability by modulating the tendencies with a digital filter.

Background fields for the next assimilation cycle are generated by a so-called predictor step that integrates the GEOS atmospheric model for an extra 6–hour period past the IAU corrector with the analysis tendency terms set to zero.

The GEOS Atmospheric General Circulation Model (AGCM) relies on a non-hydrostatic version of the cubed-sphere finite volume hydrodynamics (see Putman and Lin 2007). Its current physical processes include the short- and long-wave components of the Rapid Radiative Transfer Model for GCMs (RRTMG; Clough et al. 2005; Iacono et al. 2008); the deep convection parameterization of Freitas et al. (2018); a catchment land-surface model consistent with the level-4 GMAO SMAP products (Reichle et al. 2018); and gravity wave drag follows McFarlane (1987) and Garcia and Boville (1994). The single-moment cloud physics are based on Bacmeister et al. (2006). Of greatest relevance to the present work are the parameterizations of the boundary layer. More specifically, these consist of a non–local K–profile scheme driven by surface and cloud–top buoyancy fluxes (Lock et al. 2000), and a local scheme for stable conditions based on the Richardson number (Louis and Geleyn 1982). The Lock scheme releases parcels upward from the surface and downward from stratocumulus cloud top to determine the depth of an analytic profile of diffusivity. Above the well-mixed layer defined by the Lock surface-driven diffusivity, shallow cumulus convection is represented by the Park and Bretherton (2009) buoyancy-sorting mass flux scheme.

A diagnostic component of GEOS that is of relevance to the motivational part of the present work is its forecast-based sensitivity to observation impact (FSOI) tool. The GEOS FSOI implements a combination of the approaches of Langland and Baker (2004) and Trémolet (2007) that allows assessing the contribution of individual observations to reducing errors in 24–hour forecasts. For that, it employs a linearized moist global energy norm that serves to transform the impact of different quantities into units of energy (J/kg) (e.g., Errico et al. 2004). Use of a *moist* energy component in the norm requires proper representation of linearized moist processes in the model adjoint needed for the produce and the generation of 24-hour forecast sensitivities. Details of the latest version of such processes is found in Holdaway et al. (2015).

The near-real-time GEOS ADAS is a 12.5 km system that relies on a 50 km ensemble. This work employs a lower horizontal resolution version of GEOS, that runs the deterministic cycle at 25 km, and the ensemble cycle at 100 km. GEOS uses 72 vertical levels in all its components. Several factors can alter observation impact results, even the rankings of relative importance of observations. A few examples of such factors include change in the observing system, changes in the data assimilation algorithm, changes in the model, and changes in the weights given to the observations and background fields; horizontal and vertical resolution might also affect the assimilation of observations. The reduced (yet still reasonably high) horizontal resolution used in this study has been carefully chosen and frequently employed to evaluate the operational GEOS system and system upgrades. The impact of such horizontal resolution has been found to be secondary except for extreme weather conditions, and forecast skills and biases approximate those of the operational system. Even in this somewhat reduced resolution configuration the high computational resources requirements and the slow turnaround of experiments, led to the adoption of a conservative approach to conduct the data denial OSEs of this study. That is, the data denial experiments were set up to use the same ensemble backgrounds generated in the control experiment; this latter assimilates the complete set of observations and is set to exercise the entire machinery of the hybrid data assimilation system. The approach of using a given set of ensemble members is referred to as ensemble replay mode. A recent study by Duncan et al. (2021) finds that using a given (fixed) ensemble in various OSEs amounts to about 10% of the total change due to changes in the observing systems introduced in the various experiments, but still reliably represent the impact of such changes.

3. GEOS observations and FSOI

The version of GEOS ADAS used in this work is an *upgrade* to the then-current operational system when it was used to process the mid-November-December 2019 period covered here. The upgrade involves only changes to the analysis component, and more specifically to the underlying observing system, by adding: all-sky Advanced Microwave Scanning Radiometer 2 (AMSR-2), COSMIC-2 Global Positioning System Radio Occultation (GNSSRO), and the full spectral resolution (FSR) version of the Cross-track Infrared Sounder (CrIS) from both Suomi NPP and NOAA-20 (see Todling et al. (2022)). In the average, each 6-hour cycle assimilates roughly 4.5 million observations; the experiments here neglect a two–week spin–up in November 2019.

As a motivational introduction to the investigation of how the observing system affects the lower troposphere behavior in GEOS, we start by looking at diagnostics produced by exercising two different configurations of FSOI in the control experiment, which uses the complete set of

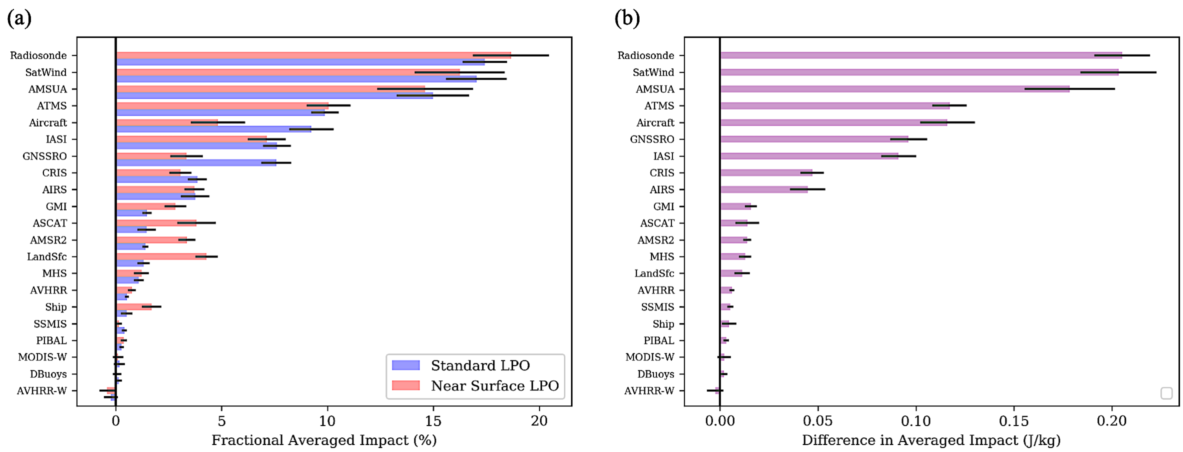

FIG. 1. Comparison of FSOI for 24-hour forecasts from all 0000 UTC analyses of December 2019 when two projections (LPO) of the norm are used in the calculations, namely, using the standard set of vertical levels, and using only near surface levels (at and below 850 hPa). The panels show: (a) fractional averaged impacts (%) in each case; and (b) the averaged impact differences obtained after subtracting the standard LPO from the near-surface LPO results (J/kg), in descending order. Error bars show 95% confidence in fractions and differences, respectively.

observations as laid out in Table 1. As mentioned in section 2, FSOI relies on a linearized total moist energy norm. The operator, $\mathbf{E}$, corresponding to this norm is enveloped by a so–called (diagonal) local projection operator (LPO) $\mathbf{S}$, as in $\mathbf{S}^T\mathbf{E}\mathbf{S}$, containing zeroes and ones along its diagonal and enabling, for example, restricting the norm to particular variables, regions, or levels. The standard LPO configuration of FSOI in GEOS avoids forecast errors within the sponge layer by excluding the top six levels of the model. For the purposes of the present work, an alternative LPO considers only forecast errors from the lower troposphere, which is set as the lowest eleven model levels (roughly at and below 850 hPa).

Evaluation of FSOI using the two LPOs described above for 24-hour forecasts for all 0000 UTC analyses in the month of December 2019 appears in Fig. 1. Panel (a) compares fractional FSOI (%) using the standard LPO (blue bars) with results when the near surface LPO (red bars) is used. Satellite winds, radiosonde observations, and MW radiances from Advanced Microwave Sounding Unit-A (AMSU-A) and Advanced Technology Microwave Sounder (ATMS) are among the top contributors in reducing forecast errors in both LPO configurations. Changing from the standard

TABLE 1. Control and OSEs Definitions

Name	Removed Observation	DA strategy
XCTL	None	Deterministic hybrid, full ensemble
XNOIR	Infrared radiances	Deterministic hybrid, ensemble-replay
XNOMW	Microwave radiances	Deterministic hybrid, ensemble-replay
XNOSATW	Satellite winds	Deterministic hybrid, ensemble-replay
XNOCONV	Conventional	Deterministic hybrid, ensemble-replay
XNOSURF	Surface	Deterministic hybrid, ensemble-replay

LPO to a near–surface LPO configuration leads to a slight reduction in the fractional impact of satellite winds, AMSU-A, Infrared Atmospheric Sounding Interferometer (IASI), and CrIS. The fractional contribution from GNSSRO is considerably reduced in comparison to what is seen in the default settings; this is similar to the reduction seen for aircraft observations. In contrast, the fractional impact of radiosondes, ATMS, Global Precipitation Measurement (GPM) microwave imager (GMI), and Advanced Very High Resolution Radiometer (AVHRR) is slightly increased when compared to the default LPO. The most noticeable increase in fractional impact is seen for land surface observations, followed by Advanced Scatterometer (ASCAT), Advanced Microwave Scanning Radiometer 2 (AMSR2), and ships observations. The error bars in panel (a) show 95% confidence levels in the fractional results. With the exception of results for AVHRR and MODIS winds and drifting buoys, which are not statistically significant, all others are within acceptable levels. Corroboration of the statistical significance of the averaged impact difference (J/kg) between the near–surface and the standard LPO configurations is provided in panel (b). Results are shown to be statistically significant for most of the components of the observing system. Since the standard LPO results have larger negative values and the near–surface LPO results have smaller negative values, where negative values indicate positive impacts, all the differences in panel (b) are positive except AVHRR Wind. The near–surface LPO removes forecast sensitivities above 850 hPa, thus FSOI derived at and below 850 hPa are only affected by the sensitivities at and below 850 hPa - this reduces quite significantly the magnitudes of FSOI relative to those derived with the standard LPO.

Even with a confined near–surface LPO, both AMSU-A and ATMS still show considerable fractional contribution to forecast error reduction. This is even more peculiar since the GEOS ADAS analysis does not assimilate window channels and very low–peaking temperature sounding channels (1–3 and 15) from AMSU-A, and corresponding channels (1–4 and 16) from ATMS. Closer examination reveals that the impact from these instruments is dominated by their low–to–mid–peaking temperature channels, namely channels 5–7 for AMSU-A, and 5–8 for ATMS, when using the standard LPO. The fact that even under a near–surface LPO these instruments contribute substantially to fractional impact is attributed to the broad weighting functions associated with these channels (not shown). The two other MW sensors, namely AMSR2 and GMI assimilated in all-sky conditions, are also seen to contribute substantially to fractional impact with the near–surface LPO. The two window channels (23.8 GHz V[1] and 36.5 GHz V) of AMSR2 and GMI show similar impact as seen from the low–to–mid–peaking AMSU-A and ATMS temperature channels. The remaining three GMI channels (166 GHz V, 183.31 ± 3 GHz V, and 183.31 ± 7 GHz V) are sensitive to water vapor and snowfall and are seen to have little impact on 850 hPa and below (not shown).

Regarding the impacts of hyperspectral infrared radiances, they are affected by the channel selections from each sensor, data usages over different surface types, and quality control procedures. Compared with AIRS radiances, CrIS radiances are used in a much more conservative way in GEOS. CrIS window channels are not used over non-water surface types, and surface-sensitive radiance observations are also excluded if brightness temperature Jacobians with respect to surface skin temperature are larger than 0.2. Therefore, it is understandable that CrIS's fractional impact in the near surface LPO is lower than in the standard LPO. The seasonal effects of the FSOI results shown in this work have not been studied. To the extent that forecast errors vary seasonally there might be some seasonality in the impacts, but experience from looking at operational impacts suggests that such effects are rather secondary.

4. Data denial OSEs

Motivated by the differences seen in the FSOI results discussed above, Table 1 lists a set of data denial experiments designed to look more closely at the impact of various observing systems with the primary focus of examining the lower troposphere response. The present work focus exclusively on the layer at and below 850 hPa, considered here to be the lower troposphere. We intend to identify the observing systems that have large impacts in the lower troposphere and that are not used effectively or have small impacts. How deep the observation impact penetrates into the lower

[1]V stands for vertical polarization.

troposphere and the length that observation impacts last with the forecast lead time are also examined. Evaluations are performed against a control (**XCTL**) experiment that uses all observations and exercises the full deterministic and ensemble ADAS machinery. The OSEs listed in Table 1 systematically remove key components of the observing system: **XNOIR**, removes all infrared radiance observations; **XNOMW**, removes all MW radiance observations; **XNOSATW**, removes all satellite-derived wind observations; **XNOCONV**, removes all so-called conventional observations[2], and finally, experiment **XNOSURF**, removes all surface observations, and is designed specifically to evaluate the response of model processes to near–surface observations. Actual evaluation of results is done either with respect to the control experiment or with respect to ERA5 analyses (Hersbach et al. 2020), as duly indicated.

a. Impact on specific humidity

According to the PBL Incubation Study Team Report (Teixeira et al. 2021), a key component to improve modeling of PBL thermodynamics processes is the ability to optimally assimilate PBL observations globally. To this extent, we start by examining the analysis of specific humidity in the model lowest levels. The difference of the control (**XCTL**) analyzed December 2019 mean, specific humidity at 850 hPa with the corresponding ERA5 monthly mean analysis appears in Fig. 2a. It shows that, at the resolution of the experiments here, GEOS ADAS is drier over Tropical oceans than ERA5; wetter over the southern oceans in the latitude band from roughly 40-60^{o}S and over the Northern Hemispheric Pacific and Atlantic storm tracks. Over South America GEOS ADAS seem wetter than ERA5 in the Amazonian rainy season; over the western African GEOS ADAS is drier than ERA5 in the area's dry season.

To facilitate comparison, the remaining panels of Fig. 2 show closeness plots of monthly mean analysis to ERA5 for each of the OSEs in Table 1 and the control experiment. That is, these panels show $|\mathbf{OSE}-\mathbf{ERA5}|-|\mathbf{XCTL}-\mathbf{ERA5}|$; hot colors indicate OSE is further away from ERA5 than the control experiment. With that, it is clear that denying microwave radiances (**XNOMW**; panel b) exacerbates the differences of the control with ERA5, turning the results further drier in the Tropical oceans. Although results in western Africa seem mixed, results in South America seem to move further away from ERA5. Only minor, mixed, changes are seen in the extratropical

[2]The wording *conventional observations* is somewhat of a misnomer as it stands for radiosondes, pilot bollons, aircraft and a host of truly traditional observations.

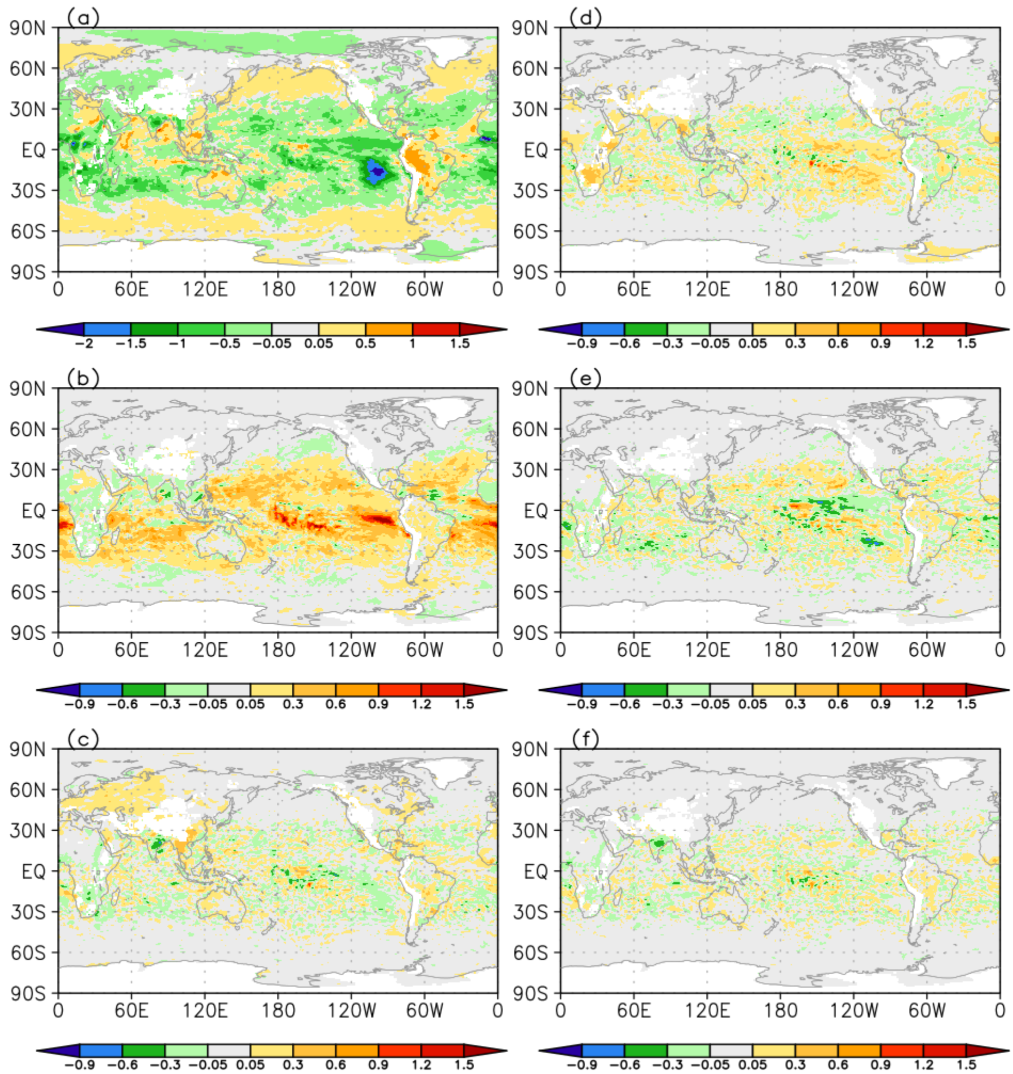

Fig. 2. Panel (a): Difference of analyzed December 2019 mean specific humidity (g/kg) analysis, at 850 hPa, and corresponding ERA5 analysis for **XCTL** experiment. Panels (b)-(f): closeness of given OSE and CTL to ERA5 monthly mean analysis, that is, difference of the absolute difference of given OSE with ERA5 and the absolute difference of control with ERA5: (b) **XNOMW**, (c) **XNOCONV**, (d) **XNOIR**, (e) **XNOSATW**, and (f) **XNOSURF**, at 850 hPa. Notice different color scale in panel (a). Locations below the surface are marked as white.

oceans. Generally, all other OSEs suffer considerably less from their corresponding observing system denial than when microwave is denied. Some noticeable exceptional differences can be seen in Europe when conventional observations are removed (panel c), and a slight move in the opposite direction to that of microwave, in the tropical oceans, when satellite winds are denied (panel e). No clear signals are seen when surface observations are denied (panel f).

Forecast skill scores for the control experiment and all OSEs have also been calculated, and are discussed here when verified against ERA5 analyses. Each panel of Fig. 3 shows globally–averaged root mean square error (RMSE) differences from the control with boxes representing 95% confidence interval for the associated RMSE difference, at selected levels. The largest increase in RMSE is due to denying microwave observations (red curves), with results being statistically significant. Loss of skill due to microwave is felt throughout the 5-day forecast at all levels displayed in the figure, though its significance decreases with increased forecast lead time. To a lesser extent than when denying microwave, loss in skill due to denying conventional (blue curves) and IR (green curves) observations is also statistically significant with the effect lasting throughout the 5-day forecast. The significance of denying conventional observations becomes more comparable with that of denying microwave observations as we approach the surface. The impact on specific humidity from denying satellite winds (purple curves) and surface (yellow curves) observations is insignificant at 850 hPa. At lowest levels, satellite winds are seen to have small positive impact in the short forecast lead times, but turn slightly negative at longer forecast lead times; the influence of surface observations is small, with neutral to slightly positive impact observed toward the end of the forecast.

The regional influence on specific humidity at low levels from the denial experiments of Table 1 is shown for the Northern Hemisphere (NHE) and Southern Hemisphere (SHE) in Figs. 4–5. In the Northern Hemisphere, conventional observations are the most influential data type (Fig. 4a) in terms of mean forecast RMSE. Their impact stretches from the surface to 200 hPa throughout the 5-day forecasts (Fig. 4b), with the largest impact in the lower to mid- troposphere. At 925 hPa and below, microwave radiances have much smaller impact, and infrared radiances can be negligible. This may be partially because very limited surface-sensitive radiances are assimilated over land. As expected, microwave radiances contribute the most in the Southern Hemisphere, followed by infrared radiances (Fig. 5a), but the large impact observed from 925 hPa to above 800 hPa decreases

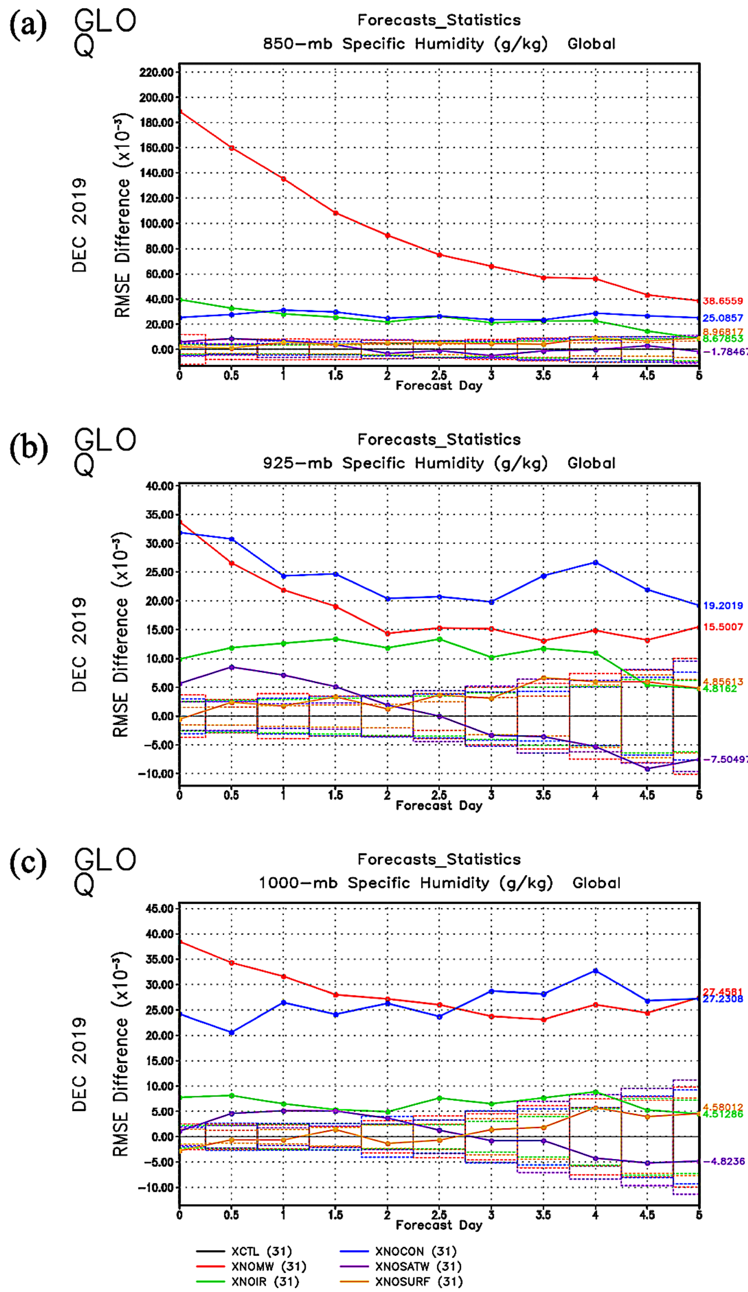

FIG. 3. Globally–averaged RMSE difference from the control for all 0000 UTC 5-day forecasts for December 2019 of specific humidity at (a) 850, (b) 925 and (c) 1000 hPa, with boxes representing 95% statistical confidence interval for each of the RMSE difference curves. The RMSE for the control and each OSE is calculated wrt ERA5 analyses. Curves are for control (black), **XNOMW** (red), **XNOIR** (green), **XNOCONV** (blue), **XNOSATW** (purple), and **XNOSURF** (yellow).

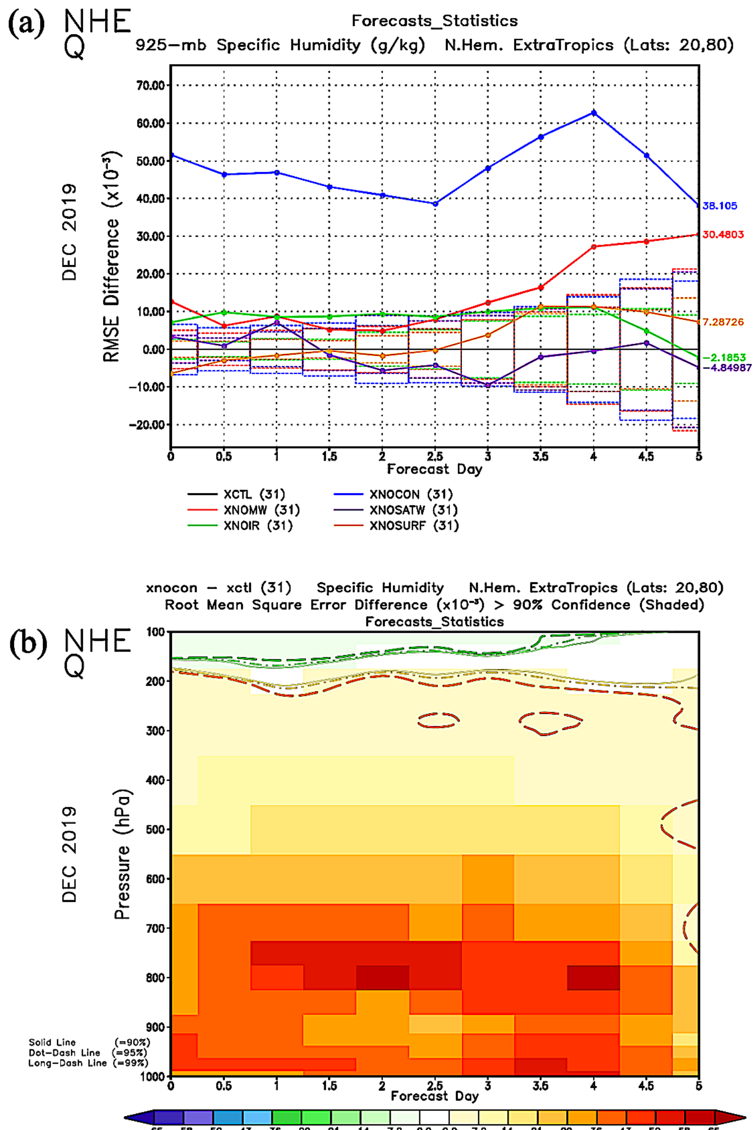

FIG. 4. Panel (a), as in Fig. 3b, but for Northern Hemisphere. Panel (b), globally-averaged specific humidity RMSE difference (g/kg) between **XNOCONV** and **XCTL**, as a function of pressure levels, for the Northern Hemisphere 0000 UTC forecasts of December 2019. Shaded areas highlight results that are statistically significant with 90% confidence. The darker the shading corresponds to larger differences. The solid, dot-dash, and long-dash lines correspond to the confidence intervals of 90%, 95%, and 99%.

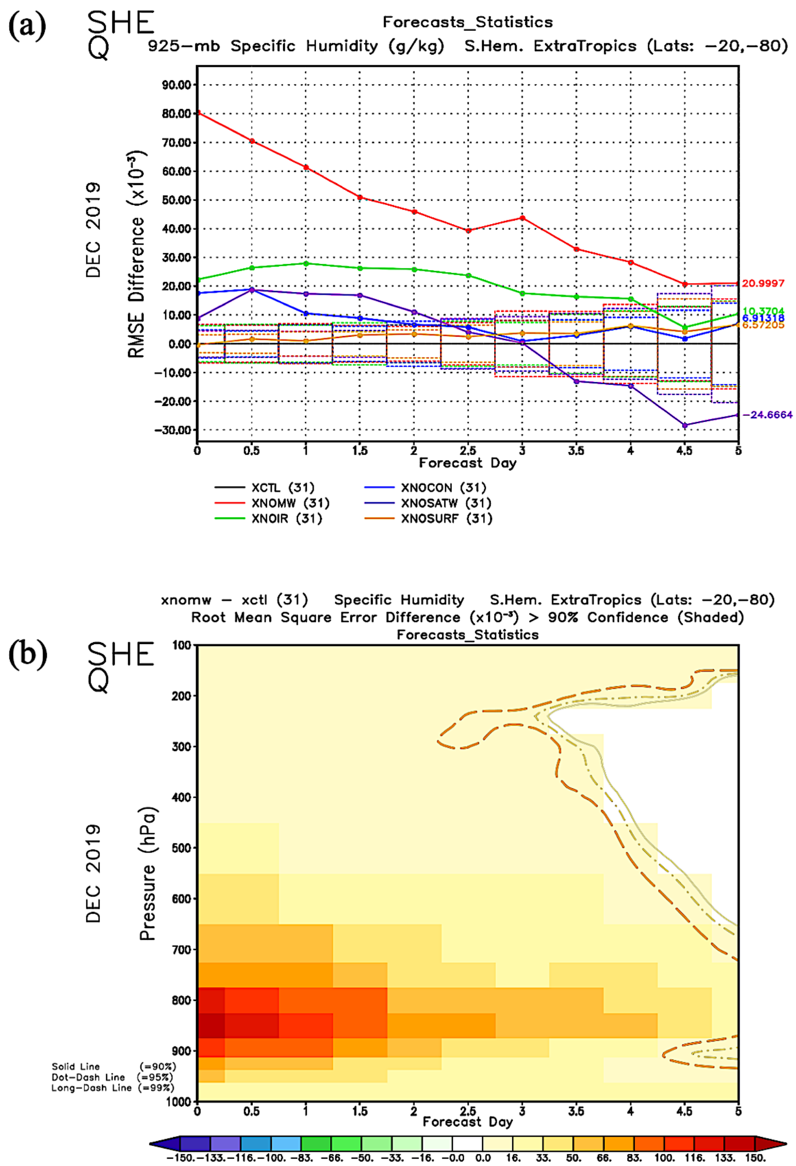

FIG. 5. As in Fig 4, but for Southern Hemisphere; and for panel (b) comparing **XNOMW** with **XCTL**.

with the forecast lead times (Fig. 5b). In the Tropics (figure not shown), while conventional data have the largest impact at 925 hPa and below, microwave radiances dominate at 850 hPa followed by infrared radiances and conventional data, and the effect of microwave radiances is most noticeable between 875 and 700 hPa but much smaller or neutral impact below. Overall, the impact from MW radiances tends to decrease quickly as forecast lead time increases while the impact of denying conventional observations stands well into the 5-day forecast.

b. Impact on temperature field

As displayed in Fig. 6a, the impact of microwave radiances on temperature forecasts can be seen more clearly in the Tropics, where its impact is neutral to negative below 925 hPa throughout the 5-day forecast lead times. Their impact below 900 hPa in the Northern and Southern Hemisphere is also negligible in the first day or two of the forecast.

Similarly, Fig. 6b suggests that infrared radiances mostly contribute to temperature forecast error reduction above 925 hPa in the Tropics (and also Southern Hemisphere). Loss of skill in temperature forecast due to denying infrared observations lasts up to about day 4, whereas considerable losses due to removal of microwave observations are felt throughout the 5-day period between 600 and 925 hPa (Fig. 6a).

Like its impact on specific humidity, conventional data dominate the observation impacts on the lower troposphere in the Northern Hemisphere. When they are removed, forecasts degrade for the 5-day forecast period from surface to 100 hPa with the most severe degradation in the lower troposphere. In Tropics, considerable impacts from conventional data are also observed around 900 and 300-550 hPa (figures not shown).

c. Impact on wind field

Although wind is not the focus of the PBL Incubation Study Team report (Teixeira et al. 2021), some of the denial experiments of Table 1 show non-negligible consequences to predicting winds in the lower troposphere. Conventional data and microwave radiances are the two most effective observation sources for reducing wind forecast RMSE globally from the surface to 850 hPa. Infrared radiances have little or neutral impact in this layer. The most interesting result is the positive impact from microwave radiances. While some radiances are affected by surface wind

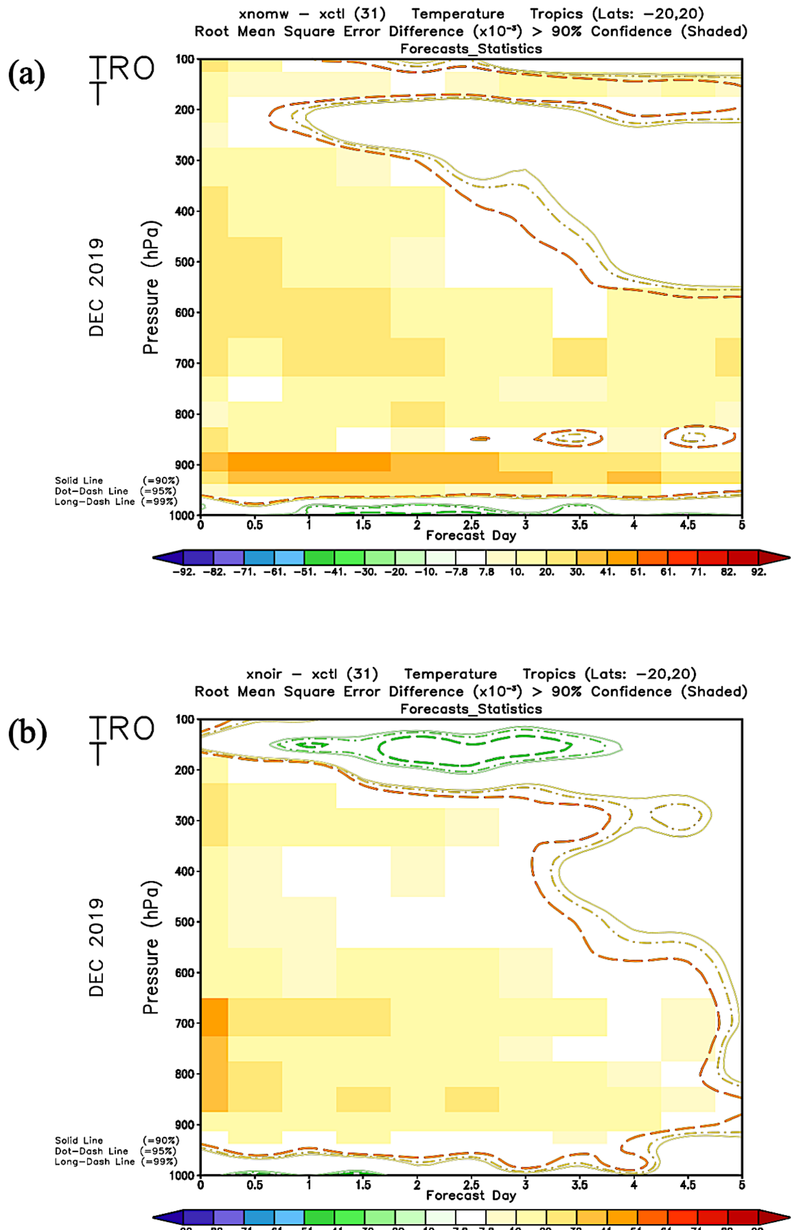

FIG. 6. Panel (a): similar to Fig. 5b, but for temperature (K) in the Tropics; panel (b): similar to panel (a), but comparing experiment **XNOIR** with **XCTL**.

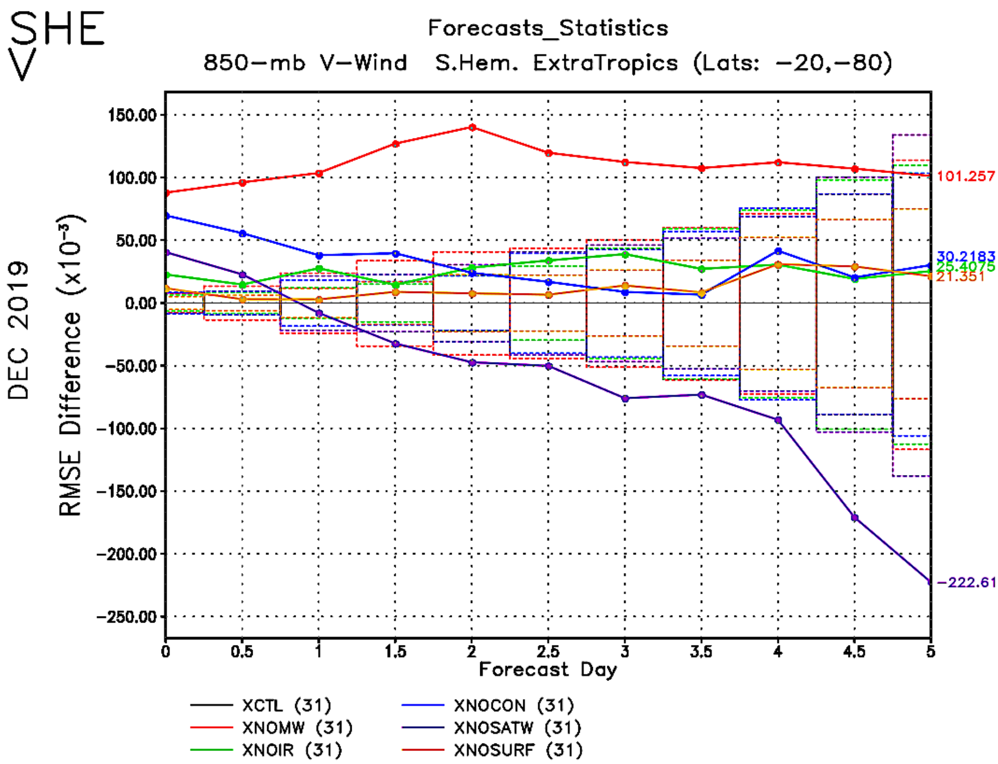

FIG. 7. As in Fig. 5a, but for meridional wind (ms^{-1}) at 850 hPa.

speed, a considerable impact in the lower troposphere may be through background error covariance and tracer effect when thermodynamic structure is improved by using the microwave radiances. For example, Fig. 7 illustrates that in the Southern Hemisphere, microwave radiances have the largest impact in reducing forecast errors at 850 hPa, followed by conventional data.

However, one of the most unexpected results obtained from examining the wind fields is the improvement in forecast scores when satellite winds are removed. This is seen in Fig. 7 (**XNOSATW**; purple curve), with the improvement being retained and increasing throughout the length of the 5-day forecast. Careful examination reveals this to be the case at other levels, as well, especially along the jet–stream level (not show). Another illustration of the undesirable improvement obtained when satellite winds are removed, from the version of GEOS ADAS used in this work is seen when looking at the ratio of observation–minus–background standard deviations between **XNOSATW** and **XCTL**, for various instruments. Figure 8 displays this ratio in the Southern Hemisphere for (a) ATMS and (b) IASI radiance observations; improvements appear as the solid black curve falls below 100% (grey, vertical curve). A careful evaluation reveals this unexpected result to be a consequence of a bug in the GEOS system with no thinning being applied to GOES-R satellite wind observations; results improve dramatically after the bug is fixed with proper thinning being applied.

d. Impact on PBL height

The height of the PBL is a relevant parameter used in air quality studies and mixing of aerosols in the turbulence layer. Though careful evaluation of the consequences in changes to PBL height and its diurnal cycle are beyond the scope of the present work, it is still worth seeing how the various OSEs here change this quantity. In the GEOS system, the PBL height uses different definitions over land and ocean. Over land, it is based on the bulk Richardson number with a critical value of 0.25, while over ocean it is based on the profile of diffusivity from the turbulence parameterizations, with a threshold of 10% of the maximum diffusivity. Figure 9 shows the diagnosed December 2019 averaged PBL height for the control experiment (panel a), and how PBL height differs for a given OSE from the control (remaining panels). Denial of microwave observation (panel b) tends to lower PBL height in the Southern Hemisphere oceans at this time of the year, and slightly increase it in the Tropics and Northern Hemisphere. This is somewhat similar to what happens when infrared observations are denied (panel d) though an increase of PBL height is observed over tropical lands in this case. The effect from denying conventional observations (panel c) is more visible in the Northern Hemisphere and Tropics with the inner portions of North America and Asia showing a decrease in PBL height and the coastal areas showing an increase. The effects from denying satellite winds is just as large as that of denying microwave radiances, especially in certain areas: noticeably off the Pacific coastal areas of North and South America where stratocumulus clouds play a significant role and are considerably affected by the winds in those regions in this case. It should be noted that some of the effects from satellite winds seen here are caused by the improper handling of GOES-R satellite winds in the GEOS system as discussed in Section 4c. Comparatively speaking, the absence of surface observations (panel f) amounts to relatively small changes in PBL height, which is perhaps indicative of how little the present data assimilation uses such observations to constrain this variable.

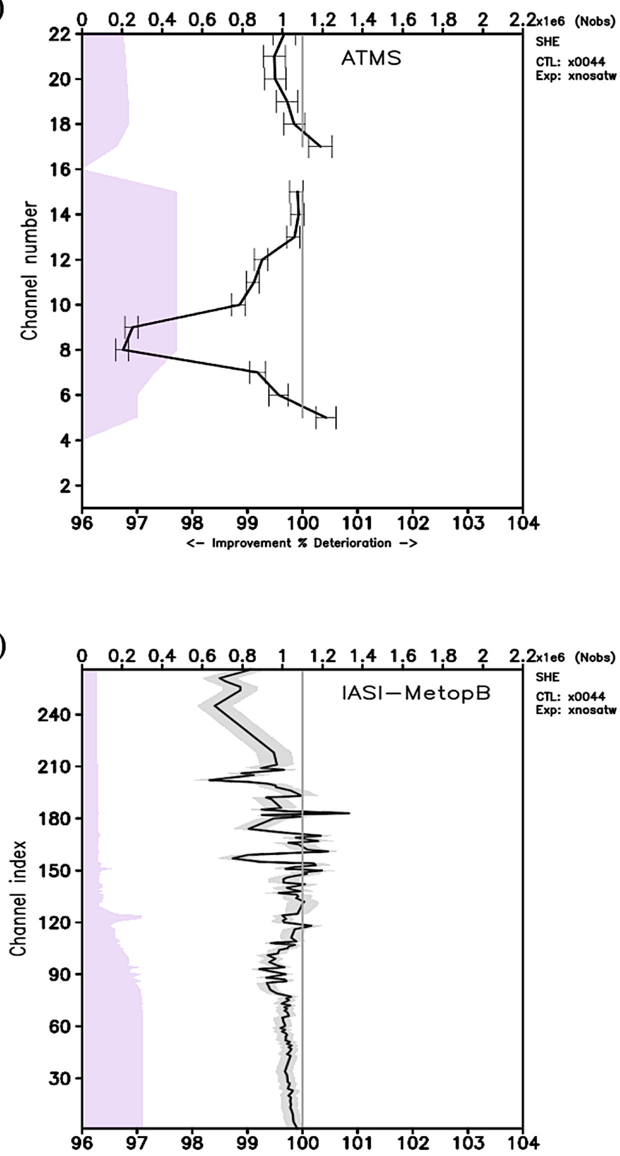

FIG. 8. Observation–minus–background standard deviation ratio of experiment (**XNOSATW**) and control for (a) ATMS and (b) IASI in the Southern Hemisphere, for December 2019. Bottom x–axis show percentage, magenta shade and top x–axis shows observation count, and bars and grey shades represent 95% significance.

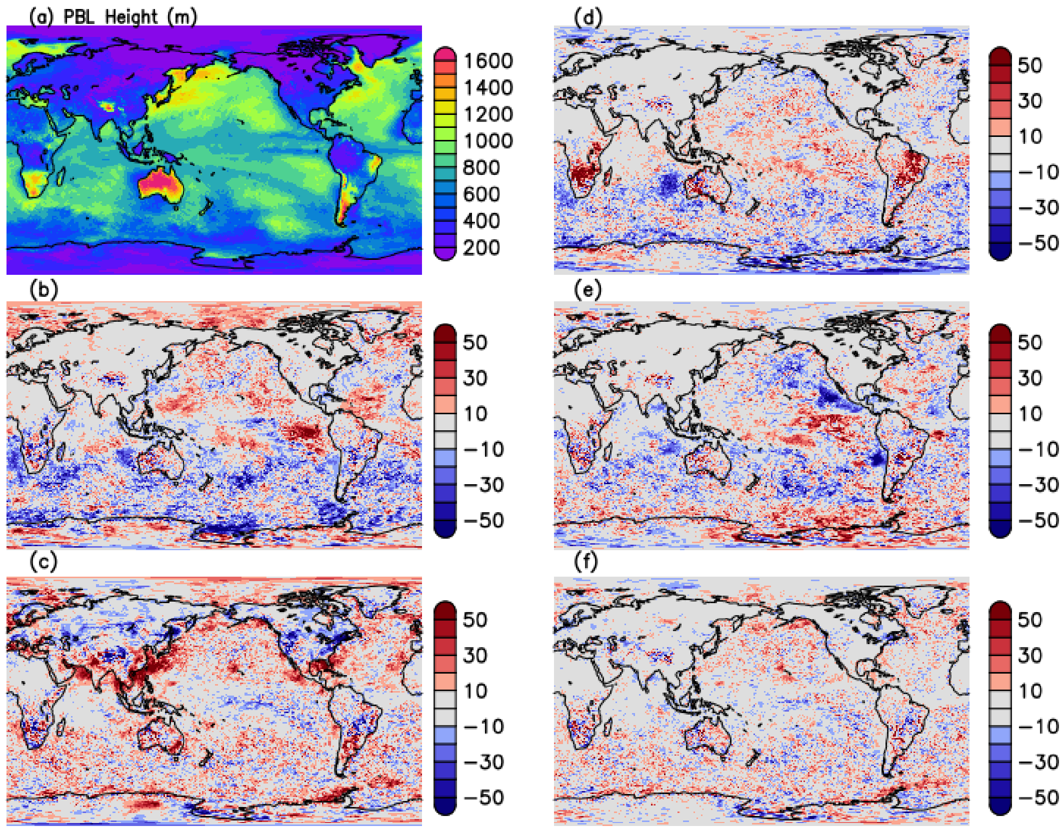

FIG. 9. Averaged December 2019 PBL height (m) for the control experiment (panel a) and difference for each OSE: (b) **XNOMW**, (c) **XNOCONV**, (d) **XNOIR**, (e) **XNOSATW**, and (f) **XNOSURF**.

e. Impact of surface observations

Early forecast models had low vertical resolution and simply had surface heat exchange, momentum drag and vertical diffusion, with consequent highly dissipative surface layers. Thus, forecasts in these layers were relatively insensitive to changes of fields aloft which would overwhelm the mechanisms for error growth and vice verse. Modern PBL schemes (e.g. Lock et al. 2000) are more sophisticated, incorporating vertically non-local forcing and effects upward from the surface as well as downward from the layer's top. As removing the surface observations in experiment **XNOSURF** introduces changes mainly to the analyses in the lowest levels, it is worth examining how the atmosphere responds to such changes.

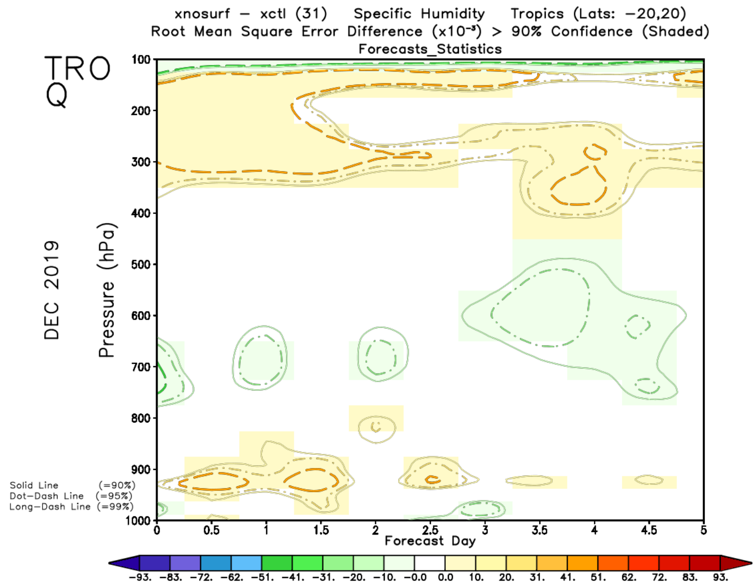

FIG. 10. As in Fig. 4b, but for comparing **XNOSURF** with **XCTL** in the Tropics.

Figure 10 shows specific humidity RMSE difference between **XNOSURF** and **XCTL** in the Tropics for the month of December 2019. It is seen that removing surface observations degrades upper atmosphere forecast skill from about 150 to 350 hPa in the Tropics persistently. Similar degradation is also observed from 100 to 200 hPa in the Northern and Southern Hemisphere up to day 4. Although the degradation magnitude is small, they are statistically significant at confidence value of 90%. The surface observations have mixed impact on the mid- and lower atmosphere. Removing surface observations makes forecast skill worse below 800 hPa but better in the mid-atmosphere in the Tropics, and the forecast skills below 900 hPa become better for up to one and a half days in the Northern Hemisphere. Removing surface observations, however, has no statistically significant impact on temperature and wind fields. In this experiment where surface observations are excluded, the results indicate that changes at or near surface do affect mid- and upper- atmosphere in the GEOS system. As more PBL observations will become available in the future, more studies should be performed to investigate how they affect not only the PBL simulation but also the above free atmosphere.

5. Model tendency responses to the IAU forcing

In a traditional data assimilation framework, analysis increments are used to intermittently correct the model's initial condition at every cycle. When the increments change the initial state in a way that is not consistent with the model hydrodynamical and physical processes the model tends to reject the observational information. Geostrophic adjustment, happening in time scales of minutes to hours, is one such well–known dynamical process the model goes through to accommodate to undesirable changes in its state induced by the analysis. When components of the initial state prove inconsistent with the physical processes, the model may reject the analysis information and return to its desired state in a single time step. One such example is discussed in Zhu et al. (2016, Fig. 16 there), where parts of the cloud analysis increment are simply rejected after the first pass through the physics term in the model integration. In this example, it was found that an inconsistency between the model clouds and the relative humidity analysis partially explained why the model readily ignores the cloud analysis increments. More generally, the way analysis increments are presented to the model, and whether they are dynamically and physically consistent with the model underlying processes determines how the model retains information from the observations. This consistency is well understood when it comes to dynamical balances, such as geostrophic balance, and it has led to careful development of initialization algorithms (e.g., Kleist et al. 2009, and references therein), and it has also led to procedures to develop dynamically consistent background error covariances (see Bannister 2008, and references therein).

As mentioned in Section 2, instead of correcting the model through traditional intermittent initial condition updates, the GEOS data assimilation system employs an IAU procedure that presents the analysis corrections as tendency terms that are continuously applied to the model during the 6-hour assimilation window around the analysis time, [-3h, +3h]. Whether in 3D, or its 4D formulation used in the current GEOS hybrid system, a digital filter modulates the analysis tendencies (Takacs et al. 2018) in a way that guarantees a smooth transition from one 6-hour assimilation cycle to the next. After the first pass of the IAU tendency, model hydrodynamics and physics tendencies start to evolve and depart from the original free model forecast path where IAU forcing is not added. In this section, the changes of model tendencies induced by the IAU forcing in the assimilation window and beyond are examined, and hopefully this information will be helpful for a future study on improving observation retention in model forecast.

For a simplified and idealized forecast model, as shown in the Appendix, negative value of the cross–covariance between total model tendency and IAU tendency causes smaller increase of model prediction error from time t_{n-1} to t_n during the IAU [-3h, +3h] assimilation window. Such negative cross-covariance values are also observed over large areas in the complex and nonlinear GEOS system (figure not shown). A closer examination of the two cross-covariance terms between model hydrodynamics/physics tendency (i.e., $\left(\frac{\partial x}{\partial t}\right)_d$, $\left(\frac{\partial x}{\partial t}\right)_p$) and IAU tendency $\frac{\delta x}{\tau}$ are conducted using an ensemble forecast with the IAU forcing for temperature state (right columns of Figs 11–12). δx represents the analysis increment, and τ represents the scaling parameter used to convert increments into tendencies. The cross-covariances using the ensemble forecast without the IAU forcing are also calculated (left columns of Figs 11–12) to help illustrate how the model hydrodynamics and physics tendencies have changed in response to the IAU forcing in the assimilation window. It is shown in Fig. 11 that the patterns of $cov\left(\left(\frac{\partial x}{\partial t}\right)_d, \frac{\delta x}{\tau}\right)$, the cross-covariances between model hydrodynamics tendency and IAU tendency, exhibit little change at 1000 and 925 hPa, but become increasingly negative with height at and above 850 hPa due to the use of the IAU forcing. In contrast, $cov\left(\left(\frac{\partial x}{\partial t}\right)_p, \frac{\delta x}{\tau}\right)$, the cross-covariances between model physics tendency and IAU tendency (Fig. 12), indicate relatively small changes at higher vertical levels but substantial changes at 1000, 925 and 850 hPa. When IAU forcing is off, these levels show no clear pattern with mixed positive and negative values, but cross-covariance values become strongly negative over large areas when IAU forcing is on. These results indicate that physics tendency tends to evolve much more in response to the IAU forcing than hydrodynamics tendency in the lower troposphere, while the change of hydrodynamics tendency dominates in the middle and upper troposphere. While negative cross-covariance values are usually viewed as beneficial to reduce the state error covariance in Equation (A4), they suggest that analysis affects hydrodynamics tendencies more in the desired direction at mid-levels of the troposphere than it does at low levels; conversely, the analysis drives physics tendencies more toward desirable values at low levels than at mid levels of the troposphere.

Similar patterns are also noticed in the cross-covariance terms between model hydrodynamics/physics tendency and IAU tendency for specific humidity and wind (figures not shown). The only noticeable pattern variation is seen in the cross-covariance between model hydrodynamics tendency and IAU tendency for specific humidity. While it is still becoming more negative at 850

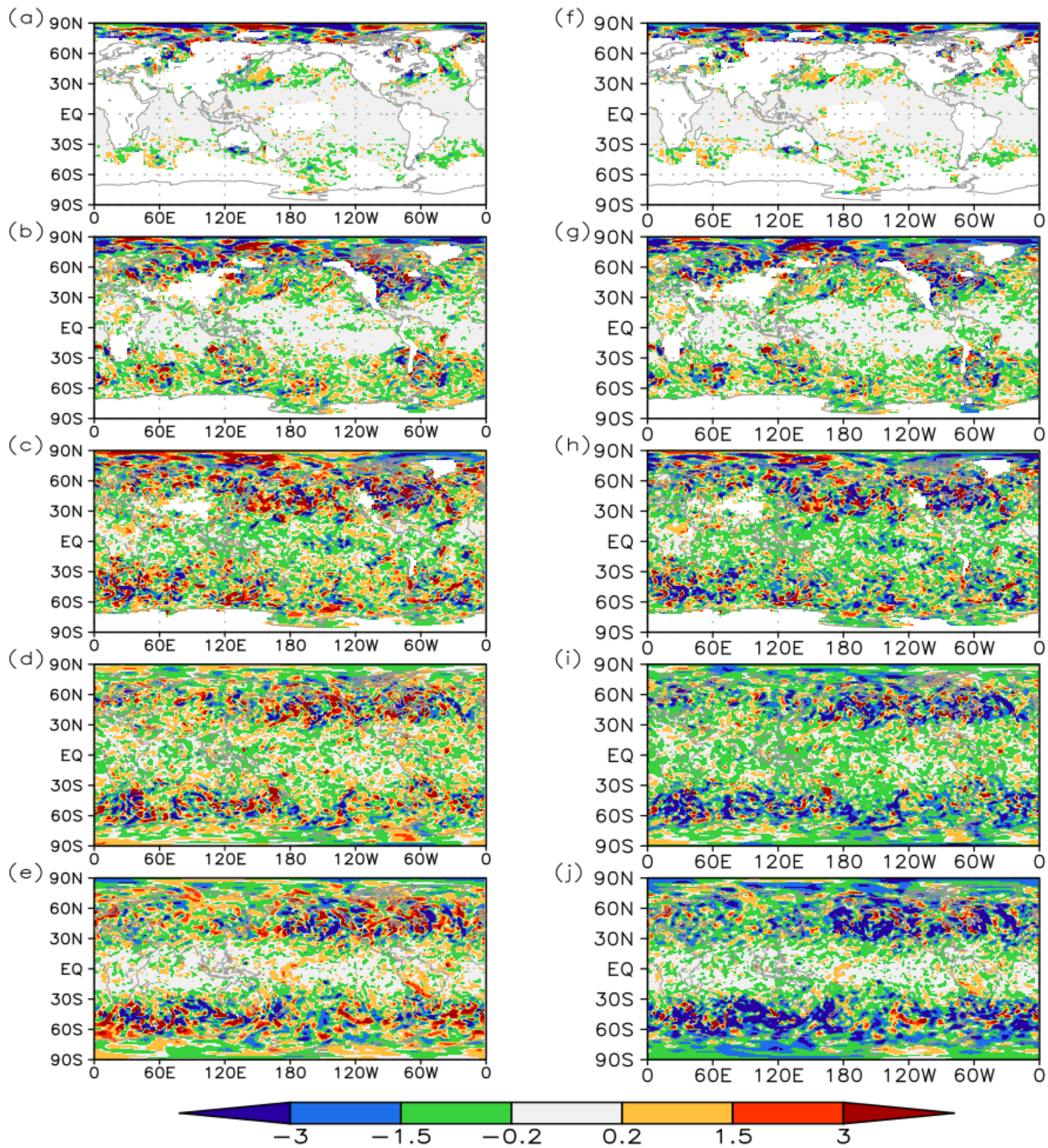

FIG. 11. Cross-covariance ($\times 1.0e^{-10}K^2s^{-2}$) between ensemble model temperature hydrodynamics tendency and ensemble IAU tendency at 1000 (a, f), 925 (b, g), 850 (c, h), 500 (d, i), and 200 (e, j) hPa at 00UTC December 16, 2019 for the ensemble forecasts with (right columns) and without (left columns) the IAU tendency forcing.

hPa when the IAU forcing is turned on, the changes between with and without IAU forcing are much smaller than those for temperature and there is little difference at 500 and 200 hPa.

The changes of hydrodynamics and physics tendencies in response to the IAU forcing are also reflected in the spatial correlation coefficients of the model tendencies between the ensemble

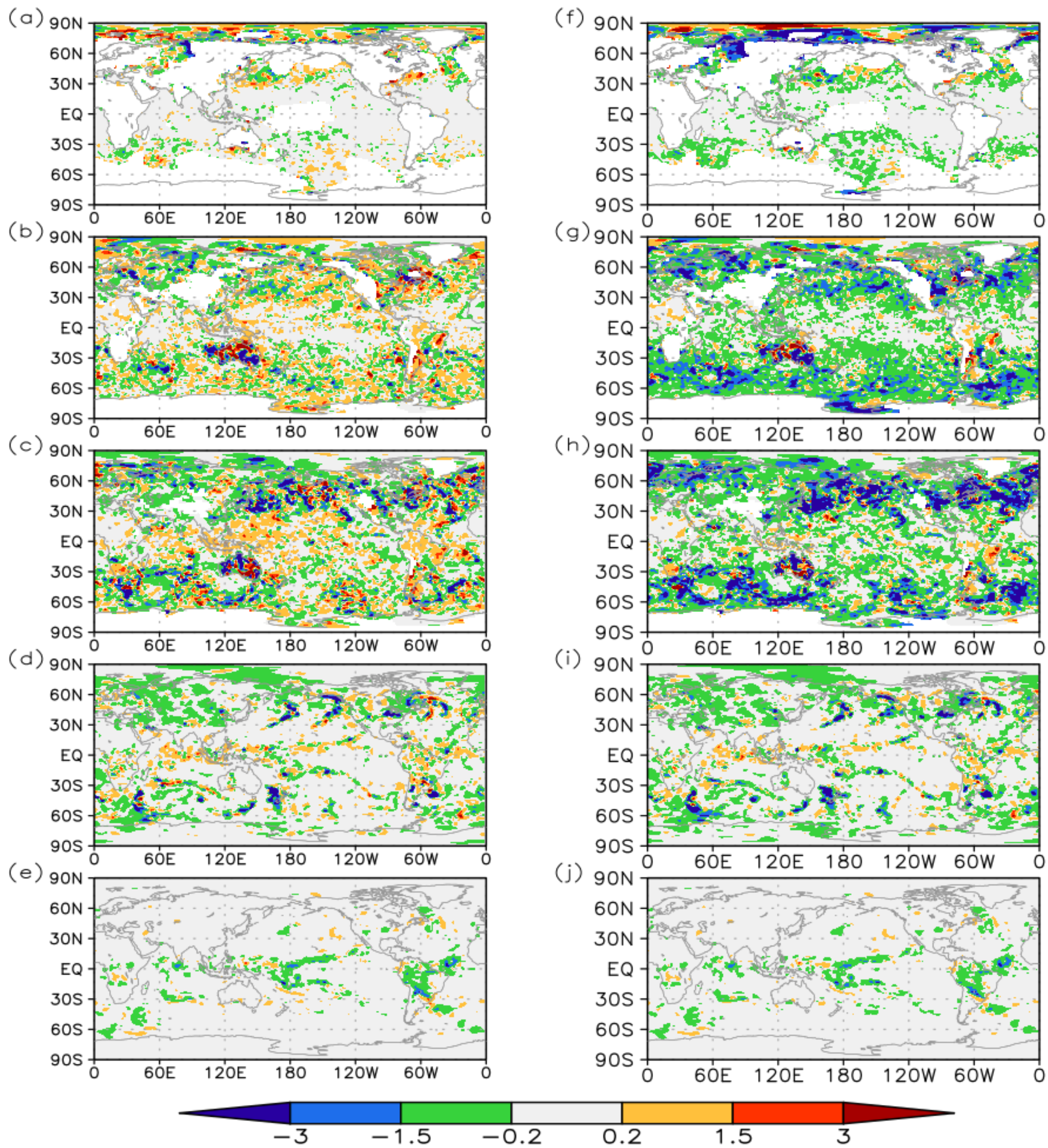

FIG. 12. The same as Fig. 11 but for cross-covariance between ensemble model temperature physics tendency and ensemble IAU tendency.

forecast with the IAU forcing and the original free ensemble forecast. A higher correlation coefficient indicates a model tendency is more similar to that of the free model forecast. With the use of IAU forcing, the model tendencies deviate from those of the original free forecast, and the departures are expected to increase with forecast hours. The behaviour of the model tendencies and the degree of similarity to the original free forecast are examined at selected model vertical levels.

The global correlation coefficients for temperature hydrodynamics and physics tendencies are displayed in Fig. 13 during the 27-hour ensemble forecast, which includes a 6-hour IAU assimilation window (i.e., [-3h, +3h] relative to the analysis cycle time) followed by a 21-hour (i.e., [3h, 24h]) free forecast. In the [-3h, +3h] IAU assimilation window, the correlation coefficient of hydrodynamics tendency (left panel of Fig. 13) dipped slightly at about 100, 300, and 850 hPa, but overall it has relatively small vertical variation globally, and experiences more rapid changes in the second half of the assimilation window than the first half of the window (except for 1000 hPa) and the following free forecast lead hours. In the subsequent [3h, 24h] forecast when the IAU forcing is turned off, the correlation decreases with the forecast lead hours and with the height globally, reaching a minimum of about 0.7 at 300 hPa at 24h, then increases to 0.81 at 200 hPa. At and below 850 hPa, the correlation coefficient has the slowest rate of change after the IAU is turned off.

On the other hand, the pattern of the physics tendency correlation coefficient (right panel of Fig. 13) is quite different from that of hydrodynamics tendency. Little departure is observed in physics tendency due to the use of IAU forcing at and above 100 hPa. In the [-3h, 3h] IAU assimilation window, the correlation coefficient of physics tendency experiences relatively smaller vertical variation below 100 hPa and above 850 hPa. However, at and below 850 hPa, it decreases rapidly as the physics tendency deviates from the original forecast model, and the model physics evolves much more significantly in the first half of the assimilation window than the second half. After the IAU forcing is off, two minimum correlation coefficient levels are noticed, one is at the lowest model levels 850–925 hPa, the other is located around 300–400 hPa. While the correlation coefficient still roughly decreases with the forecast lead hours in the mid- and upper troposphere, such pattern doesn't exist in the lower troposphere due to the complexity of physical processes, for example, the correlation coefficients at 6h, 9h, and 12h are higher than 3h, the correlation at 24h is higher than some shorter forecast lead times, and all correlation coefficients of different forecast lead hours cluster together at 1000 hPa.

Comparing the correlation coefficients in the left and right panels of Fig. 13, it is clearly shown that the lower troposphere physics tendencies diverge more quickly than the hydrodynamics in both

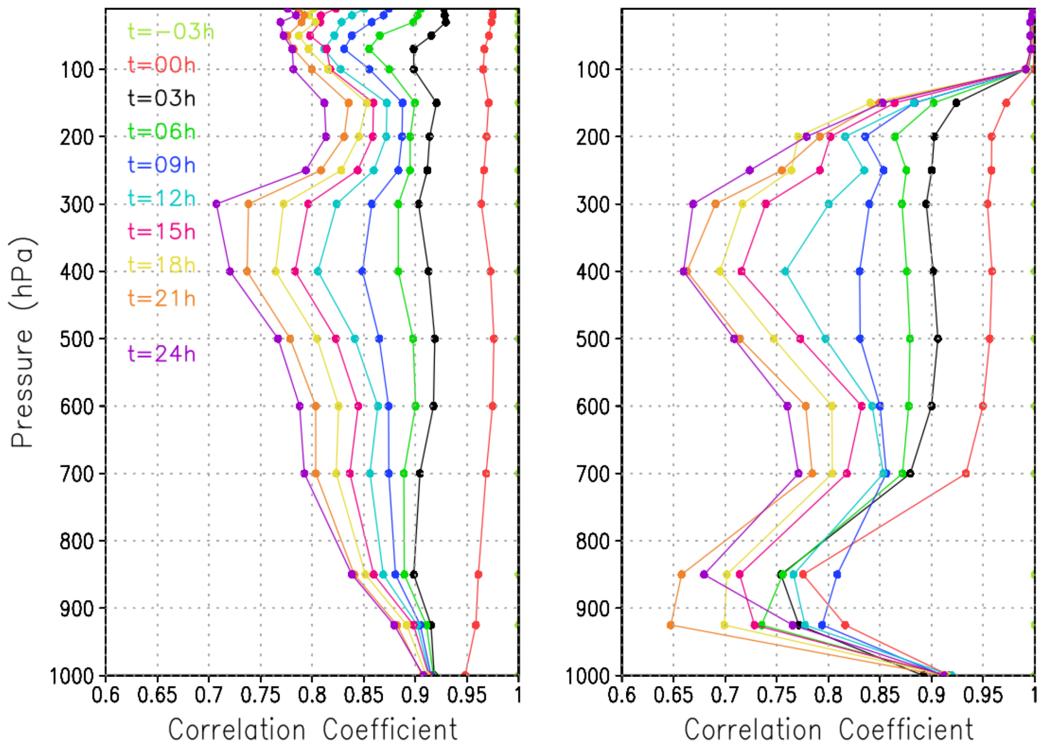

FIG. 13. Globally-averaged correlation coefficients of model temperature hydrodynamics (left) and physics (right) tendencies between the ensemble forecast with IAU forcing and original free ensemble forecast without IAU forcing during a 27-hour ensemble forecast. The light green line is for the correlation coefficient at the beginning of the assimilation window and therefore the coefficient value is 1. The red/black line is for the coefficient at the middle/end of the assimilation window.

the lower and upper troposphere, which is consistent with the idea that the PBL physics responds strongly to the IAU.

The global correlation coefficients of hydrodynamics and physics tendencies between the ensemble forecast with IAU forcing and the original free ensemble forecast are also evaluated for specific humidity and wind (figures not shown). Like the temperature hydrodynamics tendency correlation coefficient, the correlation coefficients for specific humidity and wind decrease significantly in the IAU assimilation window, then continue to decrease gradually with forecast lead times with the smallest change rate at the lower troposphere and reaching the minimum around 150 hPa for specific humidity and 200–500 hPa for wind. As to the physics tendency correlation coefficients, the patterns for temperature and specific humidity are similar. The correlation coefficients for specific humidity increase at 24h and are even comparable or larger than those at 3h in the lower troposphere, while they decrease gradually with forecast lead times in the mid- and upper troposphere with a minimum correlation at 50 hPa. The correlation coefficients for wind decrease the most at 400 hPa in the IAU assimilation window, and decrease slightly with forecast lead times at 925 hPa but there is no clear pattern at other levels in the subsequent forecast when the IAU forcing is off.

6. Conclusions

The GEOS data assimilation system has provided the critical capability of combining model physics and a wide range of observations with various data coverage and spatial/temporal resolutions coherently to improve the global thermodynamic structure and numerical forecast. To prepare for future observing systems of the next decade, we have assessed the effectiveness of the use of existing observing systems in the lower troposphere in the GEOS data assimilation system.

The full set of observations assimilated in the global GEOS data assimilation system is first assessed using the FSOI with the forecast error norm integrated from surface to about 850 hPa. Radiosonde data and microwave radiances from AMSU-A and ATMS are among the top contributors to the GEOS system. Both lower peaking temperature sounding channels from AMSU-A and ATMS and window channels from AMSR2 and GMI are found to be effective in reducing model forecast errors.

Given the FSOI results from the control experiment, a set of data denial experiments are conducted with selected observing systems of interest being removed. The results of the data denial experiments show that, microwave radiances and conventional data are the two most important data types to improving model forecast skills in the lower troposphere, which is in agreement with the FSOI results. Microwave radiances usually have large positive impact in the Southern Hemisphere and Tropics ocean, but most of the large impact is observed above 925 hPa and at early forecast lead times as the impact dissipates with longer lead times. Microwave radiances are also shown to contribute much more to specific humidity field than temperature field. Infrared radiances collectively have much smaller impact in the lower troposphere, and they have difficulty to influence model levels below 925 hPa. In contrary, conventional data have the largest contribution in the Northern Hemisphere on both specific humidity and temperature fields, and their impact ranges from surface to 100 hPa, depending on the field evaluated. Meanwhile, data denial experiment results also reveal that changes at and near the surface by assimilating surface observations can affect not only lower but also mid- and upper troposphere. They have small but persistently positive impact on specific humidity forecast skill in the upper troposphere but mixed impact in the lower and mid troposphere. However, their impacts on temperature and PBL height are small or negligible.

One surprising finding of this study is the obvious inconsistent results between the FSOI and the data denial experiment on satellite winds. FSOI results suggest high positive impact of the satellite winds data, but the data denial experiment indicates negative impact of these data in the GEOS system. Close examinations of the OmF of ATMS and IASI agree with the data denial experiment results. This inconsistency may be partially attributable to the choice of forecast aspect in the FSOI, which may lead to incomplete assessment of the observing system (Todling 2013). The negative impact from the satellite winds is most likely due to the missing of the thinning procedure for a subset of satellite winds. Further testing and tuning of the satellite winds in the GEOS system have been underway in a separate effort.

The responses of model hydrodynamics and physics tendencies to the IAU tendency forcing are also investigated in this study. In the GEOS data assimilation system, IAU procedure is employed so piecewise analysis increment tendency is introduced into the forecast model during the 6-hour assimilation window. As the model re-adjusts to the IAU forcing, the IAU tendency is found to contribute to the reduction of state error covariance, mostly noticeable through the interaction with model hydrodynamic tendency in the mid- and upper troposphere and the interaction with model physics tendency in the lower troposphere. The correlation coefficients of temperature hydrodynamics and physics tendencies between ensemble forecast with IAU forcing and free ensemble forecast further illustrate the changes of the model tendencies in the assimilation window and reveal their evolution behaviors in the subsequent free forecast. In the IAU assimilation window, physics tendency correlation coefficient exhibits significant and rapid decrease in the lower troposphere as physics tendency re-adjusts quickly there to the IAU forcing, but hydrodynamics tendency correlation coefficient has much less vertical variation. In the following forecast when

IAU forcing is off, both hydrodynamics and physics tendencies see low correlations around 300-400 hPa. In the lower troposphere, while the hydrodynamics tendency has the slowest change rate, the physics continues to diverge, but the correlation coefficient of physics tendency at 24h becomes larger, instead of smaller, than shorter forecast lead times in the lower troposphere, which may suggest that this tendency tends to get back to the original free forecast mechanism.

Overall, there are still a lot of room to improve observation usages in the lower troposphere in the GEOS data assimilation system. GMAO plans to significantly increase the vertical resolution of its GEOS model in the near future, making corresponding adjustments to the model physics and data assimilation components. This near-future upgrade will aim, among other things, to maximize the benefits of using various observations in the lower troposphere through improved first guess thermodynamics. As the complexity and short timescale of PBL processes pose additional difficulties in response to the IAU tendency forcing and retention of information content of PBL observations, future research is also much needed to further improve representation of processes by the analysis and model parameterization schemes and parameters.

Acknowledgments. We would like to thank Ronald Errico and Daniel Holdaway for their helpful discussions. This study is supported by the Global Modeling and Assimilation Office core funding from NASA.

Data availability statement. The GEOS model and data assimilation code and supporting software are maintained by the GMAO and publicly available through the NASA Open Source Agreement (NOSA) on github (https://github.com/GEOS-ESM). The data generated in this study includes output from the GEOS model and data assimilation system. All the unprocessed model output were automatically archived and can be accessed on the NCCS Dirac Mass Storage System.

APPENDIX

A brief derivation of state error covariance with IAU forcing

A brief back of the envelop, idealized, calculation provides the rationale to justify why it is thought that the correlations between the total model tendency and the IAU tendency should be negative.

In a very simplified context, and under the most straightforward flavor of IAU, a state variable x_n evolves in a single time step, Δt , from time t_{n-1} to t_n following

$$x_n = x_{n-1} + \Delta t \frac{\partial x}{\partial t} + \Delta t \frac{\delta x}{\tau}, \tag{A1}$$

where the differential in the second term in the rhs represents the total model tendency, which includes hydrodynamics and physics tendencies, $\left(\frac{\partial x}{\partial t}\right)_d$ and $\left(\frac{\partial x}{\partial t}\right)_p$. δx represents the analysis increment, and τ represents the scaling parameter used to convert increments into tendencies.i

Subtracting the true state from the equation above, multiplying the resulting expression by itself, taking the expectation of the result, and assuming that x_{n-1} is not correlated with either the total tendency (i.e., linear dynamics) and the IAU term, we get

$$\begin{aligned} P_n = P_{n-1} + (\Delta t)^2 Q_x + (\Delta t)^2 Q_i \\ + (\Delta t)^2 cov\left(\frac{\partial x}{\partial t}, \frac{\delta x}{\tau}\right) \end{aligned} \tag{A2}$$

In the above, P is the state error covariance, $Q_x = cov\left(\frac{\partial x}{\partial t}\right)$, $Q_i = cov\left(\frac{\delta x}{\tau}\right)$, and the symbol $cov()$ is used to represent either an autocovariance when only a single entry is in its argument, or a cross-covariance when two arguments are present. In the ideal world, $P_n - P_{n-1} > 0$, that is the error in the predicted state grows in time, therefore,

$$cov\left(\frac{\partial x}{\partial t}, \frac{\delta x}{\tau}\right) > -(Q_x + Q_i) \tag{A3}$$

meaning the cross–covariance must be larger than a certain negative value. Any value large than a minimum negative value implies a larger error in the prediction, that is, a large value for $P_n - P_{n-1}$. Therefore the smaller the value of the cross–covariance or the smaller the sum of the cross-covariance and Q_i, the smaller the prediction error.

If splitting total model tendency into hydrodynamics and physics tendencies, similar to Equation (A2) we also have

$$P_n = P_{n-1} + (\Delta t)^2 Q + (\Delta t)^2 Q_i + (\Delta t)^2 cov\left(\left(\frac{\partial x}{\partial t}\right)_d, \left(\frac{\partial x}{\partial t}\right)_p\right)$$
$$+ (\Delta t)^2 cov\left(\left(\frac{\partial x}{\partial t}\right)_d, \frac{\delta x}{\tau}\right) + (\Delta t)^2 cov\left(\left(\frac{\partial x}{\partial t}\right)_p, \frac{\delta x}{\tau}\right) \qquad \text{(A4)}$$

where $Q = cov\left(\left(\frac{\partial x}{\partial t}\right)_d\right) + cov\left(\left(\frac{\partial x}{\partial t}\right)_p\right)$. The last three cross-covariances terms in the rhs affect P_n simultaneously in Equation (A4).

References

Bacmeister, J. T., M. J. Suarez, and F. R. Robertson, 2006: Rain reevaporation, boundary layer-convection interactions, and Pacific rainfall patterns in an AGCM. *Journal of Atmospheric Science*, **63**, 3383–3403, https://doi.org/10.1175/JAS3791.1, URL https://doi.org/10.1175/JAS3791.1.

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, **134 (637)**, 1971–1996, https://doi.org/10.1002/qj.340, URL https://doi.org/10.1002/qj.340.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Monthly Weather Review*, **124 (6)**, 1256–1271, https://doi.org/10.1175/1520-0493(1996)124<1256:dauiau>2.0.co;2, URL https://doi.org/10.1175/1520-0493(1996)124<1256:dauiau>2.0.co;2.

Clough, S., M. Shephard, E. Mlawer, J. Delamere, M. Iacono, K. Cady-Pereira, S. Boukabara, and P. Brown, 2005: Atmospheric radiative transfer modeling: a summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **91 (2)**, 233–244, https://doi.org/10.1016/j.jqsrt.2004.05.058, URL https://doi.org/10.1016/j.jqsrt.2004.05.058.

Duncan, D. I., N. Bormann, and E. V. Hólm, 2021: On the addition of microwave sounders and numerical weather prediction skill. *Quarterly Journal of the Royal Meteorological Society*, https://doi.org/10.1002/qj.4149, URL https://doi.org/10.1002/qj.4149.

El Akkraoui, A., Y. Tremolet, and R. Todling, 2013: Preconditioning of variational data assimilation and the use of a bi-conjugate gradient method. *Quarterly Journal of the Royal Meteorological Society*, **139**, 731–741, https://doi.org/10.1002/qj.1997, URL https://doi.org/10.1002/qj.1997.

Errico, R. M., K. Raeder, and M. Ehrendorfer, 2004: Singular vectors for moisture-measuring norms. *Quarterly Journal of the Royal Meteorological Society*, **130 (598)**, 963–987, https://doi.org/10.1256/qj.02.227, URL https://doi.org/10.1256/qj.02.227.

Freitas, S. R., G. A. Grell, A. Molod, M. A. Thompson, W. M. Putman, C. M. S. e Silva, and E. P. Souza, 2018: Assessing the grell-freitas convection parameterization in the nasa geos modeling system. *Journal of Advances in Modeling Earth Systems*, **10**, 1266–1289, https://doi.org/10.1029/2017MS001251, URL https://doi.org/10.1029/2017MS001251.

Garcia, R. R., and B. A. Boville, 1994: Downward control of the mean meridional circulation and temperature distribution of the polar winter stratosphere. *J. Atmos. Sci.*, **51**, 2238–2245.

Gelaro, R., R. H. Langland, S. Pellerin, and R. Todling, 2010: The THORPEX observation impact intercomparison experiment. *Monthly Weather Review*, **138 (11)**, 4009–4025, https://doi.org/10.1175/2010mwr3393.1, URL https://doi.org/10.1175/2010mwr3393.1.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, https://doi.org/10.1002/qj.3803, URL https://doi.org/10.1002/qj.3803.

Holdaway, D., R. Errico, R. Gelaro, J. G. Kim, and R. Mahajan, 2015: A linearized prognostic cloud scheme in NASA's goddard earth observing system data assimilation tools. *Monthly Weather Review*, **143 (10)**, 4198–4219, https://doi.org/10.1175/mwr-d-15-0037.1, URL https://doi.org/10.1175/mwr-d-15-0037.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**.

Kelly, G., and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the global observing system through observing system experiments. *ECMWF Newsletter*, **113**, 16–28, https://doi.org/10.21957/ct50muxpx4, URL https://www.ecmwf.int/node/10434.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009: Improving incremental balance in the GSI 3dvar analysis system. *Monthly Weather Review*, **137 (3)**, 1046–1060, https://doi.org/10.1175/2008mwr2623.1, URL https://doi.org/10.1175/2008mwr2623.1.

Langland, R. H., and N. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189–201, https://doi.org/10.3402/tellusa.v56i3.14413, URL https://doi.org/10.3402/tellusa.v56i3.14413.

Lawrence, H., N. Bormann, I. Sandu, J. Day, J. Farnan, and P. Bauer, 2019: Use and impact of arctic observations in the ECMWF numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, **145 (725)**, 3432–3454, https://doi.org/10.1002/qj.3628, URL https://doi.org/10.1002/qj.3628.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, , and R. N. B. Smith, 2000: A new boundary layer mixing scheme. part i: Scheme description and single-column model tests. *Monthly Weather Review*, **128**, 3187–3199, https://doi.org/10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2, URL https://doi.org/10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

Lord, S., G. Gayno, and F. Yang, 2016: Analysis of an observing system experiment for the joint polar satellite system. *Bulletin of the American Meteorological Society*, **97 (8)**, 1409–1425, https://doi.org/10.1175/BAMS-D-14-00207.1, URL https://doi.org/10.1175/BAMS-D-14-00207.1.

Louis, J., and J. Geleyn, 1982: A short history of the pbl parameterization at ecmwf. *Proc. ECMWF Workshop on Planetary Boundary Layer Parameterization*, Reading, United Kingdom, ECMWF.

McFarlane, N. A., 1987: The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.*, **44 (14)**, 1775–1800.

National Academies of Sciences, E., and Medicine, 2018: *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space*. The National Academies Press, Washington, DC, https://doi.org/10.17226/24938, URL https://www.nap.edu/catalog/24938/ thriving-on-our-changing-planet-a-decadal-strategy-for-earth.

Park, S., and C. S. Bretherton, 2009: The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the community atmosphere modell. *Journal of Climate*, **22**, 3449–3469.

Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, **227 (1)**, 55–78, https://doi.org/10.1016/j.jcp.2007.07.022, URL https://doi.org/10.1016/j.jcp.2007.07.022.

Reichle, R. H., and Coauthors, 2018: Soil moisture active passive (SMAP) project assessment report for version 4 of the l4_sm data product. Nasa tech. memo., vol. 52, NASA, 72 pp pp. URL https://gmao.gsfc.nasa.gov/pubs/docs/Reichle1083.pdf.

Takacs, L. L., M. J. Suárez, and R. Todling, 2016: Maintaining atmospheric mass and water balance in reanalyses. *Quarterly Journal of the Royal Meteorological Society*, n/a–n/a, https://doi.org/10.1002/qj.2786, URL https://doi.org/10.1002/qj.2786.

Takacs, L. L., M. J. Suárez, and R. Todling, 2018: The stability of incremental analysis update. *Monthly Weather Review*, **146 (10)**, 3259–3275, https://doi.org/10.1175/mwr-d-18-0117.1, URL https://doi.org/10.1175/mwr-d-18-0117.1.

Teixeira, J., and Coauthors, 2021: *Toward a Global Planetary Boundary Layer Observing System: The NASA PBL Incubation Study Team Report*. NASA PBL Incubation Study Team, URL https://science.nasa.gov/earth-science/decadal-pbl.

Todling, R., 2013: Comparing two approaches for assessing observation impact. *Monthly Weather Review*, **141 (5)**, 1484–1505, https://doi.org/10.1175/mwr-d-12-00100.1, URL https://doi.org/10.1175/mwr-d-12-00100.1.

Todling, R., and A. El Akkraoui, 2018: The GMAO hybrid ensemble-variational atmospheric data assimilation system: Version 2.0. GMAO Office Note, NASA Goddard Space Flight Center, Greenbelt, MD, USA. URL https://gmao.gsfc.nasa.gov/pubs/docs/Todling1019.pdf.

Todling, R., and Coauthors, 2022: A brief report on recent upgrades to the atmospheric analysis of the GMAO hybrid 4denvar system: GEOS-5.27 and GEOS-5.29. NASA Tech Memo, *in preparation*, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

Trémolet, Y., 2007: First-order and higher-order approximations of observation impact. *Meteorologische Zeitschrift*, **16**, 693–694.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, **136 (2)**, 463–482, https://doi.org/10.1175/2007MWR2018.1, URL https://doi.org/10.1175/2007MWR2018.1.

Zhu, Y., and Coauthors, 2016: All-sky microwave radiance assimilation in NCEP's GSI analysis system. *Monthly Weather Review*, **144 (12)**, 4709–4735, https://doi.org/10.1175/mwr-d-15-0445.1, URL https://doi.org/10.1175/mwr-d-15-0445.1.