

The Framework for Assimilating All-Sky GPM Microwave Imager Brightness Temperature Data in the NASA GEOS Data Assimilation System

MIN-JEONG KIM

Goddard Earth Science and Technology Research, Morgan State University, Baltimore, and Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

JIANJUN JIN AND AMAL EL AKKRAOUI

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Science Systems and Applications, Inc., Lanham, Maryland

WILL MCCARTY AND RICARDO TODLING

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

WEI GU

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Science Systems and Applications, Inc., Lanham, Maryland


RONALD GELARO

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 5 April 2019, in final form 15 January 2020)

ABSTRACT

Satellite radiance observations combine global coverage with high temporal and spatial resolution, and bring vital information to NWP analyses especially in areas where conventional data are sparse. However, most satellite observations that are actively assimilated have been limited to clear-sky conditions due to difficulties associated with accounting for non-Gaussian error characteristics, nonlinearity, and the development of appropriate observation operators for cloud- and precipitation-affected satellite radiance data. This article provides an overview of the development of the Gridpoint Statistical Interpolation (GSI) configurations to assimilate all-sky data from microwave imagers such as the GPM Microwave Imager (GMI) in the NASA Goddard Earth Observing System (GEOS). Electromagnetic characteristics associated with their wavelengths allow microwave imager data to be highly sensitive to precipitation. Therefore, all-sky data assimilation efforts described in this study are primarily focused on utilizing these data in precipitating regions. To utilize data in cloudy and precipitating regions, state and analysis variables have been added for ice cloud, liquid cloud, rain, and snow. This required enhancing the observation operator to simulate radiances in heavy precipitation, including frozen precipitation. Background error covariances in both the central analysis and EnKF analysis in the GEOS hybrid 4D-EnVar system have been expanded to include hydrometeors. In addition, the bias correction scheme was enhanced to reduce biases associated with thick clouds and precipitation. The results from single observation experiments demonstrate the capability of assimilating all-sky microwave brightness temperature data in GEOS both when the model forecast produces excessive precipitation and too little precipitation. Additional experiments show that hydrometeors and dynamic variables such as winds and pressure are adjusted in physically consistent ways in response to the assimilation.

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Min-Jeong Kim, min-jeong.kim@nasa.gov

DOI: 10.1175/MWR-D-19-0100.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

1. Introduction

Numerical weather prediction (NWP) has evolved at a relatively rapid pace during recent decades, leading to significantly improved forecasts of essential weather variables such as wind, temperature and pressure (Bauer et al. 2015). Forecasting the intensity and spatial distribution of precipitation, on the other hand, remains a significant challenge even on large scales. The reasons for this are numerous, but generally involve difficulties in both modeling the underlying moist physical processes (Sellers 1965; Hartmann 1994; Pedder et al. 2000) and providing adequately constrained initial conditions in cloudy and precipitating regions. With regard to the latter, millions of observations are now actively assimilated during each analysis cycle in most modern NWP systems. These data play a critical role in constraining initial conditions globally, but especially over oceanic regions where limited conventional observations are available. Until recently, however, data affected by clouds and precipitation have been discarded in large numbers due to the additional challenges affecting their use, including large systematic errors in the background forecast state, complexity of the required observation operators (i.e., radiative transfer models) and non-Gaussianity of the model and observation errors involved in assimilating these data directly (Bauer et al. 2010).

The above challenges notwithstanding, several forecast centers are pursuing efforts to assimilate radiances affected by clouds and precipitation from microwave sensors, with the expectation that these data can provide critical constraints on meteorological parameters in dynamically sensitive regions and thus lead to improved forecast accuracy for precipitation (Geer et al. 2018). The European Centre for Medium-Range Weather Forecasts (ECMWF) assimilates Advanced Microwave Scanning Radiometer 2 (AMSR2), GPM Microwave Imager (GMI), and Special Sensor Microwave Image/Sounder (SSMIS) radiance data at frequencies from 19 to 90 GHz over ocean and Microwave Humidity Sounder (MHS), SSMIS, and Microwave Humidity Sounder (MWS) radiance data at 183 GHz over ocean, land, snow-covered land, and sea ice. These data have become one of the main observational constraints to define moisture initial conditions for the forecasts (Bauer et al. 2010; Geer et al. 2010, 2014; Baordo and Geer 2016). The Environmental Modeling Center (EMC) at the National Centers for Environmental Prediction (NCEP) assimilates AMSU-A radiances in nonprecipitating cloudy conditions over ocean in addition to other clear-sky satellite radiance data and report improvements in analyzed lower-tropospheric

temperature and humidity off the west coasts of the continents as well as reduced bias in the depiction of stratus clouds (Zhu et al. 2016). The operational Met Office Unified Model system also showed significant improvements in the short-range forecasts by assimilating AMSU-A data in nonprecipitating cloudy conditions (Migliorini and Candy 2019). The Japan Meteorological Agency (JMA) assimilates microwave imager radiances in clear-sky conditions and retrieved precipitation from the radiances in rainy conditions in their mesoscale NWP system (Kazumori 2014). An all-sky radiance framework for microwave imagers such as AMSR-2, GMI, and SSMIS is also under development in the JMA 4D-Var system (Kazumori 2016).

This article describes the effort undertaken by the Global Modeling and Assimilation Office (GMAO) to assimilate cloud- and precipitation-affected microwave radiances in NASA's Goddard Earth Observing System (GEOS) data assimilation system (Rienecker et al. 2008; Todling and El Akkraoui 2018). The key components of the system in this regard are the GEOS atmospheric model (Molod et al. 2012, 2015) and the Gridpoint Statistical Interpolation (GSI) atmospheric analysis scheme (Derber and Wu 1998; Kleist et al. 2009). Initial efforts have focused on all-sky assimilation of GMI radiances, which became operational in the GEOS real-time production system in July 2018. This paper focuses on the framework developed for assimilating all-sky GMI radiance data in GEOS and evaluates the performance of various components of the system in a simplified experimental setting. Results from a more comprehensive assessment of the impact of GMI all-sky radiances on GEOS analyses and forecasts, including examination of statistics of background departures and analysis increments, forecast skill scores, and forecast sensitivity observation impact calculations, are presented in a follow-on paper (M.-J. Kim et al. 2020, unpublished manuscript).

The paper is organized as follows. Section 2 provides a brief overview of the GEOS data assimilation system, while section 3 introduces the GPM satellite and GMI data. In section 4, extensions to GEOS to assimilate all-sky microwave radiance data are described. These extensions include the selection of water control variables and definition of background error covariances for these control variables, the observation operator, the characterization of observation error, the development of quality control procedures and bias corrections. Section 5 examines the performance of the all-sky microwave radiance assimilation in GEOS using single-observation experiments, while section 6 shows how the initial wind, temperature, and pressure fields undergo significant dynamic adjustments in response to the

analyzed cloud and precipitation features. Finally, section 7 presents a discussion and conclusions.

2. Assimilation system

The GEOS atmospheric general circulation model is a weather- and climate-capable model used for weather analysis and forecasting, reanalysis, and climate simulations at various horizontal resolutions. The model has 72 vertical layers and a top at 0.01 hPa. It employs a finite-volume cubed-sphere dynamical core (Putman and Lin 2007) for hydrodynamics and includes parameterization schemes for moist processes, radiation, turbulent mixing and surface processes (Molod et al. 2015).

The GEOS data assimilation algorithm employed in this study is based on the hybrid four-dimensional ensemble-variational (4D-EnVar) configuration of GSI currently operational at GMAO (Todling and El Akkraoui 2018). The hybrid 4D-EnVar assimilation method introduces flow dependency to the background error covariance by using a hybrid combination of a fixed climatological model of background error covariances with localized covariances from an ensemble of current forecasts (Lorenc et al. 2015). The 4D data assimilation methods seek to find a 4D incremental correction $\delta \underline{\mathbf{x}}$ to a 4D background state $\underline{\mathbf{x}}_b$ using observations $\underline{\mathbf{y}}^o$ covering the time interval associated with the 4D problem by minimizing a cost function written as

$$J(\delta \underline{\mathbf{x}}) = \frac{1}{2} \delta \underline{\mathbf{x}}^T \underline{\mathbf{B}}^{-1} \delta \underline{\mathbf{x}} + \frac{1}{2} (\underline{\mathbf{y}} - \underline{\mathbf{y}}^o)^T \underline{\mathbf{R}}^{-1} (\underline{\mathbf{y}} - \underline{\mathbf{y}}^o), \quad (1)$$

where an underline was used here to extend the standard notation of Ide et al. (1997) to four dimensions following Lorenc et al. (2015). In Eq. (1), the first term represents the guess fit to the backgrounds taking into account prior and background information, and the second term represents the observational penalty measuring the guess fit to the observations. The $\underline{\mathbf{B}}$ and $\underline{\mathbf{R}}$ are 4D weighting error covariance matrices associated with backgrounds and observations, respectively. The background error covariance matrix $\underline{\mathbf{B}}$ in a hybrid 4D-EnVar method can be expressed as a weighted combination of the climatological covariance matrix $\underline{\mathbf{B}}_c$, which is based on the National Meteorological Center (NMC) method (Parrish and Derber 1992), and a localized ensemble-based covariance matrix $\underline{\mathbf{B}}_e$:

$$\underline{\mathbf{B}} = \beta_c^2 \underline{\mathbf{B}}_c + \beta_e^2 \underline{\mathbf{B}}_e, \quad (2)$$

where β_c and β_e are weighting coefficients for static background error covariances and flow dependent ensemble background error covariances, respectively.

Currently, the hybrid configuration consists of two data assimilation systems cycling in parallel: a variational part (called central) uses the hybrid 4D-EnVar formulation to define background error covariances and produce the main analysis called “central analysis,” and an ensemble part consisting of 32 ensemble members used for the square root ensemble Kalman filter (EnSRF) analysis (Whitaker and Hamill 2002). The weighting coefficients β_c and β_e are vertically varying such that they are both set to 0.5 in the troposphere and lower stratosphere, and smoothly transition to 1 and 0, respectively, above 5 hPa. The assimilation window is 6 h. Hourly analysis increments are applied to the model forecasts through a nudged incremental analysis update (IAU) procedure (Takacs et al. 2018). Additional details are provided in the context of the all-sky development described in section 4, while a comprehensive description of GMAO’s hybrid 4D-EnVar scheme can be found in Todling and El Akkraoui (2018).

3. GPM Microwave Imager (GMI) radiance data

Launched in 2014, the Global Precipitation Measurement Mission (GPM) uses a dual-frequency radar and microwave imager in a non-sun-synchronous orbit. Its 65° orbit inclination allows to measure precipitation at different times of day from the tropics to the Arctic and Antarctic circles (Skofronick Jackson et al. 2017). Compared to its predecessor Tropical Rainfall Measuring Mission (TRMM), GPM offers broader spatial coverage and enhanced capability to measure light rain and falling snow, among other enhancements (Hou et al. 2014). This study focuses on the use of observations from the GPM Microwave Imager (GMI), a conically scanning microwave radiometer with 13 channels and a swath width of 885 km and fields of view ranging from 5 to 25 km, depending on frequency. The radiance data used are the GMI level 1 1C-R, version 4 data product that coregisters the pixels of the GMI low-frequency and high-frequency channels using a nearest-neighbor approach. More details about GPM and its instruments can be found in Hou et al. (2014) and Skofronick-Jackson et al. (2017).

Owing to the difficulties of modeling surface emissivities over land and ice, the initial implementation described here is limited to the use of GMI data over ocean surfaces. The relatively low microwave surface emissivity of the ocean allows low-frequency (<89 GHz) microwave radiance data to have large sensitivity to clouds and precipitation compared to data over land (Bauer et al. 2010). The GEOS all-sky system uses brightness temperature (TB) observations from GMI

TABLE 1. Comparison of clear-sky and all-sky microwave TB assimilation framework in GEOS-5 ADAS (T : atmospheric temperature, q : specific humidity, T_{skin} : skin temperature, Ps: surface pressure, oz: ozone mixing ratio, u : zonal wind, v : meridional wind, ql: liquid cloud mixing ratio, qi: ice cloud mixing ratio, qr: rainwater mixing ratio, and qs: snow water mixing ratio, Ψ : streamfunction, $X_{\text{unbalanced}}$: unbalanced velocity potential, $T_{\text{unbalanced}}$: unbalanced temperature, $P_{\text{unbalanced}}$: unbalanced surface pressure, RH: relative humidity).

	Clear sky	All sky
State variables	$T, q, \text{Ps}, \text{oz}, T_{\text{skin}}, u,$ and v	$T, q, \text{Ps}, \text{oz}, T_{\text{skin}}, u, v, \text{ql}, \text{qi}, \text{qr},$ and qs
Analysis variables	$\Psi, X_{\text{unbalanced}}, T_{\text{unbalanced}}, P_{\text{unbalanced}}, \text{RH},$ $\text{oz},$ and T_{skin}	$\Psi, X_{\text{unbalanced}}, T_{\text{unbalanced}}, P_{\text{unbalanced}}, \text{RH},$ $\text{oz}, T_{\text{skin}}, \text{ql}, \text{qi}, \text{qr},$ and qs (see section 4a)
Background error	$\Psi, X_{\text{unbalanced}}, T_{\text{unbalanced}}, P_{\text{unbalanced}}, \text{RH},$ $\text{oz},$ and T_{skin}	$\Psi, X_{\text{unbalanced}}, T_{\text{unbalanced}}, P_{\text{unbalanced}}, \text{RH},$ $\text{oz}, T_{\text{skin}}, \text{ql}, \text{qi}, \text{qr},$ and qs (see section 4b)
Observation operator	CRTM (version 2.2.3)	CRTM (version 2.2.3) with a reconstructed cloud coefficient file of this study (section 4c)
Observation error	Constant (and inflated during quality control process)	Symmetric error model (Geer and Bauer 2011) (see section 4d)
Quality control	Screen out cloud affected radiance	Keep cloud and precipitation affected radiance. Screen out data over ocean if SST < 278 K. Screen out data if absolute value of first-guess departure is greater than 2. (see section 4d)
Bias correction predictors in VarBC	Constant, lapse rate, square of lapse rate, cosine of the zenith angle, and cloud liquid water path	Constant, lapse rate, lapse rate2, cos (zenith angle), Clavg, Clavg ² (see section 4e)

channels 3 (18.7 GHz V), 5 (23.8 GHz V), 6 (36.5 GHz V), 10 (166 GHz V), 12 (183.31 ± 3 GHz V), and 13 (183.31 ± 7 GHz V), where V and H indicate vertical and horizontal polarization, respectively. It is noted that the raw observations were used without any superobbing procedures. Channels 1 (10.65 GHz V) and 2 (10.65 GHz H) are excluded in this study and all-sky data from these channels will be tested as part of ongoing efforts to improve sea surface temperature analysis in GEOS in future. In addition, Channels 4 (18.7 GHz H), 7 (36.5 GHz H), 8 (89 GHz V), 9 (89 GHz H), and 11 (166 GHz H) are excluded owing to relatively large first-guess departures and identified non-Gaussianity that may be related to a known error in the Community Radiative Transfer Model (CRTM) version 2.2.3 (Han et al. 2006). The known error is caused by missing the off-diagonal terms of the reflective matrix and results in no contribution from diffuse reflection in its multiple-scattering radiative transfer algorithm (Liu et al. 2019). A more recent version of CRTM corrects this error by including reflection correction (Liu et al. 2019) but was unavailable for use in this study.

4. Framework for all-sky microwave radiance assimilation

Various components of the GEOS data assimilation system have been modified to assimilate cloud- and precipitation-affected microwave radiance data. The modifications are summarized in Table 1, which compares the clear-sky and all-sky configurations of

these components, while detailed descriptions are provided below.

a. Extended model state and analysis control variables

To utilize data in cloudy and precipitating regions, additional variables such as ice cloud (qi), liquid cloud (ql), rain (qr), and snow (qs) mixing ratios were included in background vector. This was required for the observation operator to simulate radiances in cloudy and precipitating conditions via the CRTM. In the atmospheric model, liquid and ice clouds are prognostic variables, while rain and snow are diagnostic variables. Furthermore, the analysis control vector has been extended in both the central and ensemble components of the hybrid analysis system to update these variables as part of the assimilation solution. There is no variable transformation between the model state and analysis control variables for hydrometeors. It is noted that analysis increments of hydrometeors are not fed back to the initial conditions for subsequent forecasts. However, hydrometeor increments induce adjustments to other variables such as moisture, wind and pressure in the hybrid 4D-EnVar system and increments of these other variables are fed back to forecast model as discussed in section 6.

b. Background error covariance

The analysis control vector in the current GEOS analysis scheme includes streamfunction, unbalanced velocity potential, unbalanced virtual temperature,

unbalanced surface pressure, relative humidity, ozone mixing ratio, and skin temperature (Wu et al. 2002; Rienecker et al. 2008; Akella et al. 2017). As described earlier, the hybrid data assimilation system combines flow-dependent ensemble-based estimates of background error covariances and climatological time-independent background error covariances. For the assimilation of all-sky microwave radiances, the control vector has been extended to include cloud liquid, cloud ice, rain, and snow water mixing ratio.

With the newly added control variables, the corresponding static and flow-dependent background error covariances are generated. Climatological statistics were estimated following the NMC method (Parrish and Derber 1992) using pairs of 24-h and 48-h GEOS forecasts between 1 June 2016 and 16 January 2017. During this period, the extra variables were added as outputs of the forecast model without being assimilated. Ensemble covariances are based on the spread of the 32 ensemble forecasts from the GEOS hybrid scheme during each analysis cycle.

The panels on the left side of Fig. 1 show the vertical distribution of the climatological background error standard deviations for cloud liquid, cloud ice, rain, and snow water. It is noted that there are no background error correlations between the hydrometeor variables or other variables. Aside from the fact that the estimated error statistics are by construction meridionally invariant, they have generally smooth spatial structure due to the averaging and smoothing applied. Relatively large error standard deviations for cloud liquid are seen in the storm tracks in midlatitudes. Climatological background errors in the Southern Hemisphere are slightly larger than in the Northern Hemisphere. This is likely reflecting the fact that the statistics were drawn from Southern Hemispheric fall–winter season forecasts. Generally speaking, the maximum errors for cloud liquid water occur in the layer between 900 and 850 hPa. Climatological background error standard deviations for cloud ice show large values near the tropical tropopause, where large amounts of cloud ice exist in the anvils of convective clouds. Climatological background errors for rain and snow are larger in the tropics than other latitudes. Large background errors for rain occur in the tropical lower troposphere between sea level and 600 hPa, while large errors for snow occur in the tropical middle troposphere between 600 and 450 hPa.

The panels on the right side of Fig. 1 show example cross sections of ensemble background error standard deviations for hydrometeors taken from GEOS at 1200 UTC 12 December 2015, while Fig. 2 shows the ensemble spread of cloud liquid water at 850 hPa for the

same analysis cycle. Significant differences can be seen between the climatological- and ensemble-based estimates, with the latter unsurprisingly exhibiting more flow-dependent structures. In some regions for instance, nearly zero ensemble spread is shown, corresponding to areas where the ensemble members forecasted nearly zero clouds for that particular date. In contrast, the static background errors (left column) show nonzero values over broad ranges of latitudes regardless of the meteorological conditions of the day, which may increase the probability of generating erroneous cloud increments over broad areas. Generally speaking, climatological error statistics may not be well suited to capture the often-localized distributions and rapid intensity changes exhibited by clouds and precipitation features. As mentioned earlier, the cross correlations between moisture and other analysis variables are not considered in the current static background error covariance model. Ensemble-based statistics, on the other hand, can provide realistic looking flow-dependent structures with implicit, model-generated cross correlations between variables, but are subject to deficiencies associated with small ensembles and possibly low-resolution ensembles, too. The impacts of these different background error properties are discussed further in sections 5 and 6.

c. Observation operator

For satellite radiance observations, the observation operator consists of spatiotemporal interpolation of the background fields to observation space and the CRTM version 2.2.3. The observation operator that simulates radiance data in the GSI has been extended to include cloud and precipitation fields. Therefore, hydrometeors including profiles of q_l , q_i , q_r , and q_s from the GEOS model have been incorporated into first-guess fields and used in the radiative transfer calculations by the CRTM. When the single-scattering albedo is nonzero, the CRTM computes radiances using the adding doubling solver (Liu and Weng 2006) in which the number of streams is determined with respect to the unitless size parameter $x = 2\pi R_{\text{eff}}/\lambda$, where R_{eff} is an effective radius of clouds and λ is a wavelength. Two streams are used if x is less than 0.01; four streams are used when x is greater than 0.01 and less than 1; and six streams are used if x is greater than 1 to resolve the angular dependence of radiation propagation caused by multiple scattering. In CRTM version 2.2.3, scattering and extinction coefficients, asymmetry factor, and phase functions for microwave wavelengths are read from a lookup table that was built using the Mie calculation for various cloud types (i.e., cloud ice, cloud liquid, rain, snow, graupel, and hail) and for various effective radii assuming a Gamma size distribution.

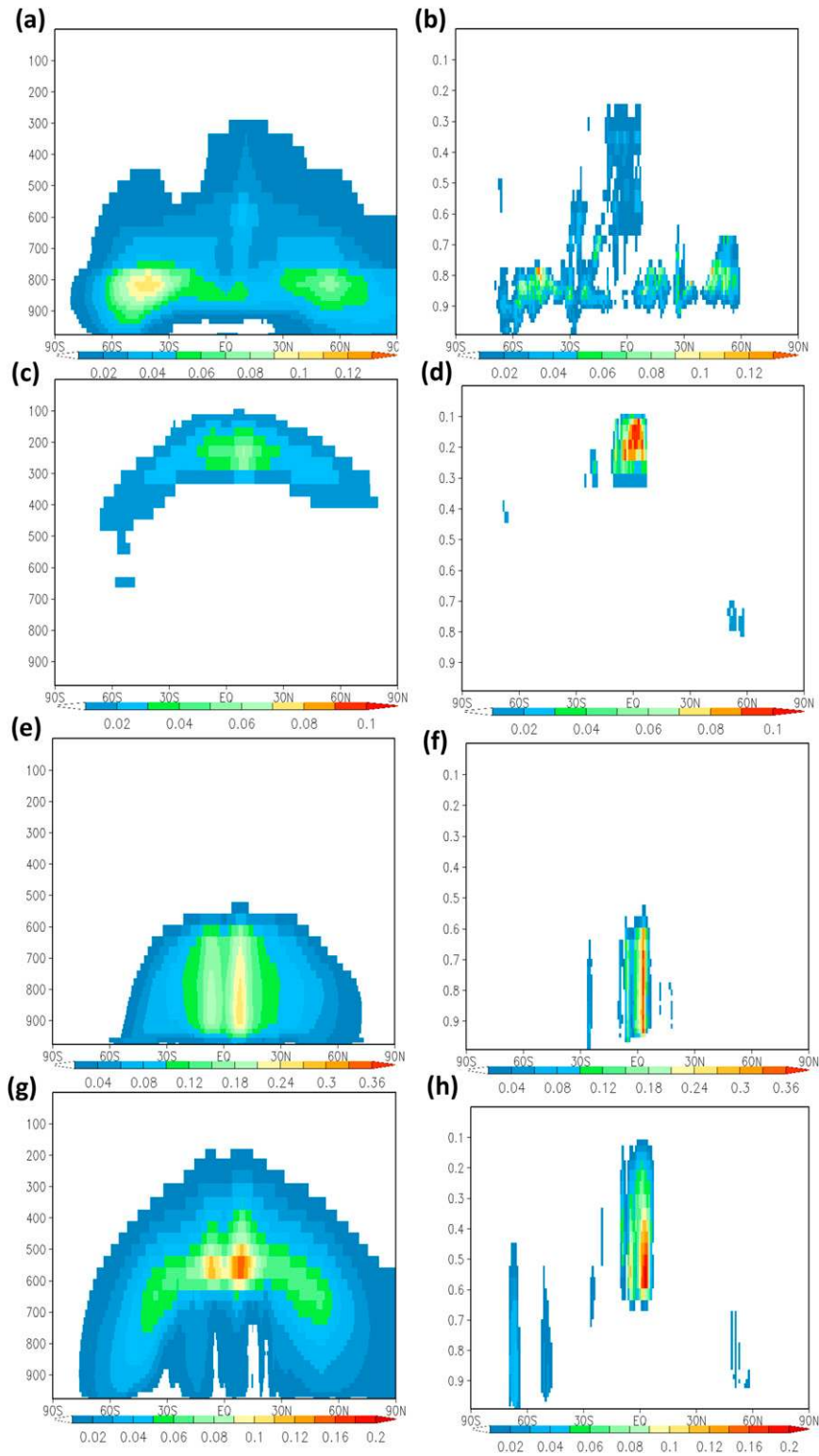


FIG. 1. Comparisons of (left) climatological background error standard deviations and (right) ensemble background error standard deviations as a function of latitude and vertical level. Ensemble background errors shown here are the cross-sectional views along the dotted line in Fig. 2. (a),(b) q_1 , (c),(d) q_i , (e),(f) q_r , and (g),(h) q_s .

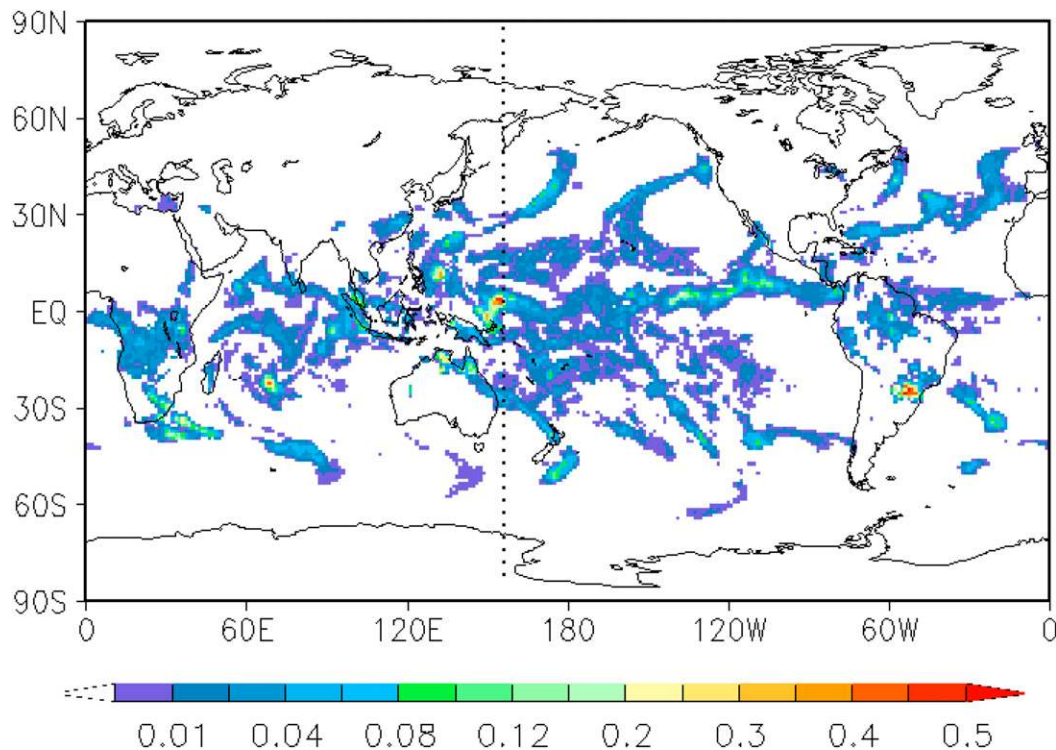


FIG. 2. Ensemble spread of cloud liquid water (g kg^{-1}) at 850 hPa calculated with 32 ensemble member backgrounds from GEOS at 1200 UTC 12 Dec 2015. Vertical cross sections of ensemble spread for hydrometeors along the dotted line are shown in Fig. 1.

Due to the known limitations of Mie scattering parameters for frozen hydrometeors, especially at high-frequency (>85 GHz) microwave channels (Kim 2006; Liu 2008; Geer and Baordo 2014), standard CRTM configuration was modified: scattering parameters for nonspherical frozen precipitation are calculated using the Liu optical properties database, which was computed using discrete dipole approximation (DDA) method (Draine and Flatau 1994) and used to construct a lookup table for the CRTM in this study. In the scattering database described in Liu (2008), optical properties of randomly oriented snow particles are tabulated as a function of frequency, temperature, and particle size for a variety of ice crystal shapes such as hexagonal ice columns, plates, rosettes, dendrites, etc.

Analogous to Geer and Baordo (2014), optical properties of 11 different nonspherical ice crystal shapes in Liu's database, in addition to scattering properties of spherical ice crystals calculated with the Mie method, were examined to find an optimal choice of ice crystal shape to reconstruct CRTM cloud scattering coefficients needed as part of the GSI observation operator. For each shape of ice crystal, a CRTM cloud coefficient lookup table was generated for 33 microwave frequencies

between 10.65 and 190.31 GHz, seven atmospheric temperatures between 243 and 303 K, and 405 effective radius sizes starting from 0.005 mm. The maximum effective radius considered for rain in the updated CRTM coefficients used in this study is 1.191 mm. For snow crystals, the maximum effective radius considered in the new CRTM cloud coefficient databases ranges from 0.664 to 1.278 mm, depending on snow crystal shape. Field et al. (2007) particle size distribution is assumed for frozen hydrometeors and Marshall-Palmer size distribution (Marshall and Palmer 1948) is assumed for liquid hydrometeors. The minimum effective radius in Liu's database ranges 0.025 to 0.033 mm depending on snow crystal shape. Mie scattering coefficients for spherical ice particles with 0.005, 0.015, and 0.030 mm effective radius were included at the beginning of optical properties for Liu(2008) nonspherical snow shape in the CRTM lookup table so that the scattering lookup table can be used both for cloud ice and snow hydrometeor. In implementing all-sky GMI data, we assumed effective radius for cloud ice as 0.005 mm. It is noted that optical properties are not sensitive to ice crystal shapes for cloud ice because of its very small particle size compared to the wavelengths of GMI channels considered in this study (Kim 2006).

The CRTM requires effective radius profiles as inputs in addition to the cloud water content (kg m^{-2}) profiles. Using the input effective radius and the input sensor's frequency, CRTM calculates optical properties by interpolating scattering coefficient values stored in the CRTM lookup table. Therefore, the input effective radius is expected to be consistent with effective radius and water content used to calculate the optical properties for the CRTM lookup table. In integrating radiative transfer processes, CRTM uses these optical properties and input water contents to compute the optical depth (τ) contributed from each cloud type in an atmospheric layer:

$$\tau = \rho \times k_e, \quad (3)$$

where ρ is integrated cloud water content for a layer (kg m^{-2}) provided as input and k_e is mass extinction coefficient ($\text{m}^2 \text{kg}^{-1}$) calculated using the CRTM lookup table.

To ensure the consistency between input cloud water content and input effective radius for a given size distribution (e.g., Field et al. (2007) for frozen hydrometeors and Marshall Palmer for rain in this study), fitting functions were derived and prescribed in a GSI code for each shape of Liu (2008)'s snow particles and rain (see appendix). Using these fitting functions, the input effective radius values are calculated depending on input cloud water contents efficiently in the GSI. This allows the CRTM to utilize Liu's DDA optical properties precalculated with Field et al. particle size distributions that are consistently applied to input water content and input effective radius relations.

Figure 3 compares spatial distribution of monthly mean $|O - F|$ of GMI channel 10 where background files were taken from GEOS-FP for a month of December 2017. Nonspherical ice crystal shapes from Liu (2008) and spherical shape were used to generate different sets of first-guess departure data. The number on top of each panel represents the spatial average of all the monthly mean absolute first-guess departure shown in each panel. In general, three-bullet rosettes produces the smallest mean of absolute first-guess departures and bias (not shown) among all shapes considered in this study. Therefore, this study employs DDA method-calculated scattering parameters of three-bullet rosettes among various ice crystal shapes in the Liu (2008) database because they generated simulated TBs close to observed TBs especially for high-frequency GMI channels.

Figure 4 compares GMI observations with CRTM-simulated GMI radiances using Mie scattering parameters and DDA method-based scattering parameters of

three-bullet rosettes near Hurricane Celia at 0000 UTC 12 July 2016. The horizontal resolution of the GEOS background fields used in the simulations is 25 km. Figure 4a shows GMI channel 10 observed TBs. The dynamic range of TBs near the storm is large, ranging from 280 K in clear-sky conditions to less than 200 K near the storm center where large precipitating particles exist. This suggests that it is critical to have reasonably accurate scattering parameters in the observation operator to be able to utilize high-frequency microwave TB observations in the analysis process. Figure 4b shows simulated GMI channel TBs based on the original CRTM cloud scattering coefficients from Mie-derived cloud optical properties. They are too warm compared to the GMI observations at high frequencies. Figure 4c shows that simulated GMI TBs are much closer to the observations in precipitating regions when the original cloud coefficients are replaced with the new cloud coefficients reconstructed using the Liu optical properties database, which was computed using DDA methods. For other frequencies such as GMI channels 12 and 13 (not shown), the comparison results were similar to channel 10. For lower-frequency GMI channels, TB differences caused by the choice of ice crystal shape were small (figure not shown) because low-frequency channels have large sensitivity to radiance emission caused by rain drops rather than ice particles. It is noted that the current version of CRTM does not consider cloud fraction. This limitation is partially accounted for by inflating the observation errors in precipitating regions and by considering cloud-dependent bias corrections. This will be discussed further in sections 4d and 4e, respectively.

d. Observation error covariance and quality control procedures

To consider non-Gaussian characteristics, relatively large representativeness errors, and complexities of observation operators for cloud- and precipitation-affected radiance data, the observation errors estimated in this study for all-sky GMI radiances ($\sigma_{\text{all-sky}}$) are based on a *symmetric observation error model*, suggested in Geer and Bauer (2011) and estimated with standard deviations of first-guess departures (observed minus calculated TBs) as a function of symmetric cloud amount. For this purpose, cloud amount is estimated with a cloud index (CI) based on the TB differences between horizontal and vertical polarizations at 37 GHz (Geer et al. 2010). CI can be calculated for the observed TB, defined as

$$\text{CI}_o = 1 - \frac{\text{TB}_{37V}^{\text{Obs}} - \text{TB}_{37H}^{\text{Obs}}}{\text{TB}_{37V}^{\text{Clear}} - \text{TB}_{37H}^{\text{Clear}}}, \quad (4)$$

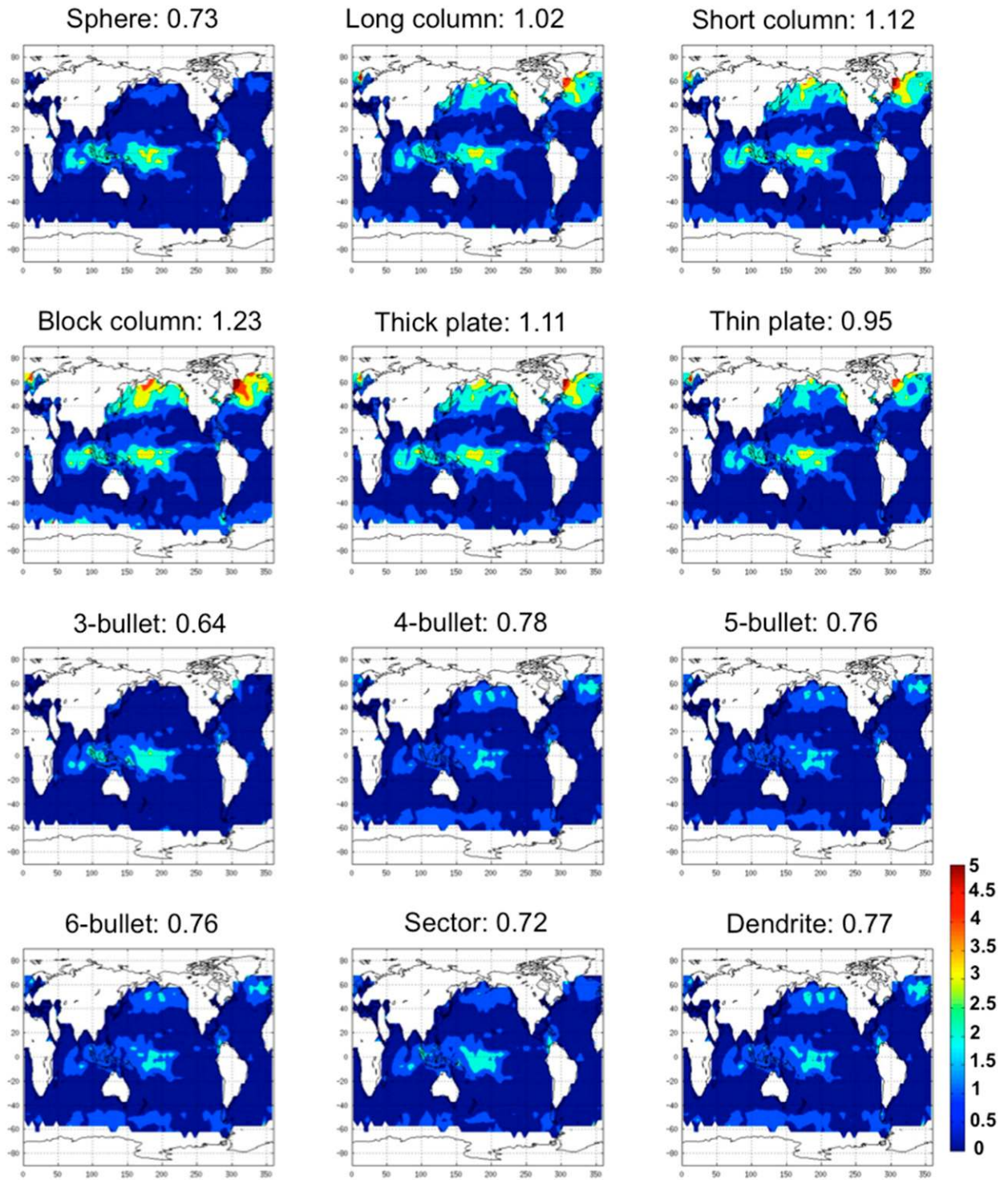


FIG. 3. Spatial distribution of monthly mean $|O - F|$ of GMI channel 10 where background fields were taken from GEOS-FP for the month of December 2017. Nonspherical ice crystal shapes from Liu (2008) and spherical shape were used to generate different sets of first-guess departures.

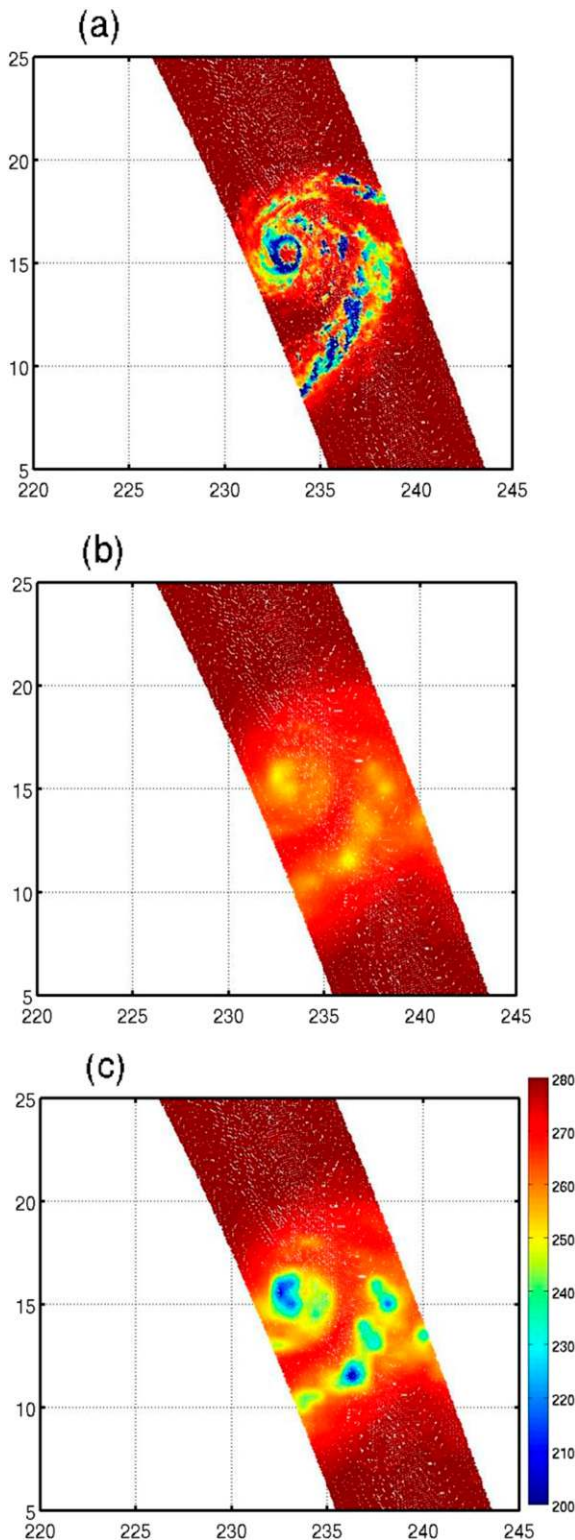


FIG. 4. Comparisons of simulated GMI 166 GHz vertically polarized TBs with the observations near Hurricane Celia at 0000 UTC 12 Jul 2016: (a) observed TBs, (b) CRTM-simulated TBs with original scattering coefficients based on Mie method, and

and for the all-sky TB derived from the background, defined as

$$CI_B = 1 - \frac{TB_{37V}^{Calc} - TB_{37H}^{Calc}}{TB_{37V}^{Clear} - TB_{37H}^{Clear}}, \quad (5)$$

where observed or simulated TB_{37V} and TB_{37H} are vertically polarized and horizontally polarized 37 GHz TBs, respectively. Both TB_{37V}^{Clear} and TB_{37H}^{Clear} are simulated TBs from the same first-guess profiles but without clouds or precipitation. The symmetric cloud amount is defined as the average of the observed and simulated cloud amounts:

$$CI_{avg} = \frac{CI_o + CI_B}{2}. \quad (6)$$

Although we consider only observation error variance and ignore observation error correlations in this study, it should be noted that [Bormann et al. \(2011\)](#) examined the error correlations (i.e., off-diagonal terms in the observation error covariance matrix) in all-sky radiance data assimilation and found considerable interchannel and spatial correlations especially in cloudy and rainy situations.

[Figure 5](#) shows the standard deviation of GMI first-guess departures binned as a function of CI_{avg} for selected channels. The background fields were taken every 6 h from a cycled data assimilation experiment between 1 December and 31 December 2015 and GMI first-guess departures were calculated offline. For the cycle experiment, where the background fields were taken from, GEOS was initialized at 2100 UTC 15 November 2015 and run in a hybrid 4D-EnVar configuration with a horizontal resolution of 0.5° for the analysis and 0.25° for the forecast. All the data used routinely in GEOS such as conventional data, clear-sky satellite radiance data from microwave and infrared sounders, GPS radio occultation data, and satellite wind data were assimilated. Standard deviations of all GMI observations before quality control (open circle), which will be described in the later part of this section, show much larger standard deviations than those of only assimilated GMI observations (closed circle) for high-frequency channels. The observation error model used in this study was fitted to the standard deviation of FG departures before quality control procedures and is indicated

←

(c) CRTM-simulated TBs with the DDA method calculated scattering properties of three-bullet rosette snow crystals. The color bar shown in (c) works for (a) and (b) as well.

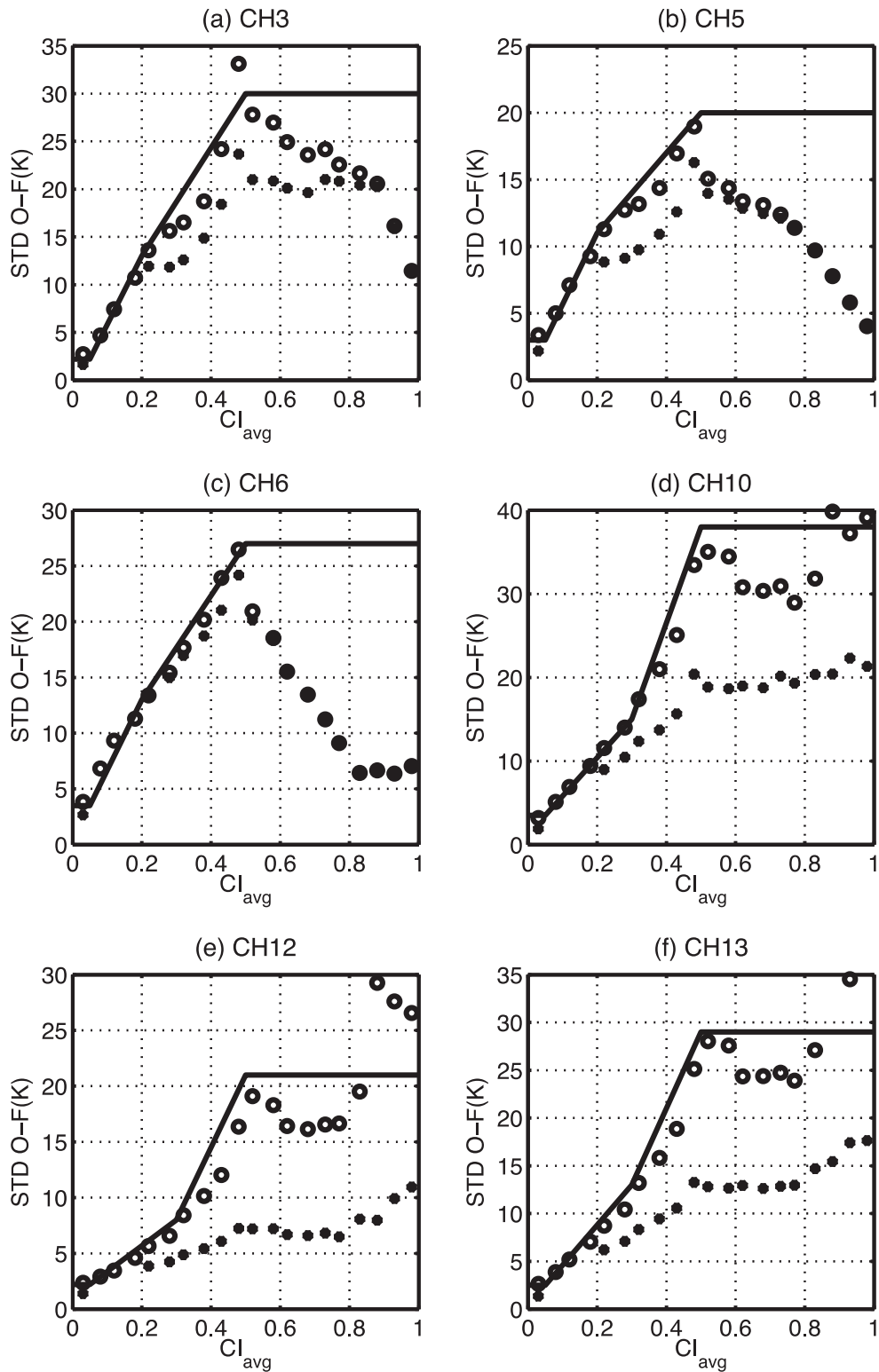


FIG. 5. Standard deviation (STD) of first-guess departures of all GMI observations (open circles) and assimilated GMI observations only (closed circles) binned as a function of averaged cloud index for observations between 1 and 31 Dec 2015. Bin size is 0.05. The observation error model applied in this study is shown as a solid line.

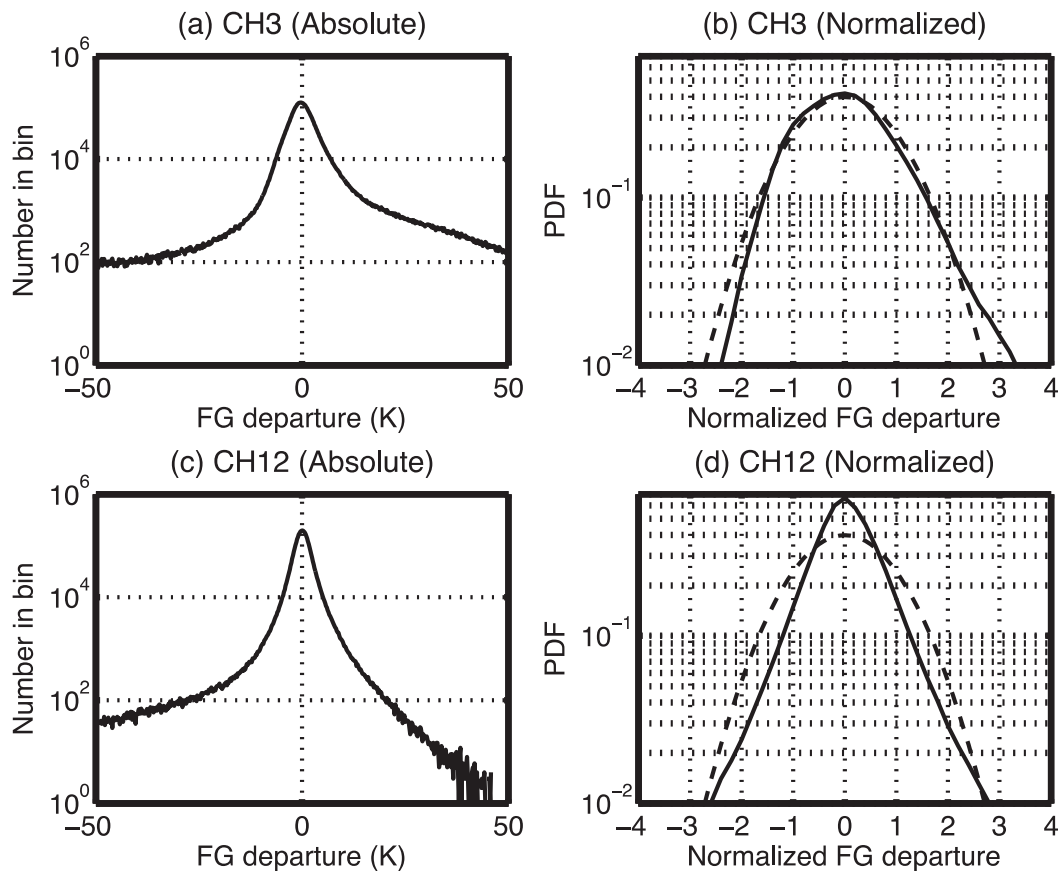


FIG. 6. Histograms of (left) absolute first-guess departures and (right) normalized first-guess departures for all GMI channels 3 and 12 observations between 1 and 31 Dec 2015. The dashed lines in the right column show the Gaussian distribution with corresponding standard deviations for comparison.

by the solid line in Fig. 5. Current implementation of all-sky GMI data assimilation in the operational GEOS using these observation error functions made significant positive impacts as briefly described in section 7 and we are planning to tune even further and to consider correlated observation errors.

To utilize all-sky GMI radiance data, quality control procedures were modified to keep as many cloud and precipitation affected observations as possible. All-sky data over ocean with SST lower than 278 K were screened out to avoid sea ice affected radiance data. In addition, gross check in quality control procedures screen out all-sky GMI data with absolute value of first-guess departure is greater than 2 times of observation error. About 24% of data were screened out due to the quality control procedures.

Variances of the GMI first-guess departures provide an estimate of the total error that is the sum of the observation and forecast error variance in observation space. The standard deviation of the first-guess departures is considered as the upper bound of the observation error

for unbiased observations under the assumption that observation errors are uncorrelated. To account for residual biases and correlated observation errors, the modeled observation errors for initial implementation in GEOS are inflated as denoted by the solid lines in Fig. 5. Figure 6 shows the probability distribution function of first-guess departures normalized by the modeled observation error values for channels 3 and 12. For comparison, the dashed line shows a Gaussian function. Normalizing with respect to the symmetric observation error model results in a more Gaussian distribution of first-guess departure values as compared with the constant observation error values used for clear-sky satellite radiances. This is consistent with the results shown in Geer and Bauer (2011).

e. Bias correction

As with clear-sky radiances in GEOS, bias correction for all-sky microwave radiances is performed using a variational bias correction scheme (VarBC; Derber and Wu 1998) that estimates bias correction coefficients as

part of the variational assimilation. For clear-sky microwave radiances, the bias predictors include a constant, the scan angle, a second-order polynomial of the atmospheric temperature lapse rate weighted by the radiance weighting function, and the retrieved cloud liquid water path. For the all-sky implementation, three changes were made to VarBC. First, the retrieved cloud liquid water path is excluded as a predictor. Second, only near-clear-sky observations with near-clear-sky background profiles are used in updating the bias correction coefficients. Data used for bias correction coefficient updates are restricted to regions where (i) the observed cloud index, CI_o , is less than 0.05; (ii) the simulated cloud index, CI_g , is less than 0.05; and (iii) the absolute difference between CI_o and CI_g is less than 0.005. The enhanced radiance bias correction scheme (Zhu et al. 2014), which adjusts the background error covariances for the bias correction coefficients automatically using an approximation of the analysis-error variances from the previous cycle, is being tested in the GEOS ADAS. The scheme was not applied yet in this study and it is expected to be included in the GEOS all-sky system in near future.

Figure 7 compares histograms of first-guess departures of all assimilated GMI TB data for two channels before (dashed line) and after (solid line) bias corrections for a low-frequency channel sensitive to emission from precipitation (channel 3), and a high-frequency channel sensitive to scattering caused by precipitation (channel 13). Roughly 35% of the assimilated observations for each channel are used to update the bias correction coefficients. Before bias correction is applied, the mode of the channel 3 first-guess departures is 4.0 K and the mode of channel 13 first-guess departures is -2.2 K. The VarBC system performs well, as indicated by reduction of the channel 3 mode to -0.4 K and channel 13 mode to 0.2 K after the correction is applied. Of the original bias correction coefficients, the scan angle bias predictor provides the largest contribution to the total bias correction (not shown). However, remaining biases associated with thick cloud and heavy precipitation were identified. In Fig. 8, the thin curves show the first-guess departure biases calculated with all assimilated data, binned as a function the CI_{avg} , for the same two channels as in Fig. 7. First-guess departure bias increases with cloud index, reaching values close to $+10$ K for channel 3 (19 GHz), which has large sensitivity to emission from rain, and values close to -10 K for channel 13, which has large sensitivity to scattering caused by precipitation. Similar behavior occurs for other low- and high-frequency channels (not shown).

Geer and Bauer (2011) showed that removing sampling bias from the observations in an all-sky assimilation

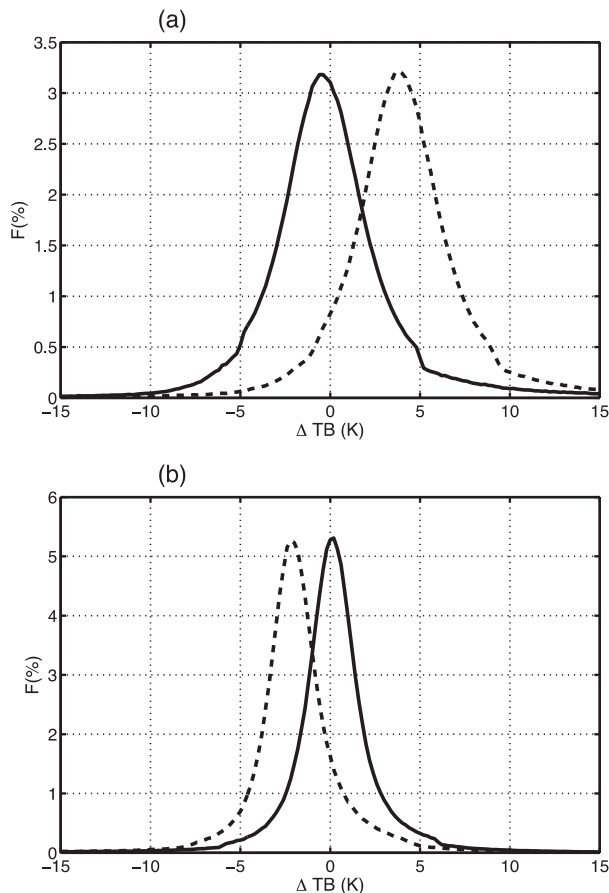


FIG. 7. Histograms of all-sky first-guess departures in GMI (a) channel 3 and (b) channel 13 before (dashed line) and after (solid line) bias correction is applied.

framework reduces the impact of the observations by removing real information, and therefore recommended using the symmetric cloud amount, CI_{avg} , as a predictor for bias correction. Similarly, Chambon et al. (2014) used an averaged scattering index as the predictor for an empirical bias correction model in assimilating all-sky SSMIS 150 GHz TB observations in the Goddard WRF Ensemble Data Assimilation System. Therefore, we made a third change to the VarBC by using CI_{avg} and CI_{avg}^2 as two additional bias correction predictors in the existing VarBC algorithm to correct the aforementioned cloud amount-dependent first-guess bias. The coefficients for these new predictors interact with coefficients for the bias predictors used to assimilate clear-sky radiance data in a way to minimize the cost function. The bias correction coefficients for these predictors are updated each cycle using only data where both the observed and simulated cloud indices are greater than 0.05 and their absolute difference is less than 0.005. After adding this CI-dependent bias correction, the first-guess departure biases are reduced to less than 2 K in all

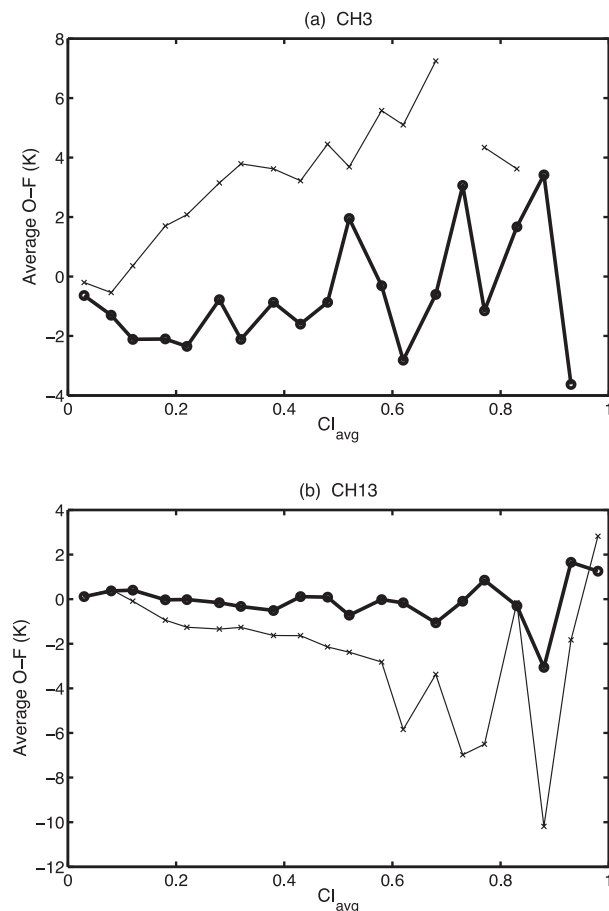


FIG. 8. Bias of first-guess departure as a function of CI_{avg} for GMI channels (a) 3 and (b) 13. Thin solid (thick solid) lines show the biases before (after) using CI_{avg} as additional predictors in Var BC. All assimilated data points between 1 and 31 Dec 2015 were used. Bin size is 0.05. Results only in the bins that have the number of data points greater than 5 are shown in this figure. The biases were computed from the same experiments using same background fields and observation error models except using symmetric cloud as an additional bias correction predictor.

CI ranges for both low- and high-frequency GMI channels, as indicated by the thick curves in Fig. 8. The oscillations in bias seen in high CI_{avg} (Fig. 8) are caused by small samples with large observed or model clouds collected each cycle. It is noted that there are some points missing above $CI_{avg} > 0.7$ in the before-cloud predictor results, but not in the “after” results because more observations were thrown out through quality control procedures when using original bias correction scheme.

5. Performance of all-sky system: Single-observation experiments

To demonstrate the performance of the all-sky radiance framework developed in this study, single observation

experiments are conducted (i.e., using the full analysis scheme but assimilating only GMI observations at a single location and time). Here, cloud- and precipitation-affected GMI TBs from channels 3, 5, 6, 10, 12, and 13 are assimilated at a selected single-observation location. All six GMI channels at each selected observation location are assimilated simultaneously. A particular focus is on examining the sensitivity of analysis increments of hydrometeors to the representation of the background error covariance matrix in its climatological, flow-dependent, or hybrid form. These experiments use the modified observation error, VarBC, and CRTM formulations as discussed earlier.

Two observation points were selected at 1200 UTC 12 December 2015 near Typhoon Melor. Melor was a powerful tropical cyclone that made landfall on the Philippines at 0543 UTC 14 December 2015. The storm developed as a low pressure area 120 km east of Chuuk on 7 December 2015. It intensified into a tropical depression on 10 December, and then into a named tropical storm south of Yap. At 0000 UTC 13 December 2015, Melor became a typhoon and made its first landfall on northern Samar. Figure 9 shows observed TB and the first-guess departures fields of GMI channel 6 near the storm and two single observation locations (marked in white star symbols) selected for two cases:

- Case A: The observations are located near the outer edge of the storm (11.1°N, 130.9°E). The first-guess profiles from GEOS have large amounts of precipitation and cloud, which results in an emission signal in the CRTM-calculated first-guess TBs in the low-frequency channels. These are warmer than the GMI-observed TBs, indicating that the observations lack the emission from cloud and rain. First-guess TBs in high-frequency channels are much lower than GMI observations due to scattering effects related to frozen precipitation. Table 2 shows the departure between the GMI observations and the first-guess TBs.
- Case B: The observations are located near the center of the storm (12.41°N, 134.0°E). GMI observes thick cloud and precipitation while the model has too little precipitation compared with the observations, as evident by the positive first-guess departures at low frequencies in Table 4. Also, GMI shows the presence of large amounts of frozen precipitation as indicated by the large negative first-guess departures at high frequencies (Table 4).

Figure 10 shows first-guess hydrometeor profiles (left column) for both cases and CRTM-calculated TB hydrometeor Jacobians (right column) for case A. There is a large amount of liquid cloud and rainwater in the first guess in case A (solid lines) compared to case B

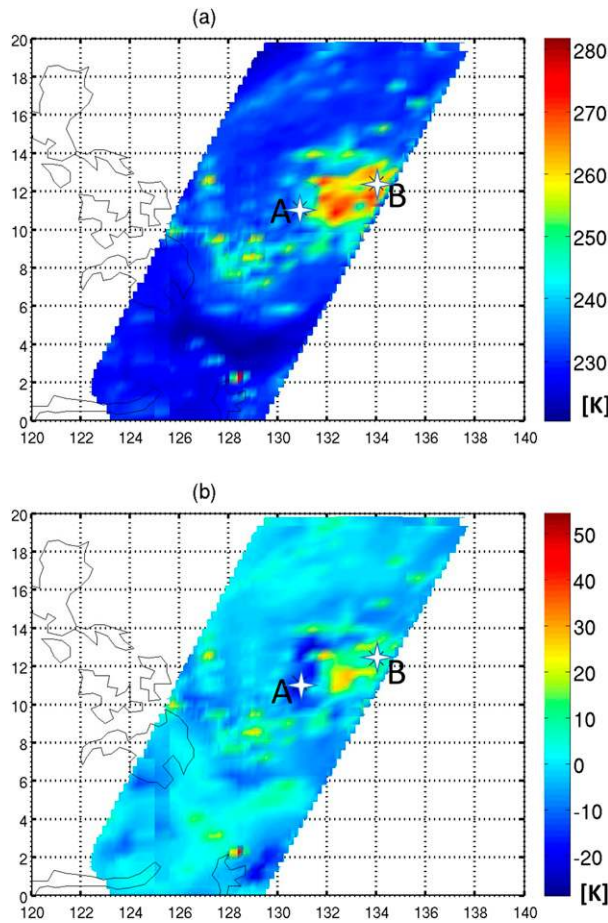


FIG. 9. (a) Channel 6 Tb observations (K) and (b) first-guess departure fields (K) near Hurricane Melor (1200 UTC 12 Dec 2015). White star symbols in the figures show the locations of case A and case B selected for single observation tests in this study.

(dashed lines). For liquid cloud (Fig. 9e), channels 3–6 show positive TB sensitivity and channels 10–13 show large negative TB sensitivity, especially in the mid-to upper troposphere. For ice cloud (Fig. 9f), high-frequency channels 10–13 show negative sensitivity. For emission channels, TB sensitivity to ice cloud is negligibly small. For rain (Fig. 9g), TBs for channels 3–5 show positive sensitivity in most layers with rain hydrometeors present in the first guess. Among the high-frequency channels used in this study, channel 10 shows the largest negative sensitivity to rain, with maximum values at 600 hPa. Channels 12 and 13 show much smaller sensitivity to rainwater than channel 10 in both case A and case B. For snow (Fig. 9h), channels 10–12 show large negative TB sensitivity that increases with altitude, while channels 3–6 show negligibly small sensitivity to snow in both case A and case B.

For each of these cases, the analysis sensitivity to the different choices of background error covariance is

TABLE 2. Brightness temperatures in single observation case A.

	Departure (K)			
	First guess	Analysis		
		Pure B_c	Pure B_e	Hybrid B
CH3	-23.3	-17.6	-13.2	-15.9
CH5	-13.0	-11.0	-9.2	-10.3
CH6	-17.7	-16.7	-16.0	-16.6
CH10	11.9	11.7	9.1	10.4
CH12	0.68	0.64	0.1	0.4
CH13	3.44	3.3	1.8	2.6

examined hereafter by applying different weights to the climatological and ensemble background error covariances. This is done by applying different sets of values for β_c and β_e in Eq. (2): (i) pure climatological background error covariance ($\beta_c = 1$ and $\beta_e = 0$), (ii) pure ensemble background error covariance ($\beta_c = 0$, $\beta_e = 1$), and (iii) hybrid background error covariance ($\beta_c = 0.5$, $\beta_e = 0.5$).

Tables 2–5 list observed TBs, departures, observed, and modeled CI, the value of total column water vapor (TCWV) and total column hydrometeor amount at the observation point and time in both the first-guess and analysis when using the three different background error configurations.

At observation location A, where first-guess departures for emission channels show large negative values, the observed CI is much smaller than first-guess CI, indicating the model has much larger amounts of hydrometeors compared to the observations. Assimilating GMI TBs at observation location A reduces the first-guess departures, which in turn leads to reduced rain, snow, and ice cloud in the background fields. As summarized in Table 2, the reductions in the first-guess departures for all assimilated GMI channels are largest when pure B_e is used and smallest when pure B_c is used. In some cases, the use of B_c only even deteriorates the analysis results. While this is not desirable, it is not surprising considering the properties of the climatological statistics, and will be discussed next. The use of hybrid covariances on the other hand offers a compromise between the two scenarios.

Figures 11 and 12 compare the vertical distribution of analysis increments of temperature, humidity, and hydrometeors from the different background error configurations at observation locations A and B respectively. The similarities in the vertical structure of the analysis increments can be explained by examining the vertical structure of error covariances themselves and shown in Fig. 13 where the vertical distribution of climatological background error standard deviations (B_c) is compared to ensemble standard deviations of

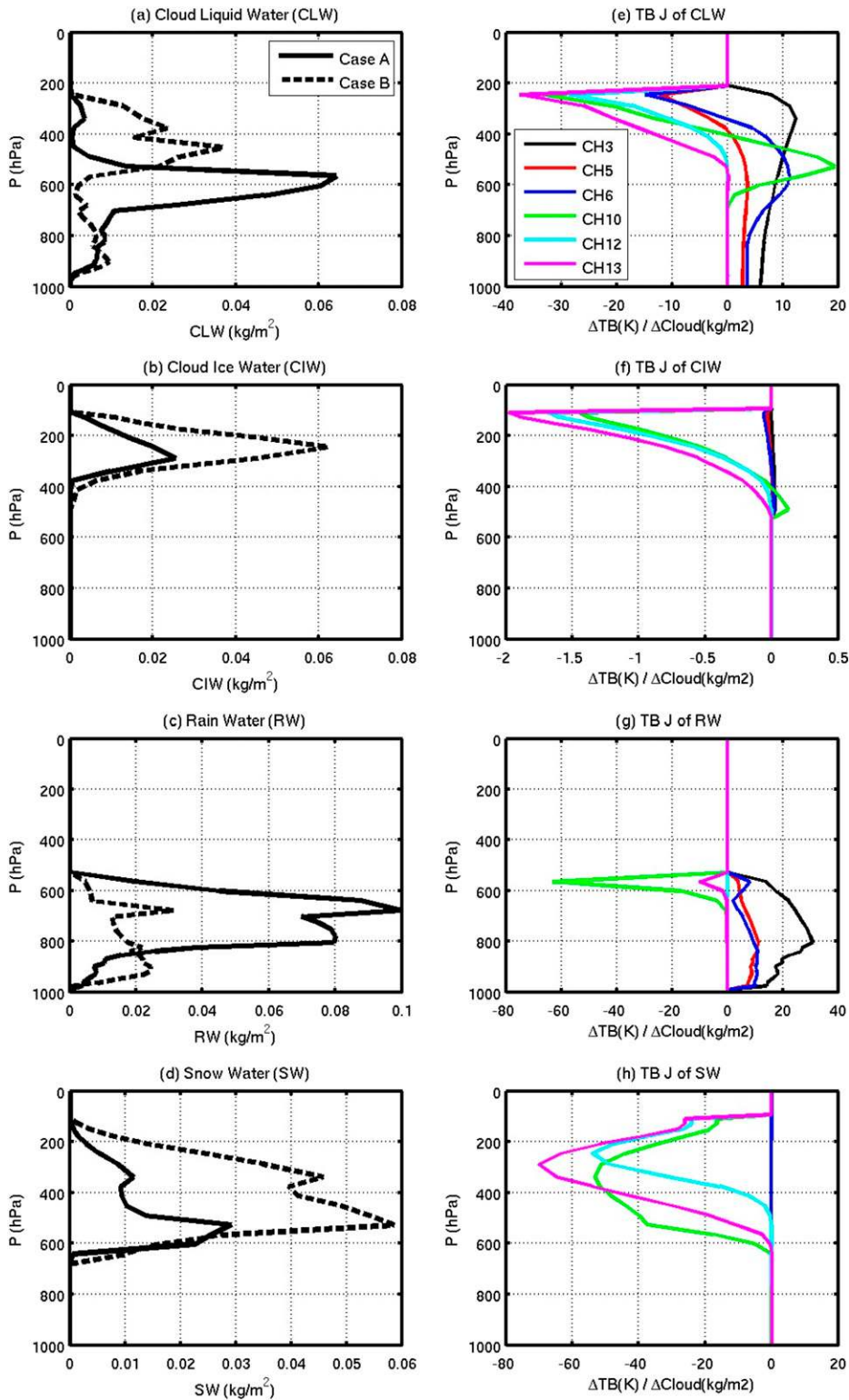


FIG. 10. (a)–(d) GEOS-5 first-guess hydrometeor profiles at locations A (solid line) and B (dashed line). (e)–(h) CRTM calculated TB Jacobian of hydrometeors at location A for GMI channels 3, 5, 6, 10, 12, and 13.

TABLE 3. Moisture variables in single observation case A (Here, CI: cloud index, TCWV: total column water vapor, CLW: cloud liquid water, CIW: cloud ice water, RW: rainwater, SW: snow water).

	First guess	Analysis		
		Pure B_c	Pure B_e	Hybrid B
CI observed: 0.91	0.87	0.79	0.70	0.76
TCWV (g kg^{-1})	330.80	330.77	325.95	328.77
CLW (kg m^{-2})	0.32	0.32	0.32	0.32
CIW (kg m^{-2})	0.091	0.090	0.084	0.088
RW (kg m^{-2})	0.091	0.60	0.48	0.56
SW (kg m^{-2})	0.091	0.15	0.09	0.11

hydrometeors (B_e) at observation location A. Although the altitude with maximum B_e for snow water is slightly higher than the altitude with maximum B_c for snow water, both B_e and B_c show generally similar vertical structures for all four hydrometeors. This is not particularly surprising as both methods represent relatively well the vertical distribution of predominance of rain, snow, liquid cloud, and ice cloud in the atmosphere. The difference is, however, more apparent in terms of magnitude, where B_e standard deviations are shown to be larger than those of B_c , especially for rainwater and snow water in the tropical cyclone location selected for the single observation case. Such differences will inevitably be reflected in the analysis increments and on the first-guess departures in each of the three scenarios. The takeaway here is that the smoothing, and time/zonal-averaging needed to produce the static B statistics would render them inadequate for representing error covariances in highly evolving weather conditions; such is the case for the localized and rapidly changing hydrometeors; This was one of our motivation in this paper to examine the all-sky analysis sensitivity in a hybrid data assimilation framework rather than a strict 3D-var for instance where the covariances are inherently static.

It is worth mentioning here that the difference in the impact is due not only to the difference in the error variances (standard deviations), but also to the localization function/correlation scales (not shown) associated with

TABLE 4. TBs in single observation case B.

	Departure (K)			
	First guess	Analysis		
		Pure B_c	Pure B_e	Hybrid B
CH3	43.1	45.1	32.3	43.5
CH5	4.4	5.3	-0.1	4.5
CH6	6.9	6.3	2.5	4.7
CH10	-53.0	-58.7	-45.4	-57.4
CH12	-24.6	-27.8	-22.2	-27.5
CH13	-39.2	-47.1	-33.8	-46.5

TABLE 5. Moisture variables in single observation case B.

	First guess	Analysis		
		Pure B_c	Pure B_e	Hybrid B
CI observed: 0.91	0.41	0.53	0.60	0.58
TCWV (g kg^{-1})	342.1	342.2	344.6	343.0
CLW (kg m^{-2})	0.25	0.25	0.25	0.25
CIW (kg m^{-2})	0.23	0.24	0.25	0.24
RW (kg m^{-2})	0.32	0.47	0.60	0.56
SW (kg m^{-2})	0.43	0.43	0.55	0.50

B_e/B_c , respectively, which control the spatial distribution of the analysis increment, and finally to the implicit cross correlations featured in ensemble covariances between the hydrometeors and other variables.

At observation location B, the first-guess departures for emission channels show large positive values and the observed CI is much larger than the first-guess CI, indicating that the model has too little precipitation (Table 3). As shown in Fig. 12, assimilating GMI TBs at observation location B increases cloud and hydrometeor amounts, which corresponds to the reduction in the analysis departures as well as in the difference between the observed and modeled CI. Again, the reductions in the first-guess departures for all assimilated GMI channels for Case B are largest when pure B_e is used and smallest when pure B_c is used, indicating the suboptimality of the latter. Accordingly, β_c will be set to zero for hydrometeors in the production version of GEOS, while β_c for all other analysis variables is set to 0.5. Note though that with such a choice, one needs to be cautious about the possibility of sharply localized analysis increment corrections that may affect the conditioning of the all-sky data assimilation problem.

6. Impacts on dynamic variables

As shown in Figs. 11 and 12, the temperature and humidity analysis increments using pure B_e have much more distinct vertical structures and larger magnitude than those using the pure B_c configuration. The former implicitly incorporates correlations between different analysis variables in GEOS and implicitly generates analysis increments for other dynamic variables such as humidity, wind, temperature, and surface pressure. The magnitudes of these increments depend on the correlations between the hydrometeors and dynamic variables, and on the magnitude of the ensemble spread.

Geer et al. (2014) report that the impact of the all-sky microwave TB data assimilation in the ECMWF 4D-Var system comes through the ‘model tracing’ effect of water vapor, cloud, and precipitation (i.e., inferring winds from the motion of humidity, clouds, and precipitation).

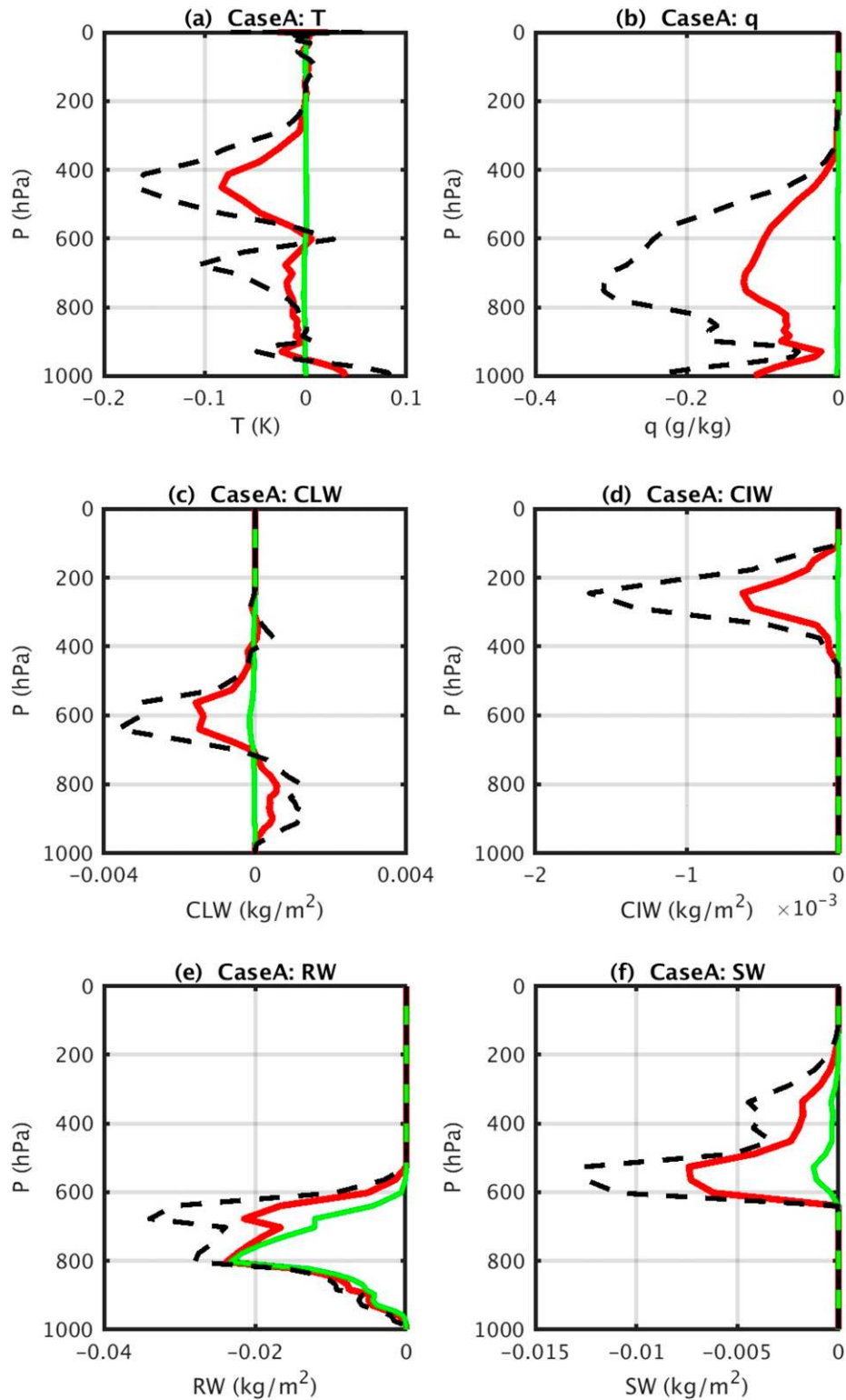


FIG. 11. Analysis increments at observation location A from experiments using three different background error configurations: pure climatological *B* (green solid), pure ensemble *B* (black dashed), and hybrid *B* (red solid). (a) temperature, (b) water vapor mixing ratio, (c) liquid cloud, (d) ice cloud, (e) rainwater content, and (f) snow water content.

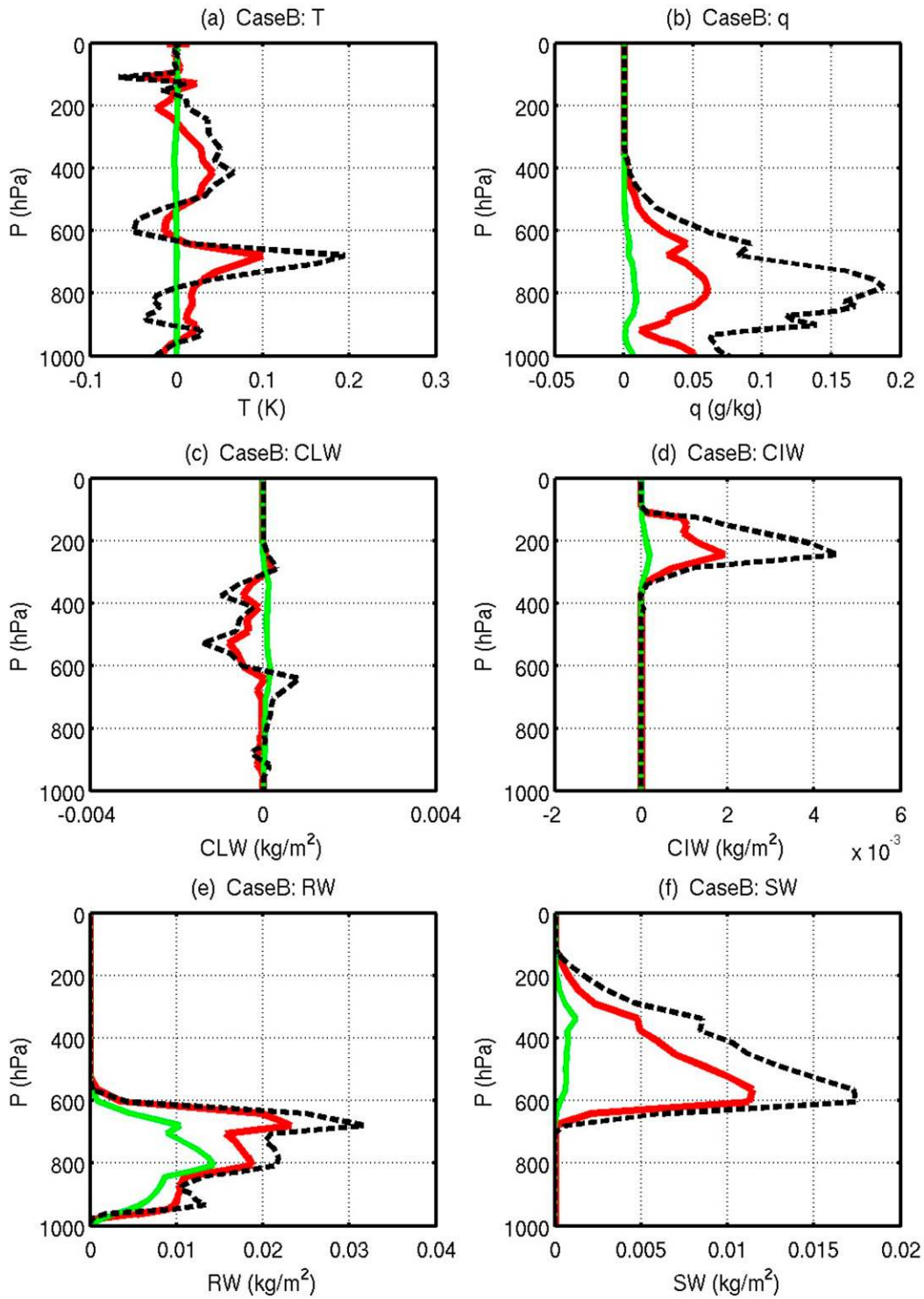


FIG. 12. As in Fig. 11, but for analysis increments at observation location B.

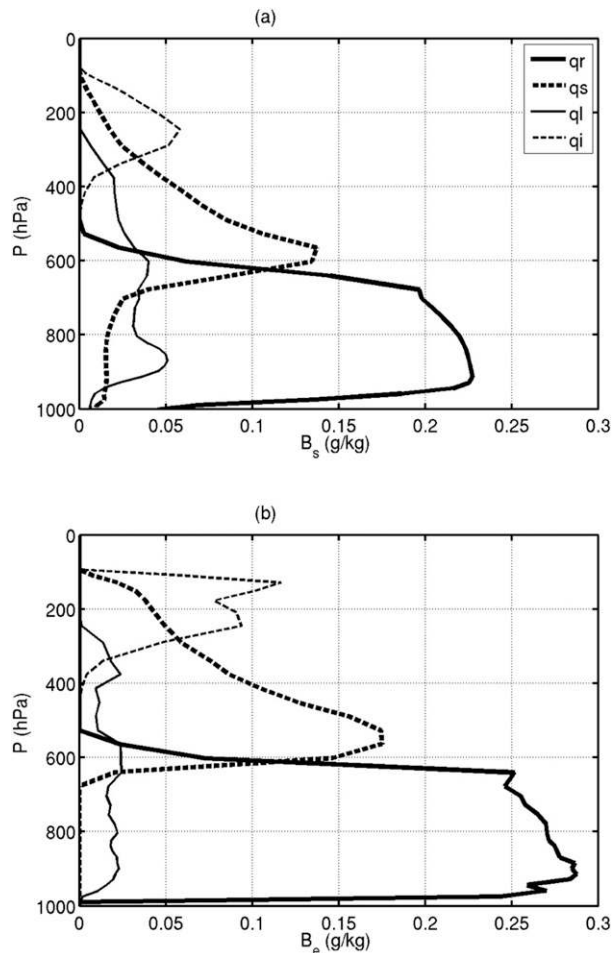


FIG. 13. (a) Static climatological background error standard deviations (B_c) and (b) ensemble background error standard deviations of rain (thick solid line), snow (thick dashed line), liquid cloud (thin solid line), and ice cloud (thin dashed line) at observation locations A.

For example, they found that the benefit of all-sky humidity sounder data was greatest in the southern mid-latitudes, where the storm tracks provide ideal conditions for model tracing. Since a hybrid 4D-EnVar system utilizes flow dependent background error distributions that evolve with time, it is worthwhile checking if similar model tracing effects can be identified here. To examine the impact of all-sky GMI TB data assimilation on wind analyses in GEOS, a case study of Hurricane Gaston in 2016 is presented. Here, the full complement of all-sky GMI data only are assimilated with hybrid 4D-EnVar GSI system during the 6-h assimilation window from 0300 to 0900 UTC 30 August 2016. Here β_c is set to zero for hydrometeors, same as in the production version of GEOS, while β_c for all other analysis variables is set to 0.5. The background fields and bias correction coefficients were taken from the cycled experiments

assimilated all the data used routinely in GEOS (conventional data, satellite radiance data from microwave and infrared sounders, GPS radio occultation data, and satellite wind data) and all-sky GMI data.

Figure 14a shows the location of Hurricane Gaston in GOES East satellite imagery while Fig. 14b shows the all-sky GMI observations assimilated in this case study. Figures 14c–e show the horizontal distribution of analysis increments of 850 hPa rainwater mixing ratio and specific humidity near the storm system at different analysis times in response to the assimilation of GMI all-sky radiances. The assimilation moves precipitation originally mislocated in the forecast toward the observations via generation and removal of precipitation near the storm (Figs. 13c,d, shaded contours). At the same time, moisture fields (lined contours) are adjusted in the analyses by generating positive (negative) analysis increments of 850 hPa specific humidity in the regions of positive (negative) rain analysis increments. Similarly, surface pressure is adjusted in the analysis (Figs. 13e,f) in conjunction with cyclonic wind increments (Figs. 13e,f, vectors) where large precipitation increments are generated near the center of the hurricane. These changes in both the analyzed moisture and dynamic variables through assimilation of GMI data make impacts on GEOS forecasts, which will be discussed in the companion paper.

7. Summary and discussion

A framework for assimilating all-sky microwave radiance data has been developed in the GEOS atmospheric data assimilation system. The performance of various components of the system was examined using reduced observing system experiments. The results demonstrated the capability of the all-sky assimilation to make appropriate changes to the model background state both when the forecast produces excessive precipitation and too little precipitation. The sensitivity of analysis increments of water variables to the representation of the background error covariance matrix in its climatological, flow-dependent (i.e., ensemble-based), and hybrid forms was also examined. Vertical distributions of analysis increments from the hybrid configuration were similar to the analysis increments from the pure ensemble-based configuration, but significant differences in the magnitude of the first-guess departure reduction and analysis increments between the climatological and ensemble-based configurations were identified: The reductions in the first-guess departures for all assimilated GMI channels are largest when pure B_e is used and smallest when pure B_c is used. Results suggested that the use of B_c only even deteriorates the

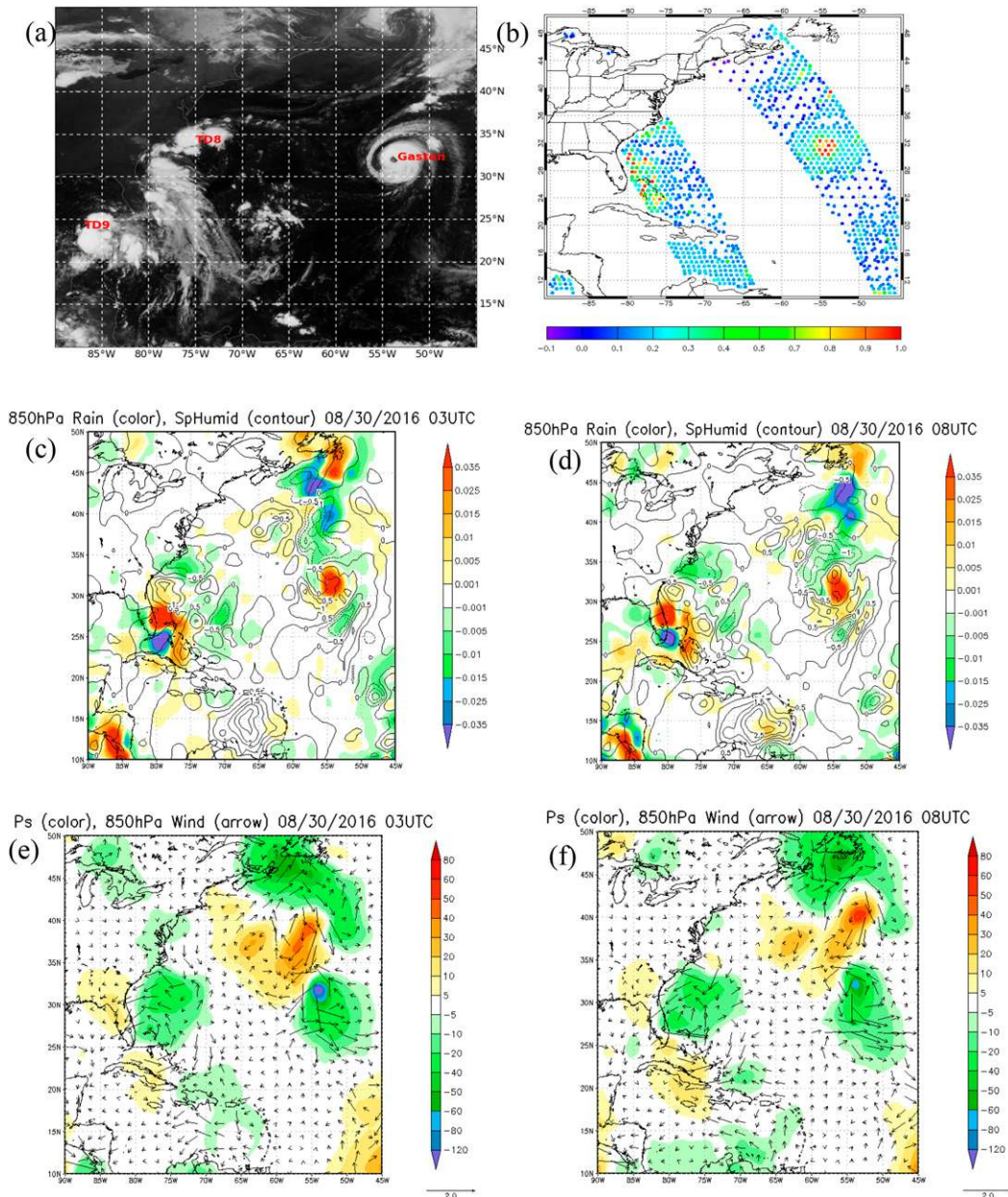


FIG. 14. (a) GOES-East IR imagery at 1015 UTC 30 Aug 2016, (b) 1 – GMI observed 37 GHz TBs polarization differences at 0300–0900 UTC 30 Aug 2016, (c),(d) analysis increments of 850 hPa rain (g kg^{-1}) and specific humidity (g kg^{-1}) at 0300 and 0800 UTC 30 Aug 2016. (e),(f) Analysis increments of surface pressure (Pa) and 850 hPa wind (m s^{-1}) generated by assimilating all-sky GMI radiance data at 0300 and 0800 UTC 30 Aug 2016. In addition to hydrometeors, dynamic variables such as wind, temperature, and pressure are adjusted by assimilation of GMI all-sky radiances in hybrid 4D-EnVar. For this specific case study, an adaptive thinning method (more data being used in the presence of clouds and precipitation) was applied as seen in (b).

analysis results in some cases. Finally, it was seen that the ensemble background error incorporates correlations between different analysis variables implicitly in the GEOS analysis system and generates analysis increments for other dynamic variables such as humidity, wind, temperature, and surface pressure.

Analysis increments of hydrometeors are not fed back to the forecast model in the current implementation of all-sky radiance assimilation in GEOS. It should be noted that hydrometeor increments induce adjustments to dynamic variables such as wind and pressure in the hybrid 4D-EnVar system. These adjustments in both

TABLE A1. Fitting coefficients for Eq. (A1) used in this study to calculate input effective radius values (μm) depending on input water contents (g m^{-3}) in GSI.

	A	B	C	D
Rain	7.93	90.86	387.81	679.94
Long hexagonal column	5.71	56.94	268.22	607.41
Short hexagonal column	4.60	45.81	215.81	488.72
Block hexagonal column	4.04	40.23	189.52	429.17
Thick hexagonal plate	4.55	45.34	213.57	483.64
Thin hexagonal plate	5.86	58.44	275.29	623.40
Three-bullet rosette	9.33	84.78	351.14	691.39
Four-bullet rosette	10.61	93.34	368.99	690.09
Five-bullet rosette	10.24	90.09	356.15	666.08
Six-bullet rosette	9.77	86.06	340.83	638.71
Sector-like snowflake	12.43	102.87	372.06	629.84
Dendrite snowflake	12.64	108.25	411.81	747.96

the analyzed moisture and dynamic variables through assimilation of all-sky GMI data contribute to GEOS forecast improvements. In the current operational GEOS-FP system, the addition of all-sky GMI radiances has the largest impact in the tropics. Specific humidity is significantly improved in the short-term (0–72 h) forecasts. Similar improvements are seen in the Tropical and lower-tropospheric temperature and winds. More detailed results on all-sky GMI data impact on GEOS forecasts and analyses will be discussed in the companion paper.

The current all-sky framework will be enhanced by various updates both in the forecast model and analysis scheme in the near future. For example, the inclusion of a two-moment microphysics scheme (Barahona et al. 2014) in GEOS will provide estimates of cloud particle size distributions to the all-sky observation operator. In addition, future versions of the CRTM will account for cloud fraction in calculating radiances. This should improve the simulation of brightness temperature compared with the current version of CRTM, which considers only clear-sky or completely overcast conditions. All these enhancements are expected to extend the scope of all-sky radiance assimilation to include more microwave measurements and, in turn, lead to improved analyses and forecasts.

Acknowledgments. We thank Dr. Gail Skofronick-Jackson at NASA HQ for having interests in our work. This work has been supported by NASA Grant NNX16AE50G for research project titled “All-sky GPM Microwave Imager (GMI) radiance data assimilation global products from the GEOS-5 system in support of the GPM mission” and by NASA Grant 80NSSC19K0682 for research project titled “Satellite Data Assimilated 4D Global Precipitation Products from the GEOS System in Support of the GPM Mission.”

APPENDIX

Coefficients Used for Fitting Effective Radius Values and Water Contents

Table A1 shows the fitting coefficients for the following fitting functions that we used in the GSI to define an input effective radius (R_{eff}) for a given input water content (WC) value to read in consistent optical properties from the CRTM lookup table:

$$R_{\text{eff}} = A(\log_{10} \text{WC})^3 + B(\log_{10} \text{WC})^2 + C(\log_{10} \text{WC}) + D. \quad (\text{A1})$$

REFERENCES

- Akella, S., R. Todling, and M. J. Suarez, 2017: Assimilation for skin SST in the NASA GEOS atmospheric data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **143**, 1032–1046, <https://doi.org/10.1002/qj.2988>.
- Baordo, F., and A. J. Geer, 2016: All-sky assimilation of SSMIS humidity sounding channels in all-sky conditions over land using a dynamic emissivity retrieval. *Quart. J. Roy. Meteor. Soc.*, **142**, 2854–2866, <https://doi.org/10.1002/qj.2873>.
- Barahona, D., A. Molod, J. Bacmeister, A. N. Gettelman, H. Morrison, V. Phillips, and A. Eichmann, 2014: Development of two-moment cloud microphysics for liquid and ice within the NASA Goddard Earth Observing System model (GEOS-5). *Geosci. Model Dev.*, **7**, 1733–1766, <https://doi.org/10.5194/gmd-7-1733-2014>.
- Bauer, P., A. J. Geer, P. Lopez, and D. Salmod, 2010: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Quart. J. Roy. Meteor. Soc.*, **136**, 1868–1885, <https://doi.org/10.1002/qj.659>.
- , A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Bormann, N., A. J. Geer, and P. Bauer, 2011: Estimates of observation error characteristics in clear and cloudy regions of microwave imager radiances from numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **137**, 2014–2023, <https://doi.org/10.1002/qj.833>.
- Chambon, P., S. Q. Zhang, A. Y. Hou, M. Zupanski, and S. Cheung, 2014: Assessing the impact of pre-GPM microwave precipitation observations in the Goddard WRF ensemble data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **140**, 1219–1235, <https://doi.org/10.1002/qj.2215>.
- Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, [https://doi.org/10.1175/1520-0493\(1998\)126<2287:TUOTCC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2).
- Draine, B. T., and P. J. Flatau, 1994: Discrete-dipole approximation for scattering calculations. *J. Opt. Soc. Amer.*, **11A**, 1491–1499, <https://doi.org/10.1364/JOSAA.11.001491>.
- Field, P. R., A. J. Heymsfield, and A. Bansemmer, 2007: Snow size distribution parameterization for midlatitude and tropical ice clouds. *J. Atmos. Sci.*, **64**, 4346–4365, <https://doi.org/10.1175/2007JAS2344.1>.
- Geer, A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037, <https://doi.org/10.1002/qj.830>.

- , and F. Baordo, 2014: Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmos. Meas. Tech.*, **7**, 1839–1860, <https://doi.org/10.5194/amt-7-1839-2014>.
- , P. Bauer, and P. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Quart. J. Roy. Meteor. Soc.*, **136**, 1886–1905, <https://doi.org/10.1002/qj.681>.
- , F. Baordo, N. Bormann, and S. English, 2014: All-sky assimilation of microwave humidity sounders. ECMWF Tech. Memo. 741, 57 pp., <https://www.ecmwf.int/sites/default/files/elibrary/2014/9507-all-sky-assimilation-microwave-humidity-sounders.pdf>.
- , and Coauthors, 2018: All-sky satellite data assimilation at operational forecasting centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1191–1217, <https://doi.org/10.1002/qj.3202>.
- Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM): Version 1. NOAA Tech. Rep. NESDIS 122, 33 pp., <https://repository.library.noaa.gov/view/noaa/1157>.
- Hartmann, D. L., 1994: *Global Physical Climatology*. Academic Press, 411 pp.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Ide, K., P. Courtier, M. Ghil, and A. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189, https://doi.org/10.2151/jmsj1965.75.1B_181.
- Kazumori, M., 2014: Satellite radiance assimilation in the JMA operational mesoscale 4DVar system. *Mon. Wea. Rev.*, **142**, 1361–1381, <https://doi.org/10.1175/MWR-D-13-00135.1>.
- , 2016: Assimilation of cloud and precipitation affected microwave radiance data using the JMA's global NWP system. *Eighth IPWG and Fifth IWSSM Joint Workshop Abstracts*, Bologna, Italy, IPWG and IWSSM, p. 74.
- Kim, M.-J., 2006: Comparisons of single scattering approximations of randomly oriented ice crystals at microwave frequencies. *J. Geophys. Res.*, **111**, D14201, <https://doi.org/10.1029/2005JD006892>.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691–1705, <https://doi.org/10.1175/2009WAF2222201.1>.
- Liu, E. H., and Coauthors, 2019: EMC contributions to CRTM development and validation. *JCSDA Quarterly*, No. 63, Spring 2019, 30 pp., <https://doi.org/10.25923/c23x-ac34>.
- Liu, G., 2008: A database of microwave single-scattering properties for nonspherical ice particles. *Bull. Amer. Meteor. Soc.*, **89**, 1563–1570, <https://doi.org/10.1175/2008BAMS2486.1>.
- Liu, Q., and F. Weng, 2006: Advanced doubling-adding method for radiative transfer in planetary atmospheres. *J. Atmos. Sci.*, **63**, 3459–3465, <https://doi.org/10.1175/JAS3808.1>.
- Lorenc, A. C., N. E. Bowler, A. M. Clayton, and S. R. Pring, 2015: Comparison of hybrid-4DVar and hybrid-4D Var data assimilation methods for global NWP. *Mon. Wea. Rev.*, **143**, 212–229, <https://doi.org/10.1175/MWR-D-14-00195.1>.
- Marshall, J. S., and W. M. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165–166, [https://doi.org/10.1175/1520-0469\(1948\)005<0165:TDORWS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2).
- Migliorini, S., and B. Candy, 2019: All-sky satellite data assimilation of microwave temperature sounding channels at the Met Office. *Quart. J. Roy. Meteor. Soc.*, **145**, 867–883, <https://doi.org/10.1002/qj.3470>.
- Molod, A., L. Takacs, M. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann, 2012: The GEOS-5 Atmospheric General Circulation Model: Mean climate and development from MERRA to Fortuna. NASA TM-2012-104606, Vol. 28, 115 pp., <https://gmao.gsfc.nasa.gov/pubs/docs/tm28.pdf>.
- , —, —, and —, 2015: Development of the GEOS-5 Atmospheric General Circulation Model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339–1356, <https://doi.org/10.5194/gmd-8-1339-2015>.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, [https://doi.org/10.1175/1520-0493\(1992\)120<1747:TNMCSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2).
- Pedder, M., M. Haile, and A. J. Thorpe, 2000: Short period forecasting of catchment-scale precipitation. Part I: The role of numerical weather prediction. *Hydrol. Earth Syst. Sci.*, **4**, 627–633, <https://doi.org/10.5194/hess-4-627-2000>.
- Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *J. Comput. Phys.*, **227**, 55–78, <https://doi.org/10.1016/j.jcp.2007.07.022>.
- Rienecker, M. M., and Coauthors, 2008: The GEOS-5 data assimilation system—Documentation of versions 5.0.1, 5.1.0, and 5.2.0. NASA/TM-2008-104606, Vol. 27, 118 pp., https://gmao.gsfc.nasa.gov/pubs/docs/GEOS5_104606-Vol27.pdf.
- Sellers, W. D., 1965: *Physical Climatology*. University of Chicago Press, 272 pp.
- Skofronick-Jackson, G., and Coauthors, 2017: The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Amer. Meteor. Soc.*, **98**, 1679–1695, <https://doi.org/10.1175/BAMS-D-15-00306.1>.
- Takacs, L. L., M. J. Suárez, and R. Todling, 2018: The stability of incremental analysis update. *Mon. Wea. Rev.*, **146**, 3259–3275, <https://doi.org/10.1175/MWR-D-18-0117.1>.
- Todling, R., and A. El Akkraoui, 2018: The GMAO hybrid ensemble-variational atmospheric data assimilation system: Version 2.0. NASA/TM-2018-104606, Vol. 50, 184 pp., <https://gmao.gsfc.nasa.gov/pubs/docs/Todling1019.pdf>.
- Whitaker, J. S., and T. H. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2).
- Wu, W., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, [https://doi.org/10.1175/1520-0493\(2002\)130<2905:TDVAWS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2).
- Zhu, Y., J. Derber, A. Collard, D. Dee, R. Treadon, G. Gayno, and J. A. Jung, 2014: Enhanced radiance bias correction in the National Centers for Environmental Prediction's Gridpoint Statistical Interpolation data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **140**, 1479–1492, <https://doi.org/10.1002/qj.2233>.
- , and Coauthors, 2016: All-sky microwave radiance assimilation in NCEP's GSI analysis system. *Mon. Wea. Rev.*, **144**, 4709–4735, <https://doi.org/10.1175/MWR-D-15-0445.1>.