All-Sky Microwave Radiance Assimilation in NCEP's GSI Analysis System

YANQIU ZHU

I.M. Systems Group, College Park, Maryland

EMILY LIU

Systems Research Group, College Park, Maryland

RAHUL MAHAJAN, CATHERINE THOMAS, DAVID GROFF, PAUL VAN DELST,[a] AND ANDREW COLLARD

I.M. Systems Group, College Park, Maryland

DARYL KLEIST

University of Maryland, College Park, College Park, Maryland

RUSS TREADON AND JOHN C. DERBER

NOAA/NWS/NCEP/Environmental Modeling Center, College Park, Maryland

(Manuscript received 21 December 2015, in final form 6 July 2016)

ABSTRACT

The capability of all-sky microwave radiance assimilation in the Gridpoint Statistical Interpolation (GSI) analysis system has been developed at the National Centers for Environmental Prediction (NCEP). This development effort required the adaptation of quality control, observation error assignment, bias correction, and background error covariance to all-sky conditions within the ensemble–variational (EnVar) framework. The assimilation of cloudy radiances from the Advanced Microwave Sounding Unit-A (AMSU-A) microwave radiometer for ocean fields of view (FOVs) is the primary emphasis of this study.

In the original operational hybrid 3D EnVar Global Forecast System (GFS), the clear-sky approach for radiance data assimilation is applied. Changes to data thinning and quality control have allowed all-sky satellite radiances to be assimilated in the GSI. Along with the symmetric observation error assignment, additional situation-dependent observation error inflation is employed for all-sky conditions. Moreover, in addition to the current radiance bias correction, a new bias correction strategy has been applied to all-sky radiances. In this work, the static background error variance and the ensemble spread of cloud water are examined, and the levels of cloud variability from the ensemble forecast in single- and dual-resolution configurations are discussed. Overall, the all-sky approach provides more realistic simulated brightness temperatures and cloud water analysis increments, and improves analysis off the west coasts of the continents by reducing a known bias in stratus. An approximate 10% increase in the use of AMSU-A channels 1–5 and a 12% increase for channel 15 are also observed. The all-sky AMSU-A radiance assimilation became operational in the 4D EnVar GFS system upgrade of 12 May 2016.

1. Introduction

Satellite radiance data have been used extensively since the 1990s in major numerical weather prediction (NWP) centers across the world. The importance of including radiance data with fields of view (FOVs) affected

[a] Deceased.

Corresponding author address: Yanqiu Zhu, IMSG, NCEP/EMC, 5830 University Research Ct., College Park, MD 20740.
E-mail: yanqiu.zhu@noaa.gov

DOI: 10.1175/MWR-D-15-0445.1

by clouds and precipitation in a data assimilation system is obvious, as these data are usually associated with meteorologically important areas. However, because of the limitations of the forecast model physics and radiative transfer model, as well as the strong nonlinearity of the observation operator under such conditions, radiances from cloud-free FOVs have primarily been used to date. Over the last decade, major efforts have been focused on the assimilation of cloudy radiance data with increasing success. These efforts have been coincident with the improvements seen in the forecast models, in particular with respect to the model physical parameterization schemes, and with the advances in radiative transfer models in modeling the radiative fluxes of cloudy regions. In the early studies, cloudy radiance data were usually preprocessed through a one-dimensional variational data assimilation (1DVAR) scheme to retrieve atmospheric properties before these radiances were assimilated into a data assimilation system. For example, in the assimilation of cloud- and precipitation-affected Special Sensor Microwave Imager (SSM/I) radiance observations, the European Centre for Medium-Range Weather Forecasts (ECMWF) derived total column water vapor (TCWV) from the 1DVAR humidity retrieval profiles and assimilated these TCWV results into the 4DVAR operational system (Bauer et al. 2006a,b). For cloud-affected infrared radiances, retrieved cloud-top pressure and effective cloud fraction were (and still are) used to constrain the radiative transfer calculation in the assimilation system (Pavelin et al. 2008). Since March 2009, all-sky microwave radiance data for SSM/I and the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) have been assimilated directly into the operational 4DVAR data assimilation system at ECMWF (Bauer et al. 2010; Geer et al. 2010). In the all-sky upgrade of 2010, the so-called symmetric observation error method was proposed by Geer and Bauer (2011b). This method was created to increase observation errors as a function of hydrometeor amount, following the observed behavior of first-guess departures, and to avoid sampling errors due to the use of only the first guess or only the observed hydrometeor amount. The direct assimilation of overcast scenes from infrared radiances has been included in the ECMWF's operational system since September 2009, where cloud-top pressure is used as a sink variable (McNally 2009). Studies related to the assimilation of overcast cloudy infrared radiances were also performed at Météo-France (Guidard et al. 2011) and Japan Meteorological Agency (JMA; Okamoto 2013). While there is no cloud control variable in the all-sky radiance assimilation at ECMWF, studies examining all-sky radiance assimilation had been conducted with total relative humidity as a control variable at the Met Office, but later it was replaced by a new humidity control variable with a nonlinear transform (Ingleby et al. 2013) in their clear-sky operational data assimilation system.

Satellite radiance data has been one of the major data sources for the Global Forecast System (GFS) at the National Centers for Environmental Prediction (NCEP). The GFS forecast system includes the GFS forecast model and the Gridpoint Statistical Interpolation (GSI) analysis system. Infrared and microwave radiances have been assimilated successfully in the original operational GFS forecast system since 1998 (Derber and Wu 1998) using a clear-sky approach. As with the clear-sky approaches used by other NWP centers, NCEP's clear-sky approach does not take into account the cloud first-guess information in the calculation of simulated radiance. In addition, the GSI analysis system (Derber et al. 1991; Parrish and Derber 1992; Wu et al. 2002; Purser et al. 2003a; Purser et al. 2003b; Kleist et al. 2009) not only utilizes radiance data with cloud-free FOVs, but the microwave radiances affected by optically thin clouds have also been used with a cloud-related bias correction term. Moreover, cloud water had been used as a control variable in the original operational analysis system.

With the improvements to the Community Radiative Transfer Model (CRTM; Liu and Weng 2006; Bauer et al. 2006c; Ding et al. 2011; Groff et al. 2013, 2014) and the forecast model, the work on all-sky microwave radiance assimilation in the GSI analysis system has progressed at NCEP over the past several years as it has on the 3D EnVar (Wang et al. 2013; Kleist and Ide 2015a) and 4D EnVar (Wang and Lei 2014; Kleist and Ide 2015b) GSI analysis systems. To incorporate the all-sky radiances, relaxations in the criteria for data thinning and quality control of radiance data have been made. Also, cloud information is taking into account in radiative transfer calculations, and brightness temperature Jacobians with respect to hydrometeors are used in projecting the radiance data information as hydrometeors are introduced into the GSI as state variables. To successfully assimilate all-sky microwave radiances, a more appropriate assignment of observation errors is necessary. The symmetric observation error method proposed by Geer and Bauer (2011b) and Geer et al. (2012) is adopted. The existing situation-dependent observation error inflation for clear-sky conditions is augmented by an inflation developed for all-sky conditions to handle the radiances with large departures from the first guess to ensure the information contained in these radiances could be used without degrading the performance of the system. A new bias correction strategy for all-sky radiances (Zhu et al. 2014b) is also applied with the current radiance bias correction scheme (Zhu et al. 2014a). This new strategy for all-sky assimilation employs only radiances with consistent cloud as determined by cloud

liquid water to generate the updates to the bias coefficients. A new static cloud water background error variance is used and the cloud variability in the dual-resolution configuration of the GFS for the cloud ensemble spread (more discussion in section 4a) is examined.

This study focuses on Advanced Microwave Sounding Unit-A (AMSU-A), which is an instrument that has a relatively large impact on analysis and forecast skill (Baker and Langland 2007; Zhu and Gelaro 2008; Gelaro and Zhu 2009; etc.). The AMSU-A microwave radiometer includes 12 sounding channels in the 57-GHz oxygen band and three window channels at 24, 31, and 89 GHz. The atmospheric optical properties for the corresponding window channel frequencies are sensitive to variability in water vapor, cloud, and precipitation. A previous study pertaining to the all-sky assimilation of AMSU-A channels 4 and 5 at ECMWF (Geer et al. 2012) was promising, but only a few additional scenes could be added in channel 5 compared to their clear-sky approach, where channels 1–4 and 15 are not used, and in the end, they found the benefit was not sufficient to justify an operational implementation. AMSU-A channels 1–13 and 15 were assimilated in the original clear-sky operational GFS, and we have been striving to include all-sky AMSU-A radiances in the GFS system. Since the first guesses of the snow and precipitation profiles are not currently available (along with the convective clouds) from the GFS forecast model output, only AMSU-A radiances affected by nonprecipitating clouds over ocean surfaces are assimilated in this study, in addition to AMSU-A radiances from cloud-free FOVs over all surfaces. Despite the challenges of modeling cloud forecasts accurately and cloud variability issues (due to the stochastic physics forcing and the GFS dual-resolution configuration discussed in section 4a), it will be seen that the all-sky radiance assimilation performs reasonably well.

This work has been included in the operational GFS system in the upgrade of 12 May 2016. The configuration of all-sky microwave radiance assimilation for this implementation is described in this paper. As we can see, this work is only the first step in assimilating all-sky radiances, and there is still much room for improvement as development continues. This paper is organized as follows. The clear-sky approach in the original operational GSI analysis system is described briefly in section 2. The GSI developments to support all-sky capabilities, including quality control, observation error, radiance bias correction, and background error covariance, are discussed in section 3. The benefits of all-sky radiance assimilation relative to the clear-sky approach are highlighted in terms of analysis increments in section 4, and its impact on the analysis and forecast skills is presented in section 5. The conclusions and future work are summarized in section 6.

2. The clear-sky approach in the original operational GSI analysis system

For the clear-sky radiance assimilation approach (Derber and Wu 1998) applied in the NCEP's original operational GFS system, preference is given to clear-sky observations in the data-thinning procedure. In this procedure, radiances are scored based on a series of criteria such as the underlying surface properties, the distance between the observation location and the center of the analysis grid box, and the cloud liquid water (CLW) amount. The retrieval formula of Grody et al. (2001) and Weng et al. (2003) is employed in the CLW calculation (in $\mathrm{kg\,m^{-2}}$) for AMSU-A observations with ocean surface FOVs:

$$\mathrm{CLW} = \cos\theta\{c_0 + c_1 \ln[285 - T_b(1)] + c_2 \ln[285 - T_b(2)]\}, \quad (1)$$

where c_0 (in $\mathrm{kg\,m^{-2}}$) is calculated as

$$c_0 = 8.240 - (2.622 - 1.846\cos\theta)\cos\theta, \quad (2)$$

$c_1 = 0.754\ \mathrm{kg\ (m^2\ K)^{-1}}$, $c_2 = -2.265\ \mathrm{kg\ (m^2\ K)^{-1}}$, θ is the zenith angle, and $T_b(1)$ and $T_b(2)$ are the brightness temperatures for AMSU-A channels 1 and 2, respectively. The sea surface temperature is approximated to be 285 K in the retrieval formula. A radiance observation with larger CLW is assigned a higher score, and the radiance with the lowest overall score in a thinning box is selected.

Additionally, model profiles of the atmospheric state are assumed to be cloud free. The radiances for cloud-free FOVs and the microwave radiances for FOVs that include optically thin clouds over ocean surfaces are assimilated. For those FOVs that include optically thin clouds, a difference term ΔCLW is calculated as a difference between CLW from Eq. (1) using the observed and the simulated brightness temperature, and is employed as a predictor in the radiance bias correction scheme (Derber and Wu 1998). This ΔCLW bias correction predictor is removed as we proceed to the all-sky approach.

Moreover, any radiances affected by thick clouds and precipitation are screened out in the GSI. The thick-cloud filtering uses the information of ΔCLW and the brightness temperature information for AMSU-A channel 4:

$$C = \begin{cases} (w_1 \Delta\mathrm{CLW})^2 + [w_2 \Delta T_b(4)]^2 & \text{over ocean} \\ (w_1 \times 0.6)^2 + [w_2 \Delta T_b(4)]^2 & \text{others}, \end{cases} \quad (3)$$

where $\Delta T_b(i)$ is the difference between the observed and the simulated brightness temperature (observed minus

first guess, OmF) for channel i, and the empirical parameters w_1 and w_2 are set to be $(0.3)^{-1}\,\mathrm{m^2\,kg^{-1}}$ and $(1.8)^{-1}\,\mathrm{K^{-1}}$, respectively. Any radiances corresponding to $C > 0.5$ are considered to be affected by thick clouds and are therefore excluded. The screening of precipitating clouds is based on the scattering index information as defined by Grody et al. (1999), and the OmF for AMSU-A channel 6:

$$P = \begin{cases} (w_3\Delta s)^2 + [w_4\Delta T_b(6)]^2 & \text{over ocean} \\ 0.8^2 + [w_4\Delta T_b(6)]^2 & \text{others}, \end{cases} \quad (4)$$

where $w_3 = 0.1\,\mathrm{K^{-1}}$, $w_4 = (0.8)^{-1}\,\mathrm{K^{-1}}$, and the scattering index difference Δs (Grody et al. 1999) is given by

$$\Delta s = [2.41 - 0.0098T_b(1)]\Delta T_b(1) + 0.454\Delta T_b(2) - \Delta T_b(15). \quad (5)$$

Any radiances with $P \geq 1$ are excluded from the GSI. Additionally, if a radiance observation from channel 4, 5, or 6 is removed, the observations peaking below this channel (including channels 1, 2, 3, and 15) will not be used.

Meanwhile, the clear-sky situation-dependent observation error inflation is applied to the channel-by-channel constant observation error e^o_{clr}, and the final observation error $e^o_{\mathrm{clr_{new}}}$ is obtained from

$$(e^o_{\mathrm{clr_{new}}})^2 = (e^o_{\mathrm{adj}})^2 + (\Delta e_{\mathrm{inf}})^2, \quad (6)$$

where the adjusted error variance $(e^o_{\mathrm{adj}})^2 = (z_s/2000)[(e^o_{\mathrm{clr}})^2/\tau]$ for AMSU-A channels 1–6 if surface elevation $z_s > 2000\,\mathrm{m}$, $(e^o_{\mathrm{adj}})^2 = (z_s/4000)[(e^o_{\mathrm{clr}})^2/\tau]$ for channel 7 if $z_s > 4000\,\mathrm{m}$, otherwise $(e^o_{\mathrm{adj}})^2 = (e^o_{\mathrm{clr}})^2/\tau$. The transmittance at model top is τ. The inflation term $(\Delta e_{\mathrm{inf}})^2$ is computed based on the sensitivities of brightness temperature to the surface emissivity ϵ and skin temperature T_s; that is, $(\Delta e_{\mathrm{inf}})^2 = (d_1|\partial T_b/\partial\epsilon| + d_2|\partial T_b/\partial T_s|)^2$, where d_1 and d_2 are empirical parameters depending on different surface types. For the radiances affected by optically thin clouds from AMSU-A channels 1–5 and 15, an additional inflation term $0.2(\beta\Delta\mathrm{CLW})^2$ is also included in Eq. (6), where β is the cloud liquid water bias correction coefficient.

A radiance observation with an OmF magnitude larger than a prespecified observation error upper limit or 3 times the observation error $e^o_{\mathrm{clr_{new}}}$, whichever is smaller, is excluded from the data assimilation system. The prespecified observation error upper limits are set to be 4.5 K for AMSU-A channels 1–3 and 13–15, 3.5 K for channel 12, 2.5 K for channels 4 and 11, and 2.0 K for the other channels.

In the GFS forecast model, cloud water (the sum of the cloud liquid water and cloud ice) is a prognostic variable. The model moist physics includes a cloud microphysics parameterization (Zhao and Carr 1997; Sundqvist et al. 1989; Moorthi et al. 2001) and parameterizations that relate to deep and shallow cumulus convection (Han and Pan 2011), but clouds due to convection are only considered through detraining the convective cloud water to the grid-scale cloud water near the convective cloud tops; thus, the cloud condensate in the convective plume is not included in the total condensate of the forecast model output, and the convective cloud is not available for use. As mentioned earlier, a cloud control variable has also been explicitly employed in the GSI. Combined with the tangent linear and adjoint of the model moist physics, the cloud water control variable was constructed to assimilate the retrieved precipitation rate product from the TRMM Microwave Imager (TMI; Treadon 1997) in the original operational clear-sky GFS forecast system. The assimilation of retrieved precipitation rate data was discontinued in 2013, but the cloud water control variable remains. Thus, even though retrieved precipitation rate data are no longer assimilated and radiance data are assimilated in the clear-sky approach, cloud analysis increments are generated in the 3D EnVar GSI analysis system via the background error cross covariances of cloud with the other variables, such as temperature, moisture, and wind. These nonzero cross covariances are introduced by the ensemble part of the hybrid ensemble–variational background error covariance.

3. The all-sky approach

Over the past few years, efforts have been made at NCEP to assimilate radiances under all-sky conditions. To enable the assimilation of radiances for FOVs with nonprecipitating clouds, elements designed for the clear-sky approach discussed in section 2, except for the precipitation screening procedure, have been either removed or turned off; this is, the preference given to clear-sky radiance data in the thinning, the thick-cloud filtering, and the bias correction term of cloud liquid water difference are no longer needed. Therefore, the radiance data for FOVs coincident with thick clouds, which are not used in the clear-sky approach, are now assimilated in the all-sky approach.

As a result of including cloud information in the inputs to the CRTM in the satellite-radiance observation operator, the simulations of satellite radiances and radiance innovation statistics are more realistic in meteorologically active weather conditions.

Furthermore, in the all-sky approach, because of the use of the cloud control variable and the newly added cloud liquid water and cloud ice state variables in the GSI, the radiance data information is directly mapped

onto not only the temperature and moisture fields as in the clear-sky approach, but also cloud fields via the brightness temperature Jacobians with respect to cloud liquid water and cloud ice. The CRTM does not calculate a cloud Jacobian for cloud-free layers, but this is overcome by providing a cloud seed of $1.001 \times 10^{-6}\,\mathrm{kg\,m^{-2}}$ to the radiative transfer model as input. The GSI all-sky approach includes capabilities to apply either individual hydrometeors or cloud water as cloud control variable(s). For the original operational clear-sky GSI, cloud water is the control variable since it is the cloud prognostic variable in the GFS forecast model. For this reason, a normalized cloud water (i.e., cloud water normalized by its background error standard deviation) control variable has been chosen in current development efforts to improve the realism of the cloud field, including reducing spurious clouds that may be generated from the background error covariance. More details are discussed in section 3e. The decomposition of cloud water (CW) into cloud liquid water and cloud ice is based on temperature T, with $f \times$ CW being cloud ice and the rest cloud liquid water, where $f = 0.05 \times (273.15 - T)$, and f is set to be 0 when f is negative and 1 when f is larger than 1. As our model physics evolves toward the use of individual hydrometeors as prognostic variables, the current GSI cloud control variable(s) capabilities will enable extension to individual hydrometeors in the future. While cloud analysis increments are produced through the background error cross covariance in the clear-sky approach, additional analysis increments are generated for temperature, moisture, and clouds from the projection of the cloudy radiance data information onto the cloud fields in the all-sky approach.

In the following subsections, we describe modifications to the GSI to implement all-sky capabilities. These modifications pertain to observation error (symmetric observation error and situation-dependent observation error inflation), quality control, bias correction, and background error covariance.

a. Symmetric observation error formulation

In the GSI clear-sky approach, a constant observation error e^o_{clr} is assigned to each individual satellite instrument channel. However, for cloudy radiances assimilated in the all-sky approach, the magnitude of OmF is seen to be closely related to cloud amount and the accuracy of the model first guess. While there is no reason that the instrument itself would perform differently for cloudy FOVs, errors in the CRTM simulation are sensitive to clouds. Therefore, it is more appropriate to specify the observation error based on the cloud amount rather than using a constant e^o_{clr}. The estimates of CLW over the ocean from either the observation, $\mathrm{CLW_{obs}}$, or the first guess, $\mathrm{CLW_{fg}}$, are calculated using the retrieval Eq. (1). Using the so-called symmetric method proposed in Geer and Bauer (2011b) and Geer et al. (2012), the standard deviation of OmF is computed as a function of the average of $\mathrm{CLW_{fg}}$ and $\mathrm{CLW_{obs}}$ (i.e., $\overline{\mathrm{CLW}}$). This method prescribes the observation error e^o as a function of $\overline{\mathrm{CLW}}$ in fitting to the OmF standard deviation:

$$e^o = \begin{cases} e^o_{\mathrm{clr}} & (\overline{\mathrm{CLW}} \le C_{\mathrm{clr}}) \\ e^o_{\mathrm{clr}} + \lambda(\overline{\mathrm{CLW}} - C_{\mathrm{clr}}) & (C_{\mathrm{clr}} < \overline{\mathrm{CLW}} < C_{\mathrm{cld}}) \\ e^o_{\mathrm{cld}} & (\overline{\mathrm{CLW}} \ge C_{\mathrm{cld}}), \end{cases} \quad (7)$$

where

$$\lambda = \frac{e^o_{\mathrm{cld}} - e^o_{\mathrm{clr}}}{C_{\mathrm{cld}} - C_{\mathrm{clr}}}; \quad (8)$$

C_{clr} and C_{cld} are the two cloud threshold values for the piecewise linear fitting, which are channel dependent; and e^o_{cld} is the observation error for radiances associated with $\overline{\mathrm{CLW}} \ge C_{\mathrm{cld}}$. In the fitting process, because of the difficulty in estimating forecast model error, the component of the forecast model error is not removed from the OmF differences as it should be. Thus, the estimates of the observation errors are probably too large. Nevertheless, observation error is often inflated at many NWP centers to account for spatial and interchannel correlations.

An example is given in Fig. 1 for AMSU-A *NOAA-19* channel 2, with the OmF standard deviation (black dots) for the period from 1 to 15 November 2013 and the assigned symmetric observation error (red dots) with

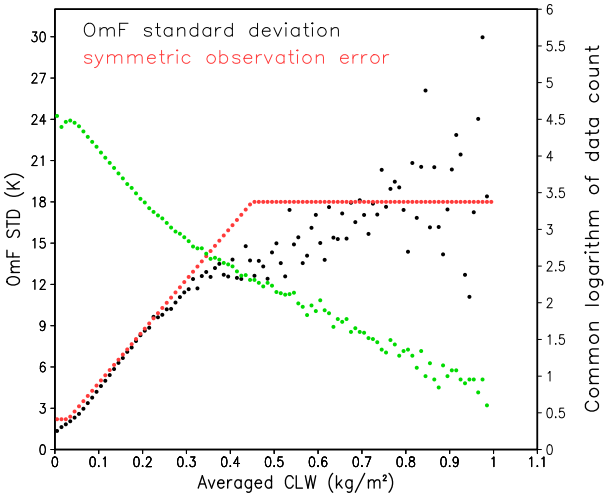

FIG. 1. The OmF standard deviation (STD; K, black dots) for the period from 1 to 15 Nov 2013 and the assigned symmetric observation error (K, red dots) of AMSU-A *NOAA-19* channel 2 with respect to $\overline{\mathrm{CLW}}$ ($\mathrm{kg\,m^{-2}}$). The common logarithm of the data count in each $\overline{\mathrm{CLW}}$ bin is shown with green dots. The $\overline{\mathrm{CLW}}$ bin width used for the plot is $0.01\,\mathrm{kg\,m^{-2}}$.

TABLE 1. A list of the parameter values used in the all-sky radiance observation error assignment for AMSU-A channels 1–4 and 15 in the GFS.

Channel	C_{clr}(kg m^{-2})	C_{cld}(kg m^{-2})	e^o_{clr}(K)	e^o_{cld}(K)
1	0.05	0.60	2.50	20.00
2	0.03	0.45	2.20	18.00
3	0.03	0.40	2.00	12.00
4	0.02	0.45	0.55	3.00
15	0.03	0.20	3.50	15.00

respect to $\overline{\text{CLW}}$. The common logarithm of data count in each $\overline{\text{CLW}}$ bin is also shown with green dots. The assigned symmetric errors follow the slope of the OmF standard deviations within $C_{clr} < \overline{\text{CLW}} < C_{cld}$, but with inflated constant errors for $\overline{\text{CLW}} \le C_{clr}$ and $\overline{\text{CLW}} \ge C_{cld}$. The values of C_{clr}, C_{cld}, e^o_{clr}, and e^o_{cld} for the observation error model [Eq. (7)] are summarized in Table 1 for channels 1–4 and 15 of AMSU-A. In this initial implementation of the all-sky approach, the cross-scan variation in observation error considered in Geer et al. (2012) is not yet included.

b. *Situation-dependent observation error inflation for all-sky radiances*

A similar strategy to observation error inflation as presented in Eq. (6) is also used in the all-sky approach, without the ΔCLW bias correction term. That is, $(e^{o*})^2 = (e^o_{adj})^2 + (\Delta e_{inf})^2$, but with $(e^o_{adj})^2$ being applied to the symmetric observation error e^o instead of e^o_{clr}. The all-sky approach experiments using the observation error e^{o*}, however, did not produce satisfactory results, and excessive analysis increments were generated. Another observation error inflation strategy was developed and is now presented.

In the all-sky approach, many useful cloudy radiance observations are associated with large OmFs. As such, the gross error check has been relaxed significantly, only bad observations or outliers with OmFs larger than 3 times the final observation error are excluded (section 3c). However, unlike ECMWF's system (Andersson and Jarvinen 1998; Bauer et al. 2010), variational quality control (VQC) is not applied to radiances in the GSI, though VQC has been applied to conventional observations. In this initial implementation of the all-sky approach, following the legacy route of situation-dependent observation error inflation in the clear-sky approach, additional inflation is constructed empirically using the physically based factors on which it is assumed the observation error (through primarily the CRTM) is dependent. That way the radiances with large OmFs can still be used in the analysis with reduced weights while not shocking the system. Testing of a new VQC algorithm (Purser 2011) is under way for conventional observations in the GSI. This algorithm will be tested on radiances in the future.

The physically based factors considered in this work include the cloud placement difference between the first guess and observation, the cloud liquid water difference between the first guess and observation, a scattering index equal to or larger than 9 (Grody et al. 1999), and the surface wind speed w_s. Differences in cloud placement are identified based on the following expressions: $(\text{CLW}_{obs} - C_{clr}) \times (CLW_{fg} - C_{clr}) < 0$ and $|\text{CLW}_{obs} - \text{CLW}_{fg}| \ge 0.005$. Cloud placement differences are commonly located at the edges of existing clouds, or in some convective locations as convective clouds are absent in the first guess. It will be seen in section 4a that we have difficulties in retaining the cloud water analysis increments in these locations. Considering the CRTM bias and forecast model errors, especially the challenge in placing clouds at the right time and location and with the right phase, we have used these observations conservatively by empirically introducing an observation error inflation as $\Delta e_1 = (1.0 - \delta)|\Delta T_b|$, where $\delta = 1.0$ if there is no cloud placement difference between the first guess and the observation; otherwise, $\delta = 0.0$.

The scattering index s over the ocean is calculated from observed brightness temperatures for channels 1, 2, and 15 using a retrieval formula from Grody et al. (1999):

$$s = -113.2 + [2.41 - 0.0049T_b(1)]T_b(1) + 0.454T_b(2) - T_b(15). \quad (9)$$

The scatterplots of OmF versus the difference between CLW_{fg} and CLW_{obs}, and OmF versus the scattering index for all ocean surface FOVs, are displayed in the top and middle panels in Fig. 2 for AMSU-A *NOAA-19* channel 1 for a 15-day period from 1 to 15 November 2013. The bottom panel in Fig. 2 shows OmF versus surface wind speed for only clear-sky FOVs. It is shown that when the scattering index is greater than 9 K the magnitudes of OmFs tend to increase with the scattering index. Although almost all radiances with scattering indices greater than 15 K are excluded by the screening of precipitating clouds (figure not shown), observation error inflation for scattering indices greater than 9 K is applied. For the top panel in Fig. 2, the standard deviation of OmF does not vary much with respect to cloud liquid water difference, but the correspondence between large OmFs and cloud liquid water difference is apparent for channel 1. Channel 2 has smaller OmF standard deviation than channel 1, but the OmF bias could be as large as ±30 K linearly varying with the cloud liquid water difference (figure not shown). It is believed that much of the OmF pattern with respect to the cloud

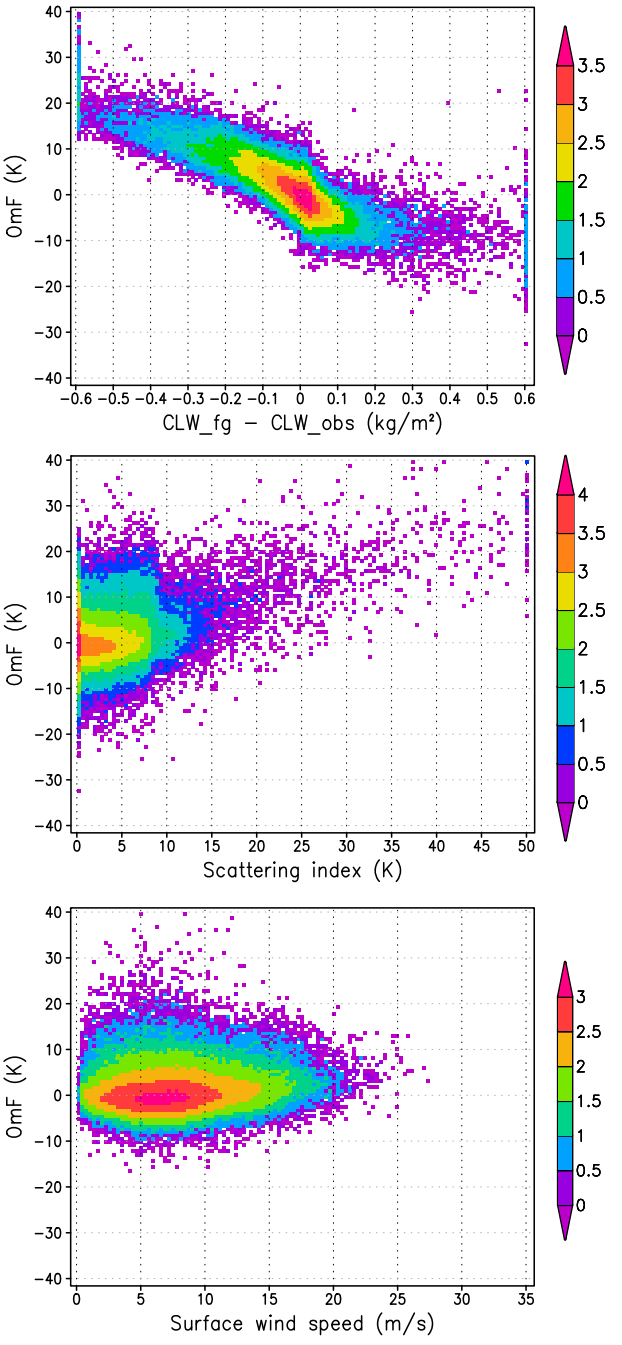

FIG. 2. Scatterplots of OmF (K) vs (top) the difference between CLW_{fg} and CLW_{obs} (kg m^{-2}) and (middle) scattering index (K);. (bottom) OmF (K) of clear-sky locations vs corresponding surface wind speed (m s^{-1}) for AMSU-A *NOAA-19* channel 1 for a 15-day period from 1 to 15 Nov 2013; the color bar shows the common logarithm of data count.

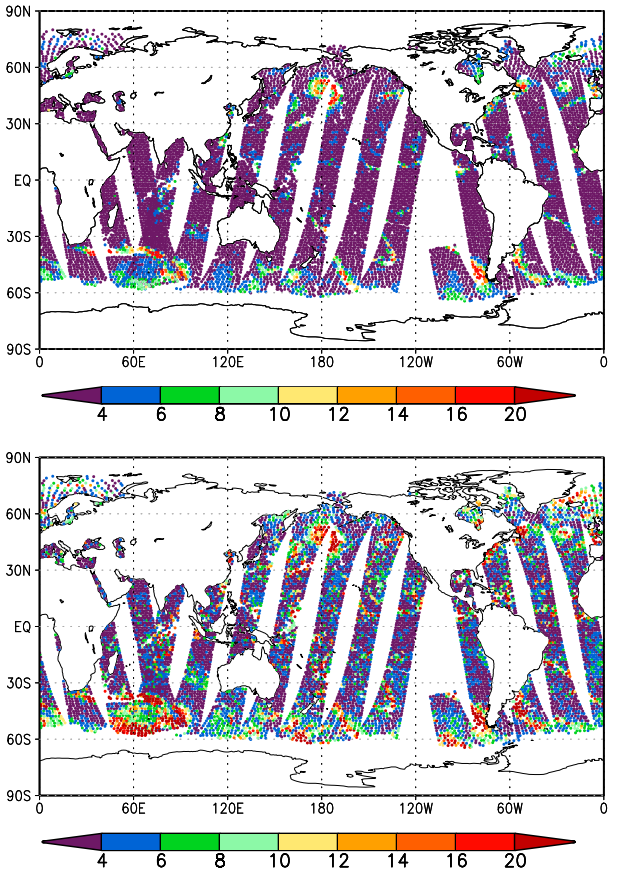

FIG. 3. (top) Symmetric and (bottom) final observation errors (K) during two analysis cycles for AMSU-A *NOAA-19* channel 1.

liquid water difference is the real observation signal we need for correcting the first guess; therefore, it is not desirable to bias correct these differences. In general, observation error inflation should only account for the components of the observation error (mainly the error of CRTM) in the OmF and the residual bias. The challenge we are facing here is to distinguish observation error from forecast error and the real observation signal we need. For simplicity, the observation error inflations due to the cloud liquid water difference and large scattering index are empirically designed as $\Delta e_2 = \gamma_1 |\mathrm{CLW_{obs}} - \mathrm{CLW_{fg}}| e^{o*}$ and $\Delta e_3 = \gamma_2 \max(0.0, s - 9.0) e^{o*}$, respectively, with two tunable parameters: γ_1 and γ_2. To a more limited extent, the OmF bias is affected by surface wind speed, and $\Delta e_4 = \gamma_3 w_s^2 e^{o*}$ is added as the last inflation term and with the smallest impact. It is aimed mainly to further reduce the weights of the observations with large OmFs over the northern Pacific Ocean and northern Atlantic Ocean, where large surface wind speeds are observed. The empirical parameters are set to be $\gamma_1 = 13.0\,\mathrm{m^2\,kg^{-1}}$, $\gamma_2 = 1.5\,\mathrm{K^{-1}}$, and $\gamma_3 = 0.002\,\mathrm{s^2\,m^{-2}}$ in this study based on the final analysis quality. Moreover, the upper limits of the error inflations due to cloud liquid water difference, larger scattering index s, and surface wind speed are set to be $3.5e^{o*}$, $2.5e^{o*}$, and $0.5e^{o*}$, respectively.

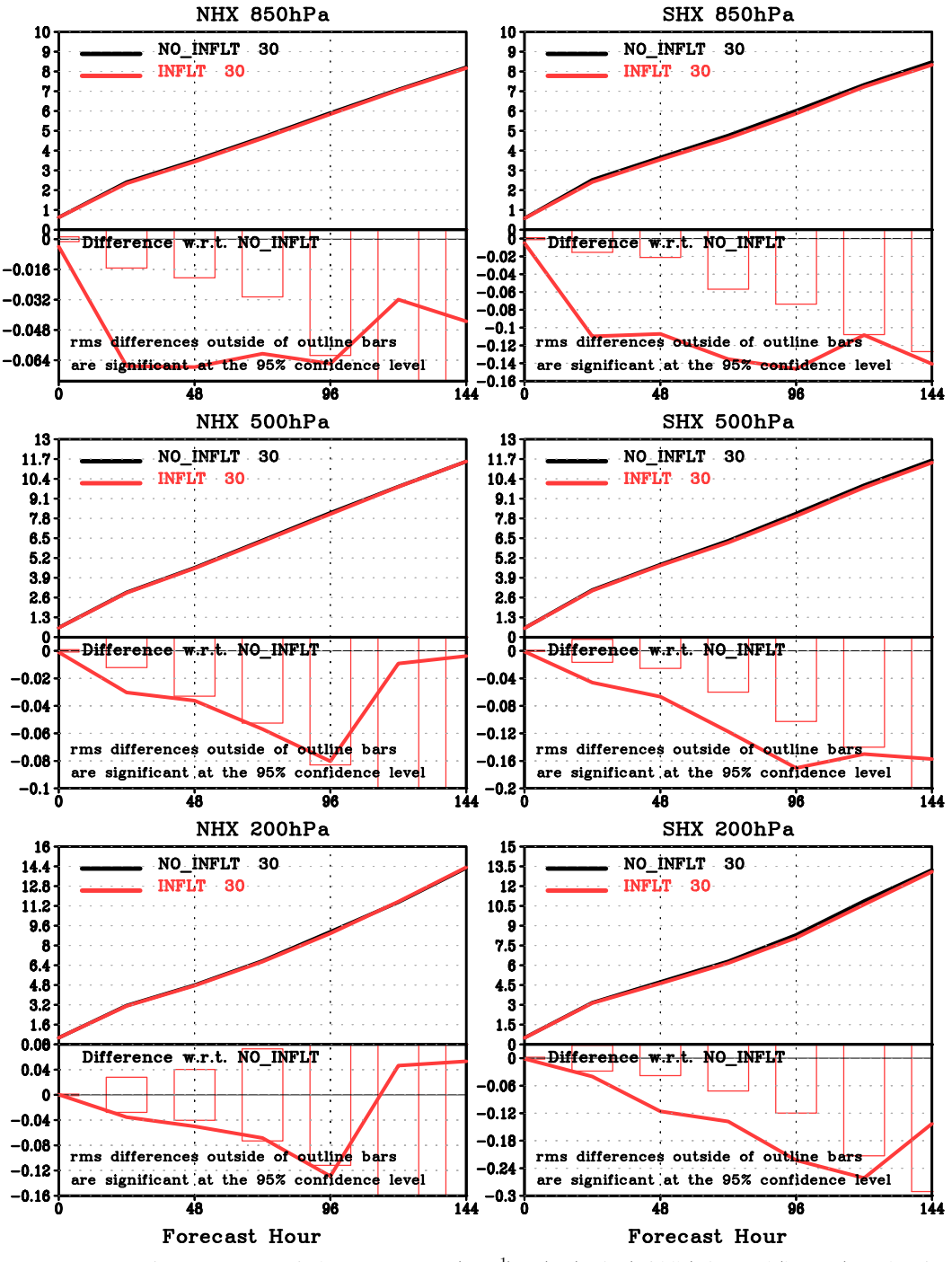

FIG. 4. Comparison of forecast wind vector RMSEs ($m\,s^{-1}$) at (top) 850, (middle) 500, and (bottom) 200 hPa in the (left) Northern and (right) Southern Hemispheres for a pair of all-sky experiments [(NO_INFLT (black) and INFLT (red)] used to examine the impact of the additional observation error inflation Δe. The observation error used in NO_INFLT does not include Δe, but the observation error used in INFLT does.

Although the inflation term Δe_2 has been examined extensively with different parameters, as the development efforts for the forecast model, GSI, and CRTM continue, it is necessary to revisit and assess the sensitivity of the assimilation results to these inflation terms (in particular the two latter added terms Δe_3 and Δe_4) in the future. A similar form as in Eq. (6) is employed to calculate the final observation error e^o_{new}:

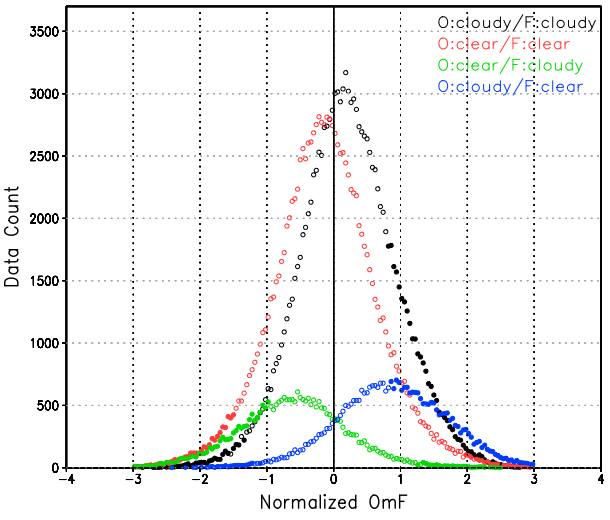

FIG. 5. Histogram of normalized OmF (normalized by symmetric observation error) for AMSU-A *NOAA-19* channel 1 for the period from 1 to 15 Nov 2013. The radiance data are grouped into four cloud categories: black circles for O:cloudy/F:cloudy, red circles for O:clear/F:clear, green circles for O:clear/F:cloudy, and blue circles for O:cloudy/F:clear. The closed circles represent the bins where the averaged ratio of the symmetric observation error to the final observation error is ≤0.5.

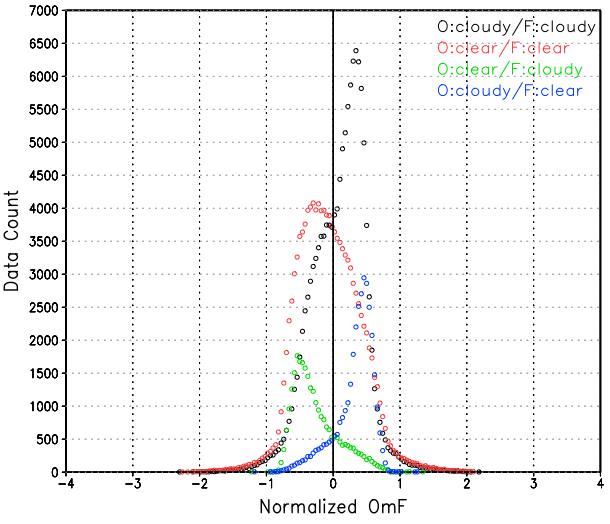

FIG. 6. As in Fig. 5, but for OmF (normalized by final observation error).

$$(e^{o}_{\text{new}})^2 = (e^{o*})^2 + (\Delta e)^2, \tag{10}$$

where $\Delta e = \Delta e_1 + \Delta e_2 + \Delta e_3 + \Delta e_4$. The effect of total observation error inflation is clearly reflected in Fig. 3, with the symmetric observation error in the top panel and the final observation error in the bottom panel at two analysis cycles for AMSU-A *NOAA-19* channel 1.

The impact of Δe is assessed using month-long samples from a pair of all-sky experiments. The experiment using the final observation error indicates significant improvement over the experiment using e^{o*} (without the additional observation error inflation for all-sky conditions), in the forecast anomaly correlation and RMSEs for temperature and wind, and the analysis and first-guess fits to rawinsonde temperature and specific humidity observations, with the bias of fits to rawinsonde specific humidity being reduced by half at 850 hPa in the tropics (figure not shown). Figure 4 shows the forecast wind vector RMSE results at 850, 500, and 200 hPa in the Northern (left panels) and Southern (right panels) Hemispheres. Significant reductions in the wind vector RMSEs are observed throughout the vertical levels due to the use of Δe.

The impact of observation error inflation for AMSU-A data is also examined for a 15-day period from 1 to 15 November 2013, with the histogram of normalized OmF (normalized by the symmetric observation error) for AMSU-A *NOAA-19* channel 1 (Fig. 5). The radiance data in Fig. 5 are grouped into four cloud categories based on the cloud information from the first guess and observation: where both the observation and first guess are considered cloud free (O:clear/F:clear, red circles); both are cloudy (O:cloudy/F:cloudy, black circles); the observation is cloud-free and the first guess is cloudy (O:clear/F:cloudy, green circles), or the observation is cloudy and the first guess is cloud free (O:cloudy/F:clear, blue circles). If $\text{CLW}_{\text{obs}} \le C_{\text{clr}}$, then the observation is considered to be cloud free. Similarly, if $\text{CLW}_{\text{fg}} \le C_{\text{clr}}$, then the first guess is considered to be cloud free. The OmF interval bins are marked by closed circles, where the mean observation weights are reduced by half or more by the total situation-dependent observation error inflation. It is shown that many more positive OmF bins in the O:cloudy/F:cloudy category are affected by the observation error inflations. Moreover, for all four categories the observations at the two ends of the histogram tails are assigned smaller weights, which resembles the weight reduction effect of VQC. The resulting OmF histograms normalized by the final observation error, however, are more non-Gaussian (Fig. 6), especially for O:cloudy/F:cloudy and the two categories with different cloud placements. For both clear-sky and all-sky radiances, the non-Gaussian error structures are currently under investigation. Since convective clouds are missing in the first guess, more radiances tend to be coincident with positive OmF biases than they should be. So an asymmetric gross error check in the GSI probably may help to improve the Gaussianity of O:cloudy/F:cloudy (and O:clear/F:clear), though it may not be desirable to discard some useful radiances. Also, different empirical parameter values of γ_1 may be necessary for positive and

TABLE 2. Summary of the radiance observation profile numbers of AMSU-A (including *NOAA-15*, *NOAA-18*, *NOAA-19*, *MetOp-A*, and *MetOp-B*) that are kept after the thinning process, over water, land, ice/snow, and mixed surface types as well as the total kept radiance numbers, for both the clear-sky (ClrSky) and all-sky (AllSky) approaches at 0000 UTC 10 Nov 2013.

Expt	Total	Over water	Over land	Over ice–snow	Over mixed
ClrSky	45 505	27 129	6117	8845	3414
AllSky	45 505	27 203	6094	8823	3385

negative cloud water differences. However, a fundamental question is whether these radiances have intrinsically a pure Gaussian distribution, and to what extent a better defined observation error can improve Gaussianity. It is very likely that many of the radiances with cloud placement differences between the first guess and the observation, and some of the radiances with large cloud liquid water differences would be affected considerably by VQC, if VQC would have been applied. VQC accounts for the non-Gaussian nature of gross errors in its formulation, and the effective observation error in VQC could be considered to be the observation error divided by the weight reduction factor. In our future investigations, we hope to gain a better understanding in regard to observation error assignment and to assess the performance of the above-mentioned new VQC algorithm on radiances.

c. *Quality control*

For the all-sky approach, since only FOVs with nonprecipitating clouds are used in this study, the thick-cloud filtering that is applied in the clear-sky approach is removed to allow more cloud-affected radiance data to be used, but the screening of precipitating clouds remains.

A surface emissivity sensitivity check is also added for channels 1–3 and 15. The ratio between the OmF and the first-order variation of brightness temperature with respect to surface emissivity is applied as a criterion for quality control. If the ratio is larger than an empirical threshold, the datum will be excluded. The threshold is channel dependent, and this check is relaxed when CLW_{obs} is larger than $0.1\,kg\,m^{-2}$. Moreover, following Geer and Bauer (2011a) and Geer et al. (2012), the cloud effect calculated from channel 5, which is the difference between the brightness temperature calculated with the cloud profile and the clear-sky brightness temperature, is employed to exclude radiance data affected by cloud ice in deep-convection regions.

The gross check is the final test for quality control. A relaxed gross check is applied to AMSU-A channels 1–5 and 15; that is, a radiance observation with an OmF magnitude larger than 3 times the final observation error is excluded from the data assimilation system. The detailed radiance data usage of AMSU-A (including *NOAA-15*, *NOAA-18*, *NOAA-19*, *MetOp-A*, and *MetOp-B*) for a typical analysis cycle at 0000 UTC 10 November 2013 is provided in Tables 2 and 3 for both the clear-sky and all-sky approaches, in terms of data over water, land, ice–snow, and mixed surface types, as well as the total. With the same total of 399 879 profiles available for the two approaches and the thinning grid box of 145 km × 145 km for AMSU-A, the profile numbers of radiance observations that are kept after the thinning process are summarized in Table 2, and the used observation numbers of channels 1–5 and 15 are presented in Table 3. AMSU-A

TABLE 3. Summary of used radiance observation numbers for channels (CH) 1–5 and 15 of AMSU-A (including *NOAA-15*, *NOAA-18*, *NOAA-19*, *MetOp-A*, and *MetOp-B*) over water, land, ice–snow, and mixed surface types, as well as the total used radiance numbers, for both the clear-sky (ClrSky) and all-sky (AllSky) approaches at 0000 UTC 10 Nov 2013.

Expt		Total	Over water	Over land	Over ice–snow	Over mixed
ClrSky	CH 1	25 986	22 921	1035	1554	476
	CH 2	26 208	23 109	1042	1579	478
	CH 3	26 205	23 109	1042	1578	476
	CH 4	26 208	23 109	1042	1579	478
	CH 5	26 208	23 109	1042	1579	478
	CH 15	25 588	22 489	1042	1579	478
AllSky	CH 1	28 760	25 350	1113	1808	489
	CH 2	28 857	25 417	1123	1823	494
	CH 3	28 957	25 517	1123	1823	494
	CH 4	28 977	25 537	1123	1823	494
	CH 5	28 981	25 541	1123	1823	494
	CH 15	28 816	25 376	1123	1823	494

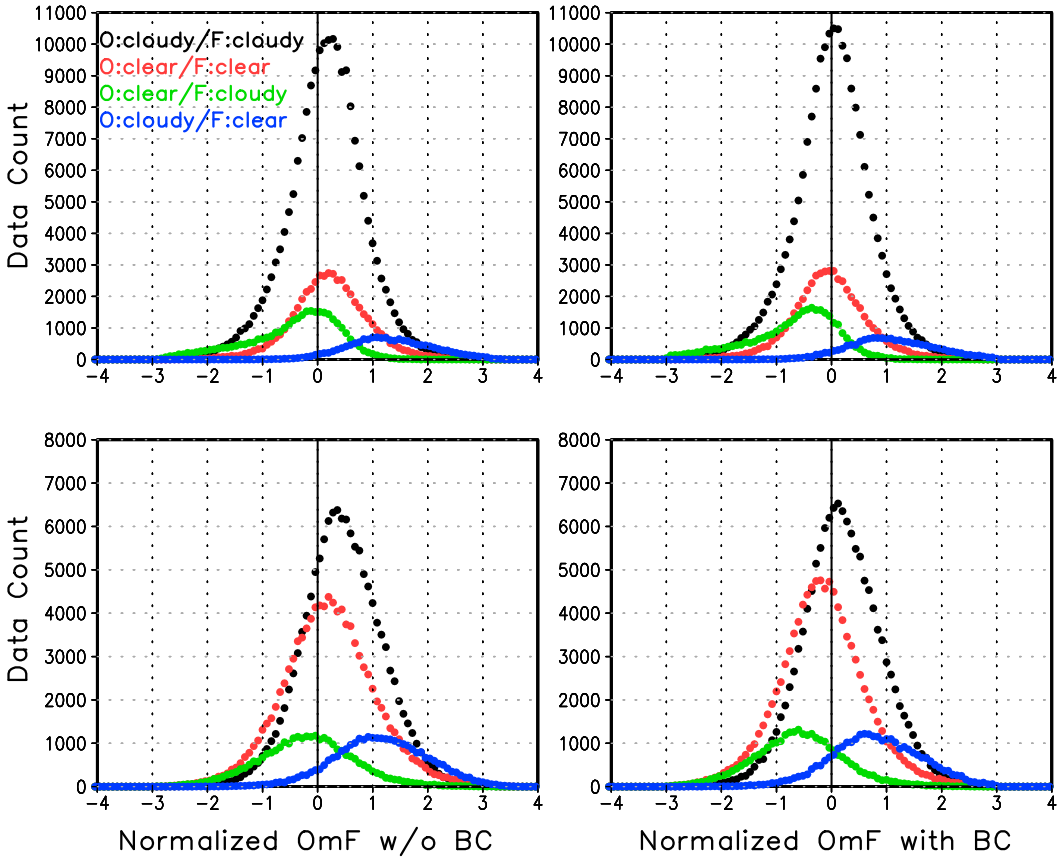

FIG. 7. Histograms of OmF (right) with and (left) without bias correction (BC) for AMSU-A *MetOp-A* channels (top) 15 and (bottom) 1 for the period from 1 to 15 Nov 2013.

AQUA radiances are not included in the tables, as AMSU-A's channels 1–5 and 15 are not used in the GSI. Overall, about 10% more radiance data for AMSU-A channels 1–5 and 12% more for channel 15 are assimilated with the all-sky approach. Our data usage results are different from those of Geer et al. (2012), where only a few additional scenes could be added in channel 5. It is not very surprising that we assimilate more radiances from window and near-window channels as clouds have more impact on these channels, but we also feel this disparity could be attributed to the differences among the quality control procedures; that is, the quality control in Geer et al. (2012) may be much more stringent than that of our GSI all-sky system. The parameter tuning for the variational quality control at ECMWF and the situation-dependent observation error inflation in the GSI would also play a role.

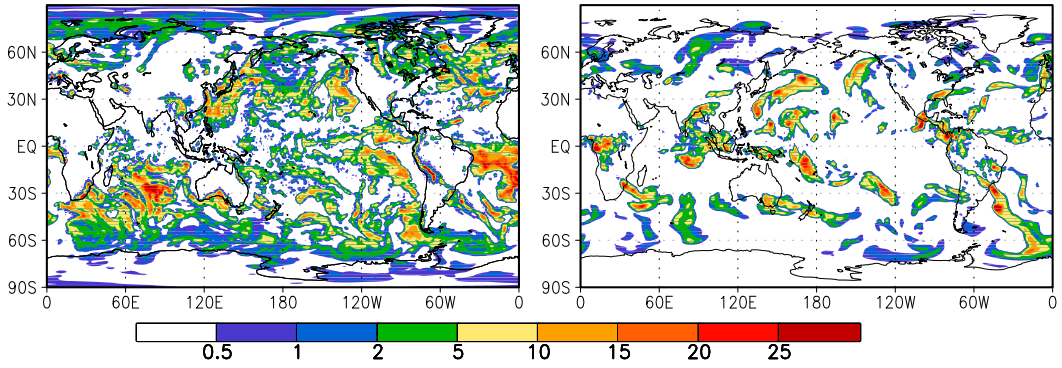

FIG. 8. Cloud water ensemble spread ($\times 10^5$ kg kg^{-1}) at levels (left) 13 (~850 hPa) and (right) 30 (~300 hPa) at 0000 UTC 22 Oct 2013.

d. All-sky radiance bias correction

As mentioned earlier, in the all-sky approach the radiance data can be grouped into four categories: O:clear/F:clear, O:cloud/F:cloud, O:cloud/F:clear, and O:clear/F:cloud. Intuitively, we would expect the two categories, where there are inconsistencies in the placement of clouds between the first guess and observations, to be associated with the generation–elimination of clouds. The two categories with mismatched cloud information contain both bias and real cloud signal information in these radiances and may affect the performance of an all-inclusive radiance bias correction, resulting in an erroneously large bias-corrected OmF. Hence, in addition to the current radiance bias correction scheme (Zhu et al. 2014a), a new strategy was proposed in Zhu et al. (2014b) for the all-sky approach. The intent of this strategy is to remove the biases from the radiance data, while preserving the useful cloud information from the OmFs corresponding to mismatched cloud information. That is, in an ideal situation, after bias correction using a proper sample of the AMSU-A radiance data, we should expect to obtain unbiased OmFs of the radiance data where the placement of clouds retrieved from the observations and those from the first guess are consistent (O:clear/F:clear and O:cloud/F:cloud), while the O:cloud/F:clear category should exhibit a distinct positive bias and the O:clear/F:cloud category a distinct negative bias. In this strategy, all quality-controlled radiance data are used to produce the analysis, but the bias correction coefficients are derived using only a selected data sample where clouds are present–absent in both the first guess and observations, and the radiance data with mismatched cloud information are bias corrected using the latest bias coefficients available.

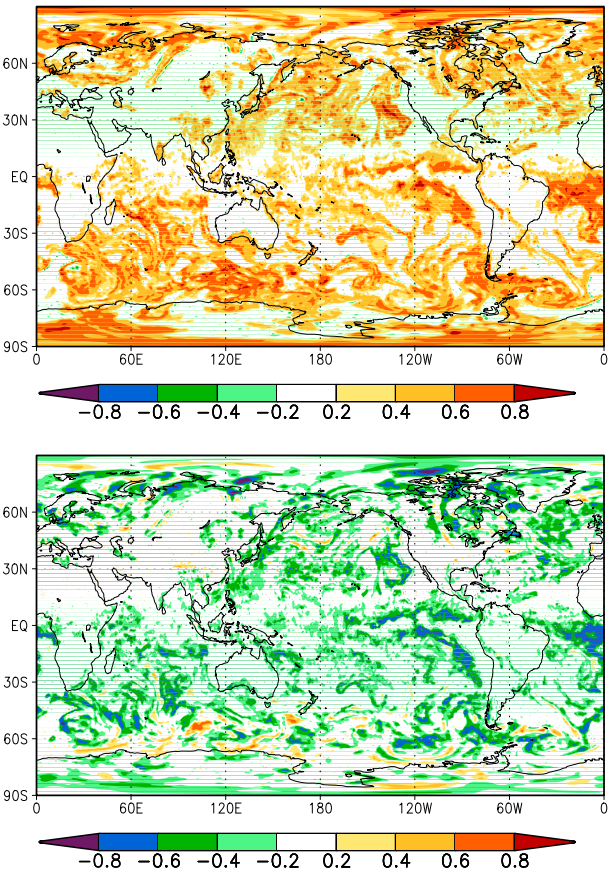

FIG. 9. Point-wise correlation coefficients between (top) cloud water and relative humidity background error and (bottom) cloud water and temperature background error at level 13 (~850 hPa) at 0000 UTC 22 Oct 2013.

Thus, the observation operator $\tilde{h}$ of the AMSU-A radiance data can be written as

$$\tilde{h}(\mathbf{x}, \Delta\boldsymbol{\beta}) = \begin{cases} h(\mathbf{x}) + \sum_{k=1}^{N} \beta_{b,k} p_k(\mathbf{x}) & \text{(if with mismatched cloud, over ocean)}, \\ h(\mathbf{x}) + \sum_{k=1}^{N} \beta_{b,k} p_k(\mathbf{x}) + \sum_{k=1}^{N} \Delta\beta_k p_k(\mathbf{x}) & \text{(otherwise)}, \end{cases} \tag{11}$$

where $\mathbf{x}$ is the model state or GSI control vector and $h(\mathbf{x})$ represents the radiative transfer model. Letting $\beta_{b,k}$ denote the latest available estimate of the predictor coefficient at each outer loop and $\Delta\beta_k$ the coefficient increment, the total bias is written as a linear combination of a set of predictors $p_k(\mathbf{x})$, $k = 1, 2, \ldots, N$, and $p_1 = 1$. In addition, $\Delta\boldsymbol{\beta}$ is also a GSI control vector and is updated along with $\mathbf{x}$ in the GSI minimization procedure of the inner loop. With the removal of the ΔCLW bias term in the all-sky approach, the current bias predictors for AMSU-A in the radiance bias correction scheme include global offset, the lapse rate convolved with the channel's weighting function, the square of the lapse rate convolved with the channel's weighting function, emissivity sensitivity, and zenith angle bias correction terms (Zhu et al. 2014a). We should be aware that, without introducing a new cloud-related bias predictor into the radiance bias correction scheme, the two categories O:clear/F:clear and O:cloud/F:cloud are assumed implicitly to have the same bias. The histograms of OmF

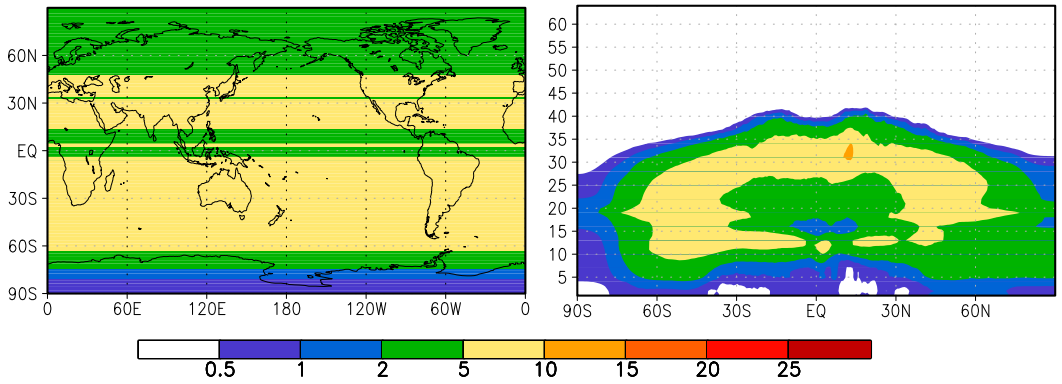

FIG. 10. Static background error standard deviation of cloud water ($\times 10^5$ kg kg^{-1}) at (left) model level 13 (~850 hPa) and (right) cross section that was used in the original operational clear-sky GFS system.

are shown in Fig. 7 for AMSU-A *MetOp-A* channels 1 (bottom panels) and 15 (top panels), respectively, for a 15-day period from 1 to 15 November 2013. The histograms of OmF before bias correction are shown in the left panels and the histograms after bias correction are presented in the right panels. The bias correction strategy works well for channel 15. After bias correction, the two cloud-consistent categories (O:clear/F:clear and O:cloud/F:cloud) are centered around a zero bias, and the other two cloud-mismatched categories are on opposite sides of the zero-bias line. However, it is observed that the bias correction is less optimal for channel 1, as the bias correction settles down somewhere between the two cloud-consistent categories. This can be explained by the histograms of OmF before bias correction for O:clear/F:clear and O:cloud/F:cloud in the left panels in Fig. 7. While the bias of O:clear/F:clear agrees well with that of O:cloud/F:cloud for channel 15, the difference between the bias of O:clear/F:clear and that of O:cloud/F:cloud is more evident for channel 1. When O:clear/F:clear and O:cloud/F:cloud have different biases, an additional cloud-related bias correction term such as $\overline{\text{CLW}}$ or an averaged scattering index may further improve the performance of the bias correction.

e. *Background error covariance*

With the use of a cloud control variable, it is important to have a properly constructed background error covariance (particularly the cloud water variance and the cross covariances among cloud water, moisture, and temperature, as well as wind for this study) in order to assimilate all-sky radiances appropriately. In the GFS 3D EnVar system, the background error covariance includes ensemble and static contributions, with an 87.5% weight given to the ensemble part and 12.5% to the static term. The ensemble is designed to represent the probability distribution of the system uncertainty, including that of the analysis and background. This is achieved through a combination of ensemble data assimilation using the ensemble serial square root filter (Whitaker and Hamill 2002) and multiplicative inflation by relaxation to the prior spread (Whitaker and Hamill 2012). To address model uncertainty, stochastic physics has been used to improve the ensemble sample's representativeness of the probability distributions. The uncertainty is simulated by the inclusion of three stochastic parameterization schemes (J. Whitaker 2015, personal communication): the stochastically perturbed parameterization tendency scheme (SPPT; Buizza et al. 1999), the stochastic kinetic energy backscatter scheme (SKEB; Berner et al. 2009; Shutts 2005), and the stochastically perturbed boundary layer humidity (SHUM; Tompkins and Berner 2008). The actual ensemble covariance is never explicitly computed but, rather, is extracted implicitly through the ensemble control variable formulation. The multivariate increment from the ensemble part is generated by taking the product of the control variable weights and the ensemble perturbations at those grid points. In contrast, the static term is generated by the National Meteorological Center (NMC, now known as NCEP) method (Parrish and Derber 1992). The NMC method estimates the forecast errors using the differences between 24- and 48-h forecasts but valid at the same time. Currently, there is no cross covariance between cloud water and other variables specified in the static term. In general, the ensemble provides flow-dependent background error covariance information, and the static term provides climatological information.

The ensemble cloud water spread (Fig. 8) at levels 13 (about 850 hPa, left panel) and 30 (about 300 hPa, right panel) reflect features of the forecast model; for example, the large magnitude of the ensemble spread is seen to be coincident with a large amount of cloud water in the first guess. However, the error covariance generated

from the ensemble also inherits the deficiencies of the forecast model (in particular, the model physical parameterization schemes) and the data assimilation system. For example, because convective clouds are not included in the cloud output from the GFS forecast model, the magnitude of the cloud water ensemble spread is small over most of the tropical regions where there are very few clouds in the first guess relative to the observations (figure not shown). The amplitudes of the ensemble cloud water spread are large, comparable to 80% of the cloud water first guess, with the spread encompassing a much larger geographic area than the clouds themselves. In terms of cross covariances, cloud water at level 13 is seen to have positive point-wise correlation with relative humidity (Fig. 9, top), and negative correlation with temperature on average (Fig. 9, bottom). The resultant analysis increment will be a function of these point-wise correlations in addition to the covariances with the nearby grid points (both horizontal and vertical) within the prescribed localization radius used in the EnVar solver. While we acknowledge that the full three- (and four-) dimensional multivariate covariances are key contributors to the analysis increments within both the clear and all-sky frameworks given the 87.5% weight to the ensemble, a full exploration is beyond the scope of this study. However, this is an area for future work in terms of continuing to improve the assimilation of all-sky radiances in addition to the more general enhancements to the prescription of multivariate correlations within the ensemble.

Regarding the static term, the cloud water background error variance is changed for the all-sky approach. Figure 10 shows the square root of cloud water background error variance at level 13 (about 850 hPa, left) and in a cross section (right) that are used in the original operational clear-sky GFS system. This static error variance is a constant for each latitude, with moderate variances observed in the tropics from levels 25 (about 500 hPa) to 35 (about 200 hPa) and in the extratropics between levels 10 and 30 (about 900–300 hPa; Fig. 10, right). A comparison of Fig. 10 with Fig. 8 indicates that the moderate static error variance may be dominant at many locations where the ensemble spread is small. For clouds, which are discontinuous, localized, and strongly constrained by temperature and moisture, applying such a static cloud water background error variance to the all-sky approach would result in the generation of spurious cloud increments at many locations that are inconsistent with the model physics and are not retained by the forecast model integration. Therefore, in the all-sky approach, to improve the Gaussianity of the errors and to reduce spurious cloud increments, normalized cloud water (normalized by its background error standard deviation) is used as the control variable, with the new static cloud water background

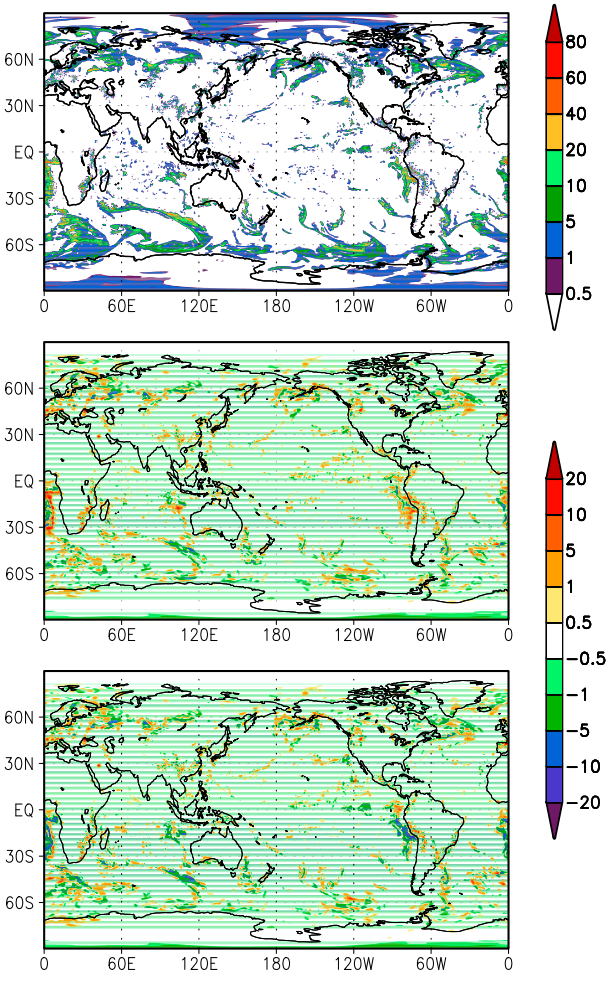

FIG. 11. Cloud water ($\times 10^5$ kg kg^{-1}) for (top) first guess; and analysis increments for the stand-alone single-analysis tests: (middle) for the clear-sky and (bottom) all-sky approaches for level 10 (~900 hPa) at 0000 UTC 3 Nov 2013.

error variance being assigned to be large only where clouds already exist. In this study, for simplicity, the static cloud water background error standard deviation is specified as 5% of the cloud water first guess from the deterministic forecast, and a small error value 5.0×10^{-12} kg kg^{-1} is assigned for cloud-free locations or locations with cloud water less than 1.0×10^{-10} kg kg^{-1}. The ensemble spread represents the uncertainty in the ensemble forecasts, and this new static term brings in flow information from the deterministic forecast. The optimal combination of the static cloud water error variance and the ensemble spread of cloud water will be explored in the future.

4. Impact of all-sky approach on analysis increments

Relative to the clear-sky approach, the impact of the all-sky approach on analysis increments is expected for

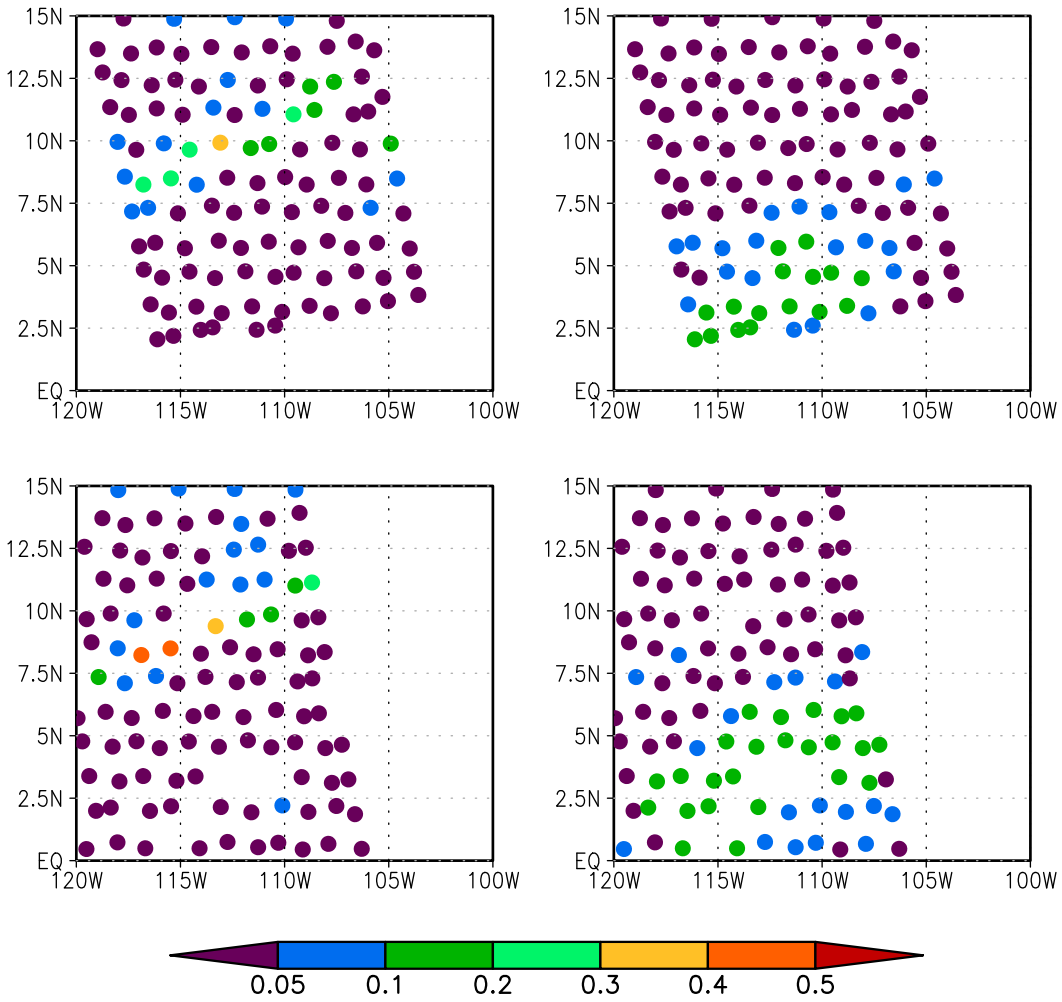

FIG. 12. An example of cloud liquid water (kg m^{-2}) of (left) observations from AMSU-A (top) *NOAA-19* and (bottom) *NOAA-18* at 0000 UTC 3 Nov 2013 and (right) the corresponding column-integrated cloud liquid water of the first guess.

the scenarios when the first guess or the radiance observation is cloudy. A pair of stand-alone single-analysis tests is examined for an arbitrary analysis cycle (0000 UTC 3 November 2013). The two tests use the same first guess from a cycled clear-sky GFS run. One test uses the clear-sky approach and the other the all-sky approach. The bias correction coefficients for each test are generated from cycled clear-sky and corresponding cycled all-sky runs, respectively, both of which are spun up from zero bias at 1800 UTC 21 October 2013. For the two tests, Fig. 11 presents the cloud water first guess (top panel) and the analysis increments (middle panel for the clear-sky and bottom panel for the all-sky approach) at level 10 (about 900 hPa). It is shown that, while the analysis increments of the two approaches are comparable for most regions, there are notable differences at a few locations (e.g., the continental western boundaries and the spot around 117°W near the equator). The GFS forecast is known to have too much cloud at the continental western boundaries probably because of issues from planetary boundary layer and moist parameterization schemes as well as feedback from the analysis. At these locations where the clear-sky and all-sky approaches differ, the all-sky approach tends to generate either much smaller positive cloud water analysis increments than the clear-sky approach or negative increments. Figure 12 displays the cloud liquid water at the location near the equator around 117°W, with the cloud liquid water of the observations shown in the left panels (top panel for AMSU-A *NOAA-19* and bottom panel for AMSU-A *NOAA-18*), and the corresponding column-integrated cloud liquid water of the first guess in the right panels. It is seen that for the area to the south of 8°N the observations (both AMSU-A *NOAA-18* and *NOAA-19*) are

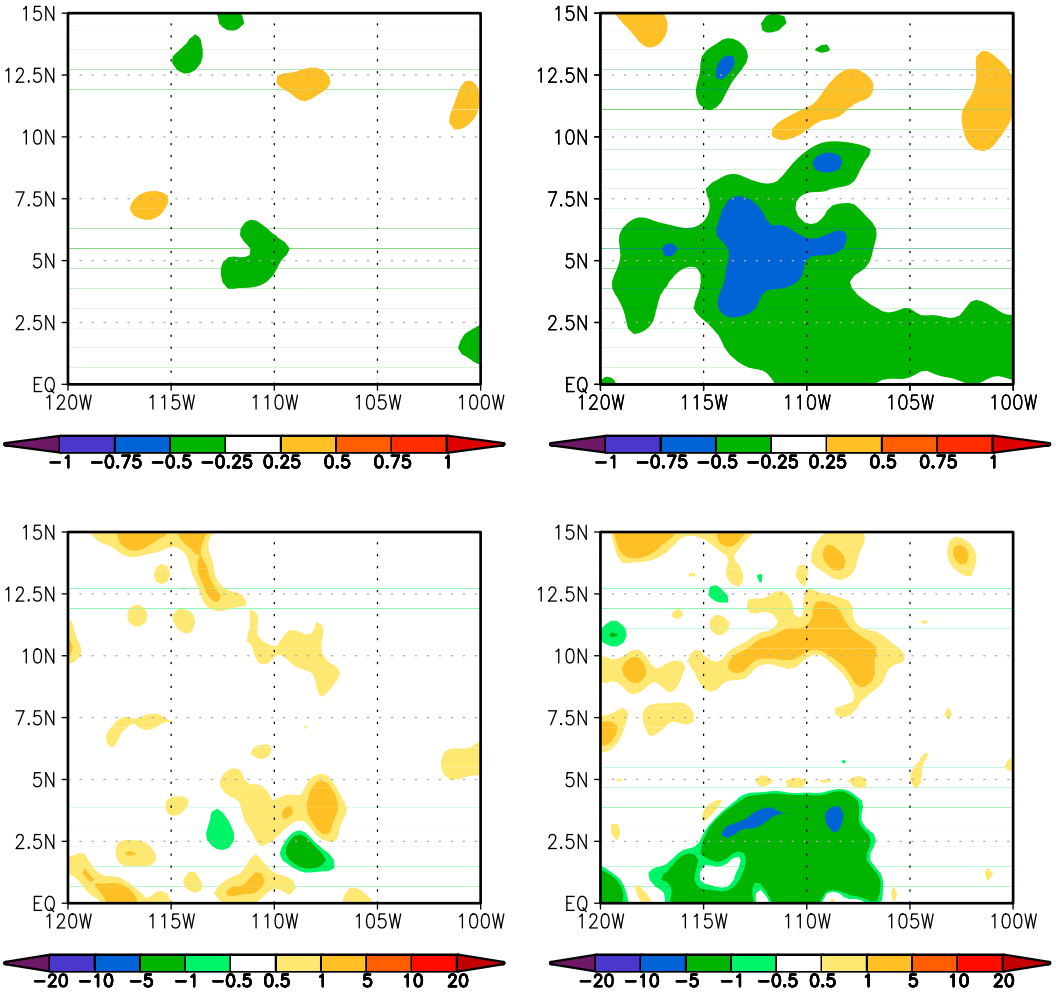

FIG. 13. Analysis increments at level 10 (~900 hPa), corresponding to Fig. 12, for the (left) clear-sky and (right) all-sky approaches at 0000 UTC 3 Nov 2013. Shown are the (top) specific humidity analysis increment ($\times 10^3$ kg kg^{-1}) and (bottom) cloud water analysis increment ($\times 10^5$ kg kg^{-1}).

found to be cloud free or with very small amounts of cloud while the first guess has clouds in this area. For this case, the cloud-free assumption in the calculation of the simulated brightness temperature in the clear-sky approach results in small OmF, but the cloudy first guess leads to large negative OmF in the all-sky approach. The resultant analysis increments at level 10, as shown in Fig. 13 (left panels for clear sky and right panels for all sky), exhibit larger reductions in specific humidity (top panels) and cloud water (lower panels) fields over a wider area in the all-sky approach than in the clear-sky approach, which is as expected.

Analysis increments from the cycled clear-sky and all-sky microwave radiance assimilation runs are also examined. The all-sky microwave radiance assimilation approach is tested within a 3D EnVar context using the same operational observation data from 21 October to 1 December 2013. Like the operational GFS forecast system, this all-sky experiment (hereafter referred to as AllSky) uses a dual-resolution configuration but with a lower resolution, which includes a 670-spectral triangular truncation (T670) horizontal resolution for the deterministic forecast model and a T254 resolution for the analysis and ensemble forecast (T670-T254) with 64 unequally spaced sigma layers (L64). The control is the corresponding clear-sky approach experiment (ClrSky). For the cycled ClrSky and AllSky runs, the patterns of the averaged analysis increments at the continental western coasts at levels 10–15 are more striking than the analysis increments for the stand-alone tests shown in Fig. 11. Fig. 14 presents the averaged cloud water first guess of the ClrSky experiment (top panel) and the averaged analysis increments for ClrSky (middle panel) and AllSky (bottom panel) at level 13 (about 850 hPa).

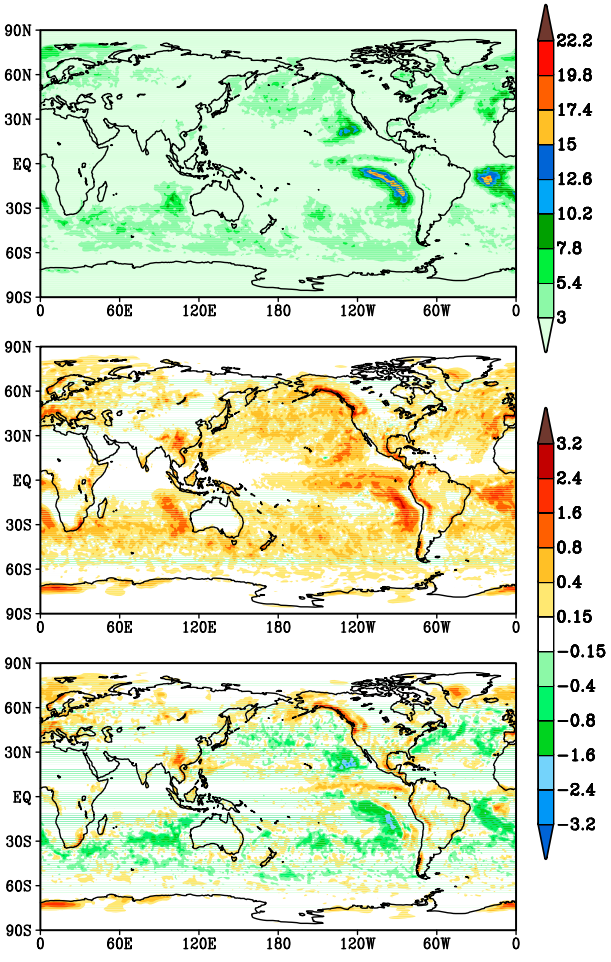

FIG. 14. (top) Averaged cloud water ($\times 10^5$ kg kg^{-1}) first guess for the ClrSky experiment; and analysis increments for the (middle) ClrSky and (bottom) AllSky experiments at level 13 (~850 hPa) during the period from 27 Oct to 1 Dec 2013.

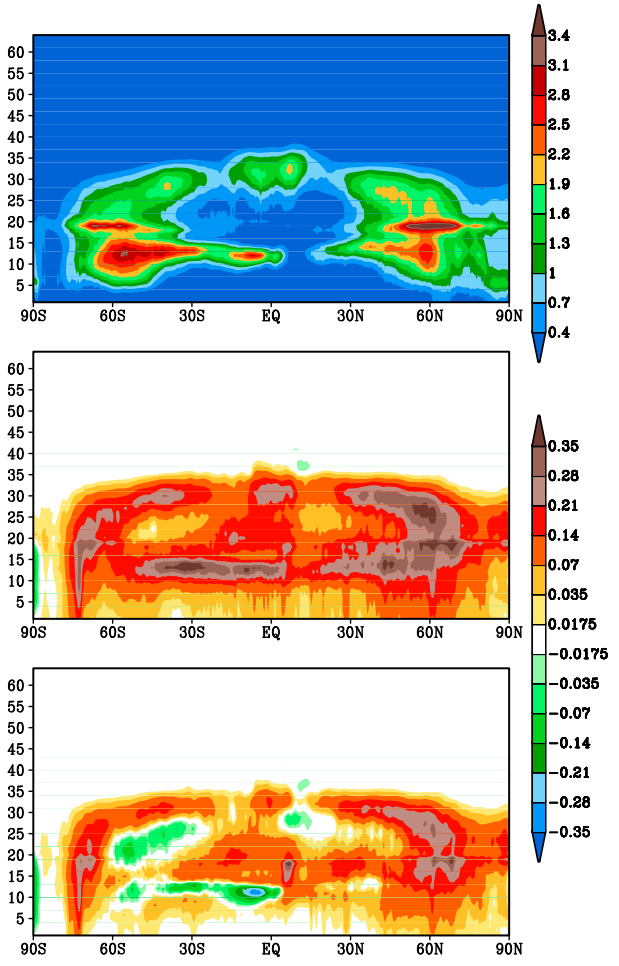

FIG. 15. Cross sections of (top) zonal mean cloud water ($\times 10^5$ kg kg^{-1}) first guess for the ClrSky experiment; and analysis increments for the (middle) ClrSky and (bottom) AllSky experiments during the period from 27 Oct to 1 Dec 2013.

For the clear-sky case, the averaged cloud water increments are uniformly positive, while in the all-sky case the mean increments show a mixture of positive and negative values. The cross sections of the zonal mean cloud water first guess and analysis increments are also presented in Fig. 15. In the clear-sky run (Fig. 15, middle), positive cloud water analysis increments extend from the surface to forecast model level 35 (about 200 hPa), and large positive increments are generated between model levels 10 (about 900 hPa) to 15 (about 800 hPa) and in the extratropics. In the all-sky run, the global mean cloud water analysis increments are smaller, and in some regions, the increments are negative. There are still some positive cloud water increments at the lowest vertical levels over ocean in the all-sky run, which may be questionable, and they may be attributed to the cloud water ensemble and ensemble spread at these levels as well as the cross covariances among cloud water, moisture, temperature, and wind.

The consistent positive mean increments seen in the clear-sky case and at some levels in the all-sky case arise primarily from a combination of the ensemble approach and the physical restriction that cloud amounts should be nonnegative. In the clear-sky case increments arise through cross covariances with the other control variables (as there is no direct observation operator for cloud) with positive and negative increments equally likely. In relatively clear regions this will result in small positive and negative cloud amounts in the analyses. However, during the minimization negative cloud amounts are set to zero at each outer loop, resulting in a net positive analysis departure. For the all-sky case the cloud increments are also affected directly through the observation operator and this effect is reduced, although it is still present at higher levels where the observations provide

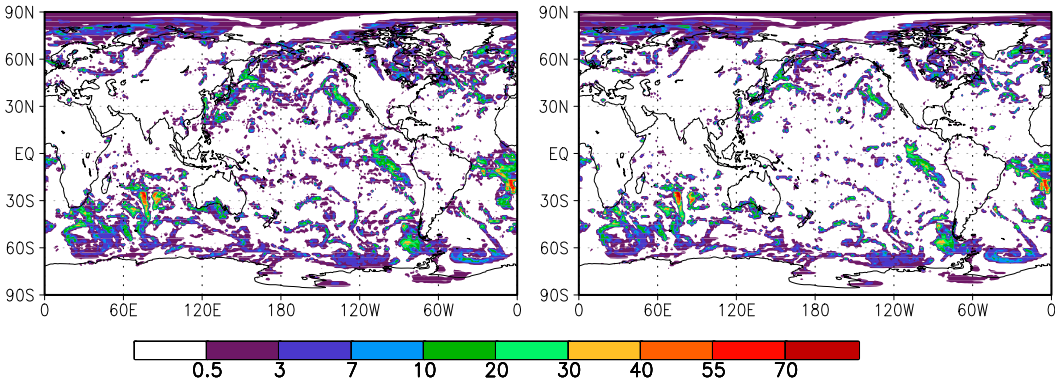

FIG. 16. (left) Cloud water analysis and (right) first model integration step of forecast ($\times 10^5$ kg kg^{-1}) at 850 hPa at 0000 UTC 22 Oct 2013.

fewer constraints on the solution. Different approaches to removing negative cloud and reducing these increments will be explored in future work.

Extraneous small cloud amounts

Over convective locations, since convective clouds are missing in the cloud water first guess, it may be challenging in the integration of the deterministic forecast model to retain the clouds that are generated in the GSI because of the presence of convective clouds information in the radiance observations. Development efforts continue including convective clouds in the cloud forecast output from the GFS model and in the GSI. The following discussion focuses on another context of cloud loss.

It is noticed that the cloud water ensemble spread is much broader horizontally than the static cloud water background error variance (figure not shown). Because of the ensemble part of the background error covariance (such as the broader cloud water ensemble spread, the spatial correlations contained within the ensemble members, and the cross covariances with other variables), the subsequent cloud water increments are realized in a larger area in the all-sky system, relative to the univariate cloud water analysis if only the static term of the background error covariance would have been used. Figure 16 depicts the cloud water analysis (left panel) and first model integration step of the forecast (right panel) at 850 hPa at 0000 UTC 22 October 2013 from the AllSky experiment. Figure 16 demonstrates that some small clouds are generated in the analysis around the edges of the existing clouds, but they may not be compatible with the physical parameterization schemes of the deterministic forecast model, and are dissipated during the integration of the deterministic forecast model (R. Sun 2014, personal communication). Further, the generation of these extraneous small clouds in the analysis does not seem to impact the convergence of the GSI system.

A close examination shows that this cloud loss also happens in the clear-sky approach (Fig. 17, top). Consistent with the above cloud water analysis increment results, the clear-sky approach is shown to generate more clouds in the analysis (green closed square) than in the all-sky approach (red closed square) and later forecasts (green and red line, respectively for the clear-sky and all-sky approaches). For both approaches, clouds are dissipated during the first model integration step and gradually spin up in later forecast hours. Therefore, cloud loss is not uniquely associated with the all-sky approach, and the biggest spindown from analysis at the first model integration step occurs in the tropics (Fig. 17, bottom), where convection may be part of the issue.

Although the above-mentioned broader cloud water ensemble spread may be partially due to the use of the stochastic physics to help represent model error, the application of the dual-resolution configuration is another cause. To verify this finding, a pair of single-cycle tests is conducted using single-resolution T574 and dual-resolution T574–T254 configurations at an arbitrary analysis time (0000 UTC 2 June 2015). Both configurations have 64 vertical levels and use 80 ensemble members. The first guesses for the analysis and ensemble spreads are examined. In the dual-resolution case, the low-resolution initial files for the ensemble are obtained by interpolating the high-resolution ensemble initial files, and the high-resolution model 6-h forecast is interpolated to the lower resolution and serves as the first guess for the analysis. The results show that, although the interpolated fields are slightly smoother at the lower resolution, the interpolation from higher resolution to lower resolution does not seem to alter the overall patterns and features of temperature, relative humidity, and cloud fields. However, after 6-h integration

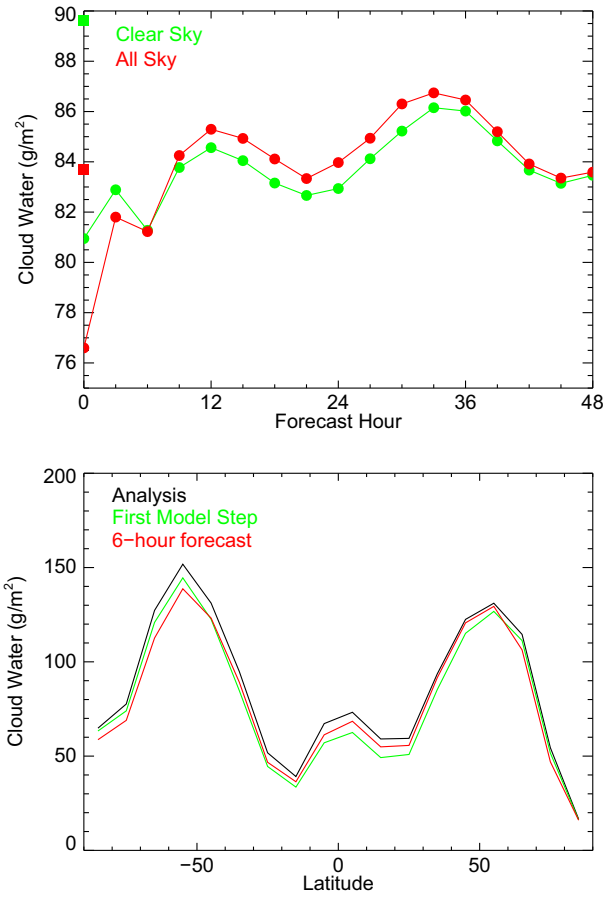

FIG. 17. (top) Evolutions of averaged global mean atmospheric column cloud water (g m^{-2}) during forecast issued from ClrSky (green line) and AllSky (red line) 0000 UTC analyses for the period from 27 Oct to 1 Dec 2013. The green and red closed squares at zero forecast hour represent the averaged global mean atmospheric column cloud water of the clear-sky and all-sky analyses, respectively. (bottom) The latitudinal dependence of the spindown between analysis (black line) and the first model integration step (green line). Also, shown in the bottom panel is the 6-h forecast (red line).

of the ensemble forecasts, Fig. 18 shows that the differences between the single-resolution (left panels) and dual-resolution (right panels) ensembles emerge. Although the ensemble spread results for temperature (Fig. 18, top) from the two configurations are comparable, in the single-resolution simulation the relative humidity ensemble spread (Fig. 18, middle panels) exhibits more locally intense values, and the higher relative humidity spread is slightly more dispersed. The most pronounced change lies in cloud water (Fig. 18, bottom panels); that is, its ensemble spread for the dual-resolution ensemble is far more spread out than in the single-resolution version, and this subsequently results in cloud water increments over a larger area in the analysis. The broader cloud water ensemble spread realized in the dual-resolution configuration may be attributable to the resolution-dependent processes in the forecast model, the stochastic physics algorithm itself, and the resolution-dependent parameter tuning in the stochastic physics. The stochastic parameters used had been specifically retuned for the low-resolution configuration to maintain the ensemble spread magnitudes of wind, temperature, and moisture realized at the higher forecast resolution. However, the three stochastic parameterization schemes (SPPT, SKEB, and SHUM) are not directly linked to cloud water variability. The extraneous small clouds that result from the broader ensemble cloud water spread compensate to some extent for the lack of clouds in the first guess during the analysis procedure, and do not seem to cause any particular issue in the forecast, as they are likely to be lost later in the high-resolution deterministic model forecast as a result of the inconsistency with the model physics.

In summary, the all-sky approach produces more realistic cloud water analysis increments; in particular, it removes some clouds off the continental western coasts. However, some outstanding issues from the clear-sky approach still remain for the all-sky approach, including the accuracy of the model cloud forecast, which is related to the resolution-dependent physical parameterization schemes and their impact on the dual-resolution configuration, the tuning of the stochastic physics for cloud variability, and the consistency of model uncertainties among temperature, moisture, and cloud variables in the stochastic physics. Moreover, since the localization employed in the EnKF may degrade the multivariable balance in the analysis (Lorenc 2003; Greybush et al. 2011), more investigations pertaining to the balance among temperature, moisture, and clouds are necessary in the future.

5. Impact of all-sky radiance assimilation on analyses and forecasts

The results discussed below are from the cycled T670–T254 ClrSky and AllSky experiments for the period from 21 October to 1 December 2013. The first 5 days of the experiment are excluded from the calculation of statistics to account for system spinup.

Compared to ClrSky, the all-sky approach reduces the mean relative humidity analysis (e.g., over the Pacific Ocean to the west of South America) at 850 hPa, and the mean temperature analysis increases correspondingly at 850 hPa but decreases slightly at 700 hPa (figures not shown). The overestimation of stratus clouds along continental western coasts has been a persistent problem in the clear-sky GFS system. The all-sky approach is seen to reduce relative humidity in these areas, a step in

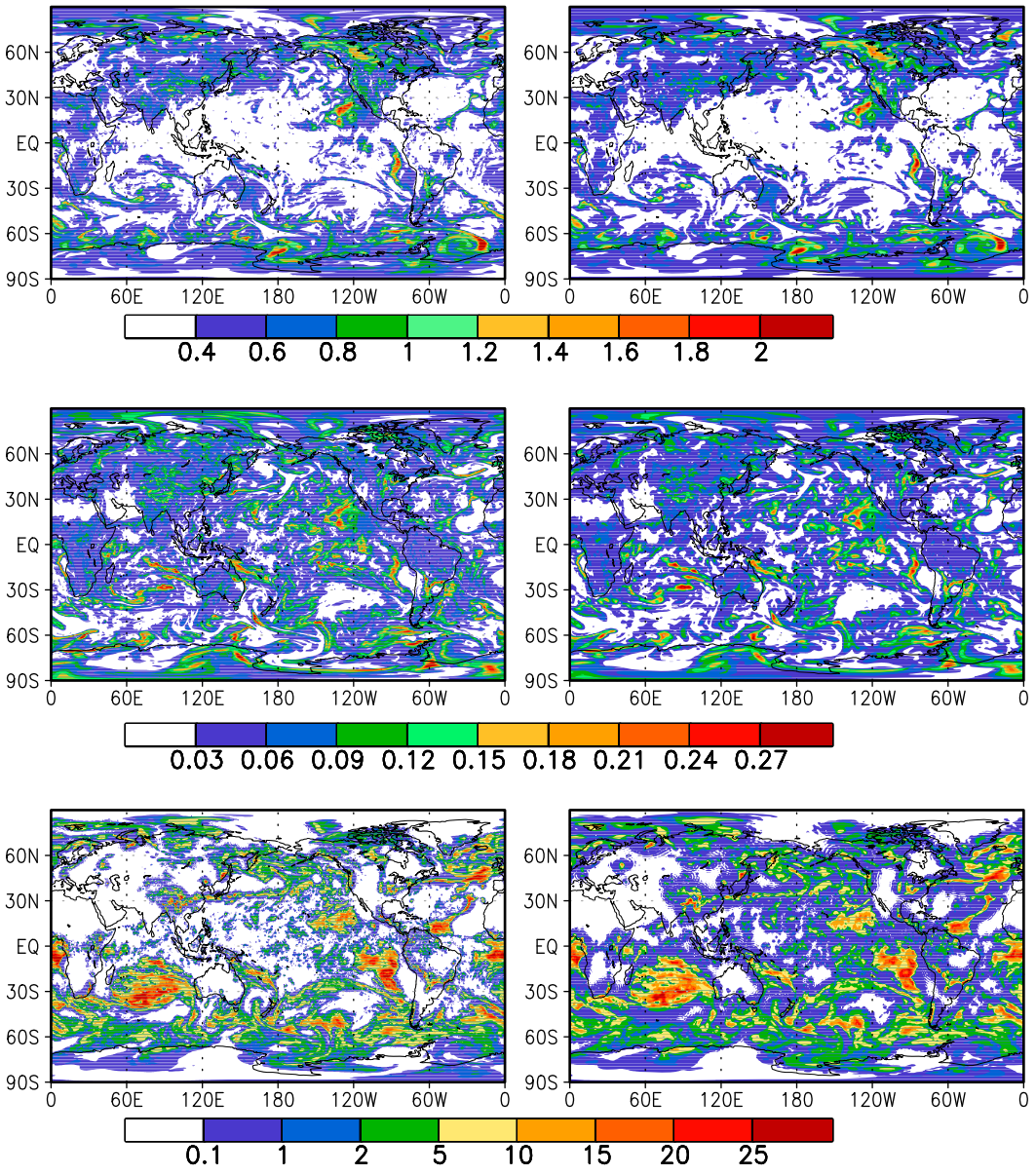

FIG. 18. Ensemble spread at level 13 (~850 hPa) that is valid at $T + 6$ h in the (left) single- and (right) dual-resolution configurations at 0000 UTC 2 Jun 2015 for (top) temperature (K), (middle) relative humidity (%), and (bottom) cloud water ($\times 10^5$ kg kg^{-1}).

the right direction. A comparison of the mean relative humidity and temperature at 850 hPa against the ECMWF analyses (Figs. 19 and 20) also indicates that the clear-sky analyses are too wet and too cold at the west of the continents, whereas the all-sky analyses are much closer to the ECMWF analyses.

Regarding the fits to other observations, as pointed out in Geer et al. (2012), it is difficult to do a "clean" comparison between the clear-sky and all-sky approaches, since our all-sky approach not only includes model cloud information in the radiative transfer calculation, but also involves many changes as described in section 3. Nevertheless, the discussion concerning the fits to other observations is presented. The RMSE differences between AllSky and ClrSky for the analysis (black line) and first-guess (red line) fits to the rawinsonde specific humidity observations are provided in Fig. 21. The results are comparable in the Northern Hemisphere and North America for the two approaches, while in the Southern Hemisphere and tropics slight degradation is observed at 850 hPa but improvement is seen at 925 hPa when all-sky AMSU-A radiances are used. The RMSE

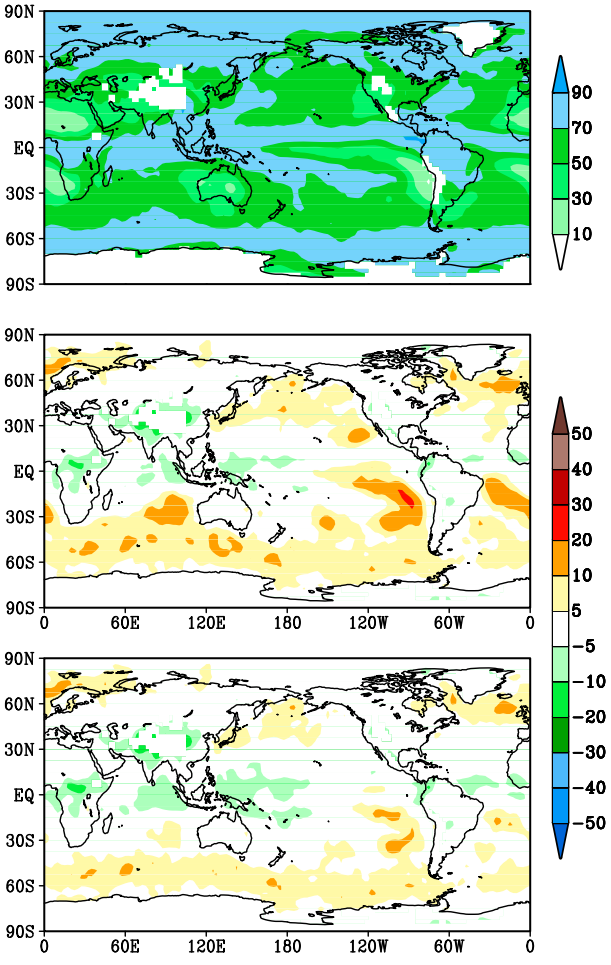

FIG. 19. (top) Averaged relative humidity analysis (%) for the ECMWF analysis, (middle) the difference between ClrSky and ECMWF (%), and (bottom) the difference between AllSky and ECMWF (%) at 850 hPa for the period from 27 Oct to 1 Dec 2013.

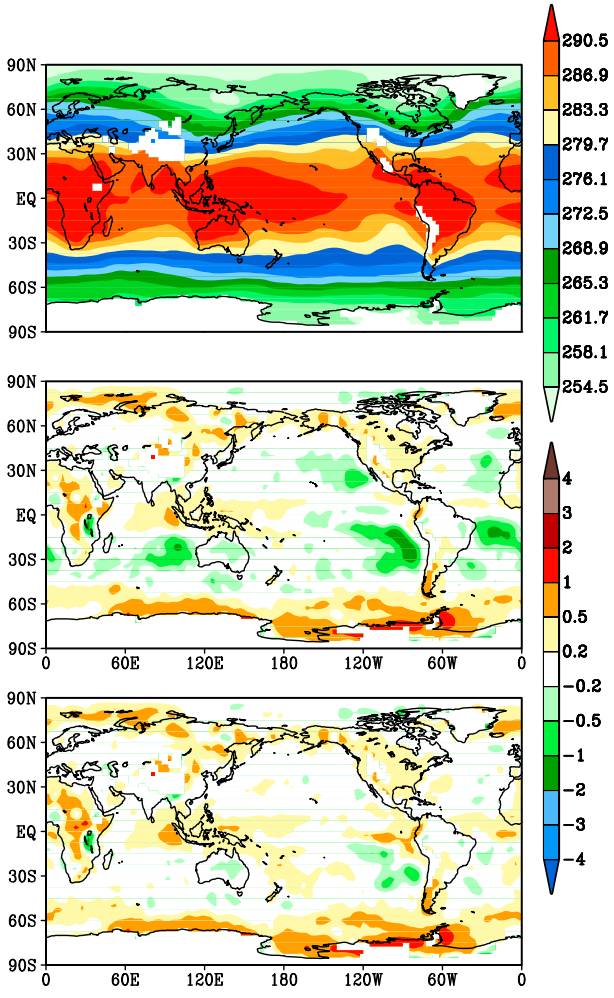

FIG. 20. As in Fig. 19, but for temperature (K).

fits to the rawinsonde temperature observations are also very similar for the two approaches, except that the RMSE analysis and first-guess fits are reduced slightly at 300 and 400 hPa in the Southern Hemisphere in the all-sky approach (figure not shown). As to the RMSE difference between AllSky and ClrSky for the first-guess fit to satellite data for the Microwave Humidity Sounder (MHS) (Fig. 22, right) degradation is found in AllSky when verifying against MHS *MetOp-B* channels 1 and 2, and to a lesser extent for MHS *NOAA-18* and *NOAA-19*. However, an approximate 1% increase in the assimilation of MHS radiances is realized in the all-sky approach relative to the clear-sky approach (Fig. 22, left). Comparable ClrSky and AllSky RMSE fit results are observed for the High Resolution Infrared Radiation Sounders (HIRS) on board *NOAA-19*. For the Atmospheric Infrared Sounder (AIRS), the RMSE fits are mixed for temperature and surface channels but are reduced slightly by up to 1.3% for water vapor channels (figure not shown). Overall, compared to the clear-sky approach, unlike Geer et al. (2012), who found that including model cloud increased the total error in the comparison between the model and observations, the results of our all-sky AMSU-A radiance assimilation experiments indicate mostly neutral impacts as they relate to the RMSE of OmF. This performance difference may be partly because more radiances from cloud-affected FOVs are assimilated in our all-sky study than in the Geer et al. (2012) study. Another reason may be that ECMWF's system possibly has a better-constrained lower-troposphere water vapor analysis through microwave imager assimilation.

The forecast anomaly correlation and RMSE, which are verified against self-analyses, are also examined for many variables, levels, and lead times for the ClrSky and AllSky experiments. The anomaly correlation results for geopotential height at 500 hPa (Fig. 23) indicate that the

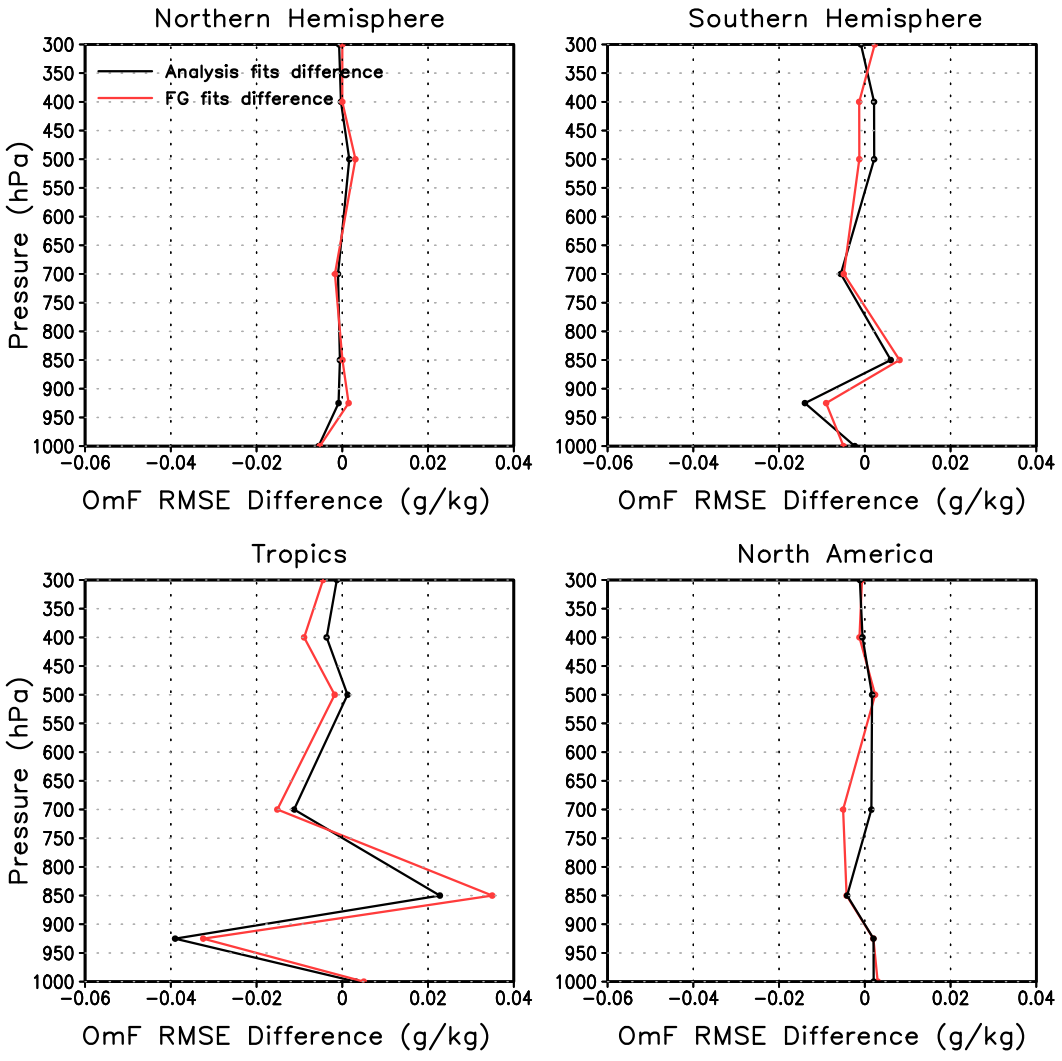

FIG. 21. RMSE differences between AllSky and ClrSky for analysis (black line) and first guess (FG, red line) fits to rawinsonde specific humidity observations in the Northern and Southern Hemispheres, tropics, and North America from 27 Oct to 1 Dec 2013.

assimilation of all-sky radiance data has a neutral impact in both the Northern (left panels) and Southern (right panel) Hemispheres. A similar pattern is also observed at 700 and 250 hPa. The vector RMSE (RMSVE) for wind forecasts is also calculated, and slight reductions of the wind RMSVE are found in both the Northern and Southern Hemispheres in the AllSky experiment (figure not shown), but they are not statistically significant at the 95% confidence level. As the anomaly correlation of 850-hPa temperature is improved significantly in the tropics and Southern Hemisphere (figure not shown), the RMSE results of temperature forecasts at 850 hPa (Fig. 24) indicate that the temperature forecast is degraded in the 24-h forecast (with a very small magnitude around 0.008 K) and is neutral at other leading times in the Northern Hemisphere (top panel) and is significantly improved beyond 48 h in the tropics (middle panel) and 24 h in the Southern Hemisphere (bottom panel). The impact of the all-sky approach on temperature is mostly neutral elsewhere. It should be noted that these results are from a roughly month-long trial; rigorous assessments of the all-sky microwave radiance assimilation for a longer period should be conducted in the future.

6. Conclusions and future work

Assimilation of all-sky microwave radiance in the GSI analysis system has been developed at NCEP. Since the original operational GFS system uses the radiance data with the clear-sky approach, necessary changes are made for all-sky conditions in data thinning, quality control, observation error assignment, bias correction, and cloud

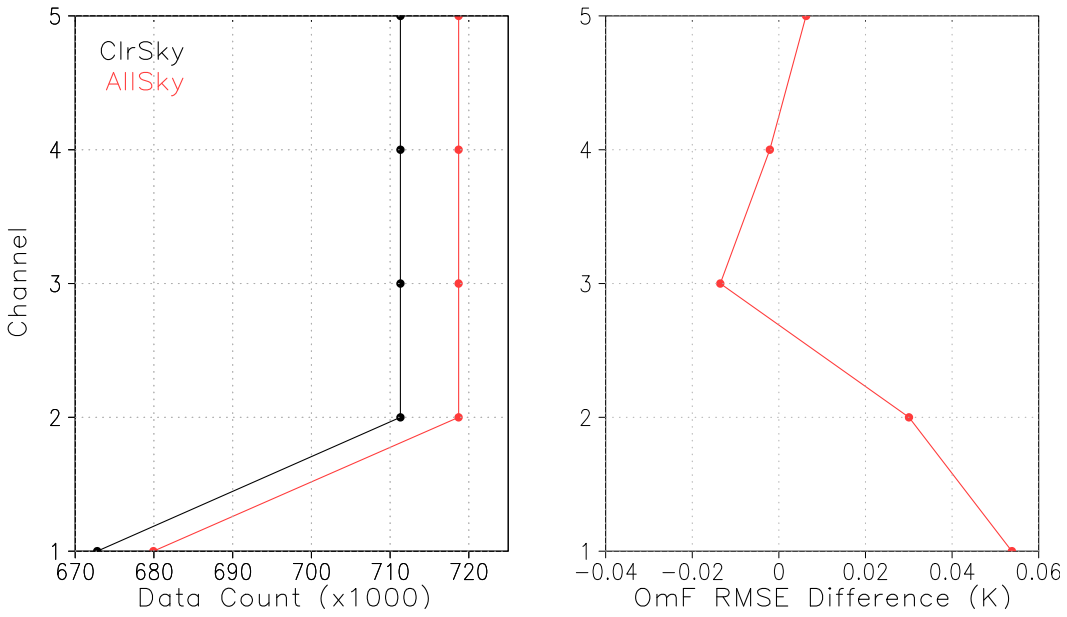

FIG. 22. (left) Used MHS *MetOp-B* radiance data counts and (right) RMSE difference between the AllSky (red) and ClrSky (black) experiments for first guess fits to MHS *MetOp-B* radiances during the period from 27 Oct to 1 Dec 2013.

background error variance. With the removal of the preference given to the radiance data of cloud-free FOVs in the data-thinning process, and the removal of the thick-cloud filtering in the quality control and cloud liquid water difference bias correction term in the radiance bias correction, nonprecipitating cloudy radiance data over ocean are now assimilated in the GSI analysis system. Despite deficiencies in the forecast model and CRTM, the symmetric observation error assignment and the additional situation-dependent observation error inflation for all-sky conditions are shown to provide reasonable weights for the all-sky radiance data. Furthermore, the new bias correction strategy for all-sky radiances, which uses only a selected data sample to derive the bias correction coefficients, is applied on top of the current radiance bias correction, and it effectively removes the bias for radiance data where clouds retrieved from the observation and those from the first guess are consistent, while preserving the cloud signal in the observation for the radiance data with mismatched cloud information. Meanwhile, within the ensemble–variational framework, the background error covariance is composed of the static term and a part generated from the ensemble forecasts. In the GSI analysis system, normalized cloud water is presently used as the cloud control variable for the all-sky approach, with the new static cloud water background error variance being assigned to be large only where clouds already exist. The issue of cloud loss at the first model integration step, which is common to both clear-sky and all-sky approaches, is investigated from the perspective of ensemble spreads in single- and dual-resolution configurations. The broader cloud water ensemble spread in the dual-resolution configuration may be related to the resolution-dependent physical processes in the forecast model, the stochastic physics algorithm that has no direct link with cloud variability, and the parameter tuning of the algorithm.

The all-sky microwave radiance assimilation, currently focused on the AMSU-A instrument, had been tested extensively in the 3D EnVar GFS system. This approach utilizes radiance data affected by thick clouds, which amounts to about a 10% increase in assimilated radiances from AMSU-A channels 1–5 and 12% from channel 15. More realistic simulated brightness temperatures are produced by considering cloud information in the radiative transfer, and radiance data information is projected directly onto not only temperature and moisture fields but also cloud fields. Furthermore, the all-sky approach produces much more realistic cloud water analysis increments; particularly, it can reduce humidity and cloud water from the analysis off the west of the continents where the clear-sky GFS system has produced too much stratus. The all-sky radiance assimilation is an important step moving toward the full utilization of radiance observations. In terms of RMSE first-guess fit to observations, while Geer et al. (2012) found with ECMWF's system that including model cloud increases the total error in the comparison

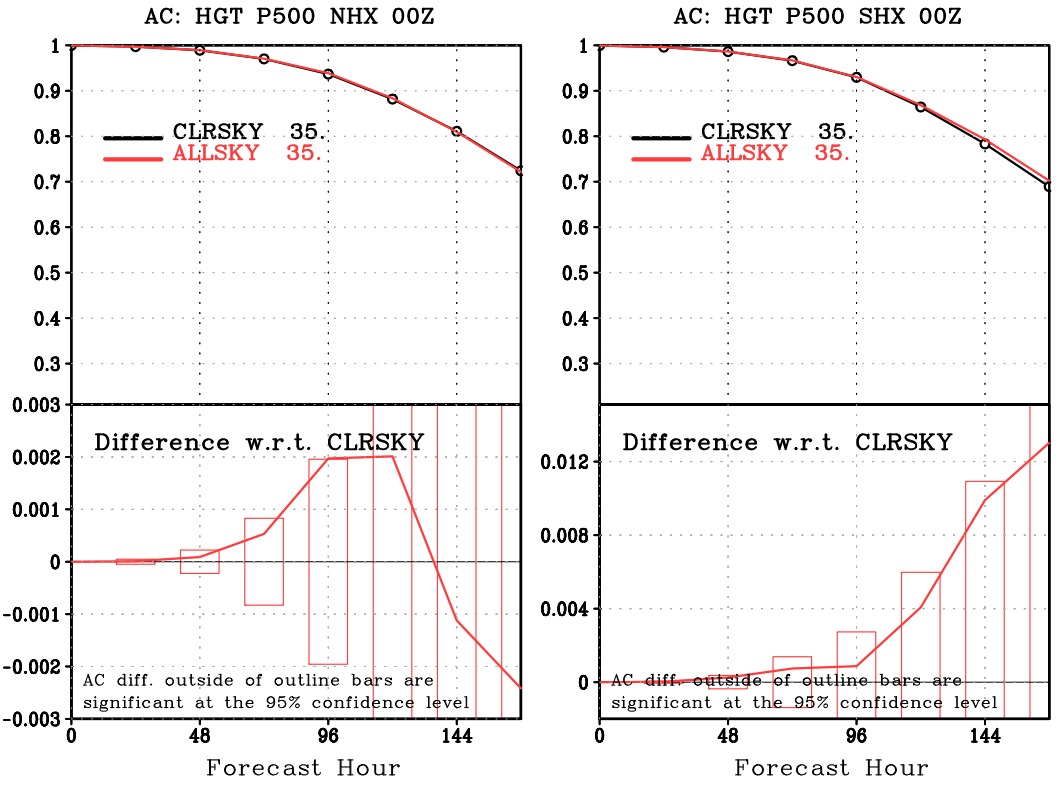

FIG. 23. (top) Geopotential height anomaly correlation at 500 hPa for the (left) Northern and (right) Southern Hemispheres during the period from 27 Oct to 1 Dec 2013 for ClrSky (black) and AllSky (red). (bottom) The differences with respect to ClrSky for the (left) Northern and (right) Southern Hemispheres.

between model and observations, our results are mixed. The RMSE of first-guess fit to rawinsonde specific humidity observations is slightly degraded at 850 hPa but improved at 925 hPa in the tropics and Southern Hemisphere, and the fits to rawinsonde temperature are improved at 300 and 400 hPa in the Southern Hemisphere. RMSE fits to MHS channels 1 and 2 increase with more MHS radiances being used, but AIRS water vapor channels see small decreases in RMSE fits. While improvements are seen in temperature analyses and forecasts at 850 hPa in the tropics and Southern Hemisphere, the impact of the all-sky approach on forecast skills is mostly neutral. The all-sky microwave radiance assimilation had been included in the May 2016 GFS pre-implementation package and tested along with other developmental components in the 4D EnVar GFS parallel experiment at NCEP. This package was implemented on 12 May 2016 and became the operational GFS system.

This work is only a starting point for all-sky radiance assimilation. More studies are necessary on observation error assignment for all-sky conditions and application of VQC on radiances. Experiments with an additional cloud-related bias correction term and sensitivity studies using different magnitudes of cloud static and ensemble background error are ongoing. The dependency of observation error on scan position, as found in Geer et al. (2012), will also be investigated. Moreover, since the all-sky GSI has the flexibility of using individual hydrometeors as the cloud control variables, the performance of the GSI all-sky radiance assimilation should be assessed with the individual hydrometeors when they later become the prognostic variables in the forecast model, as planned. As further refinements to the all-sky assimilation continue, the GSI all-sky capabilities are expected to expand to other microwave instruments. Experiments to assimilate the Advanced Technology Microwave Sounder (ATMS) instrument all-sky radiances are under way. Additionally, validation and the improvement of scattering in the CRTM is essential as we move toward the assimilation of precipitating clouds.

Finally, there are outstanding issues related to forecast model cloud accuracy and cloud variability in the stochastic physics. It is expected that further developments on these fronts will improve the all-sky radiance assimilation accordingly. Another aspect that still needs more attention is the choice of the cloud control variable(s). Although the current cloud control variable works

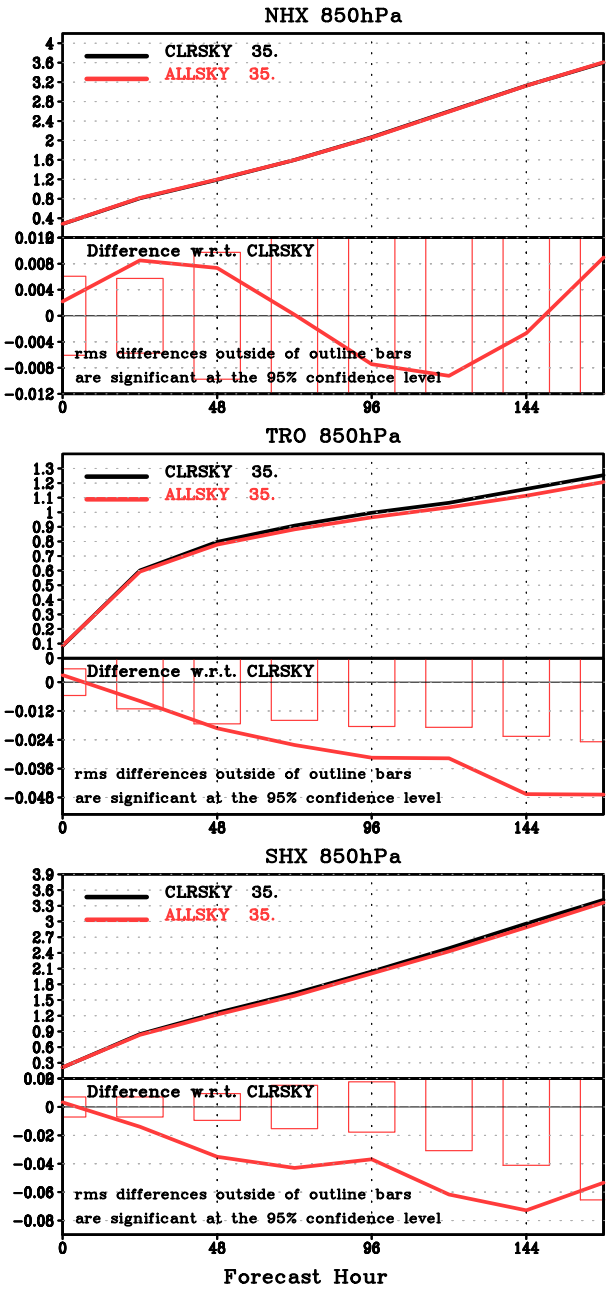

FIG. 24. The RMSE results of temperature forecasts at 850 hPa for (middle) the tropics and the (top) Northern and (bottom) Southern Hemispheres during the period from 27 Oct to 1 Dec 2013. Shown are the RMSEs of temperature forecasts for ClrSky (black line) and AllSky (red line) at the top of each panel, and the RMSE difference of AllSky with respect to ClrSky at the bottom of each panel. The RMSE is computed against each system's own analysis (in K).

reasonably well, we believe investigation of better cloud control variable(s) in the future will be beneficial, especially when combined with the consideration of balance among temperature, moisture, and clouds.

Acknowledgments. Our coauthor and colleague Paul Van Delst passed away during the revision of this paper. Paul was very enthusiastic about his work and contributed greatly to the field of radiative transfer modeling. His talent as a scientist, attention to detail as a software engineer, and generous personality will be deeply missed.

We thank Ruiyu Sun for helpful discussions and providing references on the GFS moist physics, thank Fanglin Yang for helping on plotting software package, and thank Jeffrey Whitaker for providing references for SKEB, SPPT, and SHUM in the stochastic physics. Thanks are also due to Jordan C. Alpert, James Jung, and David Parrish for their comments and suggestions during the EMC internal review process. We also would like to thank three anonymous reviewers for their very helpful suggestions and comments. This work is supported by Disaster Relief Appropriations Act of 2013 (DRA) and NCEP/EMC base funding.

REFERENCES

Andersson, E., and H. Jarvinen, 1998: Variational quality control. ECMWF Tech. Memo. 250, 31 pp. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/1998/7759-variational-quality-control.pdf.]

Baker, N., and R. H. Langland, 2007: Diagnostics for evaluating the impact of satellite observations. *Proc. Seminar on Recent Development in the Use of Satellite Observations in NWP*, Reading, United Kingdom, ECMWF, 309–326. [Available online at www.ecmwf.int/sites/default/files/elibrary/2008/7877-diagnostics-evaluating-impact-satellite-observations.pdf.]

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau, 2006a: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1D-Var. *Quart. J. Roy. Meteor. Soc.*, **132**, 2277–2306, doi:10.1256/qj.05.189.

——, ——, D. Salmond, A. Benedetti, S. Saarinen, and E. Moreau, 2006b: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **132**, 2307–2332, doi:10.1256/qj.06.07.

——, E. Moreau, F. Chevallier, and U. O'Keeffe, 2006c: Multiple-scattering microwave radiative transfer for data assimilation applications. *Quart. J. Roy. Meteor. Soc.*, **132**, 1259–1281, doi:10.1256/qj.05.153.

——, A. J. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Quart. J. Roy. Meteor. Soc.*, **136**, 1868–1885, doi:10.1002/qj.659.

Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A spectral stochastic kinetic backscatter scheme and its impact on flow-dependent predictability in the ECMWF Ensemble Prediction System. *J. Atmos. Sci.*, **66**, 603–626, doi:10.1175/2008JAS2677.1.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, doi:10.1002/qj.49712556006.

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, doi:10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

——, D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, **6**, 538–547, doi:10.1175/1520-0434(1991)006<0538:TNGOAS>2.0.CO;2.

Ding, S., P. Yang, F. Weng, Q. Liu, Y. Han, P. Van Delst, J. Li, and B. Baum, 2011: Validation of the community radiative transfer model. *J. Quant. Spectrosc. Radiat. Transfer*, **112**, 1050–1064, doi:10.1016/j.jqsrt.2010.11.009.

Geer, A. J., and P. Bauer, 2011a: All-sky data assimilation of radiances from microwave sounders at ECMWF. *Proc. 2011 EUMETSAT Meteorological Satellite Conf.*, Oslo, Norway, EUMETSAT. [Available online at http://www.eumetsat.int/website/home/News/ConferencesandEvents/PreviousEvents/DAT_2039705.html.]

——, and ——, 2011b: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037, doi:10.1002/qj.830.

——, ——, and P. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Quart. J. Roy. Meteor. Soc.*, **136**, 1886–1905, doi:10.1002/qj.681.

——, ——, and S. English, 2012: Assimilating AMSU-A temperature sounding channels in the presence of cloud and precipitation. ECMWF Tech. Memo. 670, 41 pp. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2012/9514-assimilating-amsu-temperature-sounding-channels-presence-cloud-and-precipitation.pdf.]

Gelaro, R., and Y. Zhu, 2009: Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus*, **61A**, 179–193, doi:10.1111/j.1600-0870.2008.00388.x.

Greybush, S., E. Kalnay, T. Miyoshi, K. Ide, and B. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, doi:10.1175/2010MWR3328.1.

Grody, N., F. Weng, and R. Ferraro, 1999: Application of AMSU for obtaining water vapor, cloud liquid water, precipitation, snow cover and sea ice concentration. Preprints, *10th Int. TOVS Study Conf.*, Boulder, CO, Int. ATOVS Working Group, 230–240. [Available online at http://library.ssec.wisc.edu/research_Resources/publications/pdfs/ITSC10/grody01_ITSC10_1999.pdf.]

——, J. Zhao, R. Ferraro, F. Weng, and R. Boers, 2001: Determination of precipitable water and cloud liquid water over oceans from the NOAA 15 advanced microwave sounding unit. *J. Geophys. Res.*, **106**, 2943–2953, doi:10.1029/2000JD900616.

Groff, D., Q. Liu, P. Van Delst, A. Collard, and Y. Zhu, 2013: The impact of Community Radiative Transfer Model microwave sea surface emissivity improvements on forecast skill. Preprints, *Sixth WMO Symp. on Data Assimilation*, College Park, MD, WMO, Ap-10. [Available online at http://das6.cscamm.umd.edu/program/poster_program.html.]

——, P. Van Delst, Q. Liu, A. Collard, E. Liu, J. Jung, and A. Lim, 2014: CRTM releases and ongoing CRTM development. Preprints, *12th Technical Review and Science Workshop on Satellite Data Assimilation*, College Park, MD, Joint Center for Satellite Data Assimilation. [Available online at http://www.jcsda.noaa.gov/documents/meetings/wkshp2014/posters/groff_poster_workshop2014.pptx.]

Guidard, V., N. Fourrie, P. Brousseau, and F. Rabier, 2011: Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Quart. J. Roy. Meteor. Soc.*, **137**, 1975–1987, doi:10.1002/qj.928.

Han, J., and H. L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. *Wea. Forecasting*, **26**, 520–533, doi:10.1175/WAF-D-10-05038.1.

Ingleby, N., A. Lorenc, K. Ngan, F. Rawlins, and D. Jackson, 2013: Improved variational analyses using a nonlinear humidity control variable. *Quart. J. Roy. Meteor. Soc.*, **139**, 1875–1887, doi:10.1002/qj.2073.

Kleist, D. T., and K. Ide, 2015a: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Mon. Wea. Rev.*, **143**, 433–451, doi:10.1175/MWR-D-13-00351.1.

——, and ——, 2015b: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4D EnVar and hybrid variants. *Mon. Wea. Rev.*, **143**, 452–470, doi:10.1175/MWR-D-13-00350.1.

——, 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, doi:10.1175/2009WAF2222201.1.

Liu, Q., and F. Weng, 2006: Advanced doubling-adding method for radiative transfer in planetary atmosphere. *J. Atmos. Sci.*, **63**, 3459–3465, doi:10.1175/JAS3808.1.

Lorenc, A., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203, doi:10.1256/qj.02.132.

McNally, A. P., 2009: The direct assimilation of cloud-affected satellite infrared radiances in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **135**, 1214–1229, doi:10.1002/qj.426.

Moorthi, S., H. L. Pan, and P. Caplan, 2001: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. NWS Tech. Procedures Bull. 484, 14 pp. [Available online at http://docs.lib.noaa.gov/rescue/wb_technical_procedures_bulletin/TPB_484.pdf.]

Okamoto, K., 2013: Assimilation of overcast cloudy infrared radiances of the geostationary MTSAT-1R imager. *Quart. J. Roy. Meteor. Soc.*, **139**, 715–730, doi:10.1002/qj.1994.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Pavelin, E. G., S. J. English, and J. R. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **134**, 737–749, doi:10.1002/qj.243.

Purser, R. J., 2011: Mathematical principles of the construction and characterization of a parameterized family of Gaussian mixture distributions suitable to serve as models for the probability distributions of measurement errors in nonlinear quality control. NOAA/NCEP Office Note 468, 42 pp. [Available online at http://www.emc.ncep.noaa.gov/officenotes/newernotes/on468.pdf.]

——, W. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535, doi:10.1175//1520-0493(2003)131<1524:NAOTAO>2.0.CO;2.

——, ——, ——, and ——, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1536–1548, doi:10.1175//2543.1.

Shutts, G. J., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102, doi:10.1256/qj.04.106.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1657, doi:10.1175/1520-0493(1989)117<1641:CACPSW>2.0.CO;2.

Tompkins, A. M., and J. Berner, 2008: A stochastic convective approach to account for model uncertainty due to unresolved humidity variability. *J. Geophys. Res.*, **113**, D18101, doi:10.1029/2007JD009284.

Treadon, R. E., 1997: Assimilation of satellite derived precipitation estimates within the NCEP GDAS. Ph.D. thesis, The Florida State University, Tallahassee, FL, 348 pp.

Wang, X., and T. Lei, 2014: GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. *Mon. Wea. Rev.*, **142**, 3303–3325, doi:10.1175/MWR-D-13-00303.1.

——, D. Parrish, D. Kleist, and J. S. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, doi:10.1175/MWR-D-12-00141.1.

Weng, F., L. Zhao, R. R. Ferraro, G. Poe, X. Li, and N. C. Grody, 2003: Advanced Microwave Sounding Unit cloud and precipitation algorithms. *Radio Sci.*, **38**, 8086–8096, doi:10.1029/2002RS002679.

Whitaker, J., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

——, and ——, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, doi:10.1175/MWR-D-11-00276.1.

Wu, W. S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, doi:10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Zhao, Q., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. *Mon. Wea. Rev.*, **125**, 1931–1953, doi:10.1175/1520-0493(1997)125<1931:APCSFO>2.0.CO;2.

Zhu, Y., and R. Gelaro, 2008: Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Wea. Rev.*, **136**, 335–351, doi:10.1175/2007MWR2046.1.

——, J. Derber, A. Collard, D. Dee, R. Treadon, G. Gayno, and J. A. Jung, 2014a: 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, doi:10.1002/qj.2233.

——, and Coauthors, 2014b: Variational bias correction in the NCEP's Data Assimilation System. Preprints, *19th Int. TOVS Study Conf.*, Jeju Island, South Korea, Int. ATOVS Working Group, 10.02. [Available online at http://cimss.ssec.wisc.edu/itwg/itsc/itsc19/program/papers/10_02_zhu.pdf.]