

Community Global Observing System Simulation Experiment (OSSE) Package (CGOP): Perfect Observations Simulation Validation

SID-AHMED BOUKABARA,^a KAYO IDE,^b NARGES SHAHROUDI,^{c,a} YAN ZHOU,^b TONG ZHU,^{d,a}
RUIFANG LI,^e LIDIA CUCURULL,^f ROBERT ATLAS,^f SEAN P. F. CASEY,^{g,f} AND ROSS N. HOFFMAN^{g,f}

^a NOAA/NESDIS/Center for Satellite Applications and Research (STAR), College Park, Maryland

^b Cooperative Institute for Climate and Satellites, University of Maryland, College Park, College Park, Maryland

^c Riverside Technology, Inc., Silver Spring, Maryland

^d Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

^e Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado

^f NOAA/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida

^g Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida

(Manuscript received 19 April 2017, in final form 7 September 2017)

ABSTRACT

The simulation of observations—a critical Community Global Observing System Simulation Experiment (OSSE) Package (CGOP) component—is validated first by a comparison of error-free simulated observations for the first 24 h at the start of the nature run (NR) to the real observations for those sensors that operated during that period. Sample results of this validation are presented here for existing low-Earth-orbiting (LEO) infrared (IR) and microwave (MW) brightness temperature (BT) observations, for radio occultation (RO) bending angle observations, and for various types of conventional observations. For sensors not operating at the start of the NR, a qualitative validation is obtained by comparing geographic and statistical characteristics of observations over the initial day for such a sensor and an existing similar sensor. The comparisons agree, with no significant unexplained bias, and to within the uncertainties caused by real observation errors, time and space collocation differences, radiative transfer uncertainties, and differences between the NR and reality. To validate channels of a proposed future MW sensor with no equivalent existing spaceborne sensor channel, multiple linear regression is used to relate these channels to existing similar channels. The validation then compares observations simulated from the NR to observations predicted by the regression relationship applied to actual real observations of the existing channels. Overall, the CGOP simulations of error-free observations from conventional and satellite platforms that make up the global observing system are found to be reasonably accurate and suitable as a starting point for creating realistic simulated observations for OSSEs. These findings complete a critical step in the CGOP validation, thereby reducing the caveats required when interpreting the OSSE results.

1. Introduction

Observing system experiments (OSEs) are data assimilation (DA) and forecast experiments that measure the impact of an observing system by comparing analysis and forecast results with and without the particular observing system. As described by Boukabara et al. (2016), observing system simulation experiments (OSSEs) extend the concept of OSEs to proposed future sensors by using observations simulated from the nature run (NR). OSSEs support (i) decision-makers by providing valuable quantitative information on the impact of proposed

future sensors, (ii) system engineers through trade studies to tune the design of the proposed sensor, and (iii) data assimilation scientists by allowing the development and testing of new data assimilation algorithms to optimally extract the information content of the proposed sensor before that sensor exists (Atlas 1997; Atlas et al. 2001, 2015). However, no OSSE is perfect. The OSSE system components and the entire OSSE system must be validated and calibrated, and deficiencies noted during this process must be documented to ensure the correct interpretation of the OSSE results (Hoffman and Atlas 2016). First, the development of the individual OSSE system components must include sufficient testing, validation, and calibration to ensure realistic results. Second, because OSSE systems are complex, approximations

Corresponding author: Dr. Sid-Ahmed Boukabara, sid.boukabara@noaa.gov

made in one component may have substantial impacts on other downstream components, and some iteration of the validation of the overall OSSE system may be required. Third, since even minor deficiencies or inconsistencies may be crucial in designing, conducting, and/or interpreting an experiment for a particular proposed future sensor, it is critical to document and review such deficiencies during the entire OSSE process. In addition, users should bear in mind that OSSE results are only one factor in the decision-making process concerning proposed sensors.

In section 2 of Boukabara et al. (2016), the Community Global OSSE Package (CGOP) was described in detail from a functional viewpoint. Currently, as described in Boukabara et al. (2016), CGOP includes the NOAA DA (Kleist and Ide 2015) and forecast systems. Here, this Part II reports the quantitative validation of the first component of the package: the simulation of perfect (i.e., with no added explicit errors) observations from both existing and proposed sensors. All observation types described in Boukabara et al. (2016) have been validated. This includes all infrared (IR) and microwave (MW) sensors listed in Table 4 of Boukabara et al. (2016), namely, Atmospheric Infrared Sounder (AIRS), Advanced Microwave Sounding Unit-A (AMSU-A), AMSU-B, Advanced Technology Microwave Sounder (ATMS), Cross-Track Infrared Sounder (CrIS), High-Resolution Infrared Sounder-3 (HIRS-3), HIRS-4, Infrared Atmospheric Sounding Interferometer (IASI), Microwave Humidity Sounding (MHS), sounder (SNDR), Spinning Enhanced Visible and Infrared Imager (SEVIRI), and Special Sensor Microwave Imager/Sounder (SSMIS; acronyms are also defined in the appendix). In this paper examples or summaries for each data type are presented.

The present study has two important outcomes. First, on a simplistic level, it “validates” the mechanics of simulating perfect observations—that is, it interpolates the NR and applies the forward problem procedures. Second, and more importantly, it sets goals for the modeling of explicitly added errors—differences found in comparing real and perfect simulated observations reveal the complexity, structure, and causes of errors that should be present in the error-added simulated observations. In this study, references to validation of perfect observations are to validation in these two senses: the checking of the simulation of perfect observations, and the description and characterization of the explicit simulated errors that should be added to the perfect observations.

Section 2 describes our overall approach to validate the perfect simulated observations, using the NR initial day. Section 3 summarizes the validation conducted by Gelaro et al. (2015) of the 7-km Goddard Earth Observing System Model, version 5 (GEOS-5) nature

run (G5NR). Insights identified during the validation of the G5NR are critical to our discussion of the perfect observations validation. The sections that follow describe the validation of the error-free simulated observations. There we validate the simulated brightness temperature (BT) observations from existing low-Earth-orbiting (LEO) MW and IR satellite sensors (section 4); bending angle observations from Global Navigation Satellite System (GNSS) radio occultation (RO) observations (section 5); and conventional observations from existing sensors based on land, ships, buoys, balloons, and aircraft (section 6). Following the operational convention, we include atmospheric motion vector (AMV) and satellite surface wind observations in the category of conventional observations. Then, in section 7, we present qualitative validations of simulated BT from a proposed MW sensor on a constellation of small satellites. Section 8 provides a summary, conclusions, and brief discussions of the remaining validation and calibration activities of other CGOP components and of the application of the entire OSSE system. These latter topics will be the subject of a separate paper.

2. Strategies to compare real and perfect simulated observations

Since every OSSE conducted with CGOP depends on the NR and the simulated observations, assessing the accuracy of the perfect simulated observations and explaining any differences with respect to reality are critical prerequisites before the subsequent steps of calibrating and validating the rest of the OSSE system. First, differences between the NR and reality will be mirrored in simulated and real observations. In this study, we limit these differences by restricting our comparisons to the 24 h at the beginning of the NR, that is, the NR initial day. Below, we present detailed discussions of both the principal advantage and the principal disadvantage of this approach: that one-to-one comparisons of observations are possible but that the results may not be representative of other periods. Second, differences between the forward problem and how real sensors observe the earth result in further differences between simulated and real observations. Differences of both types should be included in the errors explicitly added to the simulated perfect observations. It is important to note that there are two types of validation described here and that both of these validations of NR-based simulations are not to be confused with a thorough sensor validation, which is performed to assess the accuracy of the measurements and to establish instrumental error characteristics (typically during sensor calibration-validation activities).

In general, when conventional and satellite observations are simulated from the NR, all of these should be simulated consistently with coverage, resolution, and accuracy comparable to real observations. In addition, bias and horizontal and vertical correlations of errors with each other and with the synoptic situation should be introduced appropriately. The validation of the simulated observations should include, for example, side-by-side comparisons of coverage maps, quality-control (QC) yields, and frequency distributions (histograms) of observations and observation innovations (observations minus DA background).

The present study validates so-called perfect simulated observations, which do not include the explicit addition of random and bias components of error. When compared to real observations, perfect observations include implicit errors and representativeness differences (Errico et al. 2013).¹ Here, implicit observation error includes interpolation error and forward problem error. For example, the difference between perfect observations and reality strongly depends on the accuracy of the forward operators and the required parameters that must be provided, such as surface emissivity in the case of satellite radiances. Representativeness differences refers to scales and phenomena that in reality are not present in the NR model.²

Since the NR is a free-running numerical weather prediction (NWP) forecast, the NR atmosphere will substantially differ from the real atmosphere and as a result the simulated observations will be uncorrelated with the real observations even for the same period. The exception, of course, is at the beginning of the NR, where close agreement of the NR and reality is expected. The period used here for our assessment is the NR initial day, unless otherwise noted. The first type of validation used in this study is quantitative and involves the comparison, for each data source, of perfect simulated to actual observations that match up on a one-to-one basis. This is possible only for sensors that were functioning on 16 May 2005 and for sensors that the DA system is capable of assimilating. These one-to-one comparisons are very valuable for demonstrating the correctness of the implementation of the OSSE observation simulation software (as well as other OSSE system components). In the current CGOP, as a result of the use of quality-controlled data in PREPBUFR files [i.e., observations preprocessed and quality controlled

as described by Keyser (2017) in Binary Universal Form for the Representation of Meteorological Data (BUFR) format] and other template files, simulated and real observations agree in time and location. Observed values should match closely, since the NR is linearly interpolated in both time and location. Even in this type of quantitative validation, the G5NR and CGOP forecasts diverge over the 24-h period, and therefore we cannot expect the perfect observation increments (observation minus perfect simulated observation) to be within instrumental-noise-plus-representativeness difference levels.

The figures presented in this paper that are based on one-to-one comparisons of “perfect” simulated observations (P) to real observations (O) also show the perfect observation increments ($O - P$; increments). For existing observations, simulated observations were created for each actual observation present in the template files: PREPBUFR for conventional observations, BUFR for satellite RO observations, and BUFR thinned to match the Gridpoint Statistical Interpolation analysis system (GSI) radiance statistics and diagnostics (RADSTAT) output for satellite BT observations. PREPBUFR observations have undergone some preliminary QC. For satellite BT and RO observations, the plots include all observations during the NR initial day, that is, the four 16 May 2005 DA cycles. For existing conventional observations, the plots include all observations during the 0000 UTC 16 May 2005 DA cycle. In figures that display maps in the top half and histograms in the bottom half (Figs. 1, 3, and 9), letters identifying the panels are reused for maps and histograms, but it is always clear from the text whether a map or histogram is being referred to. In each half (maps and histograms) O , P , and $O - P$ are displayed from top to bottom. This top-to-bottom ordering is maintained in other figures (Figs. 2, 5, and 8) as well.

One-to-one comparisons are possible only for data types actually observed in reality at the start of the NR. For example, ATMS data are available after 2012, but the G5NR is available only until 31 May 2007. And the same, of course, applies to any novel proposed sensor. Therefore, the second type of validation used in this study is based on the qualitative comparison of the perfect simulated observations of sensors that are proposed, or are otherwise not available in reality, to actual observations for similar data sources on the NR initial day. For these qualitative validations, we rely on comparing maps and histograms and making sure they have similar characteristics, namely, data coverage and patterns in the case of maps and distribution shape and mode location in the case of histograms. An example is given in section 7a, but in this case, assuming the

¹Errico et al. (2013) included representativeness differences in their definition of implicit errors, but we will keep them separate in this discussion.

²We use the phrase *representativeness differences*, since there can be no representativeness errors in the OSSE system where the NR is considered to be the truth.

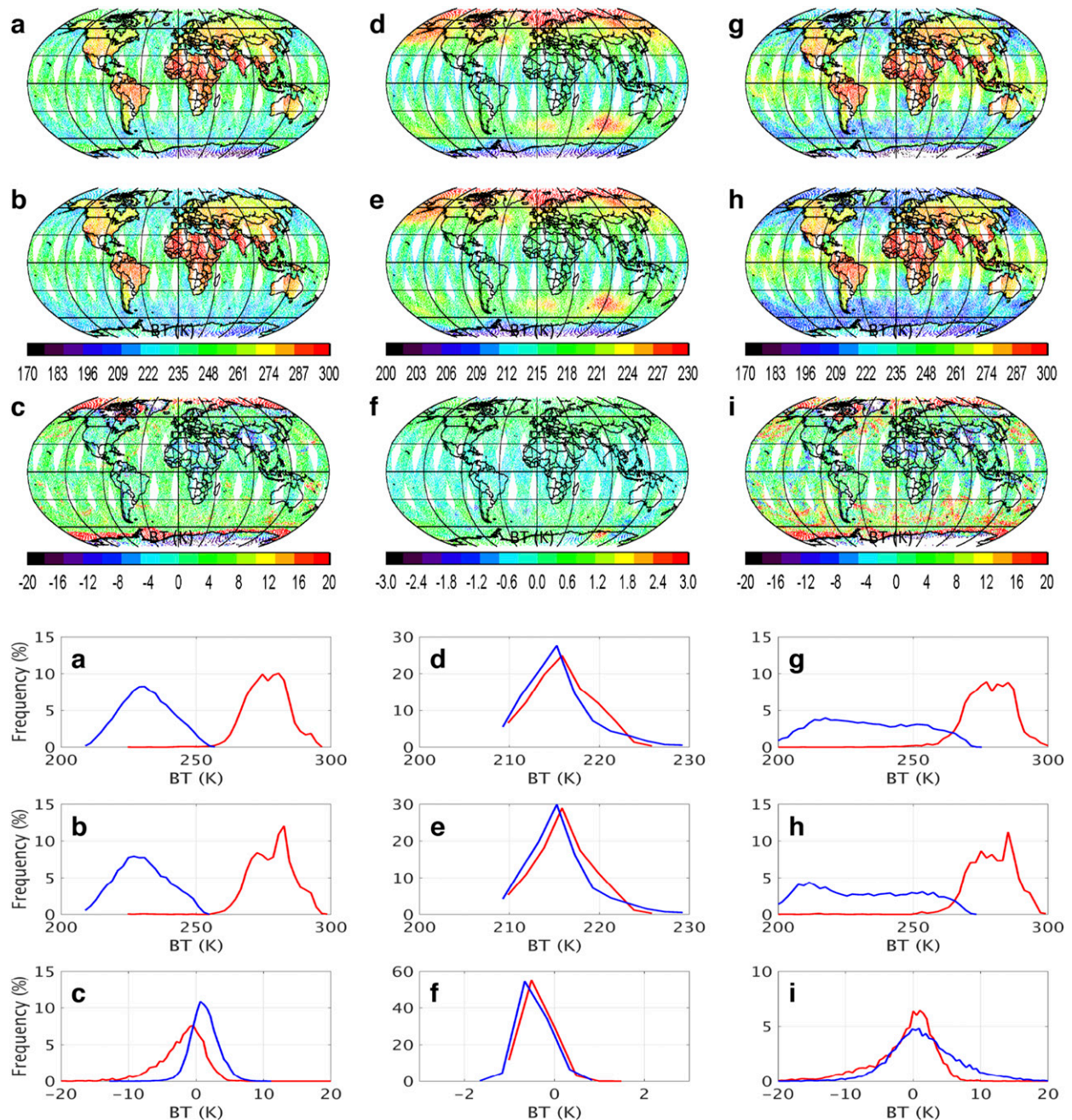


FIG. 1. (top) Maps and (bottom) histograms of AMSU-A on NOAA-15 BT (K) for (a)–(c) channels 3, (d)–(f) 8, and (g)–(i) 15 for the NR initial day, all-sky globally (i.e., for land, ocean, ice, and snow) for the maps, and clear sky (i.e., passed GSI QC) for land (red) and ocean (blue) for the histograms; and in both top and bottom for (a),(d),(g) O ; (b),(e),(h) P ; and (c),(f),(i) $O - P$. The Robinson projection is used to draw the maps. The histograms plot the frequency of observations (percentage) as a function of BT or BT difference.

proposed sensor and existing sensor are in the same orbit allows a one-to-one comparison.

In the case of totally novel sensors, proxy cross validation is employed, in which comparisons are made to proxy observations; that is, for proposed sensor channels

with no existing real observations on the NR initial day that have similar spectral and geometrical characteristics, we rely on proxy or predicted observations that are based on real observations from channels of a sensor existing at the start of the NR that have similar

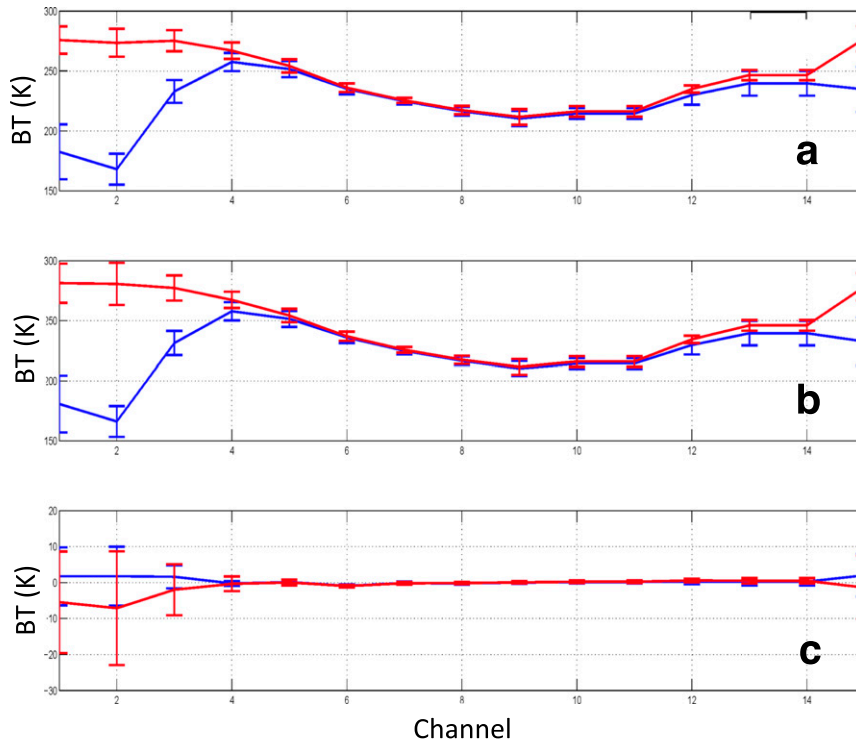


FIG. 2. Mean (lines) and standard deviation (± 1 standard deviation error bars) of AMSU-A on NOAA-15 BT (K) vs channel number (1–15) for the NR initial day, clear sky, and land (red) and ocean (blue); and for (a) *O* and (b) *P* (y axis is from 150 to 300 in increments of 50 K), and (c) $O - P$ (y axis is from -30 to 20 in increments of 10 K).

sensitivity to the environment as each proposed channel. These real observations are transformed to create proxy measurements that can then be substituted for real observations in the validation process. The relationship to predict the proxy measurements is developed by accurately calculating BTs for both existing and proposed channels for a diverse database of geophysical conditions, and then applying one of a number of estimation techniques—in the present case multiple linear regression (section 7b).

It must be noted that the start of the NR is not ideal for all validation and calibration, since 1) adjustment processes will be active and 2) only a small portion of the annual cycle will be sampled. Therefore, further validation and calibration should be done for the period(s) of interest. Except at the start of the NR, simulated observations can be validated only in terms of characteristics and statistics. Even in terms of statistics, large discrepancies between the simulated and real observations may occur. An additional diagnostic is to compare differences between (error added) simulated and perfect simulated observations for the study period to differences between real observations and perfect simulated observations for the NR initial day described in this study. If the errors added explicitly to the

simulated observations are realistic, then maps and statistics of these two differences should be similar, after accounting for seasonal changes. The reasons for large discrepancies observed between the simulated and real observations should be understood and possible effects on the applicability of the OSSE system should be clearly stated. For example, cloud and precipitation effects present in real observations are currently included in the CGOP simulated observations only in terms of data coverage; radiances that are not simulated are those that would be affected by clouds; and radiances that are simulated are simulated for clear-sky conditions. Therefore, large discrepancies are expected unless simulated observations are compared to clear-sky real observations.

3. Nature run validation

As the operational DA and forecast systems evolve, there is a continuing need for improved NRs with higher spatial and temporal resolutions and improved representation of physical processes and meteorological phenomena. With this motivation, the NASA Global Modeling and Assimilation Office (GMAO) produced the G5NR, described briefly in section 3 of Boukabara

et al. (2016) and in depth by Putman et al. (2015) and in other references cited in Boukabara et al. (2016). The G5NR is an advanced global atmospheric 7-km nonhydrostatic NR that is more than 2 yr long (16 May 2005–16 June 2007), and it includes aerosol and land surface components.

GMAO conducted extensive validation of the G5NR in comparison to reality (Gelaro et al. 2015). To paraphrase the abstract of Gelaro et al. (2015), this validation favored comparisons to observational datasets, but it made use of reanalyses and other long model simulations, and it covered many characteristics of the G5NR: the time mean fields of the model variables, the energy spectra, the water cycle, the representation of synoptic features, the surface characteristics and fluxes, the effects of clouds and radiation on the dynamics, the dynamics of the upper atmosphere, and the representation of aerosols and trace gases. While the G5NR was produced prior to the development of the CGOP, this validation is a critical step in the process of validating the CGOP, since the NR is the basis of any OSSE system. In general, the G5NR is a very detailed and realistic simulation, including representations of extreme weather events (e.g., Reale et al. 2017). However, there are definite discrepancies. Some of the differences compared to reality that have been documented by Gelaro et al. (2015) and that may be important in the OSSE context are as follows: 1) because of the very strong computational damping at the smallest resolved scales, the true resolution of the G5NR should be considered as several times coarser than its 7-km grid spacing; 2) land surface temperature biases—potentially important to the simulation of microwave observations—include a cool bias over North Africa and a warm bias in northern high latitudes; 3) tropical tropopause temperatures are warmer in the G5NR compared to reanalyses for all months; 4) specific humidities are too high above 700 hPa; 5) precipitation is overestimated over land and in the Pacific intertropical convergence zone (ITCZ); 6) nonprecipitating cloud water amounts in G5NR are high, resulting in too much reflected shortwave radiation at the top of the atmosphere; and 7) G5NR clouds tend to be biased brighter, but they also tend to be biased lower in altitude, both for high and low clouds. Of these findings, difference 1 implies that small scales that give rise to representativeness errors in reality will not be present in the perfect simulated observations and is one of the reasons that observation errors must be tuned; differences 2–4 are discussed in the comparisons presented below; and differences 5–7 are not relevant to the comparisons reported here because NR clouds and precipitation are not included in the clear-sky BTs that are simulated.

4. Validation of simulated BT observations

In this section, we compare real to perfect simulated BTs. In the CGOP, all BTs are simulated using the Community Radiative Transfer Model (CRTM; Chen et al. 2008; Ding et al. 2011). Comparisons were made for all channels of all instruments making BT observations listed in the introduction [and in Table 4 of Boukabara et al. (2016)], but here detailed comparisons are presented only for three channels of the AMSU-A and three channels of the AIRS. These selected channels are described below and provide comparisons that include sensitivity to atmospheric temperature and moisture and to the surface, and are broadly representative of a large fraction of the set of all comparisons.

In the figures of this section, the panels are ordered by channel number from left to right. The reader is cautioned that, within a figure, for the same variable (e.g., BT) the x and y axes or color scales often vary from panel to panel for different channels because of the wide variation in response with respect to frequency. The axis and color scales are yet again different for panels that plot increments. Observations are plotted in time sequence in maps, and where there is overlap only the newest observation is visible. In figures displaying both maps and histograms, the data samples displayed in the maps match the data samples used to calculate the histograms, with two exceptions for radiance observations. First, the histograms exclude locations over sea ice (ice), and over ice- and snow-covered land (snow). Second, side-by-side comparisons of cloud- or precipitation-affected BTs are expected to be poor, since the information contained in the G5NR is not sufficient to properly simulate the cloud- and precipitation-impacted BTs. This would require accurate specification of the vertical variation of the particle size distribution for each type of hydrometeor. Further, even at the start of the NR, the exact locations of cloud boundaries and precipitation areas may be different from reality. For BTs, the strategy used here is to restrict quantitative comparisons (e.g., histograms) to simulated clear-sky BTs by restricting the sample to BT locations that are considered clear sky by the GSI QC procedures, that is, with an estimated cloud fraction of less than 0.10. Of course, there may still be some signals from clouds and precipitation in the real BT observations. The simulated observations in all cases are calculated for clear-sky conditions; that is, there is no impact on the NR clouds or precipitation.

a. MW observations

AMSU-A is a 15-channel microwave instrument operating at a frequency range of 23.8–89 GHz (Rosenkranz 2001). AMSU-A channels 1 and 2 are very sensitive to

TABLE 1. Selected sensor channels for MW (AMSU-A, AMSU-B, MicroMAS-2) and IR (AIRS) instruments. The columns list the sensor name, channel number, IR subset channel index, frequency (cm^{-1} for IR channels and GHz for MW channels), and channel sensitivity. The channel subset indices are from the AIRS 281-channel subset. Channel sensitivity flags are for temperature (T), and moisture (Q); and for UT, MT, and LT, and LS. Any channel numbers used in the text or figures is the sensor channel number (n), not the subset channel index (i).

Sensor	n	i	Frequency	Sensitivity
AMSU-A	3		50.30	LT T
	8		55.50	UT-LS T
	9		57.29	UT-LS T
	10		57.29 ± 0.22	UT-LS T
	15		89.00	Surface
AMSU-B	3		183.31 ± 1	UT Q
	4		183.31 ± 3	MT Q
	5		183.31 ± 7	LT Q
MicroMAS-2	8		118.59	UT-LS T
	9		183.31 ± 1	MT Q
	10		183.31 ± 3	MT Q
	11		183.31 ± 7	MT Q
AIRS	221	94	712.73	UT T
	1252	161	1131.19	Surface
	1449	173	1330.97	Q , surface

atmospheric water vapor, and to the surface. Channel 15 is known as a surface channel because the measured radiances are mostly emitted by the surface. AMSU-A channels 3–14 are temperature sounding channels centered around the oxygen absorption band at 57 GHz. In Fig. 1 we present results from three representative AMSU-A channels—channel 3 (50.3 GHz), a lower-tropospheric (LT) channel that is also sensitive to the surface and therefore to the cloud; channel 8 (55.5 GHz), an upper-troposphere (UT), lower-stratosphere (LS) temperature sounding channel; and channel 15 (89 GHz), the surface channel—to demonstrate that the observation simulation in the MW is properly handling variations in temperature and surface properties. Table 1 lists the channels that are the main examples in this study.

In Fig. 1, and all subsequent plots, except where noted, the data sample corresponds to the NR initial day. Channel 8 shows the UT-LS is cold over the winter pole but warm over the summer pole. Southwest of Australia, there is a noticeable warm anomaly with respect to the zonal average. Channel 8 is virtually unaffected by the surface and the $O - P$ increments are small everywhere. Even the pattern of cold striping at the swath edges in the tropics as a result of increased incidence angle is correctly matched by the simulated observations. Climatologically, Gelaro et al. (2015) report that the tropical tropopause temperature is warmer in the G5NR compared to reanalyses for all months, but this is not present in this initial day

comparison. Channel 15, the surface channel, depicts the pole-to-equator and land-versus-ocean temperature gradients. While O and P agree visually, there are many differences on the order of 10 K. These differences correspond to precipitation over the ocean, which is present in reality but not included in the observation simulation. Differences over land correspond to differences in emissivities and surface temperatures between reality and what is assumed in the forward problem, especially over the deserts. Differences over snow and ice, which are very noticeable in the polar regions, occur because our knowledge of snow and ice emissivities is imperfect.

Note that the precipitation signal changes sign from cool near the equator to warm at high latitudes because the meridional gradients of sea surface temperature crosses that of the effective radiative temperature of precipitation. As a result of the low emissivity of the ocean in the MW, the surface brightness temperature varies from roughly 270 K at the equator to 200 K at high latitudes (Fig. 1h) at this frequency. Since cloud tops are higher and colder at the equator than in high latitudes, the precipitation signal is cool at the equator and increases poleward. Maps of channel 3, the LT temperature channel, are similar in character to channel 15, the surface channel, but with reduced amplitude. However, the amplitude of the snow and ice emissivity signals are as large in this channel as in the surface channel because at these high and drier latitudes, the channel weighting functions shift to lower elevation and become more influenced by the surface.

The histograms for land and ocean in Fig. 1 quantify some of the discussion of the maps. For channel 8, the UT-LS temperature channel, the increments are $O(1)$ K or less. (Basic statistics for the distributions displayed in the histograms are given in Table 2.) Differences between land and ocean are due to the differences in distribution of land with latitude. In channel 15, the surface channel, we see that ocean BT varies from 200 to 270 K (as noted earlier), while land is mostly confined to 260–300 K. However, the real observations have a noticeable number of locations over land with a BT as low as 240 K. Increments for this channel are mostly in the range of -10 to $+10$ K, with some land differences reaching -20 K (with reality appearing colder than simulation) as a result of surface emissivity impacts over deserts and some ocean differences as much as $+15$ K (with reality appearing warmer than simulation) as a result of the ice and precipitation effects at high latitudes.

Figure 2 shows the mean and standard deviation of $O - P$ for each AMSU-A channel. The largest differences occur for channels 1–3 and 15. This occurs

TABLE 2. Summary statistics for BT, bending angle (α), and conventional observations comparisons discussed in the text and corresponding to histograms plotted in the figures. For variable x , for the given platform and channel n this table reports the sample size (N), mean (μ), and standard deviation (σ) for the global domain or for land and ocean for O , P , and $\Delta = O - P$. Conventional observations (CONOBS) include surface pressure (p_s), temperature (T), specific humidity (q), wind components (u, v), and wind speed (V). Because of differences between surface pressure in reality and in the G5NR, there are small differences in N between O and P for the CONOBS domains defined by pressure. The value reported is $N(O)$.

x	Platform	n	Domain	N	$\mu_x(O)$	$\mu_x(P)$	$\mu_x(\Delta)$	$\sigma_x(O)$	$\sigma_x(P)$	$\sigma_x(\Delta)$	
BT (K)	AMSU-A	3	Land	5388	277.39	278.87	-1.48	7.59	8.95	5.67	
			Ocean	15 932	232.58	231.07	1.50	9.29	9.56	2.04	
		8	Land	5388	216.89	217.03	-0.14	3.33	3.13	0.32	
			Ocean	15 932	216.38	216.57	-0.19	3.71	3.56	0.33	
		15	Land	5388	279.15	280.01	-0.86	8.83	9.55	7.16	
			Ocean	15 932	234.35	232.41	1.95	18.56	19.72	5.28	
	AIRS	221	Land	2676	252.73	252.74	-0.01	4.71	4.56	1.44	
			Ocean	9560	253.17	252.97	0.20	5.08	4.95	1.24	
		1252	Land	2676	296.83	296.15	0.68	12.00	11.56	5.00	
			Ocean	9560	290.68	291.36	-0.68	7.46	7.08	1.65	
		1449	Land	2676	282.75	281.63	1.12	8.22	6.98	3.47	
			Ocean	9560	278.97	278.26	0.71	5.02	4.50	2.04	
		AMSU-B/ MicroMAS-2	3/9	Land	2141	249.50	248.33	1.17	6.47	6.43	4.45
				Ocean	14 974	249.89	248.53	1.36	7.41	6.94	5.13
			4/10	Land	2141	262.58	262.30	0.28	6.53	6.58	3.89
				Ocean	14 974	262.20	261.87	0.33	8.07	6.98	5.13
5/11	Land		2141	275.62	274.34	1.28	7.13	6.76	3.25		
	Ocean		14 974	272.21	271.15	1.06	9.47	7.77	5.43		
α (μ rad)	CHAMP		0-10 km	1746	12 131	12 328	-197	4344	4800	314	
			10-20 km	2231	3854	3865	-11	1636	1633	94	
			20-60 km	8862	249	250	-1	371	372	50	
p_s (hPa)	CONOBS		Land	39 107	971.6	971.4	0.2	55.6	55.3	3.5	
			Ocean	10 365	1008.5	1008.5	0.0	10.1	10.1	1.8	
T (K)	CONOBS		0-300 hPa	28 252	218.96	218.99	-0.03	7.98	7.86	1.98	
			300-1000 hPa	55 519	276.13	275.70	0.43	17.27	16.91	2.24	
			1000-1100 hPa	22 983	288.19	287.74	0.45	8.95	8.12	2.50	
q (g kg^{-1})	CONOBS		300-950 hPa	16 958	4.42	4.49	-0.07	3.59	3.49	1.42	
			950-1000 hPa	18 696	6.82	7.17	-0.35	3.37	3.14	1.46	
			1000-1100 hPa	20 518	8.93	9.10	-0.17	4.69	4.45	1.50	
u (m s^{-1})	CONOBS		200-500 hPa	73 151	14.55	14.63	-0.08	13.67	13.51	4.78	
			500-850 hPa	44 755	5.03	4.84	0.19	7.28	6.97	3.56	
			850-1100 hPa	116 259	0.38	0.27	0.11	5.35	5.10	2.55	
v (m s^{-1})	CONOBS		200-500 hPa	73 151	-0.95	-1.19	0.24	14.74	14.83	4.99	
			500-850 hPa	44 755	0.51	0.30	0.21	9.08	8.67	3.70	
			850-1100 hPa	116 259	-0.28	-0.42	0.14	5.37	4.98	2.96	
V (m s^{-1})	CONOBS		200-500 hPa	73 151	21.88	22.06	-0.18	11.75	11.46	4.87	
			500-850 hPa	44 755	10.79	10.35	0.44	6.66	6.32	3.64	
			850-1100 hPa	116 259	6.11	5.89	0.22	4.52	4.05	2.80	

because these are channels sensitive to the surface, which allows discrepancies between real and modeled emissivities to have a large impact on $O - P$. In general, for the AMSU-A channels 4-14, which are sensitive to temperature above the midtroposphere (MT), increments are very small. For LT temperature channels 1 and 2, the sensitivity to the surface, and hence the increments, is larger than for channel 15, the surface channel, with standard deviations for land >10 K, and for ocean slightly <10 K. Channel 1, at the 23-GHz water vapor resonance, is used to measure integrated water vapor, and in dry conditions is essentially a window channel. Similarly, channel 2 is the 31.4-GHz

water vapor continuum channel and the atmosphere is also very transparent at this frequency in dry conditions. Therefore, under dry conditions we expect large differences in these two channels between real and perfect simulated observations as a result of differences in emissivity in reality and in the CRTM. This is especially the case over land, but it is also a factor over the ocean, where emissivity depends strongly on surface wind speed at these frequencies. There is a noticeable difference in both O and P in the mid- and upper stratosphere (channels 11-14) between land and ocean, again as a result of the distribution of land with latitude.

b. IR observations

AIRS is a high-spectral-resolution, multiaperture, echelle-grating spectrometer with 2378 channels in the thermal infrared ($3.7\text{--}15.4\ \mu\text{m}$ or $2665\text{--}649\ \text{cm}^{-1}$) with a resolution of $13.5\ \text{km}$ at nadir (Aumann et al. 2003). AIRS channels are sensitive to the surface temperature, surface emissivity, atmospheric temperature, humidity, and other atmospheric constituents, including carbon dioxide, ozone, methane, and carbon monoxide. The operational DA system uses only about half or less (currently 120) of the 281-channel subset (Suskind et al. 2003; Le Marshall et al. 2006) of the AIRS channels that are made available to NWP centers. For example, since there is no provision to use atmospheric trace gases as CRTM inputs, those channels sensitive to these atmospheric constituents are not well simulated by CRTM and therefore are not used for NWP.

Figure 3 shows comparison results for three representative IR channels: channel 221 ($712.73\ \text{cm}^{-1}$), a far IR UT-LS CO_2 temperature sounding channel; channel 1252 ($1131.19\ \text{cm}^{-1}$), a window channel; and channel 1449 ($1330.97\ \text{cm}^{-1}$), an MT water vapor sounding channel. Channel 221 shows a fairly uniform pole-to-equator temperature gradient that is symmetric with respect to the equator, except that Antarctica, which is in darkness at this date, is considerably cooler than the Arctic. High cold clouds affect O in this channel (e.g., south of Mexico in Fig. 3a) and this is seen clearly in the negatively skewed distribution in the map of $O - P$ (Fig. 3c). Channel 1252 as a window channel is more affected by clouds and responds strongly to the high desert surface temperatures in Africa and South Asia. The $O - P$ differences are generally negative as a result of cloud effects, but there are some areas where $O - P$ reaches $+10\ \text{K}$. Over Antarctica, this may be due to differences in emissivity, while over the continental areas, daytime heating may be underestimated by G5NR. Maps of channels 1252 and 1449 are similar in many respects. Both channels are sensitive to clouds, and in dry conditions both are sensitive to the surface. But channel 1449 is also sensitive to humidity, and this is clearly seen in P along the ITCZ in the western equatorial Pacific.

In the histograms for AIRS (Fig. 3), for channel 221, differences between land and ocean again reflect the different distribution of land with respect to latitude. In this channel, residual cloud effects result in a distribution of $O - P$ that has a standard deviation of $<1.5\ \text{K}$ and that is slightly skewed negatively with a mode at $-1\ \text{K}$ and a longer fatter negative tail. For channel 1252 there is a wide range of land surface temperatures, with corresponding BTs ranging from 265 to 325 K. Ocean BT also reaches lows of 265 K, but

it has a strong high cutoff at 300 K. Again, residual cloud effects result in negatively skewed distributions for $O - P$, with standard deviations of approximately 1.5 K for ocean but 5 K for land. For channel 1449, ocean and land histograms are similar, but with the ocean cooler and a long positive tail for land. For this channel $O - P$ is not negatively skewed and has standard deviations on the order of 3.5 and 2.0 K, over land and ocean, respectively.

Figure 4 plots the mean and standard deviation of AIRS for a 2-week period in August for land and ocean versus frequency for channels used in the DA system in channel subsets for the temperature sounding channels (top), the window and ozone channels (middle), and the water vapor channels (bottom). The overall pattern for the mean and standard deviation of O and P agree very well. There is generally a larger, often 2 times larger, standard deviation channel by channel over land compared to ocean in both O and P .

5. Validation of simulated GNSS RO observations

GNSS RO observations provide highly accurate information related to atmospheric thermodynamic structure with global coverage that is not affected by clouds or precipitation (Kursinski et al. 1997). In NWP systems GNSS RO observations provide critical anchoring data for the variational radiance bias correction procedures (e.g., Cucurull et al. 2014). While profiles of atmospheric temperature and moisture can be retrieved from GNSS RO observations, NWP centers typically assimilate a more fundamental observation—refractivity, bending angle, or phase delay. In CGOP the bending angle is assimilated following Cucurull et al. (2013). The most important source of GNSS RO observations is the *Constellation Observing System for Meteorology, Ionosphere and Climate-1* (COSMIC-1; Anthes et al. 2008). However, COSMIC-1 was launched in April 2006 and RO observations were not used operationally by NCEP until May 2007. Here we validate the simulation of bending angle observations from the Challenging Minisatellite Payload (CHAMP; Wickert et al. 2001). While not available in real time in 2005, CHAMP BUFR files were routinely collected by NCEP and were used in testing RO assimilation procedures by Cucurull et al. (2007).

Figure 5 plots the bending angle observations used by the DA system in the 10–20-km height range. Each occultation provides hundreds of observations, but these are thinned for DA. In terms of horizontal locations, the CHAMP RO data are relatively sparse, with only approximately 50 bending angle profiles during the 6-h window centered at 0000 UTC 16 May

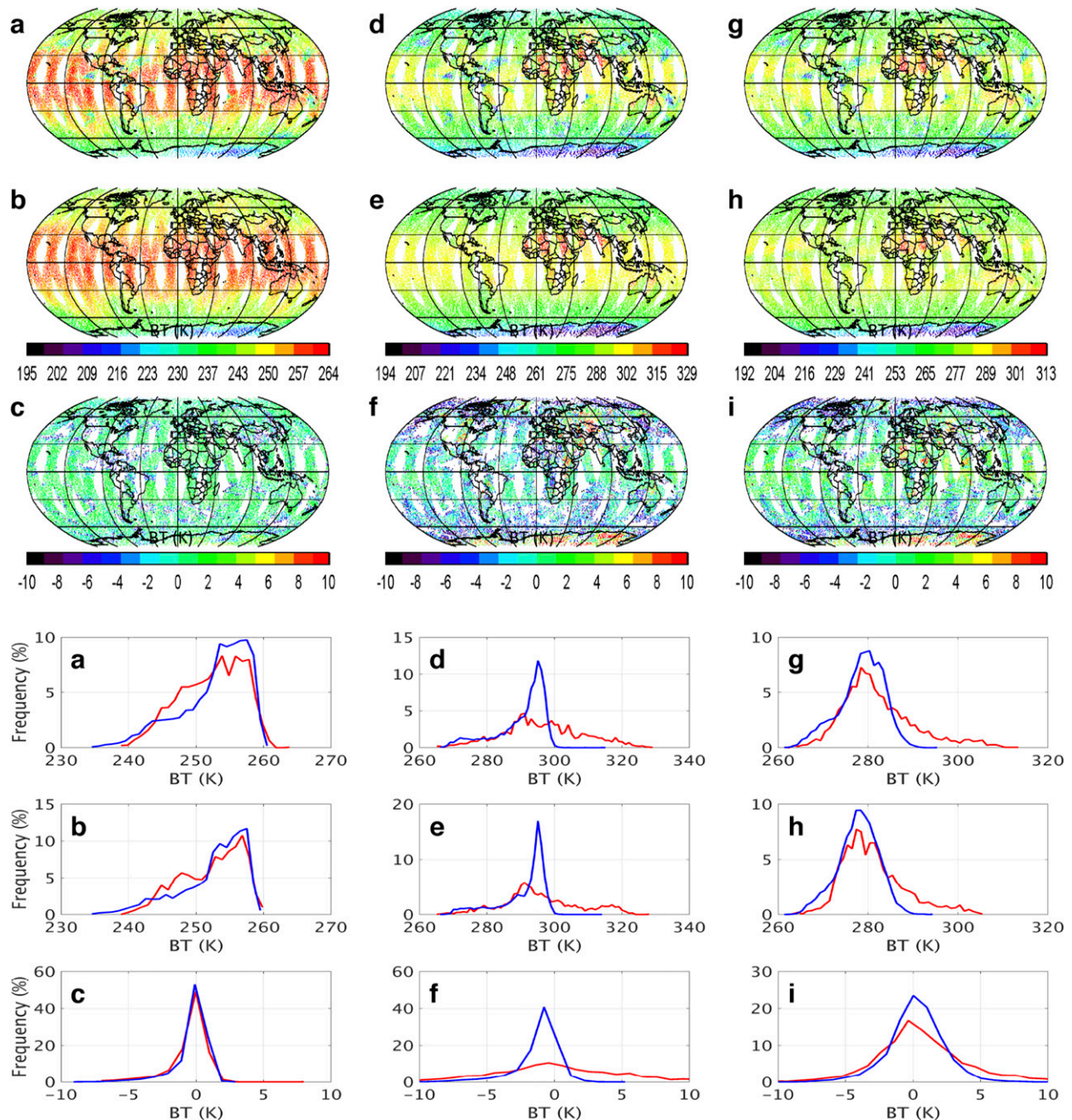


FIG. 3. As in Fig. 1, but for AIRS on *Aqua* BT channels (a)–(c) 221, (d)–(f) 1252, and (g)–(i) 1449.

2005. In the maps of Figs. 5a and 5b, there is good general agreement between O and P , respectively.

Distributions of bending angle for O and P are distinctively different for high, middle, and low elevations, but they are very similar for O and P (e.g., Figs. 5d and 5e, respectively). The $O - P$ values are all very small, with highly peaked and long-tailed distributions (e.g., Fig. 5f).

Figure 6 shows profiles of the mean and standard deviation of the bending angle. The bending angle increases exponentially with pressure and with specific humidity. As a result, so too do the standard deviations and $O - P$ statistics. The statistics for O and P match very well. To display the relative difference statistics, we also plot the mean and standard deviation of $O - P$ normalized by O in Fig. 6d. The relative difference

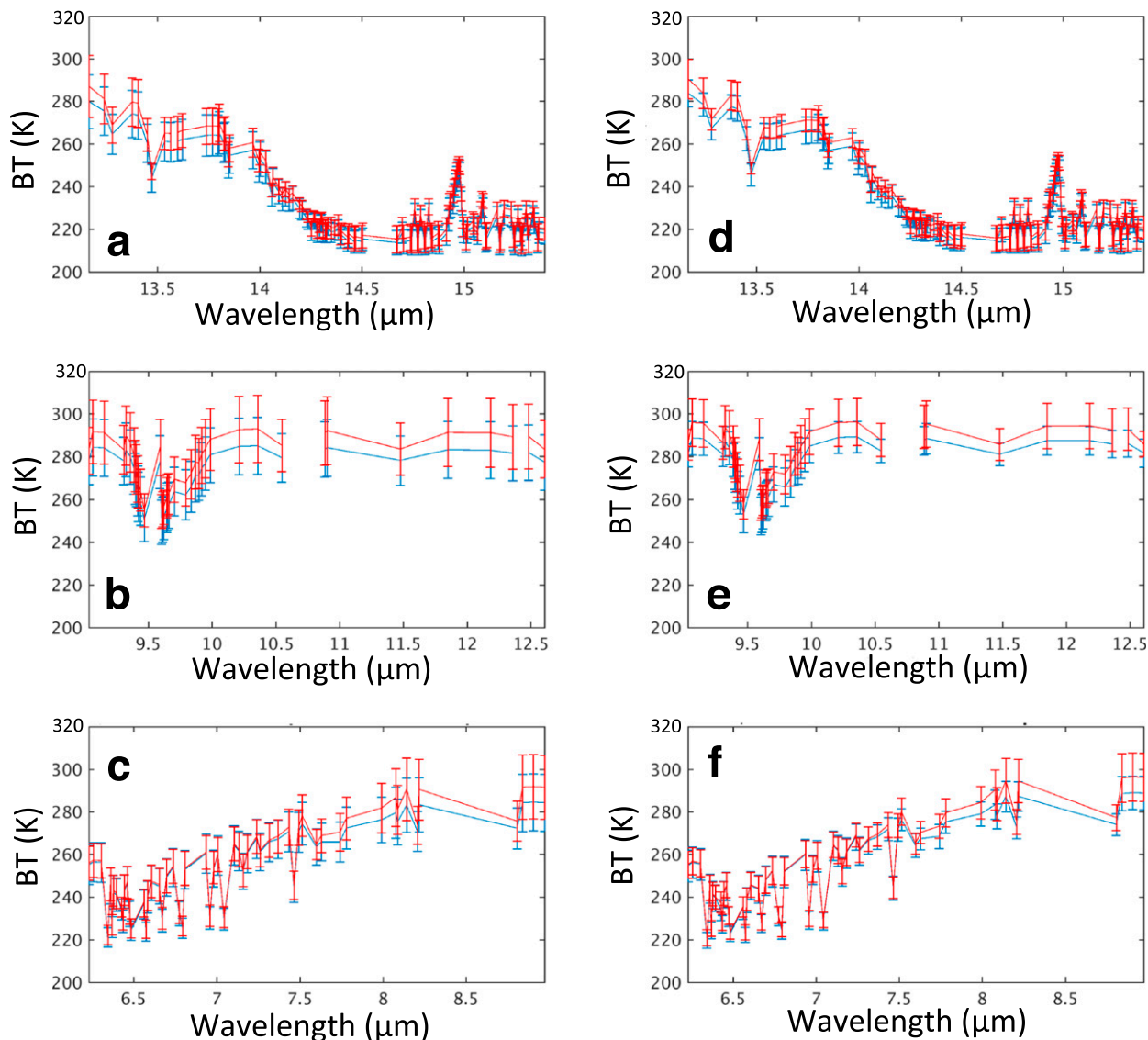


FIG. 4. Mean (lines) and standard deviation (± 1 standard deviation error bars) for AIRS on *Aqua* BT (K) vs frequency for the (a),(d) temperature sounding band, (b),(e) window and ozone bands, and (c),(f) water vapor bands for clear sky for land (red) and ocean (blue); (a)–(c) for *O* during 1–14 Aug 2014 and (d)–(f) for *P* during 1–14 Aug 2006.

statistics show that the simulation is biased high at both high and low elevations, and that relative errors are large near the surface below 10 km and very large above 40 km. These statistics are consistent with our understanding of the error structure of the real observations.

6. Validation of simulated conventional observations

Figure 7 shows that the real and perfect simulated radiosonde observations have nearly the same statistical (mean and variance) variation with pressure for

temperature, specific humidity, and wind components. As expected, temperature decreases with height through the troposphere and specific humidity and its variability decreases rapidly with height. Note that humidity is notoriously difficult to measure at high elevations and that currently humidity observations above 300 hPa are not assimilated. The *u* wind component is at a maximum at jet stream level (250 hPa) and both *u* and *v* wind variability are largest at this level.

Observations of surface pressure (Fig. 8) are influenced by synoptic features and topography and in general agree closely between reality and simulation. Even in the one 6-h period plotted in Fig. 8, there are more

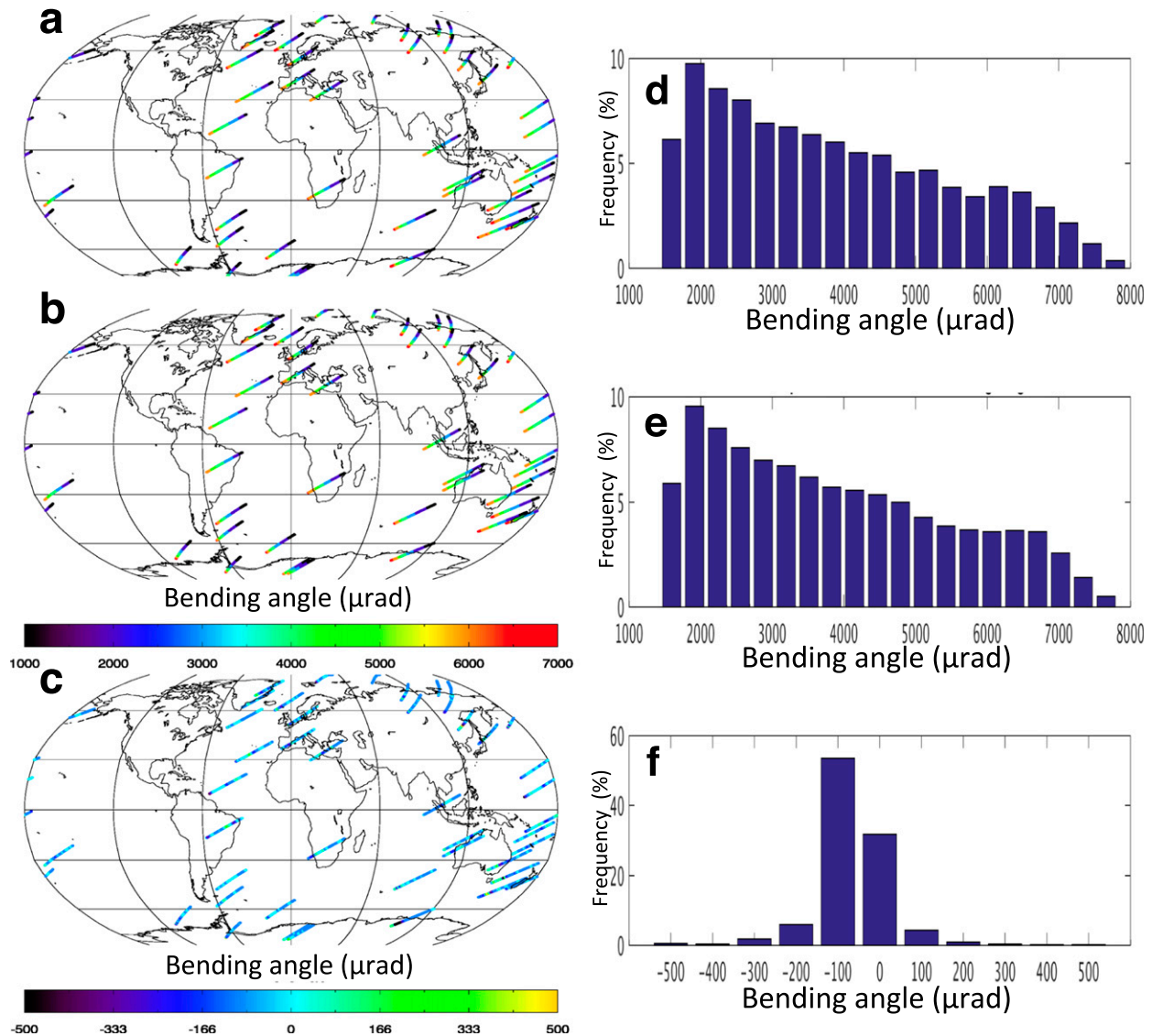


FIG. 5. (a)–(c) Maps and (d)–(f) histograms of CHAMP GNSS/RO bending angle observations (μrad) for the layer from 10 to 20 km, the 6-h window centered at 0000 UTC 16 May 2005, and globally: (a),(d) O , (b),(e) P , and (c),(f) $O - P$. Since the observations occur in a near-vertical profile, to avoid over plotting, observations are displaced toward the east-northeast by an amount proportional to $z - 10$ km.

than 50 000 observations from rawinsondes and land and ocean surface stations. Note that surface pressure is the pressure at the elevation of the observations in reality. The map of increments (Fig. 8c) shows that O pressures compared to P pressures are higher in South Africa, Turkey, Iran, and China, and that they are lower in the Rockies and Thailand. However, overall difference magnitudes are generally less than 2 hPa. Differences caused by the difference between topography at stations in reality, in the DA system, and in the NR, are minimized here because the simulated pressure observation is the G5NR pressure interpolated to the location and elevation of the real observation.

Maps and histograms were also studied for temperature, specific humidity, wind components, and wind speed. For each, the vertical domain was segmented into low, middle, and upper levels. In general, the spatial patterns of conventional observations are very similar for O and P and the distributions seen in the histograms for land and ocean are roughly similar (basic statistics for some of these distributions are given in Table 2). Specific findings are described in the following paragraphs.

For conventional temperature observations, the difference plots of $O - P$ show some areas of larger coherent difference in the troposphere. In the surface

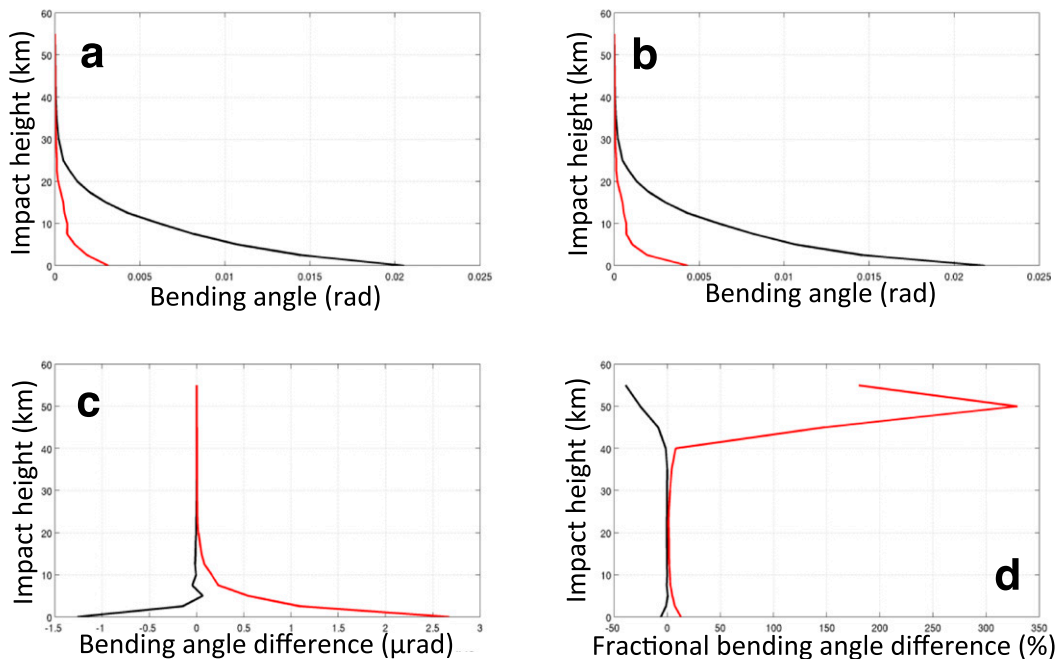


FIG. 6. Profiles of CHAMP GNSS/RO bending angle (μrad) mean (black) and standard deviation (red) for heights from 0 to 60 km, the NR initial day, and globally; for (a) O , (b) P , (c) $O - P$, and (d) the fractional difference [$100 \times (O - P)/O$; %].

and near-surface layer ($p > 1000$ hPa), O is greater than P by more than 2 K over land along the U.S. Gulf Coast. In the layer from 300 to 1000 hPa, there are a number of mountainous areas, such as the Appalachians, where O is greater than P by 2.5 K or so. The temperature difference ($O - P$) is distributed approximately normally with standard deviations increasing toward the surface.

For conventional specific humidity observations, the difference plots of $O - P$ show areas of larger coherent difference in the boundary layer. Notable areas of difference are in the high mountains of North America and Asia in the 300–950-hPa layer, in the United States east of the Mississippi River and the area of northern India in the foothills of the Himalayas in the 950–1000-hPa layer, and in the area of Siberia just north of Kazakhstan in the layer below 1000 hPa. Although divided into three layers, in all cases because humidity increases exponentially toward the surface, the difference plots, even for the 300–950-hPa layer, are dominated by the lowest observations in each layer. The $O - P$ histograms for specific humidity have standard deviations of approximately 1.5 g kg^{-1} and appear to be non-Gaussian with high central peaks and long tails.

For winds, note that the low-level winds include and are dominated by scatterometer observations. The plots of $O - P$ show many areas of smaller-scale coherent differences for both wind speed and wind

components. The histograms of wind speed show the typical Weibull distribution (Pavia and O'Brien 1986). The distribution of the differences is approximately normal for high-level wind speed, but it becomes increasingly peaked lower in the atmosphere. The histograms of $O - P$ for wind components are approximately normal with standard deviations on the order of 5, 4, and 3 m s^{-1} for high, middle, and low-level winds, respectively.

7. Validation of simulated BT observations from proposed sensors

Validation of simulated BT observations from proposed future sensors is much more difficult than the validation of existing observations described in the previous sections. First, the validation of existing observations compares forward model calculations and sensor observations that are both successfully used in operational DA. Second, for some proposed sensors, forward model calculations and/or the creation of comparison sensor observations may require entirely novel methods. Validation of simulated BT observations from proposed sensors is also critical for OSSEs. Confidence in the simulation of observations from proposed sensors is required in order that the OSSE results are reliable and actionable.

For validation of a proposed sensor, exact matchups of real and simulated observations may not be possible,

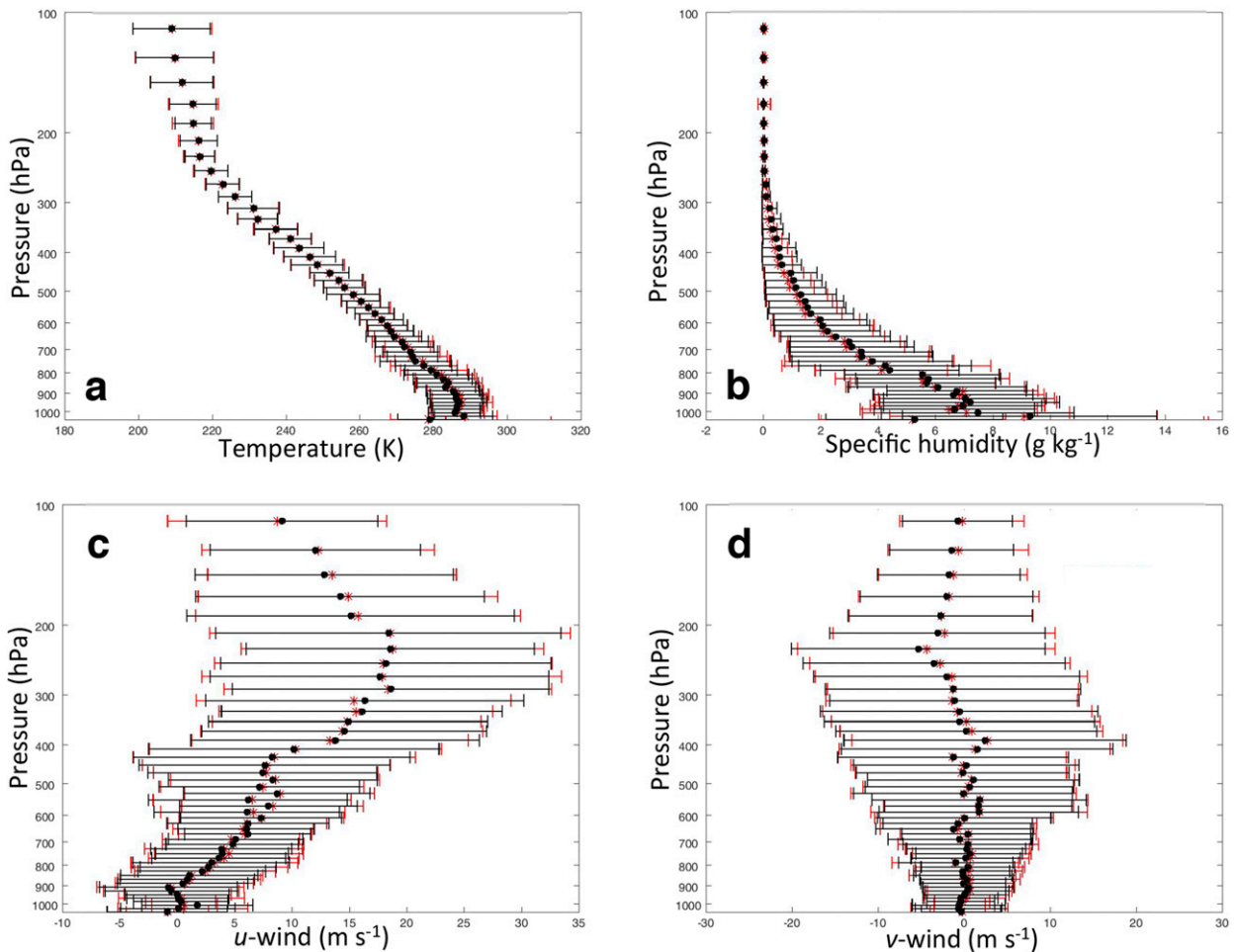


FIG. 7. Profiles of the mean (symbols) and standard deviation (± 1 standard deviation error bars) of all conventional observations (radiosonde, aircraft, and AMV observations) vs pressure (hPa, plotted on a log scale) for (a) temperature (K; $N = 106\,754$), (b) specific humidity (g kg^{-1} ; $N = 62\,920$), (c) u wind (m s^{-1} ; $N = 256\,656$), and (d) v wind (m s^{-1} ; $N = 256\,656$) for the 6-h window centered at 0000 UTC 16 May 2005 and globally; and for O (red, stars) and P (black, filled circles). Means and standard deviations are calculated for 20-hPa intervals from 1020 to 100 hPa and plotted at the midpoints of these intervals from 1010 to 110 hPa.

because the proposed sensor may have different characteristics than any existing sensor. For example, in this section we validate the simulation of BTs for the channels of the proposed Microsized Microwave Atmospheric Satellite-2 (MicroMAS-2) sensor (Blackwell 2017). MicroMAS-2 has channels both very similar to and very different from existing sensor channels. Examples of each are given in the next two sections.

In comparisons involving proposed sensors, we must usually account for differences in observing geometry and coverage. Since AMSU-A and AMSU-B are in LEO over the course of 24 h, most of Earth will be observed, except for some diamond-shaped regions around the equator. In contrast both the constellation of 12 MicroMAS-2 sensors in the Time-Resolved Observations

of Precipitation Structure and Storm Intensity with a Constellation of Smallsats (TROPICS) mission hosted on small satellites in low (30°) inclination orbits and a geostationary microwave satellite constellation (Geo-MW) will observe the entire low-latitude region frequently. AMSU-A and AMSU-B and MicroMAS-2 are cross-track scanners, so these sets of observations correspond to a wide range of scan angles (here, the term *sensor scan angle*, also sometimes referred to as *viewing angle*, is the angle between nadir and sensor telescope). However, for two different sensors, these angles generally do not match at locations on the surface because of the difference in orbits. For example, the scan angle is zero only at the five subsatellite points for geostationary Earth orbit (GEO) sensors, but it is zero all along the subsatellite track for the LEO sensors. Thus,

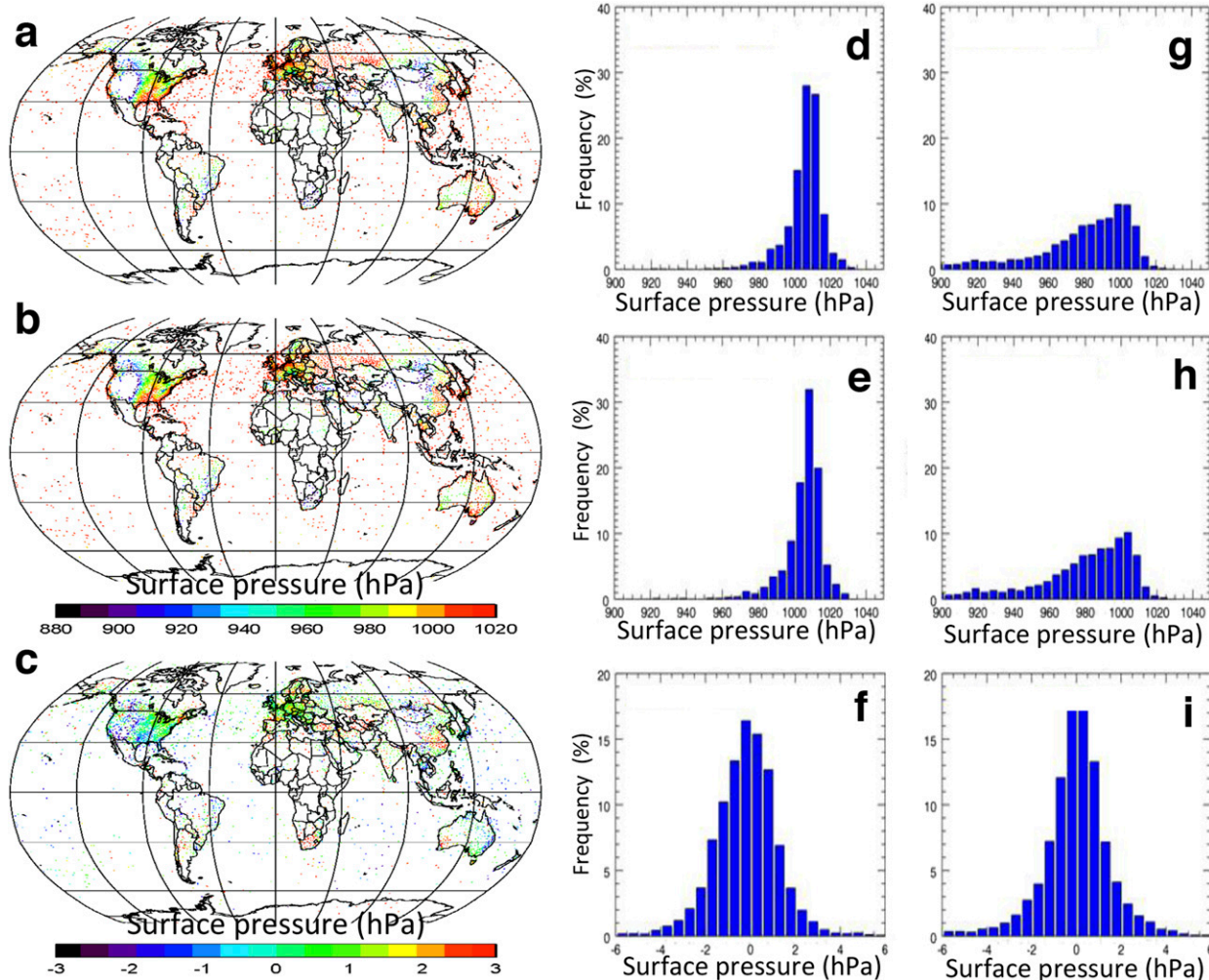


FIG. 8. (a)–(c) Maps and (d)–(i) histograms of surface pressure observations (hPa) for the 6-h window centered at 0000 UTC 16 May 2005 and globally: (a),(d),(g) *O*; (b),(e),(h) *P*; and (c),(f), (i) *O – P*. Maps are for the global sample, and histograms are for (d)–(f) ocean ($N = 12\,368$) and (g)–(i) land ($N = 32\,956$).

in the general case, validation is performed by statistical comparisons stratified by scan angle, latitude band, and land–ocean flag. For the purpose of the present MicroMAS-2 validation, we assume the observing geometry of MicroMAS-2 exactly matches that of AMSU-A or AMSU-B. In this case, for channels present in both MicroMAS-2 and AMSU, the validation applies to both (e.g., section 7a).

a. Microwave (183 GHz) water vapor channel observations

Figure 9 compares brightness temperature observed by AMSU-B and simulated for MicroMAS-2 for the NR initial day. From right to left, these channels sense the UT, MT, and LT humidity. The frequencies of these channels are all centered on the strong 183.31-GHz

water vapor spectral line. At the line center a moist atmosphere is opaque, and the farther away a channel is from the line center, the farther into the atmosphere it senses. In the maps of Fig. 9, we are essentially seeing the atmospheric temperature at one optical depth. The largest overall difference is that *P* (Fig. 9b) is much colder than *O* (Fig. 9a), indicating significantly more moisture in the G5NR UT and resulting in areas of higher *O – P* (Fig. 9c), particularly in the Indian Ocean. Since moisture is correlated vertically, we see similar spatial patterns going from left to right, but with the BT values increasing as the channels see deeper into the atmosphere and with the magnitude of the differences decreasing. In the LT, the overall moisture in the G5NR is similar to reality (Figs. 9g and 9h). In the MT, the G5NR is somewhat moister than reality (Figs. 9d and

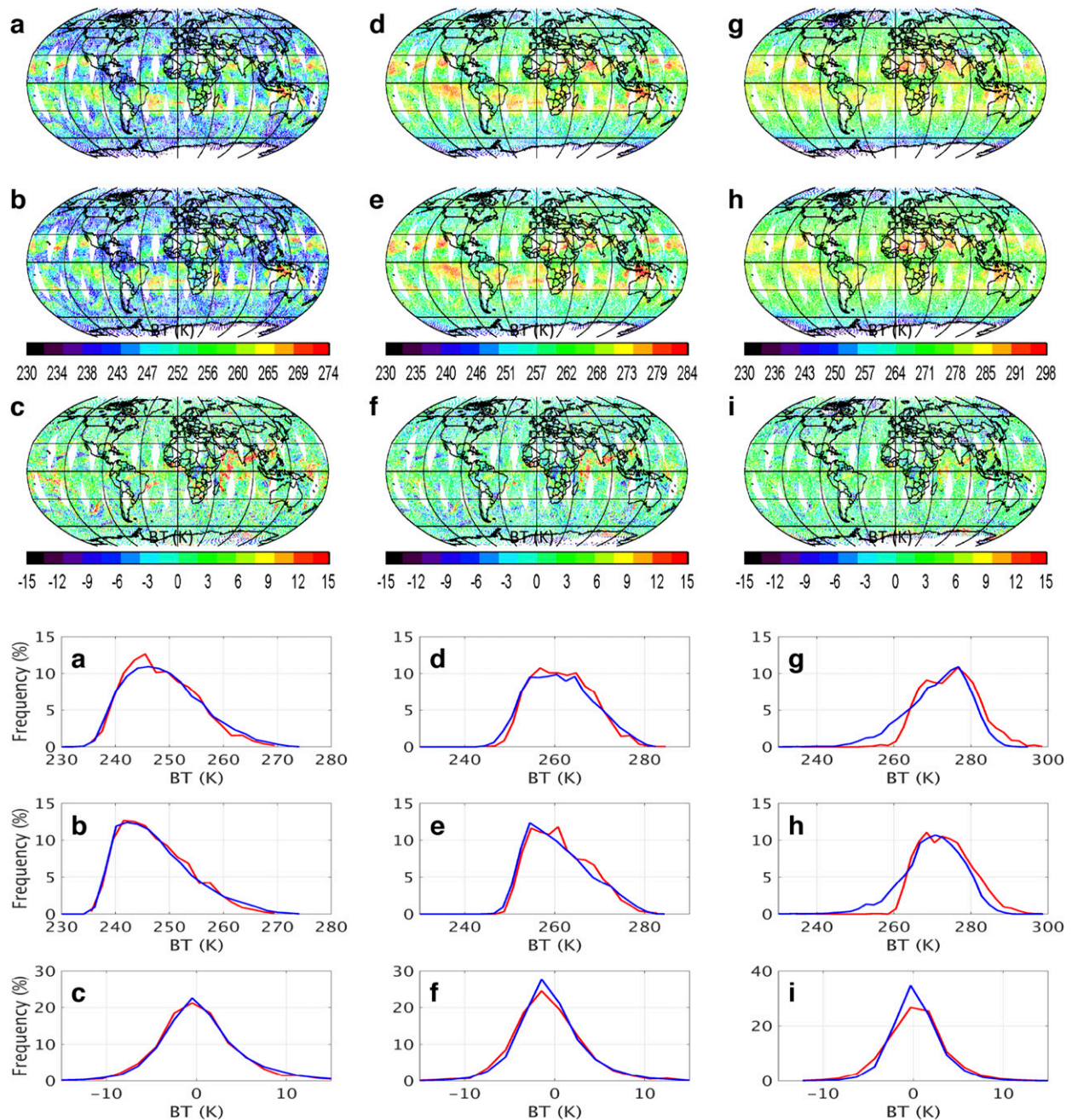


FIG. 9. (top) Maps and (bottom) histograms of MW BTs (K) for the (a)–(c) 183.31 ± 1 UT, (d)–(f) 183.31 ± 3 MT, and (g)–(i) 183.31 ± 7 LT humidity channels for the NR initial day, all-sky globally for the maps, and clear sky for land (red) and ocean (blue) for the histograms; and in both top and bottom for (a),(d),(g) AMSU-B on *NOAA-15* channels 3–5 *O*; (b),(e),(h) MicroMAS-2 channels 9–11 *P*; and (c),(f),(i) *O* – *P*.

9e). The spatial patterns match in the NR and reality, but there are substantial small-scale differences as might be expected for a short-term forecast of humidity.

In Fig. 9, the shapes of the histograms are somewhat similar, but the simulated MicroMAS-2 BTs are skewed to

lower values compared to the real AMSU-B BTs, especially for the UT (cf. Figs. 9a and 9b). Climatologically, Gelaro et al. (2015) report G5NR specific humidities are high above 700 hPa, and this is seen even in this initial day comparison.

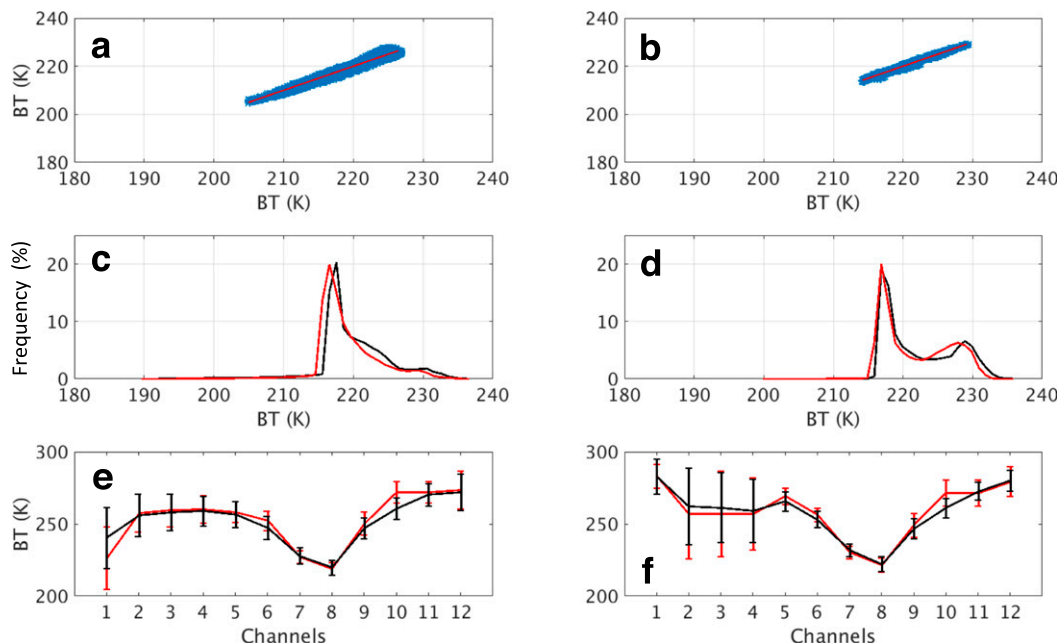


FIG. 10. (a),(b) Scatterplots of the predicted (x axis) vs the simulated (y axis) 118.59-GHz BT (K) for the training sample used in the regression. (c),(d) Histograms of the predicted (red) and simulated (black) 118.59-GHz BT (K). (e),(f) Mean (lines) and standard deviation (± 1 standard deviation error bars) of the perfect simulated observations (black) and real (channels 1, 9–11) or predicted (channels 2–8, 12) observations (red) vs MicroMAS-2 channel number. Scatterplots, histograms, and channel statistics are for (a),(c),(e) ocean and (b),(d),(f) land. Histograms and statistics are for clear sky for a 1-month period (August 2006 in the NR for the simulated observations and August 2014 in reality for the real or predicted observations).

b. Microwave (118 GHz) temperature sounding channel observations

Proposed future microwave sounding sensors such as MicroMAS-2 (Blackwell 2017) and GeoStorm³ (Lambriksen et al. 2017) exploit frequencies around the 118-GHz oxygen absorption band for temperature sounding. To date, sensor designers have preferred the 60-GHz oxygen band for temperature sounding, but the 118-GHz band is favored for GEO or for small satellites, since a smaller antenna may be used. To validate these channels, multiple linear regression making use of one or more of the 60-GHz channels is used to predict the BT of 118-GHz channels. For example, MicroMAS-2 channel 8 with a frequency of 118.59 GHz was regressed on AMSU-A channels 9 and 10 with frequencies of 57.29 and 57.29 ± 0.22 GHz. These two AMSU channels have weighting function peaks that bracket the peak of MicroMAS-2 channel 8. The training sample is perfect

simulated observations for all three channels at the AMSU-A locations during the first 6 h of the NR initial day, and the regression was determined separately for land and ocean. Figure 10 (top panels) illustrates the goodness of fit that was achieved ($R^2 = 0.996$ for land and 0.981 for ocean, where R is the correlation coefficient). Excluding the imagery channel (channel 12) over land, all values of R^2 were greater than 0.94 using either one or two predictors. Then, predicted 118.59-GHz BTs were determined using the regression relationships and real BTs from ATMS channels 10 and 11, which are identical to AMSU-A channels 9 and 10. Figure 10 (center panels) compares the distribution of both predicted (red) and simulated (black) 118.59-GHz BTs over a 1-month period. These histograms show a relatively good agreement between the predicted and simulated observations for both land and ocean. The predicted BTs are based on the August 2014 ATMS observations, and the simulated BTs are evaluated from the G5NR during August 2006 at the August 2014 ATMS observation locations. Thus, these two sets of observation agree in the distribution of the observations with respect to time of day, time of year, and geographic location. Figure 10 (bottom panels) extends the statistical

³ GeoStorm is a particular mission implementation of the Geostationary Synthetic Thinned-Aperture Radiometer (GeoSTAR) design (Lambriksen et al. 2007).

comparison to all channels and shows generally good agreement between real and simulated BT statistics. For the LT and MT temperature channels (2–6) and for the water vapor channels (9–11), simulated BTs are cool over the ocean and warm over land compared to reality. The simulated and predicted means of the surface channel (channel 1) over land and the UT temperature channels (channels 7 and 8) agree well, but over the ocean, the surface channel's simulated BTs are much cooler. Simulated and predicted standard deviations are similar over the ocean but over land the simulated BTs are generally smaller.

8. Concluding remarks

As stated in [Boukabara et al. \(2016\)](#), rigorous OSSEs require a complex infrastructure. The CGOP provides all components of such an infrastructure for global NWP applications—the NR, the simulation of existing and proposed future observations, and the entire NWP system. As part of the CGOP release process, the individual system components and the entire OSSE system are being tested, validated, and calibrated. The CGOP is by design modular, and individual components can be replaced relatively easily. However, the complexity and sensitivity of the interactions between components means that the introduction of a new component may require extensive validation and calibration of the OSSE system as a whole, including multiple OSSE components. For example, a new, more realistic NR may allow changes to make the simulated observations more realistic, to require retuning of the simulated observation errors caused by differences in representativeness, and to require modifications to the forecast model to maintain realistic predictability properties.

In this study, the simulation of observations—a critical CGOP component—is validated first by comparison of error-free simulated observations for the NR initial day to the real observations for that period. Sample results of this validation are presented here for LEO MW and IR BT observations, for RO bending angle observations, and for various types of conventional observations. For sensors not operating at the start of the NR, a qualitative validation is obtained by comparing geographic and statistical characteristics of observations for such a sensor and either an existing similar sensor or by using an existing sensor to predict the observations of the proposed sensor. Sample results of this validation are presented here for proposed MW BT observations. The comparisons agree, with no significant unexplained bias, and to within the uncertainties caused by real observation errors, time

and space collocation differences, radiative transfer uncertainties, and representativeness differences (differences between the NR model and the real world). Overall, the CGOP simulations of error-free observations for conventional and satellite-based measurements (microwave, infrared, radio occultation, geostationary, and polar platforms) are found to be reasonably accurate and suitable as a starting point for creating realistic simulated observations for OSSEs.

The inconsistencies that are observed are explained by the following factors. First, the real clear-sky BTs have some residual effects of precipitation and clouds that were not detected by the GSI QC procedures. Second, the NR representation of the radiative characteristics of land surfaces differs from reality and differs from the representation in the CRTM forward radiative model used in the CGOP simulation of BTs. These inconsistencies are acceptable if cloud- and surface-affected radiances are not assimilated, as is the current practice. In the future, when DA systems assimilating all-sky radiances are included in OSSE systems, accurate simulation of precipitation- and cloud-affected radiances will be necessary.

The remaining CGOP validation and calibration activities will be the subject of a separate paper. These activities include checking the quality control yields, tuning the explicitly added errors so that the simulated observation innovation statistics match those observed in reality, comparing forecast skill in the OSSE system and in reality, and the tuning and validating the OSSE system as a whole. This last activity includes a demonstration that the impact of existing observing systems is similar in reality and in the OSSE.

Acknowledgments. Many scientists have contributed to the development of the CGOP over a period of more than a decade. They come from numerous laboratories and centers within NOAA and NASA, the JCSDA, and the NCEP and ECMWF operational prediction centers. We thank these scientists and their institutions. The authors particularly acknowledge the contributions from other members of the NOAA QOSAP OSSE team—Krishna Kumar, Zhenglong Li, Michiko Masutani, Isaac Moradi and Jack Woollen—and our NASA GMAO colleagues—Ron Gelaro, Bill Putman, and many others—who produced the G5NR at NASA GMAO. In addition, we thank the team that developed and maintain the CRTM, including Mark Liu and the late Paul VanDelst. We thank the reviewers. Responding to their insightful suggestions improved this paper.

We gratefully acknowledge support for this work provided by NOAA—directly, under the Disaster Relief

Appropriations Act of 2013 (H.R. 152), under the Scientific and Technical Services II Federal Contracting Vehicle (DOCDG133E12CQ0020), and under the auspices of the cooperative institutes given in the author affiliations on the title page through Cooperative Agreements NA14NES4320003 for CICS, NA15OAR4320064 for CIMAS, NA14OAR4320125 for CIRA, and NA15OAR4320137 for CIRES.

APPENDIX

Definition of Acronyms

TABLE A1. Definition of acronyms.

Acronym	Definition
AIRS	Atmospheric Infrared Sounder
AMSU	Advanced Microwave Sounding Unit
AMV	Atmospheric motion vector
<i>Aqua</i>	NASA Earth science satellite mission focused on Earth's water cycle
ATMS	Advanced Technology Microwave Sounder
BT	Brightness temperature
BUFR	Binary Universal Form for the Representation of Meteorological Data
CGOP	Community Global OSSE Package
CHAMP	Challenging Minisatellite Payload
CICS	Cooperative Institute for Climate and Satellites (College Park, Maryland)
CIMAS	Cooperative Institute for Marine and Atmospheric Studies (Miami, Florida)
CIRA	Cooperative Institute for Research in the Atmosphere (Fort Collins, Colorado)
CIRES	Cooperative Institute for Research in Environmental Sciences (Boulder, Colorado)
CONOBS	Conventional observations
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
CrIS	Cross-Track Infrared Sounder
CRTM	Community Radiative Transfer Model
DA	Data assimilation
ECMWF	European Centre for Medium-Range Weather Forecasts
G5NR	GEOS-5 nature run [GMAO 7-km (1/16° × 1/16°)-resolution NR]
GEO	Geostationary Earth orbit
Geo-MW	Geostationary microwave satellite constellation
GEOS-5	Goddard Earth Observing System Model, version 5
GeoSTAR	Geostationary Synthetic Thinned-Aperture Radiometer
GeoStorm	Geostationary severe-storm observatory
GHz	Gigahertz
GMAO	Global Modeling and Assimilation Office
GNSS	Global Navigation Satellite System
GSI	Gridpoint Statistical Interpolation analysis system

TABLE A1. (Continued)

Acronym	Definition
HIRS	High-Resolution Infrared Sounder
IASI	Infrared Atmospheric Sounding Interferometer
IR	Infrared
ITCZ	Intertropical convergence zone
JCSDA	Joint Center for Satellite Data Assimilation
LEO	Low-Earth orbit
LS	Lower stratosphere
LT	Lower troposphere
MicroMAS	Microsized Microwave Atmospheric Satellite
MHS	Microwave Humidity Sounder
MT	Midtroposphere
MW	Microwave
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NESDIS	National Environmental Satellite, Data, and Information Service
NOAA	National Oceanic and Atmospheric Administration
NR	Nature run
NWP	Numerical weather prediction
OAR	Office of Oceanic and Atmospheric Research
OSE	Observing system experiments
OSSE	Observing system simulation experiment
PREPBUFR	Quality-controlled data in BUFR format
QC	Quality control
QOSAP	Quantitative Observing System Assessment Program
RADSTAT	Radiance statistics and diagnostics
RO	Radio occultation
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SNDR	Sounder
SSMIS	Special Sensor Microwave Imager/Sounder (NOAA)
STAR	Center for Satellite Applications and Research
TROPICS	Time-Resolved Observations of Precipitation Structure and Storm Intensity with a Constellation of Smallsats
UT	Upper troposphere
UTC	Coordinated universal time

Acronyms in Table A1 used as units (e.g., UTC) and proper names (e.g., names of specific institutions) are not expanded in the text.

REFERENCES

Anthes, R. A., and Coauthors, 2008: The COSMIC/FORMOSAT-3 mission: Early results. *Bull. Amer. Meteor. Soc.*, **89**, 313–333, <https://doi.org/10.1175/BAMS-89-3-313>.
 Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan*, **75**, 111–130, https://doi.org/10.2151/jmsj1965.75.1B_111.

- , and Coauthors, 2001: The effects of marine winds from scatterometer data on weather analysis and forecasting. *Bull. Amer. Meteor. Soc.*, **82**, 1965–1990, [https://doi.org/10.1175/1520-0477\(2001\)082<1965:TEOMWF>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<1965:TEOMWF>2.3.CO;2).
- , and Coauthors, 2015: Observing system simulation experiments (OSSEs) to evaluate the potential impact of an optical autocovariance wind lidar (OAWL) on numerical weather prediction. *J. Atmos. Oceanic Technol.*, **32**, 1593–1613, <https://doi.org/10.1175/JTECH-D-15-0038.1>.
- Aumann, H., and Coauthors, 2003: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products and processing systems. *IEEE Trans. Geosci. Remote Sens.*, **41**, 253–264, <https://doi.org/10.1109/TGRS.2002.808356>.
- Blackwell, W. J., 2017: New small satellite capabilities for microwave atmospheric remote sensing. *Seventh Conf. on Transition of Research to Operations/13th Annual Symp. on New Generation Operational Environmental Satellite Systems*, Seattle, WA, Amer. Meteor. Soc., J11.2, <https://ams.confex.com/ams/97Annual/webprogram/Paper314487.html>.
- Boukabara, S.-A., and Coauthors, 2016: Community Global Observing System Simulation Experiment (OSSE) Package (CGOP): Description and usage. *J. Atmos. Oceanic Technol.*, **33**, 1759–1777, <https://doi.org/10.1175/JTECH-D-16-0012.1>.
- Chen, Y., F. Weng, Y. Han, and Q. Liu, 2008: Validation of the Community Radiative Transfer Model by using CloudSat data. *J. Geophys. Res.*, **113**, D00A03, <https://doi.org/10.1029/2007JD009561>.
- Cucurull, L., J. C. Derber, R. Treadon, and R. J. Purser, 2007: Assimilation of global positioning system radio occultation observations into NCEP's global data assimilation system. *Mon. Wea. Rev.*, **135**, 3174–3193, <https://doi.org/10.1175/MWR3461.1>.
- , —, and R. J. Purser, 2013: A bending angle forward operator for global positioning system radio occultation measurements. *J. Geophys. Res. Atmos.*, **118**, 14–28, <https://doi.org/10.1029/2012JD017782>.
- , R. A. Anthes, and L.-L. Tsao, 2014: Radio occultation observations as anchor observations in numerical weather prediction models and associated reduction of bias corrections in microwave and infrared satellite observations. *J. Atmos. Oceanic Technol.*, **31**, 20–32, <https://doi.org/10.1175/JTECH-D-13-00059.1>.
- 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, <https://doi.org/10.1016/j.jqsrt.2010.11.009>.
- Errico, R. M., R. Yang, N. C. Privé, K.-S. Tai, R. Todling, M. E. Sienkiewicz, and J. Guo, 2013: Development and validation of observing-system simulation experiments at NASA's Global Modeling and Assimilation Office. *Quart. J. Roy. Meteor. Soc.*, **139**, 1162–1178, <https://doi.org/10.1002/qj.2027>.
- Gelaro, R., and Coauthors, 2015: Evaluation of the 7-km GEOS-5 nature run. Technical Report Series on Global Modeling and Data Assimilation, R. D. Koster, Ed., Vol. 36, NASA Tech. Memo. NASA/TM-2014-104606/Vol. 36, 285 pp., <http://gmao.gsfc.nasa.gov/pubs/docs/Gelaro736.pdf>.
- Hoffman, R. N., and R. Atlas, 2016: Future observing system simulation experiments. *Bull. Amer. Meteor. Soc.*, **97**, 1601–1616, <https://doi.org/10.1175/BAMS-D-15-00200.1>.
- Keyser, D., 2017: PREPBUFR processing at NCEP. NCEP, http://www.emc.ncep.noaa.gov/mmb/data_processing/prepbuf.doc/document.htm.
- Kleist, D. T., and K. Ide, 2015: 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, <https://doi.org/10.1175/MWR-D-13-00351.1>.
- Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, **102**, 23 429–23 465, <https://doi.org/10.1029/97JD01569>.
- Lambrigtsen, B., A. Tanner, T. Gaier, P. Kangaslahti, and S. Brown, 2007: Prototyping a new earth observing sensor—GeoSTAR. *Proc. 2007 IEEE Aerospace Conf.*, Big Sky, MT, IEEE, 8.0503, <https://doi.org/10.1109/AERO.2007.352690>.
- , T. Gaier, A. B. Tanner, P. Kangaslahti, B. Lim, and C. Ruf, 2017: A geostationary microwave sounder: What, why, how, when. *Seventh Conf. on Transition of Research to Operations*, Seattle, WA, Amer. Meteor. Soc., J15.5, <https://ams.confex.com/ams/97Annual/webprogram/Paper313976.html>.
- Le Marshall, J., J. Jung, T. Zapotocny, J. Derber, R. Treadon, S. Lord, M. Goldberg, and W. Wolf, 2006: The application of AIRS radiances in numerical weather prediction. *Aust. Meteor. Mag.*, **55**, 213–217, http://www.bom.gov.au/amm/docs/2006/lemarshall_hres.pdf.
- Pavia, E. G., and J. J. O'Brien, 1986: Weibull statistics of wind speed over the ocean. *J. Climate Appl. Meteor.*, **25**, 1324–1332, [https://doi.org/10.1175/1520-0450\(1986\)025<1324:WSOWSO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025<1324:WSOWSO>2.0.CO;2).
- Putman, W. M., A. Darmanov, A. da Silva, R. Gelaro, A. Molod, L. Ott, and M. J. Suarez, 2015: A 7-km non-hydrostatic global mesoscale simulation for OSSEs with the Goddard Earth Observing System model (GEOS-5). *19th Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, Phoenix, AZ, Amer. Meteor. Soc., 3.1, <https://ams.confex.com/ams/95Annual/webprogram/Paper260701.html>.
- Reale, O., D. Achuthavarier, M. Fuentes, W. M. Putman, and G. Partyka, 2017: Tropical cyclones in the 7-km NASA global nature run for use in observing system simulation experiments. *J. Atmos. Oceanic Technol.*, **34**, 73–100, <https://doi.org/10.1175/JTECH-D-16-0094.1>.
- Rosenkranz, P. W., 2001: Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements. *IEEE Trans. Geosci. Remote Sens.*, **39**, 2429–2435, <https://doi.org/10.1109/36.964979>.
- Susskind, J., C. D. Barnet, and J. M. Blaisdell, 2003: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB in the presence of clouds. *IEEE Trans. Geosci. Remote Sens.*, **41**, 390–409, <https://doi.org/10.1109/TGRS.2002.808236>.
- Wickert, J., and Coauthors, 2001: Atmosphere sounding by GPS radio occultation: First results from CHAMP. *Geophys. Res. Lett.*, **28**, 3263–3266, <https://doi.org/10.1029/2001GL013117>.