

Performance of Radiative Transfer Models in the Microwave Region

Key Points:

- Microwave radiative transfer models generally agree with each other and observations
- Spectroscopy models largely impact simulations for window and water vapor channels
- Sensor response function and emissivity show some important effect on calculated brightness temperatures

Correspondence to:

I. Moradi,
isaac.moradi@noaa.gov

Citation:

Moradi, I., Goldberg, M., Brath, M., Ferraro, R., Buehler, S. A., Saunders, R., & Sun, N. (2020). Performance of radiative transfer models in the microwave region. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD031831. <https://doi.org/10.1029/2019JD031831>

Received 15 OCT 2019

Accepted 5 MAR 2020

Accepted article online 6 MAR 2020

Author Contributions

Conceptualization: Isaac Moradi, Mitchell Goldberg, Ralph Ferraro

Data curation: Ninghai Sun

Funding Acquisition: Isaac Moradi

Methodology: Isaac Moradi, Mitchell Goldberg, Ralph Ferraro, Roger Saunders

Validation: Isaac Moradi, Mitchell Goldberg, Manfred Brath, Ralph Ferraro, Stefan A. Buehler, Roger Saunders, Ninghai Sun

Writing - Original Draft: Isaac Moradi, Mitchell Goldberg, Manfred Brath, Ralph Ferraro, Stefan A. Buehler, Roger Saunders

Formal Analysis: Isaac Moradi, Manfred Brath

Investigation: Isaac Moradi

Project Administration: Isaac Moradi





Resources: Mitchell Goldberg

Visualization: Isaac Moradi

Writing - review & editing: Isaac Moradi, Mitchell Goldberg, Manfred Brath, Ralph Ferraro, Stefan A. Buehler, Roger Saunders, Ninghai Sun

©2020. American Geophysical Union. All Rights Reserved.

This article has been contributed to by US Government employees and their work is in the public domain in the USA.

Isaac Moradi^{1,2,3} , Mitchell Goldberg⁴, Manfred Brath⁵ , Ralph Ferraro¹ , Stefan A. Buehler⁵ , Roger Saunders⁶, and Ninghai Sun¹

¹NOAA Center for Satellite Applications and Research, College Park, MD, USA, ²Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, ³NOAA Center for Earth System Sciences and Remote Sensing Technologies, New York, NY, USA, ⁴NOAA JPSS Program Science Office, Lanham, MD, USA, ⁵Meteorologisches Institut, Centrum für Erdsystem- und Nachhaltigkeitsforschung (CEN), Fachbereich Geowissenschaften, Universität Hamburg, Hamburg, Germany, ⁶Met Office, Exeter, UK

Abstract We compared two fast radiative transfer models, Community Radiative Transfer Model (CRTM) and Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV), with the LBL model Atmospheric Radiative Transfer Simulator (ARTS). We used the measurements from Advanced Technology Microwave Sounder (ATMS) and the Global Precipitation Measurement Microwave Imager (GMI) for evaluation of the radiative transfer models. The models in comparison with the observations and each other performed very well with a mean difference less than 0.5 K for the temperature sounding channels operating near the oxygen absorption band at 60 GHz. There was a difference of up to 1 K among the models as well as compared with the observations for humidity sounding channels operating around water vapor absorption line at 183 GHz. The mean difference between the simulations and observations was up to 6 K for surface sensitive channels. Water vapor and surface sensitive channels also showed to be more sensitive than the temperature sounding channels to the spectroscopy models used to calculate the absorption coefficients. There was a small difference, less than 0.1 K, between brightness temperatures calculated using traditional boxcar and actual Sensor or Spectral Response Functions, except for a difference of 0.25 K for ATMS Channel 6. Double difference technique showed about 1 K difference between water vapor channels from ATMS instruments onboard N20 and National Polar-orbiting Partnership (NPP). However, comparison of a new version of ATMS/NPP observations recently generated using an enhanced calibration algorithm with ATMS/N20 observations showed that the differences between the two instruments are less than 0.5 K after improving the ATMS/NPP calibration.

Plain Language Summary Radiative transfer models are used to simulate satellite observations from input atmospheric profiles and surface parameters. These models have a wide range of applications, including being used as forward model to assimilate satellite observations into numerical weather prediction models or for calibration and validation of satellite measurements. Radiative transfer models are subject to errors and uncertainties and therefore need to be evaluated. We compared three microwave radiative transfer models versus each other as well as satellite measurements. The results showed that the models were consistent with each other and with observations for microwave temperature sounding channels. However, the models were less consistent with each other and with the observations for the water vapor and window channels. The difference between the models and observations was affected by several factors such as spectral response functions, spectroscopy databases, and emissivity models.

1. Introduction

Radiative transfer (RT) models are extensively used in different fields of atmospheric remote sensing, including calibration and intercalibration of satellite observations, retrieving geophysical products from satellite observations, and as forward model to assimilate satellite observations into numerical weather prediction (NWP) models. The RT models can be classified into two groups based on the way absorption coefficients are calculated: (1) line-by-line (LBL) models where the monochromatic absorption coefficients are directly calculated from the spectroscopic database and (2) fast models where the absorption coefficients are computed using look-up tables and a set of predictors such as temperature and water vapor (Saunders et al., 1999).

The look-up tables are in fact the coefficients for the regressions between the predictors as independent variables and the absorption coefficients as dependent variables (Saunders et al., 1999). The LBL models are normally more accurate than the fast models because the coefficients are directly calculated from the inputs, but computationally much more expensive than the fast models.

Both LBL and fast RT models are subject to different inaccuracies that emerge from inaccuracy in the spectroscopic databases, RT inputs such as atmospheric profiles and surface emissivity, and assumptions in the RT theories (e.g., line shapes). However, RT models are an excellent tool for modeling remote sensing systems especially once evaluated against each other or independent data sets. For instance, the evaluated RT models can be used for monitoring the performance of the spaceborne instruments or calibration and intercalibration of satellite observations to support national and international activities such as the National Oceanic and Atmospheric Administration (NOAA) Integrated Calibration/Validation System (Zhou et al., 2016) and World Meteorological Organization Global Space-based Inter-Calibration System (Goldberg et al., 2011; Zhang et al., 2016). Additionally, since the RT models play a central role in the assimilation of satellite radiances, evaluation of the models and understanding the biases can potentially help to improve the assimilation of satellite observations and help the weather agencies to provide more accurate weather forecasts.

Garand et al. (2001) compared 19 infrared models and 10 microwave models for the sounding channels of Advanced Microwave Sounding Unit (AMSU) and High-Resolution Infrared Radiation Sounder (HIRS) and reported that IR LBL models agreed within 0.05 K, but fast IR models compared with the LBL models can only achieve an accuracy of 0.25 K for some (but not all) of the HIRS channels. Fast microwave models performed well against LBL model to which they were tuned, but on the other hand large differences were reported among the microwave LBL models. Buehler et al. (2006) compared Atmospheric Radiative Transfer Simulator (ARTS) and Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV) Version 7.0 using a single day of data from ERA-40 reanalysis and reported a bias of 0.014 K and standard deviation (STD) of 0.23 K for AMSU-B Channel 18 (183 ± 1 GHz). They reported large biases for dry regions as well as northern and southern latitudes. Melsheimer et al. (2005) compared several LBL models and reported a general agreement of better than 1% among the models, but a difference of up to 10% near the center of absorption lines. Similar evaluations have been conducted for the infrared region, and generally, RT models show better results for the infrared region than for the microwave region. For instance, an intercomparison of fast and LBL models for the infrared region is given in Saunders et al. (2007). The intercomparison was conducted for 14 line-by-line and fast parameterized infrared models and using 52 diverse atmospheric profiles. They reported a difference of 0.02 K among the models and 0.2 K compared with the Atmospheric Infrared Sounder radiances.

With the emergence of a new generation of microwave instruments, for example, ATMS and the Global Precipitation Measurement (GPM) Microwave Imager (GMI), and increasing application of fast microwave RT models in the NWP modeling, it is necessary to evaluate the accuracy of fast RT models compared with observations as well as LBL models. In this study, we compared two fast models (Community Radiative Transfer Model, CRTM, and RTTOV) that are internationally used for different applications. The ARTS was selected as the reference LBL model, because it was specifically developed for the microwave region. In addition to comparing the simulated brightness temperatures (Tbs) with the observations, we also evaluated the impact of different factors such as spectroscopy and emissivity models as well as Sensor or Spectral Response Functions (SRFs) on the differences between the RT models and observations. We screened out all the cloud-contaminated observations and only evaluated the forward models in clear-sky conditions over ice-free oceans. The fast models also include the Jacobian and adjoint codes which are crucial for data assimilation, but these capabilities were not evaluated in this study. Section 2 describes the RT models, section 3 discusses the satellite observations and atmospheric data set used in the study, section 4 discusses the results, and section 6 summarizes the work.

2. RT Models

We simulated the satellite Tbs using two fast RT models and a LBL model. Some aspects of these models relevant to the current study are discussed in this section. All these models require atmospheric variables such as air pressure, temperature, and water vapor, and surface variables such as surface temperature as well as other surface characteristics such as wind speed and direction to perform RT simulations. Although,

the microwave channels operating near 183 GHz are slightly sensitive to ozone concentration, the standard coefficient sets of CRTM and RTTOV do not include ozone transmittance coefficients. The impact due to exclusion of ozone can be up to 0.5 K for the channels near 183 GHz (John & Buehler, 2004).

2.1. CRTM

The CRTM is a fast RT model developed by the NOAA Center for Satellite Applications and Research and the Joint Center for Satellite Data Assimilation that is widely used in the United States to assimilate satellite radiances (Liu et al., 2008). CRTM is also widely used by remote sensing communities to simulate satellite radiances from atmospheric inputs for different purposes such as calibration of satellite instruments (Weng et al., 2013) or retrieving geophysical variables from satellite observations (Boukabara et al., 2011). CRTM is capable of simulating microwave, infrared, and visible radiances using atmospheric profiles of pressure, temperature, humidity, and other species such as ozone. CRTM also includes capabilities to simulate satellite cloudy radiances, but this work is focused on using only clear-sky observations. CRTM contains forward, tangent linear, K-matrix, and adjoint models. These capabilities are generally required in different parts of data assimilation systems for the assimilation of satellite microwave and infrared radiances.

CRTM requires as input both layer and level pressure values as well as the mean layer averages of temperature, water vapor, and other species (when necessary) to perform clear-sky RT calculations. The top pressure level is fixed at 0.01 hPa; therefore, the profiles will be filled by climatological values up to this level, if the input profile does not reach the top pressure. The surface emissivity over ocean is calculated using provided surface information such as wind speed and direction. CRTM currently includes two versions of FASTEM, Version 5.0 (Bormann et al., 2012) and Version 6.0 (Kazumori & English, 2015), for calculating sea surface emissivity. CRTM coefficients were trained using Rosenkranz spectroscopy for both ATMS instruments onboard National Polar-orbiting Partnership (NPP) and N20 (Rosenkranz, 1975, 1998). The package that is used to train the CRTM coefficients also includes the capability to use the Millimeter-wave Propagation Model (MPM) to generate the transmittance coefficients (Liebe, 1985, 1989; Liebe et al., 1992).

2.2. RTTOV

RTTOV is a popular fast RT model for passive satellite visible, infrared and microwave radiometers, spectrometers, and interferometers. RTTOV is the forward operator in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (Matricardi, 2009). RTTOV is developed by the European Organisation for the Exploitation of Meteorological Satellites NWP Satellite Application Facility and is available on request at no cost after registration at nwpsaf.eu website.

RTTOV requires a set of atmospheric and surface information as input to simulate satellite radiances. These input variables depend on the instrument type but generally for microwave radiometers include profiles of temperature, water vapor, and pressure as well as surface skin temperature. Although, the standard coefficient sets of RTTOV do not include the ozone transmittance coefficients, ozone is accounted for in the mixed gases. Depending on the surface type (land or water) other variables such as wind speed and direction (over water), vegetation type, and several other parameters are required for calculating the surface emissivity (Saunders et al., 2018). The absorption coefficients in RTTOV are trained using modified versions of the MPM 1989 spectroscopy for water vapor lines (Liebe, 1989) and the 1993 version of the Liebe (MPM-1993) for the oxygen and nitrogen lines (Liebe et al., 1992, 1993; Tretyakov et al., 2005). Some of the modifications to the MPM model include replacement of oxygen line parameters and the parameters of air-broadened half-width for the water vapor absorption lines at 22.235 and 183.31 GHz, adding ozone lines, and modifying the dry continua (Turner et al., 2019). RTTOV absorption look-up tables for recent instruments like ATMS include 54 fixed pressure levels with the bottom and top pressure levels at 1050 and 0.005 hPa, respectively. For more information about the LBL model used to train RTTOV for microwave instruments, the readers are referred to Turner et al. (2019) and Lupu et al. (2015).

2.3. ARTS

ARTS is a line-by-line RT model for thermal radiation originally developed to perform accurate simulations for the microwave region (Buehler et al., 2018; Eriksson et al., 2011). It can process fully polarized RT calculations with and without scattering in 1-D and 3-D atmospheres. ARTS is mainly developed at Universität Hamburg and Chalmers University. It is freely available and maintained as an open-source project (<http://www.radiativetransfer.org>). ARTS requires at least, profiles of pressure, temperature, and altitude on levels to perform simulations, however other inputs such as water vapor are also necessary to accurately simulate outgoing radiances for the microwave frequencies.

The simulations were done using ARTS Version 2.3.1249. For each sideband of each ATMS channel, 12 monochromatic Planck brightness temperatures were simulated and mapped to the specific channel using the boxcar SRFs. Gas absorption of molecular oxygen, ozone, molecular nitrogen, and water vapor was taken into account. For ozone and molecular nitrogen, the HITRAN 2012 database (Rothman et al., 2013) was used. In addition to the spectral line data from HITRAN, the MT_CKD continuum absorption model (Version 2.52) was used for the absorption of molecular nitrogen (Mlawer et al., 2012). The gas absorption of molecular oxygen was processed by using the full absorption model of Rosenkranz (1998) modified by the values from Tretyakov et al. (2005). We ran ARTS simulations with three different setups for water vapor spectroscopy, including (i) A23-RTV-HIT used HITRAN database and the MT_CKD model (Version 2.52) for the continuum absorption of water vapor, (ii) A23-RTV-MPM used the MPM93 water vapor absorption model (Liebe & Hufford, 1993; Liebe et al., 1993), and (iii) A23-RTV-PWR used the PWR98 water vapor absorption model (Rosenkranz, 1998). All the three ARTS setups used the ocean emissivity values calculated by RTTOV using FASTEM V6.0. We also used a specular reflection model to calculate the radiance reflected by the surface.

3. Atmospheric Data and Satellite Observations

This section discusses ERA-5 atmospheric and surface data set and also the satellite observation used in the comparison.

3.1. ERA-5 Reanalysis

ERA-5 is the fifth generation of ECMWF atmospheric reanalysis produced using 4DVar version of ECMWF's Integrated Forecasting System coupled with ocean waves and a land model. It combines a large amount of historical satellite and in situ observations into the first guess provided by the ECMWF forecast model (Hersbach et al., 2019). All the radiance observations in ERA-5 were assimilated using RTTOV (Version 11); therefore, Tbs simulated using ERA-5 profiles and RTTOV may yield better results than other RT models when compared with observations. ERA-5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables. ERA-5 covers the globe on a 30 km grid and resolve the atmosphere using 137 levels from the surface up to a height of 80 km.

We used 3-hourly ERA-5 reanalysis fields with a resolution of 30 km, 0.28°. The variables included wind speed and direction, temperature, water vapor, surface temperature, ozone, and land/sea mask.

3.2. Satellite Observations

ATMS is the newest generation of microwave instruments developed by NOAA and is currently flying on NOAA NPP and N20. ATMS consists of 22 channels operating at a frequency range of 23 to 190 GHz (see Table 1). ATMS Channel 1 operating near 23.8 GHz is located at the water vapor absorption line at 23 GHz and is sensitive to total column water vapor; the second channel is located at the edge of the water vapor absorption line at 31.4 GHz and is mostly sensitive to the surface characteristics and clouds. Channels 3 to 15 are temperature sounding channels operating at the oxygen absorption band around 60 GHz and measure air temperature from lower troposphere to upper stratosphere. The weighting functions for Channels 3 and 15 peak in lower troposphere and 2 hPa, respectively. Channel 16 operates at 88.2 GHz and is sensitive to the surface variables and clouds. Channel 17 operates at 165.5 GHz (near the edge of the water vapor absorption line at 183 GHz) and is sensitive to surface characteristics, lower tropospheric humidity, and ice clouds. Channels 18–22 operate at the water vapor absorption line at 183 GHz and are sensitive to lower to upper tropospheric relative humidity as well as ice clouds. For more information about the ATMS instrument, the readers are referred to Goldberg et al. (2013) and Moradi et al. (2015).

GMI onboard GPM core observatory consists of 13 channels operating at a frequency range of 10 to 90 GHz (see Table 1). Most GMI channels are very sensitive to the surface, excluding Channel 5 operating at 23 GHz, which is sensitive to total column water vapor and Channels 12 and 13 operating at 183 ± 3 and 183 ± 7 , which are sensitive to middle and lower tropospheric humidity, respectively. The first five channels of GMI can be used to quantify heavy and moderate precipitation, the next four channels indirectly measure snow and ice inside clouds, and the last four channels can be used to quantify light rain and snowfall as well as to measure relative humidity in lower and middle troposphere (Skofronick-Jackson et al., 2017).

3.3. Collocations

An important step in comparing model simulated and observed Tbs is to carefully collocate the observations with atmospheric profiles. The collocation process includes two factors, that is, time and distance.

Table 1
List of ATMS and GMI Channels (Goldberg et al., 2013; Skofronick-Jackson et al., 2017)

Channel	ATMS			GMI		
	Frequency (GHz)	FOV (deg)	Pol.	Frequency (GHz)	Pol.	FOV (deg)
1	23.8	5.2	QV	10.65	V	1.75
2	31.4	5.2	QV	10.65	H	1.75
3	50.3	2.2	QH	18.7	V	1
4	51.76	2.2	QH	18.7	H	1
5	52.8	2.2	QH	23.8	V	0.9
6	53.596 ± 0.115	2.2	QH	36.5	V	0.9
7	54.4	2.2	QH	36.5	H	0.9
8	54.94	2.2	QH	89	V	0.4
9	55.5	2.2	QH	89	H	0.4
10	57.290344	2.2	QH	166	V	0.4
11	57.290344 ± 0.217	2.2	QH	166	H	0.4
12	57.290344 ± 0.3222 ± 0.048	2.2	QH	183.31 ± 3	V	0.4
13	57.290344 ± 0.3222 ± 0.022	2.2	QH	183.31 ± 7	V	0.4
14	57.290344 ± 0.3222 ± 0.010	2.2	QH			
15	57.290344 ± 0.3222 ± 0.0045	2.2	QH			
16	88.2	2.2	QV			
17	165.5	1.1	QH			
18	183.31 ± 7	1.1	QH			
19	183.31 ± 4.5	1.1	QH			
20	183.31 ± 3	1.1	QH			
21	183.31 ± 1.8	1.1	QH			
22	183.31 ± 1	1.1	QH			

Note. QV and QH stand for quasi vertical and quasi horizontal polarization.

ERA-5 data are reanalyzed values and only available in 3-hourly time frame. One fundamental difference between the time in reanalyses and satellite observations is that the reanalyses are time-integrated values, while satellite observations are almost instantaneous measurements, the integration time for satellite data is just a few millisecond. This difference in time causes difficulties in collocating reanalysis and satellite observations. The magnitude of error due to this difference is unknown but depends on the temporal variability of the atmosphere, which itself depends on the time, location, and also the variable being collocated. For instance, Moradi et al. (2016) show that over the tropical oceans, relative humidity only slightly changes over the course of the day. In the collocations, however, we considered the time presented in both satellite and reanalyses to be the actual time of measurement and therefore simply subtracted the difference between the time then excluded the data with a time difference greater than 15 min.

Previous studies have shown that 1 hr time difference between profiles and observations introduces about 1 K bias for the humidity sounding channels (Moradi et al., 2013); therefore, our expectation is that this time difference only creates a small bias in comparison for humidity sounding channels and a negligible bias for temperature sounding channels. However, the error due to the time difference is likely to be larger for the surface sensitive channels than the humidity sounding channels. It is expected that averaging the differences between model simulated and real observations will cancel out the random affects due to the time difference. We tested the statistics using different time thresholds, ranging from 15 to 60 min, but did not find any important differences between the results. The distance was calculated based on euclidean distance between the center of the ERA-5 grid and the center of satellite footprint and was limited to 20 km.

This study was limited to clear-sky only; therefore, the cloud-contaminated observations were removed from the comparison. The difference between humidity sounding channels sensitive to different layers of the atmosphere have been used in the past by numerous studies for cloud clearing (Buehler et al., 2007; Moradi et al., 2015). In clear-sky conditions, due to environmental temperature lapse rate, the brightness

Table 2
List of the Fast and LBL Models Used in the Study

Acronym	Model	Version	Spectroscopy	Emissivity	SRF
C23B-F6-PWR	CRTM	2.3	Rosenkranz	FASTEM 6.0	Boxcar
C23-F6-PWR	CRTM	2.3	Rosenkranz	FASTEM 6.0	Real SRF
C23-RTV-PWR	CRTM	2.3	Rosenkranz	RTTOV FASTEM 6.0	Real SRF
C21-F6-PWR	CRTM	2.1	Rosenkranz	FASTEM 6.0	Real SRF
R121-F6-MPM	RTTOV	12.1	MPM	FASTEM 6.0	Real SRF
R121-F5-MPM	RTTOV	12.1	MPM	FASTEM 5.0	Real SRF
A23-RTV-HIT	ARTS	2.3	HITRAN	RTTOV FASTEM 6.0	Boxcar
A23-RTV-MPM	ARTS	2.3	MPM	RTTOV FASTEM 6.0	Boxcar
A23-RTV-PWR	ARTS	2.3	Rosenkranz	RTTOV FASTEM 6.0	Boxcar

temperatures from a channel sensitive to higher altitudes (Tb_{upper}) are colder than the Tbs for a channel sensitive to lower altitudes (Tb_{lower}). However, in cloudy conditions, the channels sensitive to higher altitudes are less affected by the clouds than the channel sensitive to lower altitudes; therefore, due to scattering, the Tb difference between the two channels will be reduced. At some point depending on cloud optical depth, the Tbs from the lower channel will become even colder than the Tbs from the upper channels (Buehler et al., 2007). The satellite observations are considered clear sky when $(Tb_{upper} - Tb_{lower}) < Tb_{thr}$. We used the difference of $Tb_{22} - Tb_{18} < -15$ for ATMS and $Tb_{13} - Tb_{12} < -10$ for GMI to determine clear-sky cases (numbers show the channel number). In addition, we used the ERA-5 total column cloud liquid and ice water contents with a threshold of 100 g m^{-2} as an additional measure for cloud clearing. We evaluated the results with different thresholds for cloud liquid and ice water contents, but even using a threshold of 10 g m^{-2} did not noticeably change the statistics. Additionally, because of difficulties in estimating the emissivity over land, we limited the study to the profiles nominally over ocean only. It is also not possible to accurately calculate the emissivity over ice covered surfaces; therefore, we only used the ice-free oceanic profiles. We used the ice cover provided in ERA-5 to remove cases with frozen surface water. We did not distinguish between ascending and descending orbits and used all the field of views for each scanline.

4. Results

We simulated ATMS and GMI Tbs using different setups of the fast models (CRTM and RTTOV) and the LBL model (ARTS). Table 2 shows a list of the setups for different RT models used in the study. This includes using different spectroscopy databases, using actual instrument SRFs versus traditional boxcar SRFs, and using different emissivity databases. After applying all the filters, the total numbers of collocations were about 1.5 million for ATMS/N20, 2.1 million for ATMS/NPP, and about 650,000 for GPM/GMI.

4.1. Direct Comparison of RT Models

It is generally less challenging to directly compare the RT models with each other, because sources of errors such as collocation errors and cloud contamination in the observations are not present when directly comparing the models. The most important factor in comparing the models is to ensure that all the essential inputs are provided to the models. The units of the absorbers such as water vapor are generally different among the models; therefore, the users need to ensure that the units are provided or defined correctly. In addition, the model internal calculations may be conducted in either level or layer averaged, so the inputs should be provided accordingly. Another important factor is that some models lack internal modules to calculate essential variables from other inputs; therefore, it is the user's responsibility to ensure that all the variables required by the models are provided as input. For instance, while RTTOV is able to perform all the surface emissivity calculations if the zenith angle is provided, CRTM requires both zenith and scan angles so that if the scan angle is not provided, it will result in a scan angle-dependent bias in the CRTM simulations for the surface sensitive channels. Additionally, CRTM requires both level and layer pressures to be able to calculate the transmittance values. Note that the fast models can only be used to simulate brightness temperatures for a given instrument; therefore, it is not possible to compare the models for a given frequency range.

Table 3
Difference Between ARTS (ARTS-RTV-MPM) and Simulations by the Fast Models: RTTOV (R121-F6-MPM) and CRTM (C23-F6-PWR)

	Channels										
	1	2	3	4	5	6	7	8	9	10	11
RTTOV bias	0.90	-0.44	0.38	0.59	0.47	0.28	0.17	0.12	0.07	0.03	-0.01
CRTM bias	0.49	0.49	1.39	1.07	0.67	0.71	0.34	0.33	0.20	-0.05	-0.10
RTTOV STD	0.78	0.93	1.20	0.89	0.39	0.11	0.11	0.09	0.06	0.04	0.04
CRTM STD	1.17	1.25	1.26	0.90	0.37	0.16	0.13	0.13	0.13	0.08	0.09
Channels	12	13	14	15	16	17	18	19	20	21	22
RTTOV bias	0.00	0.05	0.06	-0.05	-3.21	-0.80	0.89	0.85	0.76	0.62	0.24
CRTM bias	0.01	-0.03	-0.04	-0.13	-2.19	-0.58	0.98	1.01	0.96	0.80	0.63
RTTOV STD	0.04	0.04	0.07	0.14	1.27	1.55	0.32	0.20	0.23	0.27	0.28
CRTM STD	0.05	0.06	0.06	0.05	1.35	1.57	0.40	0.24	0.22	0.20	0.20

Note. Positive values show that the fast models' simulations are higher than ARTS simulations.

The difference between the fast models (R121-F6-MPM and C23-F6-PWR) and the LBL models ARTS (ARTS-RTV-MPM) is shown in Table 3. ARTS showed to perform generally better than fast models when run with the MPM spectroscopy so we have considered it as the reference model. The values are fast model simulations minus ARTS simulations; therefore, positive values show that the Tbs simulated using the fast models are larger than ARTS simulated Tbs. Overall, the largest differences are observed for Channel 16 with RTTOV showing a difference of -3.2 K and CRTM -2.2 K. We used RTTOV calculated emissivity values to run ARTS, so the difference between RTTOV and ARTS is due to difference in total transmittance. As shown later in Table 4, RTTOV calculated emissivity for Channel 16 is about 0.009 less than that for CRTM. Therefore, if we use CRTM emissivities to run ARTS then ARTS simulated Tbs will be larger compared to when we use RTTOV emissivities to run ARTS. Since the Tbs simulated using the fast models are already lower than ARTS Tbs calculated using RTTOV emissivities, larger difference would exist between the fast models and ARTS for Channel 16, if we run ARTS using CRTM calculated emissivities. Both RTTOV and CRTM Tbs are larger than ARTS Tbs for Channels 1-4, except that RTTOV shows a negative bias compared with ARTS for Channel 2. There is a difference of about 1 K between RTTOV and CRTM for Channels 2 and 3. The differences for the temperature sounding channels are relatively smaller than the differences among the models for the surface sensitive and water vapor channels. Although, CRTM and RTTOV are largely consistent for water vapor channels (Channels 18-22), there is a difference of more than 0.5 K between the fast models and ARTS for most channels.

Figure 1 shows the difference between Tbs simulated using C23-F6-PWR and R121-F6-MPM versus the Tbs simulated using R121-F6-MPM for some selected ATMS channels. ATMS Channel 1 is positioned near

Table 4
RTTOV Mean Emissivity Values Calculated Using FASTEM-6 for Different ATMS Channels as Well as the Mean Differences Between RTTOV and CRTM Emissivity Values ($\delta\epsilon = \text{RTTOV Minus CRTM}$) Calculated Using the Same Input Profiles and FASTEM-6 Emissivity Model

	Channels										
	1	2	3	4	5	6	7	8	9	10	11
RTTOV ϵ	46.89	50.04	55.12	55.52	55.81	56.02	56.24	56.38	56.53	56.99	56.99
Mean $\Delta\epsilon$	-0.39	-0.50	-0.84	-0.86	-0.86	-0.87	-0.88	-0.88	-0.89	-0.90	-0.90
STD $\Delta\epsilon$	0.27	0.30	0.36	0.36	0.37	0.37	0.37	0.37	0.37	0.38	0.38
	Channels										
	12	13	14	15	16	17	18	19	20	21	22
RTTOV ϵ	56.99	56.99	56.99	56.99	64.62	71.12	72.10	72.10	72.10	72.10	72.10
Mean $\Delta\epsilon$	-0.90	-0.90	-0.90	-0.90	-0.87	-0.61	-0.46	-0.46	-0.46	-0.46	-0.46
STD $\Delta\epsilon$	0.38	0.38	0.38	0.38	0.45	0.40	0.36	0.36	0.36	0.36	0.36

Note. All the values are multiplied by 100.

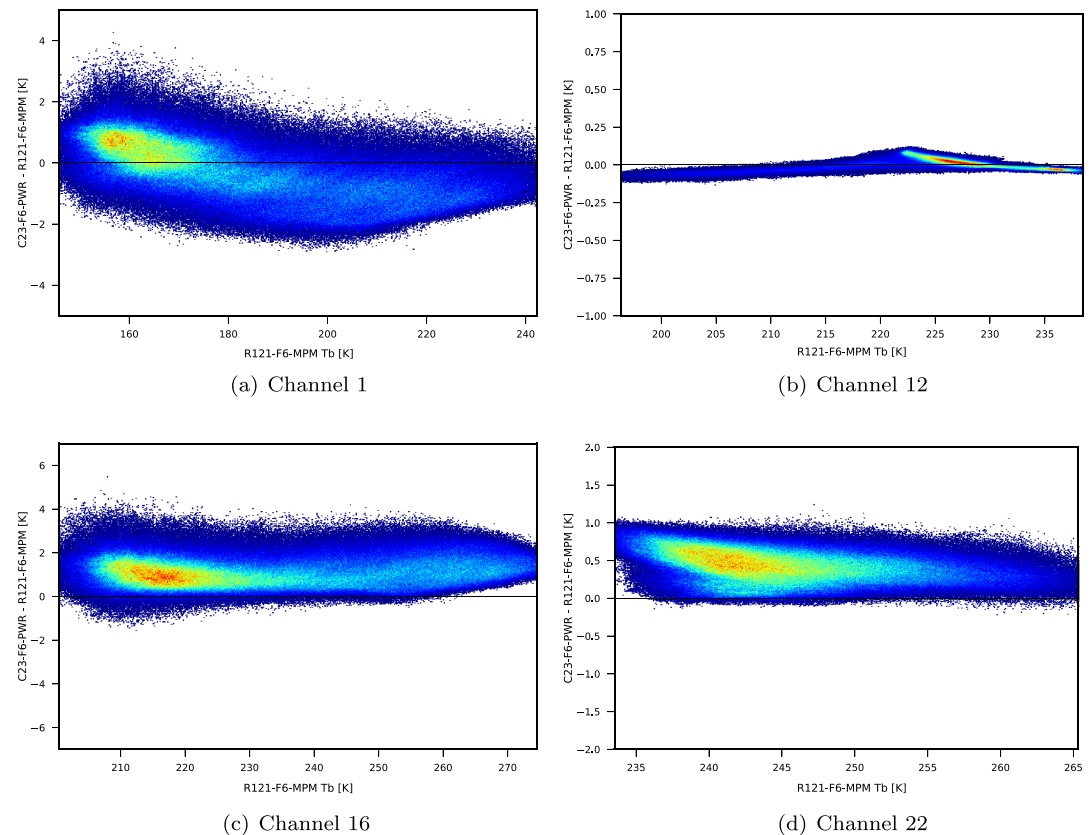


Figure 1. Brightness temperatures simulated using C23-F6-PWR versus Tbs simulated using R121-F6-MPM.

the water vapor absorption line at 23 GHz and sensitive to the total column water vapor as well as to the surface conditions. Channel 12 is sensitive to stratospheric and upper tropospheric temperature, Channel 16 is mostly sensitive to the surface conditions, and Channel 22 is positioned near water vapor absorption line at 183 GHz and is sensitive to upper tropospheric humidity. In the case of Channel 1, the differences are largely uniform and the variance is only slightly larger for the lower Tbs than for the higher Tbs. The differences between the models for Channel 12 are generally very small with the differences being slightly larger for the Tbs between 220 to 230 K. The variance between the differences for Channel 16 decreases as the Tb increases and overall the differences range between -1 and 4 K. In the case of Channel 22, the differences between the two models are largely uniform ranging mostly between 0 and 1 K.

4.2. Comparing Simulations and Observations

Direct comparison of the RT models shows the immediate differences among the models, but such differences do not represent the actual models' errors, because RT models share a lot of methods and parameters so that similar biases exist in models' simulations, which cannot be detected based on the differences among the models. Therefore, it is always required to compare the simulations with measured observations to be able to evaluate the performance of the models. Satellite data are also prone to biases and uncertainties (Berg, 2017; Moradi et al., 2015), but data from well calibrated instruments can be used to evaluate (not validate) the performance of the RT models. The confidence in such comparisons can be increased by using data from multi instruments measuring in the same frequency range. We used ATMS and GMI observations to evaluate the performance of the RT models. ATMS is currently flying on both NPP and N20; however, the statistics from the NWP analysis as well as this study show that the ATMS onboard N20 provides more accurate observations. GMI is also a well calibrated instrument owing to the work performed by a dedicated intercalibration team (Berg, 2017); therefore, we chose ATMS/N20 and GMI to show the models' biases versus observations. We later in section 5 use ATMS/NPP observations to show how the RT models can be used to detect relative instrument errors using double difference technique.

There are three sources of errors when comparing RT simulations with real observations, including error in ERA-5 profiles used as input, RT error, and collocation error. The ERA-5 is reanalysis and its accuracy

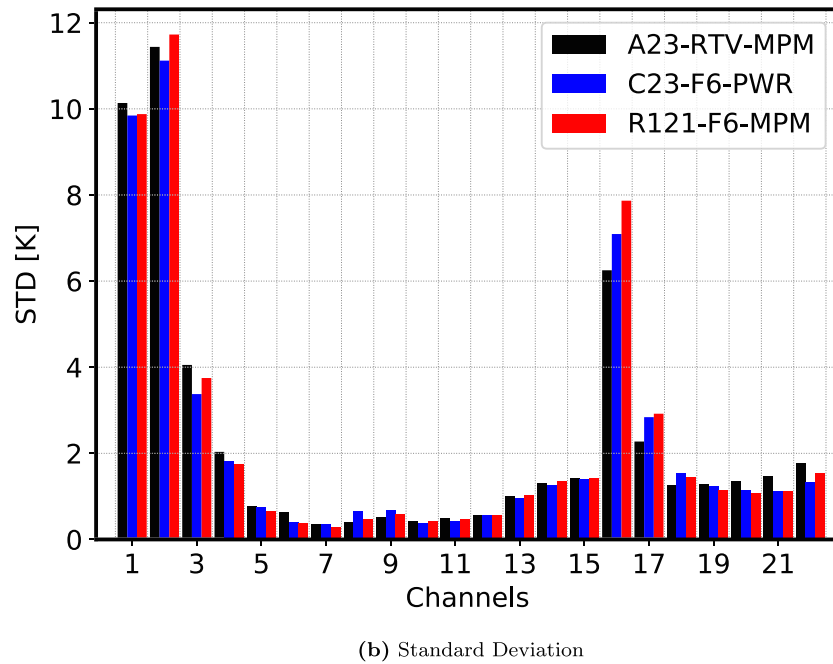
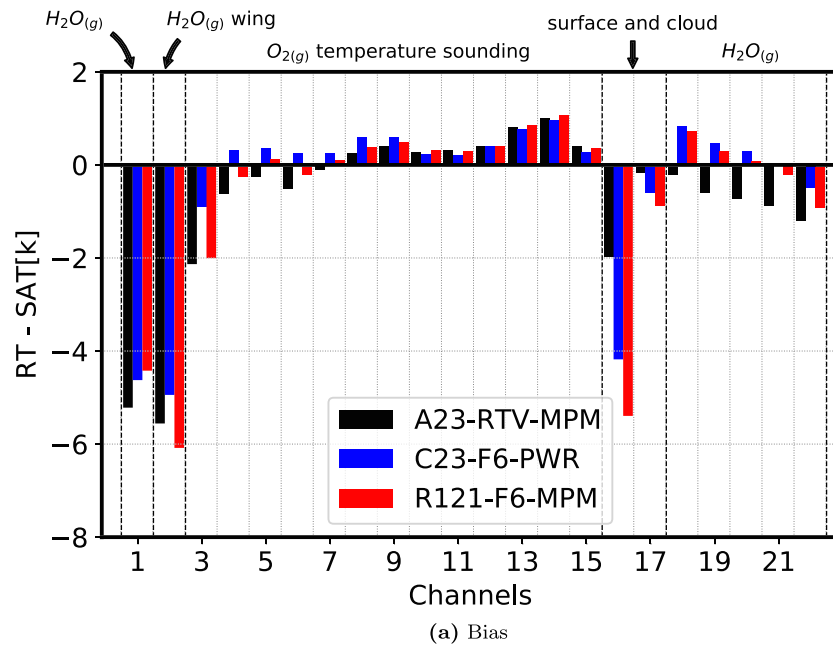


Figure 2. (a) Mean (bias) and (b) standard deviation of the differences between ATMS/N20 real observations and Tbs simulated using ERA-5 and different RT models.

is limited by factors such as the first guess provided by the NWP model, error in observations assimilated into the system, and limitations in data assimilation techniques to properly account for observation and background errors. In addition, RTTOV is used in ECMWF as forward model; therefore, ERA-5 profiles are expected to yield better results when comparing RTTOV simulated Tbs with real observations. There are also several sources of error in RT calculations, including error in spectroscopy models, error in emissivity models, and assumption in RT theories, for example, absorption line shapes. The collocation error includes error in interpolating from the model grid to satellite footprint, especially not considering the instrument actual footprint, time difference between observations and input atmospheric profiles, and the possibility of cloud-contaminated observations being used in the comparison.

The differences between ATMS/N20 real observations versus the T_b s simulated using ERA-5 profiles and different RT models are shown in Figure 2. In addition to C23-F6-PWR and R121-F6-MPM, we have also included the results for the LBL model ARTS. The difference between the RT simulations and ATMS observations is much smaller for both temperature and humidity sounding channels than for the surface sensitive channels. The differences are also relatively smaller for the temperature sounding channels than for the humidity sounding channels. We ran ARTS with three different spectroscopy models, but Figure 2 only shows the results for the MPM spectroscopy (A23-RTV-MPM), because as discussed later, the MPM resulted in the lowest biases compared with ATMS/N20 observations. Nevertheless, the difference between different ARTS setups is later discussed in section 4.3.1. ARTS showed a difference of about 2 K for Channel 16 compared with almost 4 K for C23-F6-PWR and about 5.5 K for R121-F6-MPM. The performance of the models was different for the humidity sounding channels. R121-F6-MPM bias was about -1 K for Channel 17 (RT-ATMS/N20), but A23-RTV-MPM showed a very negligible bias for this channel (about -0.1 K). The CRTM bias for this channel was about -0.8 K. CRTM and RTTOV showed a bias of about 0.8 to 0.9 K compared to ATMS/N20 observations for Channel 18, whereas ARTS showed -0.1 K. The fast models (CRTM and RTTOV) showed a very small difference with the observations from ATMS Channels 19 and 20, but ARTS showed a difference of more than -0.5 K for these channels. ARTS also showed larger differences for Channels 21 and 22 compared with CRTM and RTTOV. The disagreement among the humidity sounding channels shows the challenge for validating microwave channels sensitive to tropospheric humidity. It should be noted that in dry conditions most water vapor channels can potentially be affected by the surface emissivity; therefore, screening out the simulated T_b s that are affected by surface emissivity may reduce the differences between simulated and observed T_b s for the water vapor channels.

The temperature sounding channels that peak higher in the atmosphere, Channels 13–15, show a larger bias than the channels that peak lower in the atmosphere (Channels 5–12). This may be related to the quality of temperature profiles in ERA-5. Global Positioning System (GPS) radio occultation (GPS-RO) are used as reference in reanalysis (Poli et al., 2010), but GPS-ROs are only assimilated up to around 40 km. However, the weighting functions for Channels 13–15, to a large extent, reside above 40 km (Moradi et al., 2015). Although, AMSU-A Channel 14 and ATMS Channel 15, which peak in stratosphere, are assimilated in addition to GPS-RO, the effect of Zeeman splitting on the oxygen lines can introduce small biases if not included in the RT models. In addition, the model coarse vertical resolution in the upper levels may play a role in the quality of ERA-5 temperature profiles in stratosphere. Figure 2 also shows the STDs of the differences between ATMS/N20 observations and RT simulated T_b s. As it was expected, the STDs are much larger for the window channels than water vapor and temperature sounding channels. The STDs are as large as 12 K for the window channels, which is due to a larger variation in the surface conditions than in the atmospheric conditions. The STDs are less than 1 K for most temperature sounding channels and less than 1.5 K for water vapor channels.

In summary, Channels 1 and 2 showed the largest biases among all the ATMS/N20 channels, which is mainly due to large biases along the coastlines affected by the large fields of view (FOVs) for these channels. Channel 16 is also a window channel and shows a relatively large bias compared with the temperature and humidity sounding channels; the bias for Channel 16 ranges from 2 K for ARTS to more than 5 K for RTTOV. Note that although both ARTS and RTTOV used the same emissivity model (FASTEM V6.0), we still see a large difference between them for the surface sensitive channels. The measured T_b is not only affected by the surface emissivity (ϵ) but also depends on total transmittance, in a simplified form $T_b = \epsilon \cdot T_s \cdot \Gamma$, where T_s is surface temperature and Γ is the transmittance integrated from the surface to the top of atmosphere. Therefore, for instance, if we assume $T_s = 300$ and $\epsilon = 0.7$, then 1% change in Γ ($\Gamma = 0.99$ vs. $\Gamma = 0.98$) would introduce a difference of 2.1 K in measured T_b s. The actual impact will be even larger because the reflected radiance, which is not included in the above calculations, passes through the atmosphere twice.

Figure 3 shows the spatial distribution of the differences between ATMS N20 T_b s and C23-F6-PWR simulated T_b s. The largest differences for Channel 1, which is a surface sensitive channel, are observed along the coastlines. The reason for these large differences along the coastlines, which also tend to be negative (RT – Obs), is that the profiles are considered to be over ocean in the RT simulations, but in the real measurements due to the large instrument footprint, a mixture of land and ocean is observed by the instrument. The FOV for Channels 1 and 2 is 5.2° along track, which is almost 2.5 times wider than the 2.2° FOV for Channel 16; see Table 1. In addition, the emissivity for the coastal water is different from the emissivity over deep waters due to difference in the salinity and other factors that affect the sea surface emissivity. If we

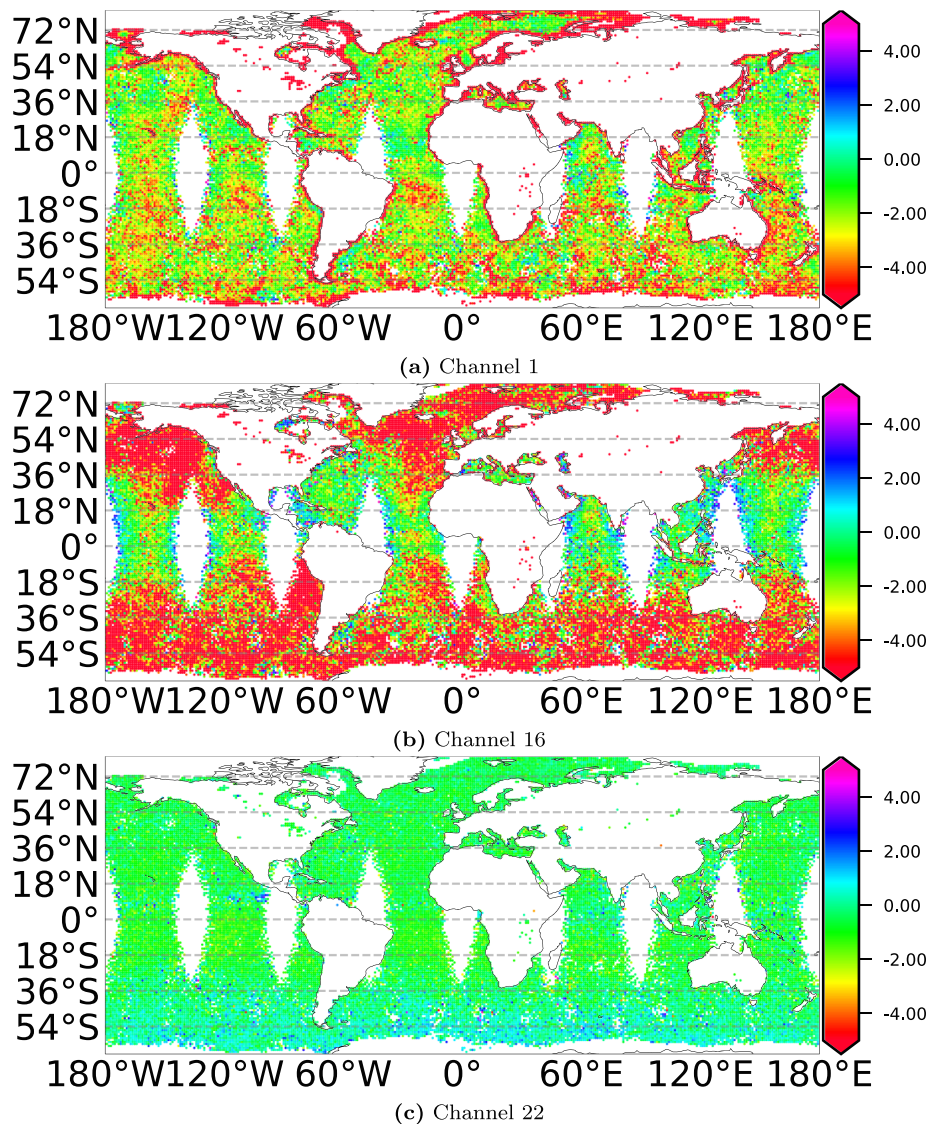


Figure 3. Spatial distribution of the difference between C23-F6-PWR simulated Tbs and ATMS observations from Channels (a) 1, (b) 16, and (c) 22.

ignore these coastal cases, then the average of the differences between the observations and the RT model simulations for Channel 1 would be very small. In the case of Channel 16, the differences outside tropical band are also systematically larger and tend to be negative (RT – Obs). The RT minus observation values for Channel 16 within the tropical band tend to be less than 2 K.

The analysis of the RT minus observation differences showed that these differences tend to be larger for the profiles with lower total precipitable water vapor. This is because in dry conditions the weighting functions for the window channels peak much closer to the surface; therefore, the brightness temperature become much more sensitive to errors in surface emissivity calculations. In the case of Channel 22, the differences for the latitude band between 36°S to 65°S tend to be larger than the differences for the rest of the globe. This is likely because the relatively dry atmosphere of the Southern Hemisphere winter causes the Channel 22 weighting function peaking low in the atmosphere exerting a greater surface influence. The differences for the temperature sounding channels were generally very small and did not show any patterns and thus are not included.

In order to investigate whether the differences between ATMS and models are largely due to error is simulation or observations, we also compared the CRTM and RTTOV simulations with GPM/GMI observations.

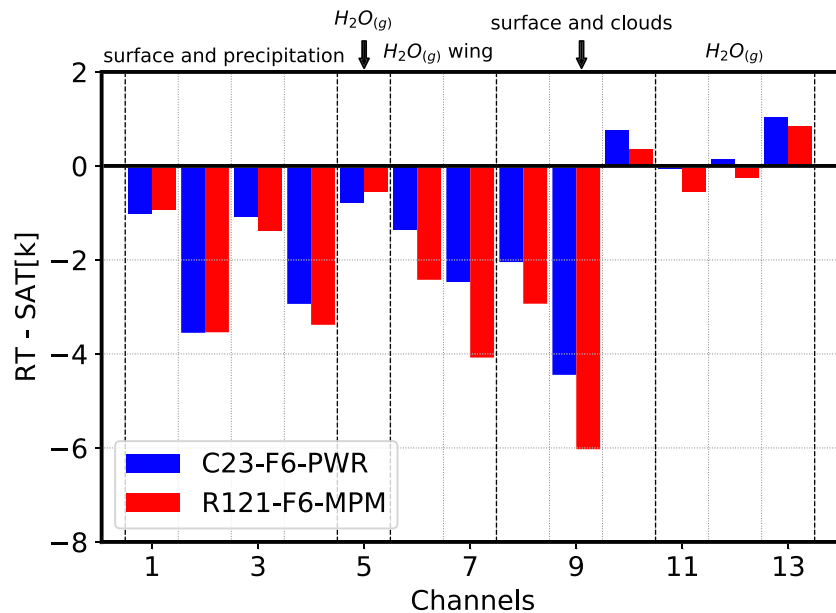


Figure 4. Mean difference (bias) between GMI/GPM observations and Tbs simulated using different RT models.

The differences between GMI/GPM measured Tbs and the model simulated Tbs are shown in Figure 4. GMI only has two channels sensitive to tropospheric relative humidity, Channels 12 and 13, and to a lesser extent Channels 10 and 11. The rest of GMI channels with H and V polarization are sensitive to the surface as well as clouds and precipitation. The surface sensitive channels again show some large negative biases (model vs. satellites) affirming negative biases in model simulated brightness temperatures for the surface sensitive channels. CRTM generally performs better than RTTOV for the surface sensitive channels. The difference between model simulated and GMI observations for the channels sensitive to relative humidity are relatively small, less than 0.5 K. The STDs of the differences between the GMI observations and simulations were also very large for the window channels (as large as 12 K similar to ATMS), but less than 1 K for the water vapor channels.

4.3. Impact of Inputs on RT Simulations

There are several factors that affect the simulations by RT models including spectroscopy models used to calculate the atmospheric transmittance, surface emissivity used to calculate the surface emitted and reflected radiance, and sensors response function used to average the spectral radiance. This section shows the impact of these parameters on the differences between the simulated and measured Tbs.

4.3.1. Impact of Spectroscopy Models

Information provided in spectroscopy models, which includes absorption and emission lines as well as continuum, is used to calculate the absorption coefficients (Clough et al., 1989). In microwave frequencies, accuracy of the absorption lines is especially important for the channels that are sensitive to water vapor (183 and 23 GHz) and oxygen absorption lines (60 GHz). However, previous studies generally show better agreement between simulated and measured Tbs for the microwave channels operating at the oxygen absorption band at 60 GHz than water vapor channels operating at 183 GHz (Moradi et al., 2015). Although, it is clear that NWP models represent the three dimensional temperature fields more accurately than water vapor fields, the uncertainty in the water vapor spectroscopy plays a role here as well.

We evaluated ARTS simulations conducted using three different spectroscopy models versus ATMS/N20 observations; see Figure 5. All the ARTS setups used the same spectroscopy for the oxygen absorption band at 60 GHz, so the differences between the ARTS setups for the channels operating near 60 GHz originate from the water vapor continuum, which turned out to be negligible. The difference between Tbs simulated using different spectroscopy models were especially very large for the surface sensitive and water vapor channels. The difference between simulated and real observations for Channel 16 ranges between almost 2 K for A23-RTV-MPM to 9 K for A23-RTV-HIT. The A23-RTV-HIT and A23-RTV-PWR generally showed larger differences than A23-RTV-MPM compared with observations for the surface sensitive and water vapor

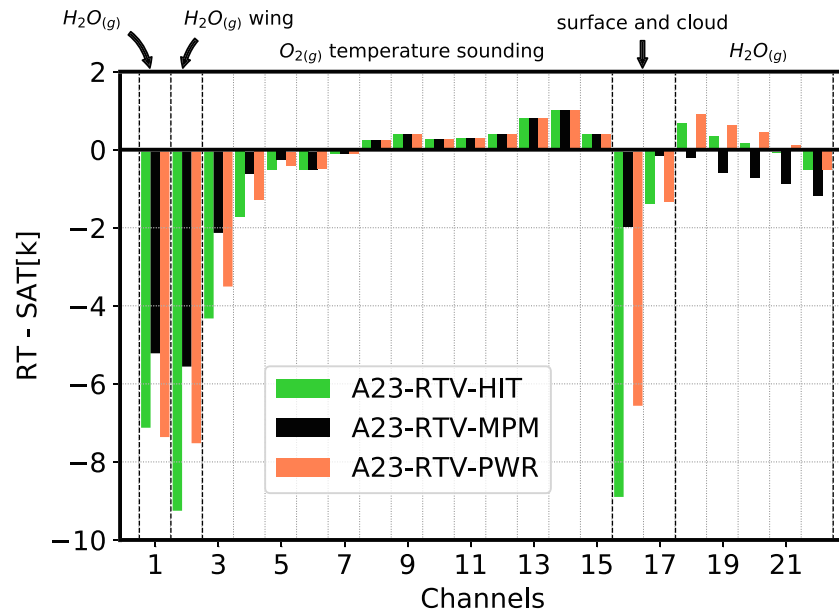


Figure 5. Difference between ATMS/N20 observations and Tbs simulated using ARTS with different spectroscopy models.

channels. Tbs from A23-RTV-MPM simulations showed negative biases compared with ATMS/N20 observations for water vapor channels, but two other spectroscopy models showed mostly positive biases, except for Channel 22. As explained before, the reason for the large effect of spectroscopy models on the window channels is that the measured or simulated Tb is a function of both surface emissivity and total transmittance so that a small change in total transmittance can introduce a large difference in calculated Tbs for the window channels.

4.3.2. Impact of Sensor Response Function

Many RT models traditionally have assumed boxcar SRFs. In recent years there has been some efforts to measure the actual SRFs for some of the new instruments including ATMS and GMI. Figure 6 shows the differences between CRTM simulations for ATMS/N20 using boxcar and real SRFs. The differences ultimately depend on how much the real SRFs are different from the expected boxcar SRFs. The difference between Tbs simulated using boxcar and real SRFs for the temperature sounding channels is mostly around 0.1 K, except for Channel 6 with more than 0.25 K difference. Figure 7 shows the SRFs for Channels 6 and 11 with the largest and smallest biases among temperature sounding channels.

Obviously, the difference between boxcar function and actual SRF is much larger for Channel 6 than for Channel 11, especially the right side of real SRF for Channel 6 deviates from the boxcar function. Note that as shown in Figure 7 and Table 1, both Channels 6 and 11 have two passbands. It was also expected to see a larger difference between Channel 22 boxcar and real SRFs, because of an unsymmetrical real SRF for this channel, but the impact of SRF for the water vapor channels is very small.

4.3.3. Impact of Surface Emissivity

All the components of the Earth system including land and ocean surfaces emit radiation according to Planck's law as a functions of their temperature times surface emissivity. Depending on the total transmittance of the atmosphere, the surface emitted radiance may completely be absorbed by the atmosphere between the surface and the sensor in which case the measured radiances are not affected by the surface emissivity at all. However, if surface emitted radiance contributes to the measured radiance, which is the case for the surface sensitive channels, then error in surface emissivity and surface temperature would contribute to the error in simulated Tbs. In both CRTM and RTTOV, the ocean emissivity is calculated using the FASTEM model which requires parameters such

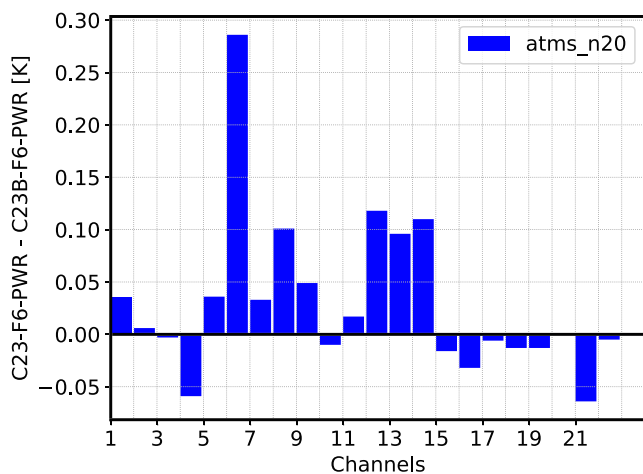


Figure 6. The differences between Tbs simulated using C23B-F6-PWR and C23B-F6-PWR, which shows the impact of using boxcar SRFs versus real ATMS/N20 SRFs.

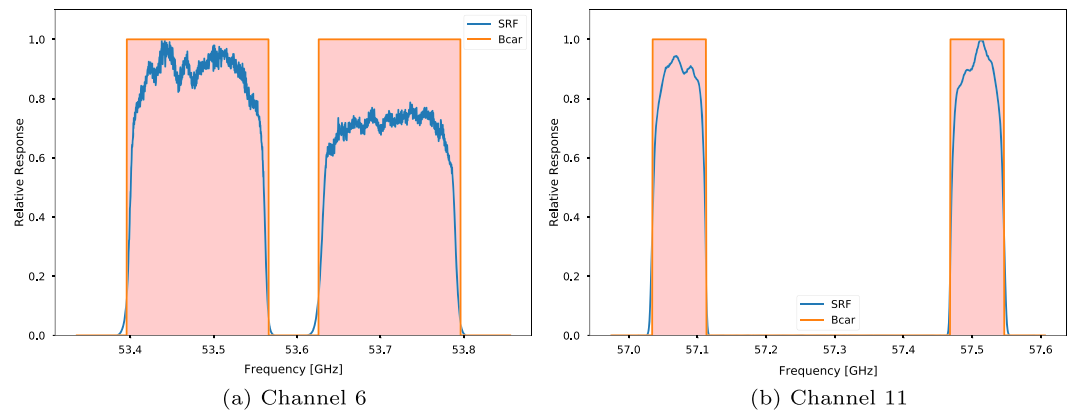


Figure 7. The actual and boxcar SRFs for ATMS/N20 Channels (a) 6 and (b) 11.

as wind speed and direction as input (Bormann et al., 2012; Kazumori & English, 2015). Error in the surface emissivity model or inputs to the emissivity model would contribute to the error in calculated emissivities. Although, Sea Surface Temperature (SST) in reanalysis is known to have good accuracy owing to many observations assimilated into the model, the surface wind as well as the surface emissivity models to a large extent lack the accuracy required, for instance, for the assimilation of observations from surface sensitive microwave channels into NWP models. CRTM and RTTOV currently have different versions of the FASTEM model implemented into the RT code. The only major difference between the two recent versions of the FASTEM model used in this study (V5.0 and V6.0) is in wind direction (Kazumori & English, 2015). We used RTTOV calculated surface emissivity values (FASTEM V6.0) to run ARTS.

Another factor that may contribute to the surface emissivity error is how the reflectivity is calculated. The downwelling radiance is reflected by the surface by a factor of $(1 - \text{emissivity})$; therefore, error in emissivity will contribute to the error in calculated reflectance (ρ) as well. The reflectivity itself is modeled using either Lambertian Bidirectional Reflectance Distribution Function (BRDF) also known as Diffuse Reflection

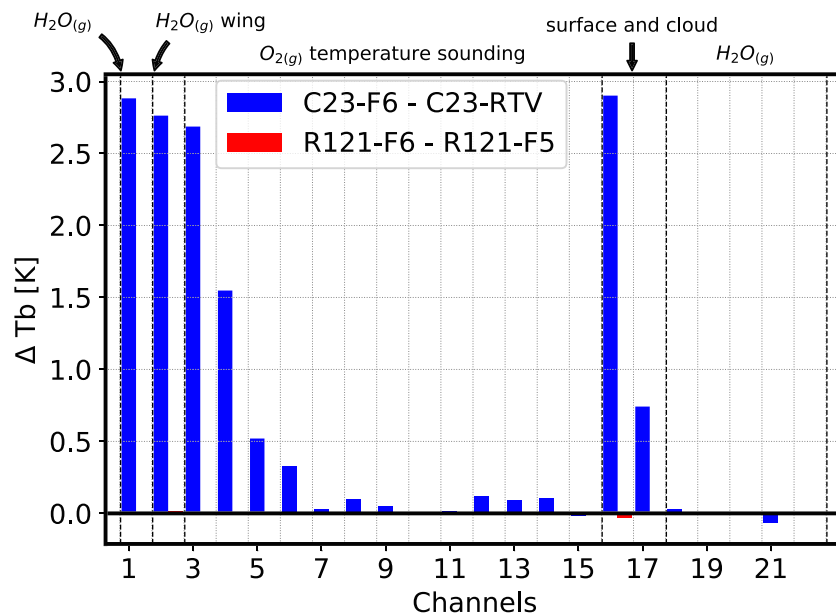


Figure 8. Impact of emissivity on the calculated Tbs. C23-F6 – C23-RTV is the difference between CRTM calculations with internally calculated FASTEM-6 (C23-F6-PWR) and RTTOV calculated FASTEM-6 emissivity (C23-RTV-PWR), and R121-F6 – R121-F5 shows the difference between RTTOV Tb with FASTEM-6 (R121-F6-MPM) and FASTEM-5 (R121-F5-MPM).

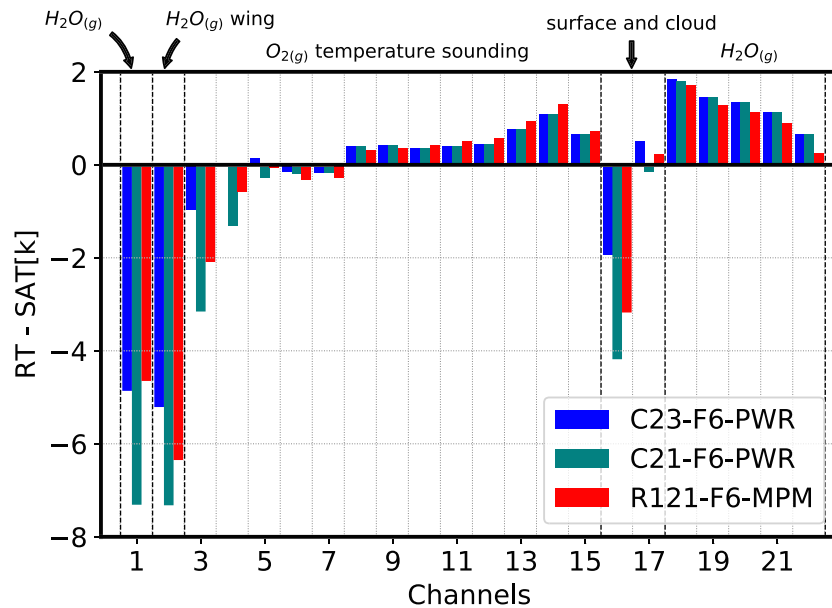


Figure 9. Same as Figure 2 but for ATMS/NPP.

(ρ/π) or Specular Reflection also known as Mirror BRDF, which is a double delta function. RTTOV uses Lambertian BRDF for the reflectivity model, but CRTM and ARTS use a Specular Reflection model.

Figure 8 shows the difference between CRTM calculations using internally calculated emissivities and RTTOV calculated emissivity manually supplied as input to CRTM. We also evaluated the impact of FASTEM-5 and FASTEM-6 on RTTOV calculations but no significant difference was observed between the two runs. The difference in CRTM calculations for the surface sensitive channels is about 2.5 K for Channels 1–3 and 16 but smaller for the other surface sensitive channels.

5. Double Difference Technique

The double difference is calculated using the biases (difference) between the model simulated and satellite observed values (Alsweiss et al., 2015). In the case of ATMS observations from NPP and N20 satellites, we calculated the double difference as (RT – ATMS/NPP) minus (RT – ATMS/N20), which become equal to ATMS/N20 minus ATMS/NPP. We evaluate in this section whether double difference technique can be used to investigate relative differences in satellite observations from different instruments. Although, there are several sources of errors when comparing model simulated Tbs versus observed values, it is expected that some of these errors cancel out when calculating the double differences.

The differences between ATMS/NPP observations and model simulated Tbs are shown in Figure 9. We have only included three models in the comparison including C21-F6-PWR, C23-F6-PWR, and R121-F6-MPM. The results for the temperature sounding and surface sensitive channels are very similar to the results for ATMS/N20. Similarly, the difference between C23-F6-PWR and R121-F6-MPM for temperature sounding channels is less than 1 K. However, the results for the humidity sounding channels are different between ATMS/NPP and ATMS/N20. In Figure 9, the differences between the measured and the model simulated Tbs for the humidity sounding channels are less than 1.5 K, but the bias is larger for the channels near the edge of the water vapor absorption line at 183 GHz, for example, Channel 18 compared with Channel 22. Although, Channel 17 is located at the edge of the water vapor absorption line at 183 GHz but shows the lowest bias (0.1 K) compared with other humidity sounding channels. The results for ATMS/NPP are largely in agreement with Bobryshev et al. (2018) who reported positive biases (RT – ATMS) for Tbs simulated using ARTS and radiosonde profiles compared with ATMS/NPP. However, Bobryshev et al. (2018) show that carefully matching radiosonde profiles with satellite observations (e.g., using only radiosonde profiles with less than 15 km horizontal drift) and also using cloud masks from visible satellite instruments can potentially reduce the differences between the simulated and observed Tbs. However, the biases shown for water vapor

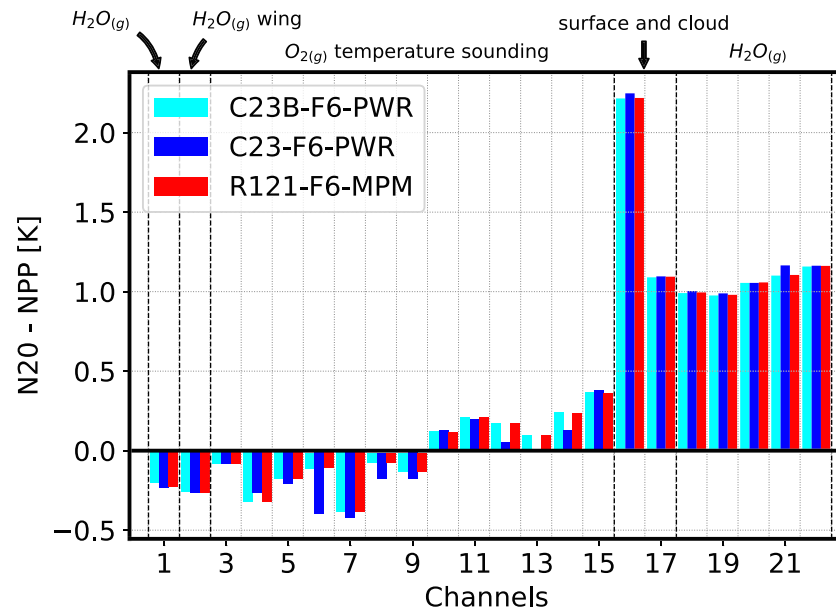


Figure 10. Double difference between RT simulated Tbs and real observations from ATMS onboard NPP and N20, which indirectly shows ATMS/N20 minus ATMS/NPP.

channels in Figure 9 as well as in other studies, for example, Bobryshev et al. (2018) and Brogniez et al. (2016), are still considerably larger than the biases shown in Figure 9 for the temperature sounding channels.

Figure 10 shows the double difference between simulated and real ATMS/NPP and ATMS/N20 brightness temperatures. The difference between the two instruments is generally less than 0.5 K for Channels 1–15. Channel 16 shows the largest difference among all the channels (about -2.0 K). The difference for the humidity sounding channels is about -1.0 K and both C23-F6-PWR and R121-F6-MPM double differences are generally consistent. Figure 11 shows the differences directly calculated using real observations. We used observations from two different time periods, June and July 2018 and February 2019, to assess possible seasonal effect on the differences. We first read the data for each time period and binned ATMS/NPP and ATMS/N20 observations separately into a common grid with a resolution of 1° , then calculated the difference by subtracting the binned values and averaging the differences between the two instruments. The values shown in Figure 11 are clear-sky global oceanic averages, because limiting the data to a certain latitude band (e.g., tropical region only) did not affect the differences. Overall, the differences for Channels 1–15 are less than 0.5 K and the differences from the two time periods for these channels are also consistent. The differences between the observations and the RT double differences are also consistent for Channels 1–15. The differences are larger for the water vapor channels but less consistent than the double differences for other channels. For the channels operating near the center of absorption line at 183 GHz, there is a difference of about 0.5 K between the two time periods. Note that NPP and N20 are essentially on the same orbit, with NPP following the N20 with a roughly 50 min delay. Therefore, the difference in overpass times is not expected to have a large effect on the difference between clear-sky ocean only measurements from the two instruments.

This work led to reprocessing of ATMS/NPP observations using an enhanced calibration algorithm and a new data set was released in mid-October 2019. The two major improvements in the ATMS/NPP calibrations include, correcting antenna emissivity and also an enhanced modeling of antenna pattern correction. The ATMS/NPP antenna emission was derived using pitch-over maneuver measurements and was estimated to be as large as 2.5 K for the calibration space view (Leslie et al., 2013; Yang et al., 2016). In ATMS/NPP prelaunch calibration, the relatively low antenna beam efficiency for water vapor channels was derived from the raw antenna pattern measurements, which caused issues in the antenna pattern correction for ATMS/NPP and prevented generating high-quality sensor data record. In ATMS/N20 prelaunch calibration, the lower noise floor in antenna pattern measurements for water vapor channels provided a better estimate of channel beam efficiency. In addition, a more accurate channel beam efficiency was derived by introducing

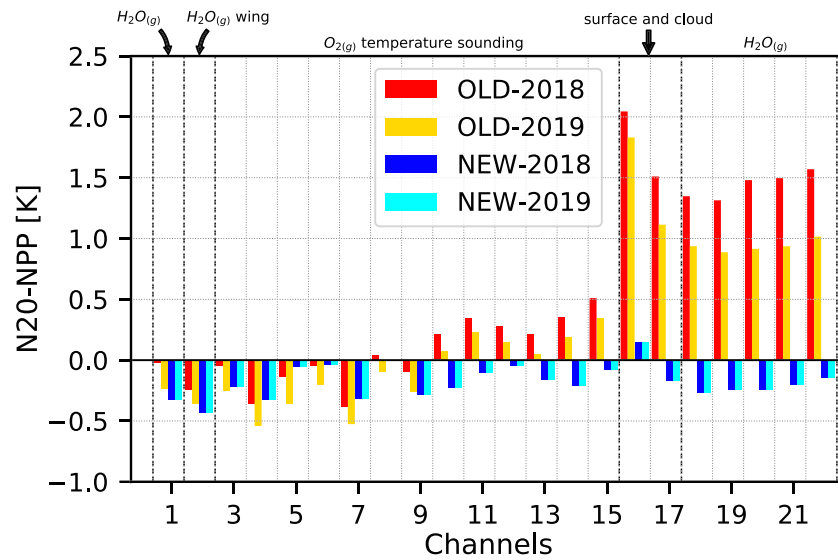


Figure 11. Difference between real observations of ATMS/N20 and ATMS/NPP for two different time periods: June and July 2018 and February 2019. NEW is reprocessed ATMS/NPP data, and OLD indicates data that were distributed until mid-October 2019 and were calibrated using ATMS/NPP original calibration algorithm.

an antenna pattern simulation model as well as using the hybrid antenna pattern data sets. In the reprocessing algorithm, the same technique was adopted for ATMS/NPP to derive more accurate antenna beam efficiency coefficients (Yang et al., 2016). The statistics from NWP centers, for example, ECMWF observation minus first guess statistics, already shows a great improvement in ATMS/NPP observations after enhancing the calibration algorithm. Figure 11 also shows the difference between ATMS observations from NPP and N20 satellites before (OLD) and after (NEW) reprocessing. As it is shown, the differences between the two instruments, especially for water vapor channels as well as Channel 16, are largely decreased after reprocessing and the two instruments are largely consistent for all the channels. After reprocessing, the differences between the two instruments are less than 0.25 K for most of the channels.

6. Conclusions

RT models have a wide range of applications in satellite remote sensing including being used as the forward model and adjoint to assimilate satellite observations into NWP models, and simulating satellite observations from input atmospheric profiles for product retrieval or calibration and validation of satellite observations. Given the wide range of applications of fast models, it is necessary to evaluate these models against the observations as well as the LBL models. We selected CRTM and RTTOV because these models are widely used all over the world for different purposes. We also included in the comparison the LBL model ARTS that was specifically developed to perform accurate simulations for the microwave region. We evaluated the models versus each other as well as against observations from ATMS onboard N20 and GMI onboard GPM.

The models were generally very consistent with each other and observations for the temperature sounding channels operating near 60 GHz oxygen band. The difference between the models as well as simulations and observations were less than 0.5 K for most ATMS temperature sounding channels. However, there was a larger difference between the models as well as simulations and observations for the water vapor channels (ranging between 1 and -1 K). The biases for window channels were much larger than the biases for the water vapor and temperature sounding channels and consistently negative (ranging between -0.9 and 6 K for different ATMS and GMI window channels). However, in the case of Channels 1 and 2 with a large FOV, the differences between the observations and the models were especially larger for the coastal cases. The models assume these profiles to be over ocean, but because of the large instrument FOV, these channels measure a mixture of radiances emitted by both land and ocean. Although, window channels are not currently assimilated into NWP models, work is underway to especially include these channels in all-sky assimilation systems. This study shows that assimilation of observation from these channels not only depends on

the improvement of the emissivity models but highly depends on the water vapor spectroscopy database as well. Double differences showed that, although the differences between the RT models and observations may not be used to evaluate the absolute errors in the observations, it can be used to assess the relative differences among the instruments. For instance, the double difference technique used in this study revealed some noticeable differences between ATMS/N20 and ATMS/NPP water vapor channels. The direct comparison of the observations from the two instruments also confirmed the inconsistency among the observations for water vapor channels. An enhanced calibration algorithm was developed by the instrument calibration team for ATMS/NPP. Direct comparison of the ATMS/NPP and ATMS/N20 observations showed that the two instruments are generally in good agreement after reprocessing of ATMS/NPP data.

We evaluated several spectroscopy models including HITRAN database (Rothman et al., 2013), MPM (Liebe, 1985; Liebe et al., 1993), and PWR (Rosenkranz, 1998). ARTS results showed some relatively large differences for the water vapor and especially window channels when using different spectroscopy models. The MPM showed lower biases compared with the observations, ranging between -2 and -6 K for window channels and less than -1 K for the water vapor channels. The reason that the spectroscopy models impact the surface sensitive channels is because the simulated T_b is a function of both surface characteristics (emissivity and temperature) and total atmospheric transmittance, so that even one percent change in total atmospheric transmittance can have a large impact on simulated T_b s for surface sensitive channels.

We also compared the T_b s simulated using actual ATMS/N20 and boxcar SRFs. The difference between the two was less than 0.1 K for most channels, except for Channel 6 with a difference of about 0.25 K between T_b s calculated using actual and boxcar SRFs. Finally, we evaluated the impact of emissivity models on calculated T_b s using two different versions of FASTEM (V5.0 vs. V6.0) as well as by using RTTOV calculated emissivities to run CRTM. The difference between T_b s calculated using different versions of FASTEM was negligible; however, running CRTM using emissivities calculated by RTTOV introduced a large impact on calculated T_b s.

Overall, although the fast and LBL microwave RT models perform satisfactorily for the temperature sounding channels, some relatively large differences still exist for the water vapor and surface sensitive channels. Two main factors affecting these differences include spectroscopy and emissivity models. Although, while the spectroscopy models were largely in agreement for the oxygen absorption band at 60GHz, they still show some noticeable differences for the channels operating around water vapor absorption line at 183 GHz. The emissivity models are still unable to simulate the surface emissivity with enough accuracy, even over ocean; therefore, more accurate surface emissivity models are required. This study focused on clear-sky observations; however, given the increasing application of all-sky RT models, work is required to evaluate all-sky RT models.

Acknowledgments

This study was funded by the NOAA Joint Polar Satellite System (JPSS) program under NOAA Cooperative Agreement. The RTTOV development has been funded by the EUMETSAT NWP SAF. ERA-5 data are available from ECMWF (<https://apps.ecmwf.int/data-catalogues/era5/?class=ea>), ATMS data are available from NOAA CLASS (<https://www.class.noaa.gov>), and GMI data are available from NASA DISC (<https://disc.gsfc.nasa.gov>).

References

- Alsweiss, S. O., Jelenak, Z., Chang, P. S., Park, J. D., & Meyers, P. (2015). Inter-calibration results of the advanced microwave scanning radiometer-2 over ocean. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(9), 4230–4238. <https://doi.org/10.1109/JSTARS.2014.2330980>
- Berg, W. (2017). Towards developing a long-term high-quality intercalibrated TRMM/GPM radiometer dataset, 2017 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 248–250). Fort Worth, TX, USA: IEEE. <https://doi.org/10.1109/IGARSS.2017.8126941>
- Bobryshev, O., Buehler, S. A., John, V. O., Brath, M., & Brogniez, H. (2018). Is there really a closure gap between 183.31-GHz satellite passive microwave and in situ radiosonde water vapor measurements? *IEEE Transactions on Geoscience and Remote Sensing*, 56(5), 2904–2910. <https://doi.org/10.1109/TGRS.2017.2786548>
- Bormann, N., Geer, A. J., & English, S. (2012). Evaluation of the microwave ocean surface emissivity model FASTEM-5 in the IFS 667. <https://doi.org/10.21957/1936s9df>
- Boukabara, S.-A., Garrett, K., Chen, W., Iturbide-Sanchez, F., Grassotti, C., Kongoli, C., et al. (2011). MiRS: An all-weather 1DVAR satellite data assimilation and retrieval system. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(9), 3249–3272.
- Brogniez, H., English, S., Mahfouf, J.-F., Behrendt, A., Berg, W., Boukabara, S., et al. (2016). A review of sources of systematic errors and uncertainties in observations and simulations at 183 GHz. *Atmospheric Measurement Techniques*, 9(5), 2207–2221. <https://doi.org/10.5194/amt-9-2207-2016>
- Buehler, S. A., Courcoux, N., & John, V. O. (2006). Radiative transfer calculations for a passive microwave satellite sensor: Comparing a fast model and a line-by-line model. *Journal of Geophysical Research*, 111, D20304. <https://doi.org/10.1029/2005JD006552>
- Buehler, S. A., Kuvatov, M., Sreerekha, T. R., John, V. O., Rydberg, B., Eriksson, P., & Notholt, J. (2007). A cloud filtering method for microwave upper tropospheric humidity measurements. *Atmospheric Chemistry and Physics*, 7(21), 5531–5542. <https://doi.org/10.5194/acp-7-5531-2007>
- Buehler, S. A., Mendrok, J., Eriksson, P., Perrin, A., Larsson, R., & Lemke, O. (2018). Arts, the Atmospheric Radiative Transfer Simulator—Version 2.2, the planetary toolbox edition. *Geoscientific Model Development*, 11(4), 1537–1556. <https://doi.org/10.5194/gmd-11-1537-2018>

- Clough, S., Kneizys, F., & Davies, R. (1989). Line shape and the water vapor continuum. *Atmospheric Research*, 23(3-4), 229–241. [https://doi.org/10.1016/0169-8095\(89\)90020-3](https://doi.org/10.1016/0169-8095(89)90020-3)
- Eriksson, P., Buehler, S., Davis, C., Emde, C., & Lemke, O. (2011). ARTS, the Atmospheric Radiative Transfer Simulator, version 2. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(10), 1551–1558. <https://doi.org/10.1016/j.jqsrt.2011.03.001>
- Garand, L., Turner, D. S., Larocque, M., Bates, J., Boukabara, S., Brunel, P., et al. (2001). Radiance and jacobian intercomparison of radiative transfer models applied to HIRS and AMSU channels. *Journal of Geophysical Research*, 106(D20), 24,017–24,031. <https://doi.org/10.1029/2000JD000184>
- Goldberg, M., Kilcoyne, H., Cikanek, H., & Mehta, A. (2013). Joint polar satellite system: The United States next generation civilian polar-orbiting environmental satellite system. *Journal of Geophysical Research: Atmospheres*, 118, 13,463–13,475. <https://doi.org/10.1002/2013JD020389>
- Goldberg, M., Ohring, G., Butler, J., Cao, C., Datla, R., Doelling, D., et al. (2011). The global space-based intercalibration system. *Bulletin of the American Meteorological Society*, 92, 467. <https://doi.org/10.1175/2010BAM2967.1>
- Hersbach, H., Bell, W., & Berrisford, P. (2019). Era5: State-of-the-art global atmospheric reanalysis at ECMWF. *The title of the book*. Phoenix Convention Center in Phoenix, AZ: American Meteorological Society (AMS). Retrieved from <https://ams.confex.com/ams/2019Annual/meetingapp.cgi/Paper/351182>
- John, V. O., & Buehler, S. A. (2004). The impact of ozone lines on AMSU-B radiances. *Geophysical Research Letters*, 31, L21108. <https://doi.org/10.1029/2004GL021214>
- Kazumori, M., & English, S. J. (2015). Use of the ocean surface wind direction signal in microwave radiance assimilation. *Quarterly Journal of the Royal Meteorological Society*, 141(689), 1354–1375. <https://doi.org/10.1002/qj.2445>
- Leslie, R. V., Blackwell, W. J., Anderson, K., Kim, E. J., & Weng, F. (2013). S-NPP Advanced Technology Microwave Sounder: Reflector emissivity model, mitigation, & verification. *2013 IEEE international geoscience and remote sensing symposium, IGARSS 2013, melbourne, australia, july 21-26, 2013* (pp. 1927–1929). Melbourne, Australia: IEEE. <https://doi.org/10.1109/IGARSS.2013.6723182>
- Liebe, H. J. (1985). An updated model for millimeter wave propagation in moist air. *Radio Science*, 20(5), 1069–1089. <https://doi.org/10.1029/RS020i005p01069>
- Liebe, H. J. (1989). MPM—An atmospheric millimeter-wave propagation model. *International Journal of Infrared and Millimeter Waves*, 10(6), 631–650. <https://doi.org/10.1007/bf01009565>
- Liebe, H., & Hufford, G. (1993). MPM93—Propagation model of moist air at frequencies below 1000 GHz. In *Agard conf. proceedings*. Melbourne, Australia.
- Liebe, H. J., Hufford, G. A., & Cotton, M. G. (1993). Propagation modeling of moist air and suspended water/ice particles at frequencies below 1000 GHz.
- Liebe, H., Rosenkranz, P., & Hufford, G. (1992). Atmospheric 60-GHz oxygen spectrum: New laboratory measurements and line parameters. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 48(5-6), 629–643. [https://doi.org/10.1016/0022-4073\(92\)90127-p](https://doi.org/10.1016/0022-4073(92)90127-p)
- Liu, Q., Weng, F., Han, Y., & van Delst, P. (2008). Community radiative transfer model for scattering transfer and applications, *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium* (Vol. 4, pp. 1193–1196). Boston, Massachusetts, USA: IEEE.
- Lupu, C., Geer, A. J., & Bormann, N. (2015). Revision of the microwave coefficient files in the IFS 749. Retrieved from <https://www.ecmwf.int/node/10861>
- Matricardi, M. (2009). Technical Note: An assessment of the accuracy of the RTTOV fast radiative transfer model using IASI data. *Atmospheric Chemistry and Physics*, 9(18), 6899–6913. <https://doi.org/10.5194/acp-9-6899-2009>
- Melsheimer, C., Verdes, C., Buehler, S. A., Emde, C., Eriksson, P., Feist, D. G., et al. (2005). Intercomparison of general purpose clear sky atmospheric radiative transfer models for the millimeter/submillimeter spectral range. *Radio Science*, 40, RS1007. <https://doi.org/10.1029/2004RS003110>
- Mlawer, E. J., Payne, V. H., Moncet, J.-L., Delamere, J. S., Alvarado, M. J., & Tobin, D. C. (2012). Development and recent evaluation of the mt_ckd model of continuum absorption. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1968), 2520–2556.
- Moradi, I., Arkin, P., Ferraro, R., Eriksson, P., & Fetzer, E. (2016). Diurnal variation of tropospheric relative humidity in tropical regions. *Atmospheric Chemistry and Physics*, 16(11), 6913–6929. <https://doi.org/10.5194/acp-16-6913-2016>
- Moradi, I., Ferraro, R., Eriksson, P., & Weng, F. (2015). Intercalibration and validation of observations from ATMS and SAPHIR microwave sounders. *IEEE Transactions on Geoscience and Remote Sensing*, 53(11), 5915–5925. <https://doi.org/10.1109/TGRS.2015.2427165>
- Moradi, I., Soden, B., Ferraro, R., Arkin, P., & Vömel, H. (2013). Assessing the quality of humidity measurements from global operational radiosonde sensors. *Journal of Geophysical Research: Atmospheres*, 118, 8040–8053. <https://doi.org/10.1002/jgrd.50589>
- Poli, P., Healy, S. B., & Dee, D. P. (2010). Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 136(653), 1972–1990. <https://doi.org/10.1002/qj.722>
- Rosenkranz, P. (1975). Shape of the 5 mm oxygen band in the atmosphere. *IEEE Transactions on Antennas and Propagation*, 23(4), 498–506. <https://doi.org/10.1109/tap.1975.1141119>
- Rosenkranz, P. (1998). Water vapor microwave continuum absorption: A comparison of measurements and models. *Radio Science*, 33(4), 919–928. <https://doi.org/10.1029/98rs01182>
- Rothman, L., Gordon, I., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P., et al. (2013). The HITRAN2012 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 130, 4–50. <https://doi.org/10.1016/j.jqsrt.2013.07.002>
- Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., et al. (2018). An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11, 2717–2737. <https://doi.org/10.5194/gmd-11-2717-2018>
- Saunders, R., Matricardi, M., & Brunel, P. (1999). An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quarterly Journal of the Royal Meteorological Society*, 125(556), 1407–1425. <https://doi.org/10.1002/qj.1999.49712555615>
- Saunders, R., Rayer, P., Brunel, P., von Engel, A., Bormann, N., Strow, L., et al. (2007). A comparison of radiative transfer models for simulating Atmospheric Infrared Sounder (AIRS) radiances. *Journal of Geophysical Research*, 112, D01S90. <https://doi.org/10.1029/2006JD007088>
- Skofronick-Jackson, G., Petersen, W. A., Berg, W., Kidd, C., Stocker, E. F., Kirschbaum, D. B., et al. (2017). The global precipitation measurement (GPM) mission for science and society. *Bulletin of the American Meteorological Society*, 98(8), 1679–1695. <https://doi.org/10.1175/BAMS-D-15-00306.1>
- Tretyakov, M. Y., Koshelev, M. A., Dorovskikh, V. V., Makarov, D. S., & Rosenkranz, P. W. (2005). 60-GHz oxygen band: Precise broadening and central frequencies of fine-structure lines, absolute absorption profile at atmospheric pressure, and revision of mixing coefficients. *Journal of Molecular Spectroscopy*, 231(1), 1–14. <https://doi.org/10.1016/j.jms.2004.11.011>
- Turner, E., Rayer, P., & Saunders, R. (2019). AMSUTRAN: A microwave transmittance code for satellite remote sensing. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 227, 117–129. <https://doi.org/10.1016/j.jqsrt.2019.02.013>

- Weng, F., Zou, X., Sun, N., Yang, H., Tian, M., Blackwell, W. J., et al. (2013). Calibration of Suomi National Polar-Orbiting partnership Advanced Technology Microwave Sounder. *Journal of Geophysical Research: Atmospheres*, *118*, 11,187–11,200. <https://doi.org/10.1002/jgrd.50840>
- Yang, H., Weng, F., & Anderson, K. (2016). Estimation of ATMS antenna emission from cold space observations. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4479–4487. <https://doi.org/10.1109/TGRS.2016.2542526>
- Zhang, P., Holmlund, K., Goldberg, M., & Lafeuille, J. (2016). The Global Space-based Inter-Calibration System (GSICS), 2016 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 5522–5523). Beijing, China: IEEE. <https://doi.org/10.1109/IGARSS.2016.7730440>
- Zhou, L., Divakarla, M., & Liu, X. (2016). An overview of the Joint Polar Satellite System (JPSS) science data product calibration and validation. *Remote Sensing*, *8*(2), 139–151.