Aumann Hartmut, H (Orcid ID: 0000-0002-4311-7546)
Havemann Stephan (Orcid ID: 0000-0002-3259-091X)
Huang Xianglei (Orcid ID: 0000-0002-7129-614X)
Liu Xu (Orcid ID: 0000-0002-0473-3143)
Liuzzi Giuliano (Orcid ID: 0000-0003-3638-5750)
Masiello Guido (Orcid ID: 0000-0002-7986-8296)
Matricardi Marco (Orcid ID: 0000-0001-7514-9473)
Moradi Isaac (Orcid ID: 0000-0003-2194-1427)
Natraj Vijay (Orcid ID: 0000-0003-3154-9429)
Serio Carmine (Orcid ID: 0000-0002-5931-7681)
Strow L, Larrabee (Orcid ID: 0000-0001-5999-3519)

Evaluation of Radiative Transfer Models with Clouds

Hartmut H. Aumann[1], Xiuhong Chen[2], Evan Fishbein[1], Alan Geer[3], Stephan Havemann[4],Xianglei Huang[2], Xu Liu[5], Giuliano Liuzzi[6], Sergio DeSouza-Machado[7], Evan M. Manning[1], Guido Masiello[6], Marco Matricardi[3], Isaac Moradi[8], Vijay Natrai[1], Carmine Serio[6], Larrabee Strow[7], Jerome Vidot[9], R. Chris Wilson[1], Wan Wu[5], Qiguang Yang[5] ,Yuk L.Yung[1, 10]

1) Jet Propulsion Laboratory, California Institute of Technology
2) Department of Climate and Space Science, U. Michigan
3) European Center for Medium range Weather Forecasting
4) UK Met Office, Exeter, UK
5) NASA Langley Research Center, Hampton, VA
6) School of Engineering, University of Basilicata, Italy
7) Department of Physics, University of Maryland Baltimore County, MD
8) ESSIC, University of Maryland, College Park, MD
9) Meteo France, Toulouse, France
10) Department of Planetary Science, California Institute of Technology

Corresponding author: Hartmut Aumann (aumann@jpl.nasa.gov)

Key Points:

In the 900 cm^{-1} atmospheric window channel several Radiative Transfer Models have less than 2 K bias relative to collocated observations by the Atmospheric Infrared Sounder and a better than 0.95 correlation between the histogram derived from the observations and those derived from the calculations. Differences in the bias between observations and calculations for the 2616 cm^{-1} atmospheric window channel at night are not inconsistent with results at 900 cm^{-1}. For day time data the differences are much larger due to differences in the way scattering of the solar reflected light is treated.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2017JD028063

Differences in the cloud physics and cloud overlap assumptions between Radiative Transfer Models result in a standard deviation of the pairwise difference of between 6 and 12 K. Differences due to the cloud overlap assumption alone results in a 3 K standard deviation, an order of magnitude larger than the uncertainty of the observations.

The radiative effects of a bias in the cloud model in the European Center for Medium range Weather Forecasting are much less than the scatter in the differences between the AIRS observations and Radiative Transfer Models calculations or the differences between calculations by different Radiative Transfer Models for the same clouds.

Abstract

Data from hyperspectral infrared sounders are routinely ingested worldwide by National Weather Centers (NWCs). The cloud-free fraction of this data is used for initializing forecasts which include temperature, water vapor, water cloud and ice cloud profiles on a global grid. Although the data from these sounders are sensitive to the vertical distribution of ice and liquid water in clouds, this information is not fully utilized. In the future, this information could be used for validating clouds in NWC models and for initializing forecasts. We evaluate how well the calculated radiances from hyperspectral Radiative Transfer Models (RTMs) compare to cloudy radiances observed by AIRS and to one another. Vertical profiles of the clouds, temperature and water vapor from ECMWF (European Center for Medium-range Weather Forecasting) were used as input for the RTMs. For non-frozen ocean day and night data, the histograms derived from the calculations by several RTMs at 900 cm^{-1} have a better than 0.95 correlation with the histogram derived from the AIRS observations, with a bias relative to AIRS of typically less than 2 K. Differences in the cloud physics and cloud overlap assumptions result in little bias between the RTMs, but the standard deviation of the differences ranges from 6 to 12 K. Results at 2616 cm^{-1} at night are reasonably consistent with results at 900 cm^{-1}. Except for RTMs which use full scattering calculations, the bias and histogram correlations at 2616 cm^{-1} are inferior to those at 900 cm^{-1} for daytime calculations.

1. Introduction

Clouds are a key component of the Earth's weather and climate system. The data from hyperspectral infrared sounders have the information content to sense the vertical distribution of temperature and water vapor in clear air and of ice and liquid water inside semi-transparent clouds. The data from four hyperspectral sounders in polar orbit are routinely ingested by the National Weather Centers (NWCs) (e.g. Collard and McNally, 2009): The Atmospheric Infrared Sounder (AIRS, Aumann et al., 2003) on the Earth Observing System (EOS) Aqua satellite, the Crosstrack Infrared Sounder (CrIS, Glumb et al., 2003) on the Suomi National Polar-orbiting Platform (SNPP) satellite, and the Infrared Atmospheric Sounder Interferometer (IASI, Blumstein et al., 2008, Hilton et al., 2012) on MetOp A and B each makes more than two million observations of the state of the atmosphere and the clouds each day. The NWCs predominantly use the cloud-free portion of these data to initialize forecasts that provide temperature, water vapor, water cloud and ice cloud profiles on a global grid every three hours. Using cloudy observations in forecast models is difficult (Errico et al., 2007, Bennartz and Greenwald, 2011) and although all-sky microwave radiances are now used (e.g. Geer et al., 2017), the use of infrared radiances represents a harder problem. NWCs make use of some cloudy scenes, such as low-level cloud or fully overcast scenes, but the cloud information is still not used to initialize forecasts (Guidard et al., 2011, Lavanant et al., 2011). A number of NWCs and university research groups have developed fast and accurate Radiative Transfer Models (RTMs) for infrared sounders, which include the effects of cloud and aerosol scattering. The names and associated organizations of the RTM developers are summarized in Table 1. Summaries of the RTMs are found in Appendix 3. While each RTM has been subject to its own validation, our paper is the first to compare results from major RTMs for cloudy hyperspectral infrared applications on the same data set to collocated observations and to each other.

The objective of our paper was to evaluate the degree to which the radiative effects of clouds in NWC models agree with collocated hyperspectral observation. The availability of RTMs with a high degree of radiometric fidelity relative to observation, or at least the availability of tools to assess this fidelity, are expected to lead to the increased utilization of hyperspectral sounder data in the forecast. We selected AIRS observations and AIRS RTMs for our analysis to follow the Saunders et al., (2007) RTM analysis under cloud-free conditions.

2. Data, Participants and Evaluation

2.1. Data

We selected data provided by the European Center for Medium-range Weather Forecasting (ECMWF, 2009) as representative for the definition of the atmospheric states with clouds. The ECMWF description of the atmospheric state (temperature, water vapor and cloud vertical profiles, and surface temperature) has been widely documented and validated (e.g., Tiedtke 1989, Tiedtke 1993, Tompkins et al., 2007, Köhler et al., 2011, Kazumori et al., 2016). Details are in Appendix 1.

For the inter-comparison of RTMs we used AIRS observations from March 1, 2009 and the matching atmospheric state defined by ECMWF. A subset of this data was created using the difference between the

ECMWF estimate of the surface temperature (stemp) and the brightness temperature measured in the 1231 cm^{-1} window channel (bt1231), (stemp-bt1231). This difference is a measure of the radiometric effect of clouds. Under clear conditions the difference is less than 2 K, but the difference can increase to as much as 100 K in the presence of cold clouds in the tropics. We limited the size of this dataset to control the magnitude of the computational effort involved in scattering calculations by using stratified sampling. This method selected a representative mix of cloudy conditions from the AIRS data, which resulted in 7377 unique cases. The surface emissivity and surface reflectance were obtained from a monthly climatology (Zhou et al., 2012). The surface reflectance was assumed to be Lambertian. Details of the selection algorithm are given in Appendix 2.

2.2. Participants and RTM methodology

Table 1 summarizes the affiliation of the developers of cloud-capable RTMs at major NWCs, government and university facilities. Six RTMs were used: (1) SARTA, (2) RTTOV, (3) HT-FRTC, (4) PCRTM, (5) CRTM, and (6) σ-IASI. Largely based on discussions at the 2016 AGU meeting, every RTM team, except the RTTOV teams, submitted revised results. Five of the RTM developers generated variants related to details of how cloud overlap, cloud type and scattering were handled. Details on the individual RTMs are summarized in Appendix 3.

All RTMs calculated cloudy radiances using a linear combination of clear sky calculations and scattering calculations for one or more cloud columns. The results of the clear sky column calculations from all these RTMs were nearly identical, consistent with Saunders et al., (2007). The cloudy spectra were calculated as the linear combination of clear and cloudy columns based on the cloud fraction. The Maximum Overlap (MO) model is the simplest case:

$$R_{MO} = CF\, R_{FO} + (1 - CF)\, R_{CLR} \tag{1}$$

R_{MO} is the spectrum calculated with the MO assumption and CF is defined as the maximum cloud fraction in the cloud coverage profile. Some RTMs set CF equal to the total cloud cover (tcc) specified in the ECMWF record. R_{CLR} is the clear sky spectrum, and R_{FO} is the spectrum assuming full overcast (i.e., clouds fill the entire satellite footprint). Some RTMs allow the user to make more complicated overlap assumptions. The Maximum Random Overlap (MRO) assumption states that any continuous vertical cloud profile is maximally overlapped, and the discontinuous parts of the vertical cloud profile are randomly overlapped (Hogan and Illingworth, 2000). If two cloud slabs are used, the MRO radiance is:

$$R_{MRO} = CF_1\,(1 - CF_2)\, R_{C1} + CF_2\,(1 - CF_1)\, R_{C2} + CF_1\, CF_2\, R_{CLD} + (1 - CF_1)\,(1 - CF_2)\, R_{CLR} \tag{2}$$

where CF_1 is the maximum cloud fraction of the first cloud slab, CF_2 is the maximum cloud fraction of the second cloud slab, R_{C1} is the calculation where only the first cloud slab is included, R_{C2} is the calculation where only the second cloud slab is included, R_{CLD} is the calculation where both clouds are included, and R_{CLR} is the clear sky calculation. There are several variants of the overlap assumption, including Maximum Overlap (MO), Exponential Random Overlap (ERO) and the Random Overlap (RO). We indicated these variants in the names of the models, e.g. CRTM_mro is the cloudy spectrum calculated using CRTM with the MRO assumption. Most RTM developers submitted results with a number of variants.

2.3. Evaluation

The inter-comparison of the RTMs used three methods:

1) The pairwise comparison of the observed AIRS spectra with the calculated spectra: We calculated the mean and standard deviation (stddev). This comparison is complicated by several factors: a) The collocation error: The location and local time of the AIRS data obtained with a 12-km footprint (effectively 1/8 degree lat/lon in the tropics) is not a good match to the temporal (3 hour) and spatial grid (approximate 25 km) of the ECMWF data available to this study; b) The total cloud cover is specified in the ECMWF data, the cloud fraction, liquid water and ice water content are specified for each level, but the cloud overlap is not specified. Each RTM can handle the cloud overlap with different assumptions. c) The ECMWF description of the cloud in 91 levels is itself subject to random and systematic errors. d) The liquid water and ice cloud particle size distributions are not directly specified. For a sufficiently large data set, factor (a) should have zero bias, but will cause a large stddev. Factors (b), (c) and (d) may create a bias as well as a large stddev.

2) Characterization of the radiometric effect of clouds using histograms of (stemp-bt): Here stemp is the surface temperature from ECMWF and bt is the AIRS observed or RTM calculated brightness temperature in an atmospheric window channel. In the absence of a solar reflected component, (stemp-bt) increases from near zero under clear conditions to 100 K with increasing cloudiness. Under ideal conditions of a perfect matchup between AIRS and ECMWF, perfect clouds and thermodynamic profiles in the ECWMF model, and a perfect RTM, the two histograms will be identical. We evaluate the closeness of the match between observations and calculations by calculate the histogram correlation. Small residual biases that result from compensatory large positive and negative differences between AIRS and the RTM calculations under different conditions of cloudiness (or cloud types) are revealed as distortions of the histograms, resulting in a lowered correlation with the observations. This approach is not sensitive to random errors in the ECMWF cloud forecasts (e.g. the miss-location of clouds) but remains vulnerable to systematic errors; nevertheless, known systematic errors in ECMWF cloud forecasts are globally infrequent and limited to specific meteorological conditions (Kazumori et al., 2016). Infrequent ECMWF cloud errors are not likely to impact the histogram correlation because of the wide variety of cloud conditions in our data set.

3) The pairwise comparison of results from different RTMs: This approach has the advantage that it sidesteps matchup uncertainties with ECMWF. All RTMs use the same cloud model description. The comparison reveals the radiometric effect of differences between RTMs in cloud microphysics assumptions, cloud overlap assumptions, and scattering algorithms.

3. Results

Figure 1 illustrates typical spectral patterns in the mean of the pairwise difference between AIRS spectra and the spectra calculated using six representative RTMs for 1377 night and 1437 day non-frozen ocean cases, respectively. For wavenumbers lower than 1700 cm^{-1} there is a relatively day/night independent, spectrally correlated pattern in the mean of the difference between AIRS and different RTMs. For wavenumbers above 2200 cm^{-1} the pattern for the different RTMs is inconsistent even for the nighttime data. This inconsistency is even larger for the daytime data due to the differences in the way the RTMs deal with scattering and solar reflected radiation. For the detailed evaluation of the RTMs we focus on the two representative atmospheric window channels at 900 and 2616 cm^{-1}.

Figure 2 illustrates the comparison between AIRS and three RTMs at 900 cm^{-1} for the same two cases as Figure 1 using histograms of (stemp-bt900). The peak of the histogram in all cases is near +10 K, i.e., relatively little cloudiness or low clouds. The coldest cloud tops are 100 K colder than stemp. The black trace is derived from the AIRS observation. HT_CRO and HT_SRO results are nearly identical, even though HT_CRO uses Chou scaling, while HT_SRO uses a full scattering calculation (both RTMs use the RO assumption). SARTA traced the AIRS histogram somewhat better than either HT variant for nighttime cases with (stemp-bt900) between 40 and 70 K, with SARTA finding many more cases than AIRS, while the HT RTMs had less cases. SARTA uses Chou scaling and the RO assumption similar to HT_CRO. However, SARTA takes multi-leveled clouds and converts them into two single layer clouds, one for ice clouds and one for water clouds.

For daytime cases (Figure 2b), both SARTA and HT RTMs deviated from the AIRS trace for (stemp-bt900) cases between 15 and 70 K. Similar to the nighttime case, SARTA had more cases than AIRS in this range, while both versions of HT had less cases than AIRS between 15 and 30 K.

Figure 3 illustrates the histogram comparison using the same three RTMs and the same day and night non-frozen ocean cases as Figure 2, but at 2616 cm^{-1}. In this figure the agreements between the RTMs and AIRS at night is much better than during the day. In fact, the histograms from SARTA and both HT RTMs are more closely matched to the AIRS histogram for cases between 15 and 70 K than the histograms at 900 cm^{-1} (Figure 2a). During the day the full scattering calculations used by HT_SRO result in histograms without the long tail for high clouds (stemp-bt2616 > 60 K) seen in the HT_CRO histogram.

3.1. Numerical summary of histogram correlations, bias and standard deviations

Table 2 summarizes the histogram correlations and the bias relative to AIRS for the day and night non-frozen ocean cases at 900 cm^{-1}. The results shown are separated into six groups, with each group representing the six RTM developers. The histogram correlations with AIRS exceeded 0.95 for five of the six groups for day and night cases. The histogram correlations were slightly higher for the night cases than for the day cases. This observation may be related to a day/night dependence of the cloud structure or microphysics properties. A more in depth study of his observation is beyond the scope of this paper. The row labeled “clear column RTM” used SARTA without clouds. The mean bias relative to AIRS for all RTMs was +0.3 K (range -4.2 K to +2.2 K) for the day cases, and +0.8 K (range -3.6 K to +2.8 K) for the night cases. The number following the ± in Table 2 is the standard deviation of the differences, typically 22 K for the day, 20 K for the night cases. Using the typical standard deviation for nighttime, the probable error in the mean is about $22/\sqrt{1377} = 0.6$ K (assuming random sampling and uncorrelated errors). Bias differences of more than three times the probable error, 2 K, are significant.

When the RTM calculations are compared to one another the effect of the collocation error is eliminated, since we are comparing calculations for the same cloud conditions. Results are shown in Table 3 for the 1437 day non-frozen ocean cases using four RTMs as references: CRTM_mro, HT_SMRO, PCRTM_MRO4 and RTTOV_MRO. For the “5 group” summary in the last row we used only the first entry from each group, excluding the RTM used as the reference. The bias between RTMs was -0.2 K, +1.5 K, +0.3 K and -3.4 K, and the standard deviations were 12.3 K, 7.6 K, 8.7 K and 6.5 K, respectively for the four reference RTMs.

In order to explore the extent to which a small bias on a global scale may be the result of compensating biases, we divided the data into latitude zones. Results are summarized in Table 4 for the tropical zone (|lat| < 30 degree) with 3644 cases; 2662 cases for the extratropical zone (|lat| > 30) degree, limited to non-frozen surface cases using stemp > 275 K; and 1070 cases from the polar zone (|lat| > 60) degree. Based on the last column in Table 4, which shows the difference between the mean tropical and the mean mid-latitude bias, some RTMs show a latitude dependence in the bias of several degree K.

Table 5 summarizes the results at 2616 cm^{-1} for day and night non-frozen ocean. The bottom row summarizes the results in terms of a mean bias and stddev, excluding RTMs with less than 0.9 histogram correlation with AIRS. At night 16 of the 18 RTMs had a mean bias relative to AIRS of -1.8 K with 20 K stddev and 5 of the six RTM teams produced results with histogram correlation with AIRS better than 0.95. During the day only two of the six RTM teams produced results with histogram correlations with AIRS better than 0.95. For these cases the mean bias was +2.4 K, and the mean stddev was 16 K.

4. Discussion of the Results

Under clear conditions at night, SARTA, PCRTM and RTTOV have previously been shown to agree with each other and with AIRS within 0.05 K bias and 0.1 K standard deviation (Saunders et al., 2007). The current versions of SARTA, PCRTM and RTTOV, including CRTM, σ-IASI and HT-FRTC under clear conditions all have shown the same level of agreement. Under cloudy conditions the agreement is not as close and the magnitudes of the differences are wavenumber dependent.

4.1. The longwave region

We define the longwave region as portions of the spectrum with wavenumbers lower than 1700 cm^{-1}. In this region, the differences between AIRS and RTMs are relatively day/night independent (Figure 1) and range between ±4 K. The correlation between the histogram calculated from the AIRS observations and the histograms calculated for several of the RTMs at 900 cm^{-1} (Table 2 and Figure 2) exceeds 0.97. The bias averaged over 17 of 18 RTMs at 900 cm^{-1} for the day and night non-frozen oceans was +0.3 K and +0.8 K respectively (Table 2). However, Table 4 shows that the low bias for the non-frozen oceans for some RTMs was due to the compensating effects of a bias for the tropical zone balanced by a bias of the opposite signs in the mid-latitude and polar zones. A high correlation between the observed and calculated histograms and a zone independent low bias are a measure of the skill of the RTM and the statistical fidelity of the ECMWF specification of the atmospheric state, including clouds. The bias between the clear column RTM calculations and AIRS observations is more than -25 K (Table 2). The typically 20 K standard deviation of the difference between AIRS and the RTMs is essentially the same with and without clouds. This indicates that the high standard deviation of (AIRS-RTM) is dominated by the mismatch between the clouds observed by AIRS and the clouds specified by the ECMWF model.

RTTOV_MRO is biased about 4 K high relative to AIRS (Table 2). As shown in Table 3, where we compare the RTMs to reference RTMs, all RTMs are biased low relative to RTTOV_MRO. RTTOV_MRO used the OPAC cumulus cloud type option. RTTOV_CMSS used the RO overlap scheme and was optimized for mid to upper tropospheric sounding channels, not for window channels. RTTOV_CMSS shows much less bias relative to AIRS than RTTOV_MRO in the global analysis (Table 2), but shows a large bias of opposite signs for the zonal bias (Table 4). These observations suggest that a combination of cloud types (derived from the cloud, temperature or water vapor profiles) may produce a closer match to observations than choosing one cloud type. Future work will examine the impact of cloud type assumptions on the RTMs' match to observations.

The typical standard deviation of the RTMs relative to AIRS, 20 K, decreases to a range of 6 to 12 K when the RTMs are compared to one another (Table 3), excluding siblings within the same RTM group. Since the RTMs used the same cloud input profile, the decrease from 20 K to 12 K (or less) is dominated by the elimination of the collocation error. The standard deviations of the differences are in this case related to differences in the way the ECMWF cloud description is converted to cloud microphysical parameters and then to the radiances calculated by the RTMs. When this conversion is identical, as in the case of CRTM_mro and CRTM_2col, the standard deviation of the difference was 3 K, and was due to the difference in the overlap assumption alone.

4.2. The shortwave region

We define the shortwave region as portions of the spectrum with wavenumbers greater than 2000 cm^{-1}. By inspection of Figure 1 we already noted that the spectral patterns for the different RTMs are less consistent in the shortwave region than in the longwave region, even for the night data. However, at night three of the six RTMs (Figure 1) have less than 2 K bias relative to AIRS and five of the six RTMs have histogram correlations better than 0.95 (Table 5). At night the histogram correlations at 2616 cm^{-1} are not inconsistent with those at 900 cm^{-1}. During the day the results from only two of the six RTM teams reached histogram correlations better than 0.95, and the bias values relative to AIRS were much larger than those at night and the result at 900 cm^{-1}. If cloud scattering parameters at 2616 cm^{-1} were too weak, it would have impacted the night calculations as well, but the night calculations (for three of the RTMs) agree reasonably well with AIRS. The degradation of the results during the day is probably related to the use of Chou scaling, which was not designed for shortwave calculations (Chou et al., 1999). This is clearly shown in the comparison of HT_CRO (with the Chou approximation) and HT_SRO (using full scattering) in Figure 3. The tail of the histogram of HT_CRO at (stemp-bt2616) extends all the way to 100 K for the coldest clouds, while the tail of the histogram of HT_SRO stops at 60 K. The PCRTM_MRO also uses full scattering. SARTA and RTTOV used Chou scaling, while CRTM used the advanced doubling adding method (Liu and Weng, 2006). Full scattering calculations usually are assumed to be costly in terms of computation time, but this need not be the case. For example, PCRTM performs full scattering calculations with multiple streams and multiple scatterings performed offline to generate lookup tables. However, even by employing full scattering, the HT_SRO histogram at 2616 cm^{-1} (Figure

3b) showed fewer clear or low cloud cases (stemp-bt2616 < 0 K) calculated by the RTMs than were observed by AIRS. This suggests that the reflectance from the Earth's surface has a stronger angular dependence than Lambertian scattering.

In the discussion of the histograms calculated from the RTMs at 900 cm^{-1} we noted the differences between AIRS observations and the RTMs for 40<(stemp-bt900)<70 K cases (Figure 2). This effect is much less pronounced at 2616 cm^{-1}, particularly at night (Figure 3a). We believe that in this region a significant portion of the scenes contains multiple cloud types. This may amplify any systematic bias that exists in the assumptions about cloud microphysics and their assumed spectral dependence. The interpretation of these differences in terms of cloud types is outside of the scope of this paper.

4.3. ECMWF cloud bias and RTM cloud bias

The bias in the RTMs relative to AIRS has two components:

1) The ECMWF cloud description is vulnerable to systematic errors.
2) The methodologies used by the RTMs to convert the cloud description into radiances are likely to contain assumptions which lead to systematic biases.

In order to quantify this bias we assume that the results from the six RTM teams represent plausible and sufficiently independent radiometric realizations of the cloud effects. At 900 cm^{-1} the RTMs show a bias in the range from +1.5K to -3.4K relative to one another (Table 3). Relative to AIRS the RTMs have a +0.3 K (+0.8 K) mean bias for the day (night) non-frozen oceans (Table 2). These results indicate that the radiative effect of a bias in the ECMWF clouds could be of the order of 1 K.

The difference between the mean tropical and mean mid-latitude bias (Table 4, last column) for each RTM could reveal a cloud type dependence in the ECMWF clouds or in the RTM cloud algorithms. Several of the six RTM groups have a zonal bias lower than 2 K. The low zonal bias seen in the results from these RTMs is consistent with the radiative effect of a zonal ECMWF cloud bias of less than 2 K. The observation that several of the RTMs achieve a low zonal bias and a high histogram correlation relative to the observations indicates that the low global bias is not the result of compensatory much larger cloud-type dependent biases.

We interpret the zonal bias seen in the three other RTM groups, which ranges from 4 K to 8 K, as a cloud-type dependence in those RTMs. The RTTOV_CMSS has a 5 K zonal bias, compared to 0.2 K for RTTOV_MRO (Table 4). On the other hand, RTTOV_CMSS shows much less bias relative to AIRS than RTTOV_MRO (Table 2). Both use a single (but different) cloud type. A combination of cloud types (derived from the cloud, temperature or water vapor profiles) may produce a closer match to observations than choosing one cloud type. A future analysis of cloud-type effects on the RTMs could include data acquired in cloudy conditions from other instruments.

5. Summary

The objective of our paper was to evaluate the degree to which the radiative effects of clouds in NWC models agree with collocated hyperspectral observations. We selected AIRS observations and AIRS RTMs for our analysis. We selected data provided by the European Center for Medium-range Weather Forecasting (ECMWF, 2009) as representative for the definition of the atmospheric states with clouds. . We used the bias and histogram correlations relative to AIRS observations for the 2616 and 900 cm^{-1} atmospheric window channels as performance metrics. For some RTMs the histogram calculated at 900 cm^{-1} has a correlation of better than 0.95 with the histogram derived from the AIRS observations, with a bias relative to AIRS of less than 2 K for non-frozen ocean day and night data. However, several of the six RTM groups showed between zero and 2 K bias between the tropical zone and the mid-latitude zone at 900 cm^{-1}, while others had a bias between 4 K and 8 K. This observation and the high histogram correlation with AIRS shows that the ECMWF cloud prescription may have a bias, but the radiative effect of the bias at 900 cm^{-1} is most likely less than 2 K, relatively insignificant compared to the bias introduced by some RTMs. The results for the 2616 cm^{-1} window channel are consistent with day and night results at 900 cm^{-1} only when full scattering calculations were used. For these cases the correlation between the histogram deduced from the AIRS observations and the histograms calculated by the RTMs exceeds 0.95 and the bias at night is less than 2 K relative to AIRS. During the day the AIRS observations at 2616 cm^{-1} are 2 to 4 K higher than the RTM calculations with full scattering. This suggests that the reflectance of the surface has a steeper angular dependence than Lambertian.

Our study created a testable dataset, baseline results, and testing methodology to support continuing RTM development, with the goal of increasing the utilization of hyperspectral observations in the forecast. As illustrated in Figure 1, there is no need (with the current state of the art RTMs) to make these calculations for all channels, since just one or two surface channels will provide valuable insights. However, the selection of a shortwave channel requires an RTM with full scattering. The choice of the RTM and how many channels to use comes down to computer resource requirements.

Acknowledgments

The work described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Work at JPL and UMBC was funded by NASA ROSES, with the longtime support of Dr. Ramesh Kakar of NASA Headquarters. Mr. Steve Broberg helped with technical editing. The work at LARC was supported by NASA NAST-I and CLARREO projects. The work at NASA/GMAO was funded under NASA Grant NNX17AE79A Goddard Space Flight Center Cooperative Agreement. The work at the University of Michigan was supported by NASA grant NNX15AC25G. The work at Meteo France was supported by the EUMETSAT NWP-SAF program. The work at the UK Met Office was supported as part of the science program theme “Improved use of Satellite Data”. The 7377 atmospheric states and associated AIRS spectra used for this paper can be found at ftp://thunder.jpl.nasa.gov/hha/Cloudy_RTA. The file readme.20160518.txt defines various parameters. The file AIRS_SRF_m140f.mat defines the AIRS SRF for each of the 2378 channels in a MATLAB Version 7.0 format.

References

Alvarado, M. J., Payne, V. H., Mlawer, E. J., Uymin, G., Shephard, M. W., Cady-Pereira, K. E., Delamere, J. S., and Moncet, J.-L. (2013). Performance of the Line-By-Line Radiative Transfer Model (LBLRTM) for temperature, water vapor, and trace gas retrievals: recent updates evaluated with IASI case studies, Atmos. Chem. Phys., 13, 6687-6711, https://doi.org/10.5194/acp-13-6687-2013

Aumann, H. H., Chahine, M. T., Gautier, C., Goldberg, M., Kalnay, E., McMillin, L., Revercomb, H., Rosenkranz, P. W., Smith, W. L., Staelin, D. H., Strow, L. and Susskind, J. (2003). AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products and Processing Systems," *IEEE Trans. Geosci. Remote Sensing*, 41.2, 253-264.

Barker, H. W. (2008). Representing cloud overlap with an effective decorrelation length: An assessment using CloudSat and CALIPSO data. J. Geophys. Res., 113, D24205, doi:10.1029/2008JD010391.

Bauer, P., Lopez, P., Salmond, D., Benedetti, A., Saarinen, S., & Bonazzola, M. (2006). Implementation of 1D+ 4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var. Q.J.R. Meteorol. Soc., 132(620). 2307-2332.

Baum, B., P. Yang, A. Heymsfield, C. Schmitt, Y. Xie, A. Bansemer, Y.-X. Hu, and Z. Zhang (2011). Improvements to shortwave bulk scattering and absorption models for the remote sensing of ice clouds, J. Appl. Meteor. Clim, 50, 1037-1056.

Baran, A.J., R. Cotton, S. Havemann, K. Furtado, F. Marenco, A. Smith, J.C. Thelen, L.C. Labonnote (2014). A self-consistent scattering model for cirrus. II. The high and low frequencies. Q.J.R.Meteorol. Soc, VOL. 140, PP. 1039-1057.

Bechtold, P., Chaboureau, J. P., Beljaars, A., Betts, A. K., Köhler, M., Miller, M., & Redelsperger, J. L. (2004). The simulation of the diurnal cycle of convective precipitation over land in a global model. Q.J.R. Meteorol. Soc., 130, (604). 3119-3137.

Bennartz, R., and T. Greenwald (2011). Current problems in scattering radiative transfer modelling for data assimilation, Q.J.R. Meteorol. Soc., 137(661), 1952-1962, doi:10.1002/qj.953.

Bhawar, R., et al., (2008). Spectrally resolved observations of atmospheric emitted radiance in the H2O rotation band, Geophys. Res. Lett., 35, L04812, doi:10.1029/2007GL032207.

Blumstein, D., Chalon, G., Carlier, T., Buil, C., Hebert, Ph., Maciaszek, T., Ponce, G., Phulpin, T., Tournier, B., Simeoni, D., Astruc, P., Clauss, A., Kayal, G., and Jegou, R. (2008). IASI instrument: Technical overview and measured performances, Infrared Spaceborne Remote Sensing XII, M.Strojnik Editor, Proc SPIE 5543, 196-207.

Bower, K.N., T.W. Choularton, J. Latham, J. Nelson, M.B. Baker, and J. Jensen, (1994). A Parameterization of Warm Clouds for Use in Atmospheric General Circulation Models, J. Atmos. Sci., **51**, 2722–2732, https://doi.org/10.1175/1520-0469(1994)051<2722:APOWCF>2.0.CO;2

Chen, X., X. Huang, and X. Liu (2013). Non-negligible effects of cloud vertical overlapping assumptions on longwave spectral fingerprinting studies, J. Geophys. Res. Atmos., 118, 7309–7320, doi:10.1002/jgrd.50562.

Chou, M.-D., K.-T. Lee, S.-C. Tsay, and Q. Fu (1999). Parameterization for Cloud Longwave Scattering for Use in Atmospheric Models, J. Climate, 12, 159–169, doi:10.1175/1520-0442- 12.1.159.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown (2005). Atmospheric radiative transfer modeling: A summary of the AER codes, Short Communication, J. Quant. Spectrosc. Radiat. Transfer, 91, 233–244.

Collard, A. D., & McNally, A. P. (2009). The assimilation of infrared atmospheric sounding interferometer radiances at ECMWF. Q.J.R. Meteorol. Soc., 135(641), 1044-1058.

Deng, A., and D. R. Stauffer (2006). On improving 4-km mesoscale model simulations, J. Appl. Meteorol. Climatol., 45(3), 361–381, doi:10.1175/JAM2341.1.

DeSouza-Machado, S., Strow, L. L., Hannon, S. E., Motteler, H. E., López-Puertas, M., Funke, B., 964 and D.P. Edwards (2007). Fast forward radiative transfer modeling of 4.3 μm nonlocal 965 thermodynamic equilibrium effects for infrared temperature sounders, Geophys. Res. Lett., 34, 01802, 966 doi:10.1029/2006GL026684.

Ebert, E. E., and J. A. Curry (1992). A parameterization of ice cloud optical properties for climate models, J. Geophys. Res., 97(D4), 3831–3836, doi:10.1029/91JD02472.

ECMWF (2009) "IFS documentation Cy33r1", available from http://www.ecmwf.int/en/forecasts/documentation-and-support

Edwards, J.M. and A. Slingo (1996). Studies with a flexible new radiation code. I. Choosing a configuration for a large scale model, Q.J.R. Meteorol. Soc., 122, 689-719.

Errico, R. M., P. Bauer, and J. F. Mahfouf (2007). Issues regarding the assimilation of cloud and precipitation data, Journal of the Atmospheric Sciences, 64(11), 3785-3798, doi:10.1175/2006jas2044.1.

Fouquart, Y. (1987). Radiative transfer in climate models, in Physically Based Modelling and Simulation of Climate and Climate Changes, edited by M. E. Schlesinger, pp. 223–284, Kluwer Acad., Norwell, Mass.

Geer, A. J., Baordo, F., Bormann, N., Chambon, P., English, S. J., Kazumori, M., Lawrence, H., Lean, P., Lonitz, K. and Lupu, C. (2017). The growing impact of satellite observations sensitive to humidity, cloud and precipitation. Q.J.R. Meteorol. Soc. Accepted, doi:10.1002/qj.3172

Glumb, R. J., Williams, F. L., Funk, N., Chateauneuf, F., Roney, A., and Allard, R. (2003). "Cross-track Infrared Sounder (CrIS) development status," Proc. SPIE 5152.

Grieco, G., Masiello, M. Matricardi, C. Serio, D. Summa, and V. Cuomo (2007). Demonstration and validation of the IASI inversion scheme with NAST-I data Q.J.R. Meteorol. Soc., 133 (S3), 217–232 http://dx.doi.org/10.1002/qj.162

Guidard, V., N. Fourrié, P. Brousseau, and F. Rabier (2011). Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes, Q.J.R. Meteorol. Soc, 137(661), 1975-1987, doi:10.1002/qj.928.

Havemann, S., (2006). The development of a fast radiative transfer model based on an empirical orthogonal functions (EOF) technique, SPIE VOL. 6405, (2006), PP.348-358.

Hilton, F. et al., (2012). Hyperspectral Earth observation from IASI, five years of accomplishments, Bulletin of the American Meteorological Society, **93**, 347-370, doi:10.1175/BAMS-D-11-00027.1

Hogan, R. J. and Illingworth, A. J. (2000). Deriving cloud overlap statistics from radar. Q.J.R. Meteorol. Soc., 126, 2903–2909. doi:10.1002/qj.49712656914

Kazumori, M., Geer, A. J., & English, S. J. (2016). Effects of all-sky assimilation of GCOM-W/AMSR2 radiances in the ECMWF numerical weather prediction system. Q.J.R. Meteorol. Soc., 142, 721-737.

Köhler, M., Ahlgrimm, M. and Beljaars, A. (2011). Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model. Q.J.R. Meteorol. Soc., 137: 43–57. doi:10.1002/qj.713

Lavanant, L., et al., (2011). Comparison of cloud products within IASI footprints for the assimilation of cloudy radiances, Q.J.R. Meteorol. Soc., 137(661), 1988-2003, doi:10.1002/qj.917.

Liu, Q. and F. Weng, (2006). Advanced Doubling–Adding Method for Radiative Transfer in Planetary Atmospheres. *J. Atmos. Sci.,* **63**, 3459–3465, https://doi.org/10.1175/JAS3808.1

Liu, Q., X. Liu, Y. Liu, N. R. Nalli, and C. Tan (2017). Fast Radiative Transfer Algorithms for Real-Time Sounder Applications, Capter in book Reference Module in Earth Systems and Environmental Sciences, Elsevier Publishing, https://doi.org/10.1016/B978-0-12-409548-9.10391-4

Liu, X., W. L. Smith, D. K. Zhou, and A. Larar (2006). Principal component-based radiative transfer model for hyperspectral sensors: theoretical concept, Appl. Opt. 45, 201–209.

Liu, X., D. K. Zhou, A. M. Larar, W. L. Smith, P. Schluessel, S. M. Newman, J. P. Taylor, and W. Wu (2009). Retrieval of atmospheric profiles and cloud properties from IASI spectra using super-channels, Atmos. Chem. Phys., 9, 9121-9142, doi:10.5194/acp-9-9121-2009.

Liu, X., Q.Yang, H. Li, Z. Jin, W. Wu, S. Kizer, D. K. Zhou, and P. Yang (2016). Development of a fast and accurate PCRTM radiative transfer model in the solar spectral region, Appl. Opt. 55, 8236-8247

Liuzzi, G., G. Masiello, C. Serio, D. Meloni, C. Di Biagio, and P. Formenti (2017). Consistency of dimensional distributions and refractive indices of desert dust measured over Lampedusa with IASI radiances, Atmos. Meas. Tech., 10, 599–615, doi:10.5194/amt-10-599-2017.

Machado, S. D., and L. L. Strow (2016). "An Accurate Two-Slab Cloud-Representation Model for Hyperspectral Infrared Radiative Transfer Codes", AGU2016 (A11P-01).

DeSouza-Machado, S., Strow, L. L., Tangborn, A., Huang, X., Chen, X., Liu, X., Wu, W., and Yang, Q. (2018). Single-footprint retrievals for AIRS using a fast TwoSlab cloud-representation model and the SARTA all-sky infrared radiative transfer algorithm, Atmos. Meas. Tech., 11, 529-550, https://doi.org/10.5194/amt-11-529-2018.

Martinet, P., Fourrié, N., Guidard, V., Rabier, F., Montmerle, T. and Brunel, P. (2013). Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances. Q.J.R. Meteorol. Soc., 139: 1402–1416. doi:10.1002/qj.2046

Matricardi, M. (2005), The inclusion of aerosols and clouds in RT-IASI, the ECMWF fast radiative transfer model for the IASI, ECMWF Tech. Memo. 474, 53 pp., ECMWF, Reading, U.K.

Mlawer, E. J., V. H. Payne, J.-L. Moncet, J. S. Delamere, M. J. Alvarado, and D. D. Tobin (2012). Development and recent evaluation of the MT_CKD model of continuum absorption, *Phil. Trans. Roy. Soc. A*, 370, 1–37, doi:10.1098:rsta.2011.0295.

McNally, A. P. and Watts, P. D. (2003). A cloud detection algorithm for high-spectral-resolution infrared sounders. Q.J.R. Meteorol. Soc., 129: 3411–3423. doi:10.1256/qj.02.208

McNally, A. P. (2009). The direct assimilation of cloud-affected satellite infrared radiances in the ECMWF 4D-Var. Q.J.R. Meteorol. Soc., 135: 1214–1229. doi:10.1002/qj.426

Nalli, N. R., P. J. Minnett, P. van Delst, C. D. Barnet, and M. D. Goldberg (2006). Developments in ocean infrared emissivity/reflection modeling: Comparisons against observations. Preprints, 14th Conf. on Satellite Meteorology and Oceanography, Atlanta, GA, Amer. Meteor. Soc., P6.28. (2016) Available online at http://ams.confex.com/ams/pdfpapers/104810.pdf.

Okamoto, K., McNally, A. P. and Bell, W. (2014). Progress towards the assimilation of all-sky infrared radiances: an evaluation of cloud effects. Q.J.R. Meteorol. Soc., 140: 1603–1614. doi:10.1002/qj.2242

Ou, S. S. C., Kahn, B. H., Liou, K.-N. (2013). Retrieval of cirrus cloud properties from the Atmospheric Infrared Sounder: The k-coefficient approach combined with SARTA plus delta-four stream approximation, IEEE T. Geosci. Remote Sens., 51, 1010-1024.

Ou, S.C., K.N. Liou, Y.Y. Takano, N.X. Rao, Q.Q. Fu, A.J. Heymsfield, L.M. Miloshevich, B.B. Baum, and S.A. Kinne (1995). Remote Sounding of Cirrus Cloud Optical Depths and Ice Crystal Sizes from AVHRR Data: Verification using FIRE II IFO Measurements, J. Atmos. Sci., 52, 4143-4158, doi: 10.1175/1520-0469.

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

Saunders, R., Matricardi, M. and Brunel, P. (1999). An improved fast radiative transfer model for assimilation of satellite radiance observations. Q.J.R. Meteorol. Soc., 125: 1407–1425. doi:10.1002/qj.1999.49712555615

Saunders, R., et al., (2007). A comparison of radiative transfer models for simulating Atmospheric Infrared Sounder (AIRS) radiances, J. Geophys. Res., 112, D01S90, doi:10.1029/2006JD007088.

Segelstein, D. (1981).The complex refractive index of water, Master's thesis, University of Missouri, Kansas City.

Stamnes K., Tsay, S. C., Wiscombe, W., Jayaweera, K. (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Appl. Opt., 27, 2502-2509.

Strow, L. L., S.E. Hannon, S. DeSouze-Machado, H. E. Mottler and D. C. Tobin (2006). Validation of the atmospheric infrared sounder radiative transfer algorithm, J.Geophys. Res, 111 D09S06 doi:10.1029/2005JD006146.

Thelen, C. and J.M. Edwards (2013). Short-wave radiance comparisons between SEVERI and the Unified Model, Q.J.R. Meterol. Soc, VOL. 139, (2013), PP. 1665-1679

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Monthly Weather Review, 117(8), 1779-1800.

Tiedtke, M. (1993). Representation of clouds in large-scale models. Monthly Weather Review, 121(11), 3040-3061.

Tompkins, A. M., Gierens, K., & Rädel, G. (2007). Ice supersaturation in the ECMWF integrated forecast system. Q.J.R. Meteorol. Soc., 133(622), 53-63.

Vidot, J., A. J. Baran, and P. Brunel (2015). A new ice cloud parameterization for infrared radiative transfer simulation of cloudy radiances: Evaluation and optimization with IIR observations and ice cloud profile retrieval products. J. Geophys. Res. Atmos., 120, 6937–6951. doi:10.1002/2015JD023462.

Wu, W., X. Liu, D. K. Zhou, A. M. Larar, Q. Yang, S. Kizer, and Q. Liu (2017). The application of PCRTM physical retrieval methodology for IASI cloud scene analysis, IEEE Trans. Geosci. Remote Sens., **55**, 5042–5056, doi:10.1109/TGRS.2017.2702006.

Yang, Q., Liu, X., Yang, P., Wang, C., (2014). A fast radiative transfer parameterization under cloudy condition in solar spectral region, AGU2014 (A23E-3293).

Zhou, D., A. Larar and Xu Liu (2012). Monitoring surface climate with its emissivity derived from satellite measurements, Proc. SPIE 8524, Land Surface Remote Sensing, 85240I (2012/11/21); doi:10.1117/12.974215; http://dx.doi.org/10.1117/12.974215 .

Appendixes

A1. Model Profiles (Alan Geer)

Model profiles were taken from the ECMWF operational global weather forecasting system on 01 March, 2009. The best available estimate of the atmospheric state was taken every 3 hrs, being a mix of analysis and very short-range forecast. At the time, cycle 35r1 of the ECMWF system was operational and full documentation is available from ECMWF (2009). This describes cycle 33r1, but there were no major changes going to 35r1. The model fields used 91 levels. An additional level at 0.005 hPa was added using the US Standard Atmosphere to avoid ambiguity when different RTMs were using the model fields. The forecast model uses T799 (roughly 25km) spatial binning. Each AIRS position and observing time from 1 March 2009 was associated with nearest forecast time and interpolated across the 35r1 grid points in time and space.

Analyses and forecasts are based on clouds and precipitation models using three main schemes: convection by a mass-flux scheme (Tiedtke 1989, Bechtold et al., 2004), large-scale cloud and precipitation, including the possibility of ice supersaturation (Tiedtke 1993, Tompkins et al., 2004), and an eddy-mass flux turbulent diffusion scheme for the boundary layer, representing stratocumulus (Köhler et al., 2011). Together these contribute to producing the vertical profile of cloud water, cloud ice and cloud fraction at every grid point. Where necessary to assume an overlap formulation for clouds in the radiation scheme, a generalized formulation was assumed, increasing from maximum to random overlap with increasing cloud layer separation (Barker, 2008).

The ECMWF analysis is a combination of short-range forecast and observational information, including satellite radiances, satellite-retrieved atmospheric motion vectors, nearsurface wind vectors from scatterometers, Global Navigation Satellite System (GNSS) radio-occultation measurements, and conventional data sources including ground stations, ships, radiosondes and aircraft. Of particular relevance to the current study is that AIRS and IASI radiances are assimilated, but in clear-sky areas only (Collard and McNally, 2009). As a consequence, the cloud description in the ECMWF model does not contain AIRS cloud information. Further, clouds and precipitation are constrained in the analysis by the assimilation of cloud and precipitation-affected microwave-imager radiances (Bauer et al., 2006).

A2. Selection of the test data set (Evan Manning)

Each day AIRS produces 3 million spectra, each with 2378 spectral channels. The locations of the spectra are biased towards the polar areas due to the high inclination of the EOS Aqua orbit. In order to test the relative performance of cloudy RTAs we created a data set with emphasis on cloud variability. This data set was created using stratified sampling: the clouds are roughly characterized by the difference between the surface temperature (stemp), provided by ECMWF, and the brightness temperatures measured in five sounding regions from AIRS, including the brightness temperatures at 1231 cm^{-1} (bt1231). We traversed the spectra in time order and assigned to each spectrum a tag which combined the following elements:

1) Day vs Night (the divide was solar zenith angle = 90) (2 bins)
2) Land/Sea + Latitude band. 30 degree bins. Non-polar bins were divided into land and sea. Sea was defined as any AIRS footprint containing less the 1% land. (10 bins)

3) bt1231 in 10 K increments between 170 and 360 K (19 bins)
4) stemp-bt1231 in 10 K increments from -40 to +210 K (26 bins)
5) We defined five broad spectral bands at [650, 800], [800, 1200], [1200, 1700], [1700, 2400], and [2400, 2700] cm^{-1}. For each band we used the mean brightness temperature in the band minus bt1231, bt_band-bt1231, to define 20 bins in 10 K increments from -110 to +110 K (20 bins per band).

This allows for up to $2*10*19*26*20^5 \sim 30$ billion bins, but most bins were empty. Each spectrum's tag was compared to the tags of the spectra previously collected. If the tag did not match any, then the new spectrum was added to the data set. This procedure created 7377 uniquely tagged spectra. The latitude, longitude, solar zenith and solar azimuth angles and the ECMWF definition of the state of the atmosphere associated with this set and the associated AIRS spectral radiances were posted on the anonymous FTP site at "ftp://thunder.jpl.nasa.gov/hha/Cloudy_RTA/atm.state". The rta7377readme.20160518.txt explains the details.

The distribution of the 7377 test cases does not match the distribution of clouds in a global grid, but the set spans the natural variability of spectra, which is dominated by clouds. The emphasis on clouds is illustrated in Figure A1. The red trace in Figure A1 is the distribution of the cloud effect (stemp-bt1231) for a global area representative random sample of the non-frozen oceans. A large fraction of the ocean is covered by relatively low or broken clouds. The peak of the random sampled distribution is at 5 K. Only 7% of the non-frozen oceans are associated (stemp-bt1231)>50 K, (roughly correponding to brightness temperatures colder than 250 K). The blue trace in Figure A1 is the distribution of the cloud effect in the test set. The peak of the distribution is at 8 K, and 30% of the test data are associated with (stemp-bt1231)>50 K.

A3. RTM model summaries.

A3.1. CRTM (Moradi and Wilson)

The Community Radiative Transfer Model, CRTM, (Nalli et al., 2016), is a fast radiative transfer model developed by the Joint Center for Satellite Data Assimilation (JCSDA) that is widely used in the U.S. (including at the NASA Global Modelling and Assimilation Office, GMAO) to assimilate satellite radiances. The current version is 2.2.3. CRTM is capable of simulating microwave and infrared radiances using atmospheric profiles of pressure, temperature, humidity, and other species such as ozone. CRTM also includes capabilities to simulate satellite cloudy radiances. The ice cloud single scattering properties are based on Baum et al., (2011). All spectra were calculated with CRTM 2.2.3. However, within this version CRTM gives the user wide flexibility for the cloud overlap assumption and the cloud composition. Six cloud types can be defined at the same time: water, ice, rain, snow, graupel, and hail. The calculations presented in this paper used only water and ice clouds based on cloud liquid and ice water content profiles.

CRTM requires pressure values at levels as well as layer-averaged, and the layer averages of temperature, water vapor and other absorbers as input to perform clear-sky calculations. The top pressure level is fixed at 0.005 hPa. For the calculation of cloudy radiances CRTM requires cloud liquid water content in g/m^2 and the effective radii of water and ice particles.

The same version of CRTM was used, but with different assumptions. The results identified as CRTM_tcc uses Eq. 1 with the MO assumption and tcc, the total cloud fraction specified by ECMWF. The effective radius of the particles was calculated using Eq. 3 in Ou et al., (1995) for ice clouds and Eq. 2 in Bower et al., (1994) for water clouds. The CRTM_2col and the CRTM_mro calculations used the 2col and MRO assumption, respectively, with the identical cloud microphysics. The parameterization for ice particle effective radius used a 4th order empirical polynomial given by Ou (1995), same as CRTM_tcc. The effective water particle radius was logarithmically interpolated from 10 microns at the surface to 45 microns at the top of atmosphere, consistent with the ECMWF documentation. Following Eq. 1, CRTM_2col uses the maximum cloud fraction specified by each ECMWF cloud coverage profile to calculate CF. This value is close to, but not exactly the same as tcc.

The CRTM results were received on 2016/09/06, revised 2017/03/09

A3.2. PCRTM (Xiuhong Chen, Xianglei Huang, Xu Liu, Qiguang Yang, and Wan Wu)

The Principal Component-based Radiative Transfer Model (PCRTM) is a fast and accurate forward model for hyperspectral instruments with thousands of spectral channels. It uses principal components to compress spectral information and reduces computational time by performing radiative transfer calculations at just a few hundred monochromatic frequencies (Liu, 2006, Liu, 2016). The molecular absorption coefficients of gases are based on a lookup table calculated off-line using a line-by-line radiative transfer model based on HITRAN2008. Both ice and water clouds were parameterized into transmittance and reflectance matrices for the isotropic thermal scattering. The ice cloud matrices were obtained using single scattering properties from Baum et al., (2011) and a 32-stream Discrete Ordinates Radiative Transfer (DISORT) (Stamnes et al., 1988). The water clouds were obtained the same way by using the refractive indices from Segelstein (1981). The anisotropic solar scattering is modeled according to Liu et al., (2016) and Yang et al., (2014). Transmittance and reflectance lookup tables were obtained under various conditions for parameters such as cloud optical depth, cloud effective size, wavelength, and solar and satellite zenith angles and azimuth angles (Yang et al., 2014). The non-LTE effect was calculated according to the parameterization described by DeSouza-Machado et al., (2007).

When generating input parameters for the PCRTM from the ECMWF fields, ice cloud optical depths were calculated from ice water content as in Ebert and Curry (1992). Warm cloud optical depth based on cloud liquid water content follows Fouquart (1987). The effect of different cloud overlapping assumptions on the simulated radiance has been discussed in Chen et al., (2013). The PCRTM_MRO and PCRTM_ERO entries in tables 2-4 represent the results obtained with a maximum random overlap (MRO) assumption and an exponential random overlap (ERO) assumption, respectively. The cloud fraction and cloud profiles are used to generate 50 sub-columns (Chen et al., 2013) for both overlapping assumptions. The PCRTM_ERO2 and PCRTM_MRO4 represent the simulation results using less sub-columns, namely 2 sub-columns for the ERO and 4 sub-columns for the MRO, respectively.

The PCRTM model has been used to perform cloud and atmospheric temperature and water vapor vertical profile retrievals from hyperspectral instruments such as IASI, CrIS, and AIRS (Liu et al., 2009, Wu et al., 2017, Liu et al., 2017). Two validations of the PCRTM under cloudy conditions were given by Chen et al., (2013). One validation case used NOAA/GFDL data with the random overlap (RO) assumption, the other one used ECMWF with MRO assumption. Validations showed satisfactory consistency between the calculated OLR and the counterparts from the GCM/analysis.

The PCRTM spectra were received 2017/02/24, revised with full scattering 2017/07/14.

A3.3. SARTA (Sergio DeSouza-Machado and L. Strow)

SARTA RTM uses a four-column RO cloud overlap assumption. The clear column calculations use SARTA V6.0 (Strow et al., 2006). The absorption coefficients of gases are from line-by-line calculations based on HITRAN2008. The ECMWF clouds are converted into two thick slabs. Typically this is an ice cloud between zi_top and zi_bottom, and a water cloud between zw_top and zw_bottom. The Mie scattering parameters for water clouds use a modified gamma droplet size distribution of effective variance 0.1 (dimensionless) and effective radius (typically) of 20 μm. The cirrus cloud scattering parameters are based on Baum et al. (2007, 2011), and the ice effective particle size is estimated from a temperature-based parametrization by Ou and Liou (1995), where the ECMWF temperature profile is used to associate the ice cloud slab top pressure with a cloud top temperature. The effective absorption due to each slab is then calculated using PCLSAM (Chou et al, 1999) scattering code and used in the SARTA TwoSlab RTM (DeSouza-Machado et. al., 2018). Each pixel is then divided into four columns:

Case 1) A clear column from the surface to the Top Of the Atmosphere (TOA).
Case 2) A clear column between the surface and zw_bottom. Between zw_bottom and zw_top the pre-calculated water cloud absorption is added.
Case 3) A clear column between the surface and zi_bottom. Between zi_bottom and zi_top the pre-calculated ice cloud absorption is added.
Case 4) The transmittance calculated from case 2) up to the zi_bottom is continued with the transmittance from there to TOA using the transmittance calculated from case 3.

A cloud fraction for each case is then chosen such that all of the ice cloud and a random portion of the water cloud is seen from TOA, such that the ECMWF specified total cloud cover (tcc) is satisfied. Details are summarized in Machado and Strow (2016) and in DeSouza-Machado et al., (2018). The difference between SARTA_TwoSlab(C) and SARTA_TwoSlab(P) is due to the difference in the way the boundaries of the thick slabs are calculated. The small differences between the results from the two SARTA versions show that the results are not very sensitive to these details.

The SARTA results were received 2016/11/08, revised to be consistent with DeSouza-Machado et al., (2018) 2017/02/27

A3.4. RTTOV (J. Vidot and M. Matricardi)

The fast radiative transfer model RTTOV (Saunders et al., 1999) is widely used by a number of NWCs to assimilate infrared radiance observations. In this study, we used RTTOV Version 12. The predictors of the fast atmospheric transmittances were calculated with the line-by-line model LBLRTM 12.2 (Clough et al., 2005; Alvarado et al., 2013) that uses the AER3.2 spectroscopic database (mostly based on HITRAN 2008 but with many improvements regarding line mixing and absorption line parameters) and MTCKD 2.5.2 (Mlawer et al., 2012). The scattering by clouds is modeled using the Chou-scaling approximation (Chou et al., 1999). The liquid and ice cloud optical properties are parameterized following the work of Matricardi (2005) and Vidot et al., (2015), respectively. Within the RTTOV version the user has the option to select a limited number of cloud type and the cloud overlap assumptions. RTTOV_mro used the OPAC cumulus cloud type and the MRO assumption (Matricardi, 2005). Additionally, a much faster, experimental version of the cloud overlap method has been tested in RTTOV. This method is named CMSS (Cloud fraction Maximum Single Stream). It simulates cloudy infrared radiances using Eq. 1 with CF set to the maximum cloud fraction in the layers above a certain pressure level (here fixed to 750 hPa). This method is optimized for mid-and upper-tropospheric sounding channels.

The RTTOV results were received 2016/12/09

A3.5. HT-FRTC (Havemann)

For the AIRS radiance simulations presented in this paper, the Havemann-Taylor Fast Radiative Transfer Code (HT-FRTC) has been specifically trained for the infrared part of the electromagnetic spectrum. The HT-FRTC does only monochromatic radiative transfer calculations. The gaseous absorption of all the trace gases included in HITRAN 2008 is included in the form of lookup tables. During the code training phase monochromatic calculations are performed at a very high spectral resolution (10^{-3} cm^{-1}) for a diverse set of 1000 atmospheric profiles and surface conditions. The training run included vertical profiles of liquid and ice cloud. The results of the radiance calculations for the training profiles at the very high spectral resolution were then used to calculate the principal components which are the eigenvectors of the covariance matrix containing the radiance spectra. The HT-FRTC works slightly different to other codes like PCRTM in that the principal components are not derived for the spectra of any particular instrument but rather at the full very high spectral resolution. This means that the spectra for any number of instruments can be calculated in a single fast code run. It requires just an offline convolution of the highly resolved principal components with the instrument response functions. For the simulations in this paper only the first 100 principal components which contain most of the variance were used. The weights of the principal components are predicted from a small number of radiance calculations at about 100 monochromatic frequencies. The optimal set of frequencies for prediction are selected by a k-means clustering algorithm which operates on all frequencies (2.5 million). A linear regression is carried out on the results on the training profiles (the dependent profiles). This regression then allows the prediction of the principal component weights for any independent profiles by calculating the radiances only at the 100 selected monochromatic frequencies.

For the simulations presented in this paper an effective radius of ten micrometers was used throughout for cloud liquid droplets. The cirrus optical properties that were used in the simulations are due to Baran (Baran et al., 2014). Baran has developed an ensemble model of cirrus particles of different shapes and sizes. The optical properties are parametrized solely in terms of cirrus cloud temperature and cirrus cloud ice water content. The same parametrization is applied to all types of cirrus. The HT-FRTC allows two different treatments of scattering. Scattering can be treated approximately as a modification to the extinction by using the Chou scaling approximation (Chou et al. 1999) or the scattering phase function can be fully accounted for. In this case a monochromatic version of the Edwards-Slingo spherical harmonics radiation code is called which has been incorporated into the HT-FRTC (Edwards and Slingo, 1996 and Thelen and Edwards, 2013). Calculations with Chou scaling are indicated by "C and full scattering calculations by "S".

The HT-FRTC has been run for three different cloud overlap assumptions (MRO, MO and RO). In all cases five cloud columns were used. The columns were constructed from the horizontal cloud fraction provided for each atmospheric level, which prescribed how many of the cloud columns would be clear and how many fully overcast at each level. The different overlap assumptions then determine the how the cloudy layers are stacked in the vertical. In the case of MO, all the cloudy layers are concentrated in the same columns as much as possible, in the case of RO the cloudy layers are distributed randomly across the columns and in the case of MRO the cloudy columns are maximally overlapped in adjacent vertical layers which are both cloudy but randomly distributed if there happens to be a cloud-free layer in between. One HT-FRTC fast code run is done per cloud column. In the tables the type of scattering treatment and the kind of overlap is indicated. As an example, SMRO indicates full scattering calculations applied to the five individual cloud columns that were generated using the MRO assumption.

The HTFRTC results were received with Chou scaling 2017/01/19, revised with full scattering 2017/04/17.

A3.6. σ-IASI-as (Liuzzi, Masiello and Serio)

The σ-IASI-as RTM is an advanced version of the σ-IASI model (Amato et al., 2002) with respect to cloud and aerosol treatment (Liuzzi et al., 2017). The model computes the Earth/atmosphere-emitted radiance in the spectral range 100–3000 cm^{-1}. In its current version, the model can generate radiances in both upwelling and downwelling modes. Although initially developed for IASI, σ-IASI-as is presently a generic radiative transfer model, which is well suited for nadir viewing satellite, airplane (Grieco et al., 2007) and ground-based (Bhawar et al., 2008) infrared sensors with a sampling rate in the range 0.1–2 cm^{-1}.

The σ-IASI-as RTM calculation of gas optical depths is based on a pseudo-monochromatic scheme, in which transmittances are calculated on an equally spaced wavenumber grid by means of a look-up table. For each atmospheric layer, atmospheric species, and wavenumbers, optical depths are pre-computed and stored. Then, they are rescaled with air pressure and temperature. The dependence on temperature is parameterized by a second-order polynomial. This allows optical depths to be generated at any wavenumber using the version 12.2 of LBLRTM (Clough et al., 2005), equipped with the spectral library AER v_3.2 (essentially based on HITRAN 2012 spectral database – with the continuum model MT-CKD v_2.5.2 (Mlawer et al., 2012).

The σ-IASI-as RTM simulates the impact of the presence of clouds and aerosols with a physically based method that computes their extinction as a function of the effective ice or liquid water particle or droplet radii and concentrations (Liuzzi et al., 2017). The model exploits an ab-initio approach embodying Mie routines which are called iteratively within the calculation of single-layer transmittances. The results of Mie calculations are manipulated according to the scheme described in Chou et al., (1999) for calculating effective aerosol and cloud optical depths taking into account the multiple scattering effects through the so-called scaling approximation. With the scaling approximation, radiative transfer equations for a cloudy/aerosols atmosphere are identical to those for a clear atmosphere, and the difficulties in applying a multiple-scattering algorithm to a partly cloudy atmosphere (assuming homogeneous clouds) are avoided. The RTM used for the calculation of the σ-IASI spectra was identified as version 2017.as.lr. The calculations used the MO assumption and tcc specified by ECMWF.

The σ-IASI results were received on 2016/11/29, revised 2017/03/17.

RTM name	Base Model Spectroscopy	Participant	Organization
SARTA	HITRAN2008	DeSouza-Machado Strow	UMBC
RTTOV	HITRAN2008	Vidot Matricardi	NWPSAF (France) ECMWF (EU)
HT-FRTC	HITRAN2008	Havemann	U.K. Met Office
PCRTM	HITRAN2008	Xianglei Huang Xu Liu	U. Michigan LARC
CRTM	HITRAN2008	Moradi, Wilson	NASA GMAO NASA JPL
σ-IASI-as	HITRAN2012	Liuzzi Masiello	U. Basilicata, Italy

Table 1. Cloudy RTM developers who participated in the comparison.

AIRS-bt900	day correlation	day bias ± stddev [K] 1437 cases		night correlation	night bias ± stddev [K] 1377 cases	day-night bias [K]
SARTA_TwoSlab(C)	0.9502	-2.33±22.1		0.9707	-0.96±19.5	-1.3
SARTA_TwoSlab(P)	0.9621	+2.23±22.9		0.9710	+2.80±20.4	-0.6
PCRTM_ERO	0.9629	+0.53±21.5		0.9792	+1.37±19.5	-0.8
PCRTM_MRO	0.9591	-0.21±21.1		0.9796	-0.36±19.1	-0.5
PCRTM_ERO2	0.9680	+0.96±22.8		0.9785	+1.37±21.1	-0.5
PCRTM_MRO4	0.9715	-0.53±22.36		0.9625	+0.38±19.85	-0.8
HT_CMO	0.9773	+1.09±22.58		0.9591	+1.12±20.36	-0.0
HT_CRO	0.9695	+2.18±22.26		0.9638	+2.53±20.04	-0.3
HT_CMRO	0.9774	+1.15±22.56		0.9613	+1.26±20.29	-0.1
HT_SMRO	0.9764	+0.60±22.54		0.9591	+1.12±20.36	-0.5
HT_SRO	0.9692	+1.69±22.22		0.9629	+2.47±20.07	-0.7
HT_SMRO	0.9765	+0.67±22.52		0.9590	+1.18±20.34	-0.5
RTTOV_MRO	0.9666	-4.22±21.42		0.9748	-3.65±19.5	-0.8
RTTOV_CMSS	0.9107	+0.69±21.62		0.9105	+1.04±20.0	-0.4
σ-IASI	0.9261	+0.75±20.9		0.9437	+2.02±19.2	-1.3
CRTM_tcc	0.8816	-0.98±21.1		0.8915	+1.46±19.97	-0.5
CRTM_mro	0.9552	-0.72±23.9		0.9819	-0.10±20.8	-0.6
CRTM_2col	0.9553	+0.12±23.9		0.9817	+0.71±21.3	-0.6
clear column RTM (SARTA)	0.4168	-29.57±22.8		0.4631	-25.78±22.1	+3.8
corr > 0.9 mean bias mean stddev		17 of 18 +0.3 22			17 of 18 +0.8 20	

Table 2. Histogram correlation and bias for day and night non-frozen ocean cases at 900 cm^{-1}.

bt900 (RTM-reference)	Reference CRTM_mro bias ± stddev [K]	Reference HT_SMRO bias ± stddev [K]	Reference PCRTM_MRO4 bias ± stddev [K]	Reference RTTOV_MRO bias ± stddev [K]
SARTA TwoSlab(C)	+1.69±12.10	2.99±11.79	+1.79±10.52	-1.90±11.63
SARTA TwoSlab(P)	-2.72±11.32	-1.58±10.69	-2.75±9.76	-6.47±10.79
PCRTM_ERO	-1.28±9.90	+0.12±6.83	-1.00±6.39	-4.78±6.59
PCRTM_MRO	-0.32±13.89	1.22±9.60	-0.06±9.80	-3.65±9.47
PCRTM_ERO2	-1.28±12.58	-0.06±9.69	-1.27±9.46	-4.94±9.64
PCRTM_MRO4	-0.13±11.44	+1.26±7.57	0.00±0.00	+3.59±7.51
HT_CMO	-1.65±12.54	-0.43±0.39	-1.68±7.69	-5.31±5.72
HT_CRO	-2.75±12.19	-1.40±1.79	-2.74±7.44	-6.41±5.87
HT_CMRO	-1.72±12.55	-0.48±0.35	-1.75±7.69	-5.38±5.73
HT_SMO	-1.17±12.41	+0.04±0.17	-1.20±7.58	-4.83±5.57
HT_SRO	-2.25±12.08	-0.91±1.82	-2.25±7.35	-5.91±5.74
HT_SMRO	-1.24±12.42	0.00±0.00	-1.26±7.57	-4.90±5.57
RTTOV_MRO	+3.65±12.70	+4.90±5.57	+3.59±7.51	0.00±0.00
RTTOV_CMSS	-1.34±10.35	-0.02±5.62	-1.23±6.93	-4.92±6.45
σ-IASI-as	-1.49±12.85	-0.10±6.85	-1.37±7.99	-5.00±4.34
CRTM_tcc	+0.20±14.02	+1.64±14.07	+0.43±11.86	-3.26±10.57
CRTM_mro	0.00±0.00	+1.24±12.42	+0.13±11.44	-3.65±12.70
CRTM_2col	-0.60±3.43	+0.47±12.22	+0.65±11.22	-4.29±12.75
5 group bias 5 group stddev	-0.2 12.3	+1.5 7.6	+0.3 8.7	-3.4 6.5

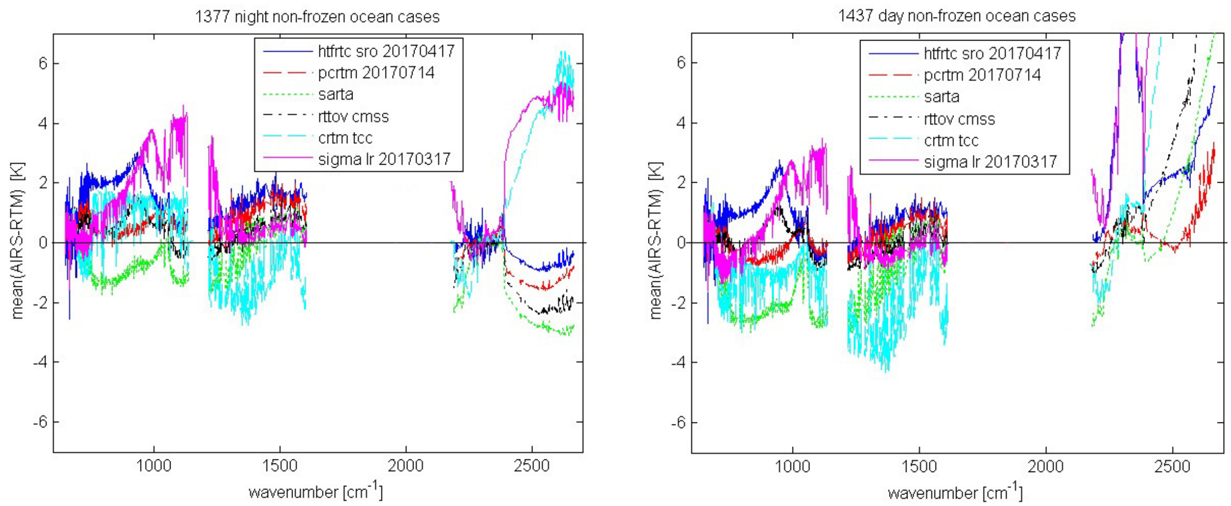
Table 3. Bias and standard deviation of bt900 calculated for 1437 day non-frozen ocean cases relative to CRTM_mro, HT_SMRO, PCRTM_MRO4 and RTTOV_MRO.

bt900 (AIRS observed – calculated) [K]	Tropical zone 3644 cases bias ± stddev [K]	Mid latitude 2662 cases bias ± stddev [K]	Polar 1070 cases bias ± stddev [K]		Tropical–midlatitude bias [K]
SARTA TwoSlab(C)	+0.41±22.59	-4.66±14.84	+3.19±10.15		+5.0
SARTA TwoSlab(P)	+4.15±23.51	-2.57±15.02	+0.79±10.06		+6.7
PCRTM_ERO	+2.63±22.48	-2.86±13.86	-0.53±9.60		+5.5
PCRTM_MRO	+0.35±22.01	-3.42±13.93	-0.56±9.63		+3.8
PCRTM_ERO2	+3.21±24.33	-2.65±14.84	-0.41±9.79		+5.9
PCRTM_MRO4	+0.94±23.19	-3.12±14.36	-0.32±9.81		+4.1
HT_CMO	+1.12±23.41	-0.84±14.99	+1.39±9.65		+2.0
HT_CRO	+2.97±23.08	-0.17±14.68	+1.48±9.58		+3.1
HT_CMRO	+1.17±23.40	-0.78±14.97	+1.40±9.65		+2.0
HT_SMO	+0.92±23.39	-1.08±14.98	+1.15±9.72		2.0
HT_SRO	+2.76±23.06	-0.42±14.68	+1.24±9.65		3.2
HT_SMRO	+0.97±23.38	-1.03±14.96	+1.16±9.72		2.0
RTTOV_MRO	-4.33±22.32	-4.50±14.49	-0.24±9.77		0.2
RTTOV_CMSS	+2.71±22.24	-2.35±14.05	-0.06±10.13		5.1
σ-IASI-as	+1.23±22.35	+0.19±14.45	+2.21±9.34		1.0
CRTM_tcc	+2.60±22.39	-2.76±15.14	+0.73±9.98		5.4
CRTM_mro	+1.95±24.62	-5.60±14.67	-1.68±10.50		7.6
CRTM_2col	+2.94±24.57	-5.24±14.73	-1.69±10.60		8.2

Table 4. Bias and standard deviation between AIRS bt900 and different RTMs separated into the tropical, mid-latitude and polar zones (as defined in the text).

bt2616 (AIRS observed –calculated)	day correlation	day bias ± stddev [K] 1437 cases		night correlation	night bias ± stddev [K] 1377 cases
SARTA TwoSlab(C)	0.8744	+5.00±14.79		0.9738	-3.06±18.40
SARTA TwoSlab(P)	0.8779	+7.05±15.57		0.9746	-0.37±19.45
PCRTM_ERO	0.9855	+1.55±15.90		0.9371	-0.29±19.23
PCRTM_MRO	0.9846	+0.73±15.84		0.9334	-1.88±18.84
PCRTM_ERO2	0.9804	+2.52±16.79		0.9563	+0.58±20.56
PCRTM_MRO4	0.9766	+0.91±16.22		0.9446	-1.25±19.47
HT_CMO	0.6669	+15.04±20.11		0.9075	-2.53±20.49
HT_CRO	0.6492	+15.97±19.76		0.9809	-1.43±20.12
HT_CMRO	0.6653	+15.13±20.10		0.9074	-2.46±20.47
HT_SMO	0.9215	+3.69±15.86		0.9791	-2.11±20.59
HT_SRO	0.9628	+3.87±15.83		0.9736	-0.83±20.14
HT_SMRO	0.9245	+3.65±15.88		0.9799	-2.02±20.56
RTTOV_MRO	0.8948	+4.81±15.86		0.9574	-6.48±19.85
RTTOV_CMSS	0.8803	+8.07±15.26		0.9605	-2.38±19.02
σ-IASI_as	0.4809	+20.94±18.94 (*)		0.8739	+5.37±19.50
CRTM_tcc	0.3991	+20.55±17.60		0.6623	+6.38±20.52
CRTM_mro	0.6680	+13.90±18.74		0.9604	-2.16±19.58
CRTM_2col	0.6410	+15.33±18.59		0.9486	-1.04±19.63
corr > 0.9 Summary mean (stdev)		7 of 18 +2.4 (16)			16 of 18 -1.8 (20)

Table 5. Histogram correlation, bias and standard deviation at 2616 cm^{-1}
for day and night non-frozen ocean. (*) solar reflected component not implemented



a) Night b) Day

Figure 1. Mean difference between AIRS and six RTM implementations for non-frozen ocean cases.

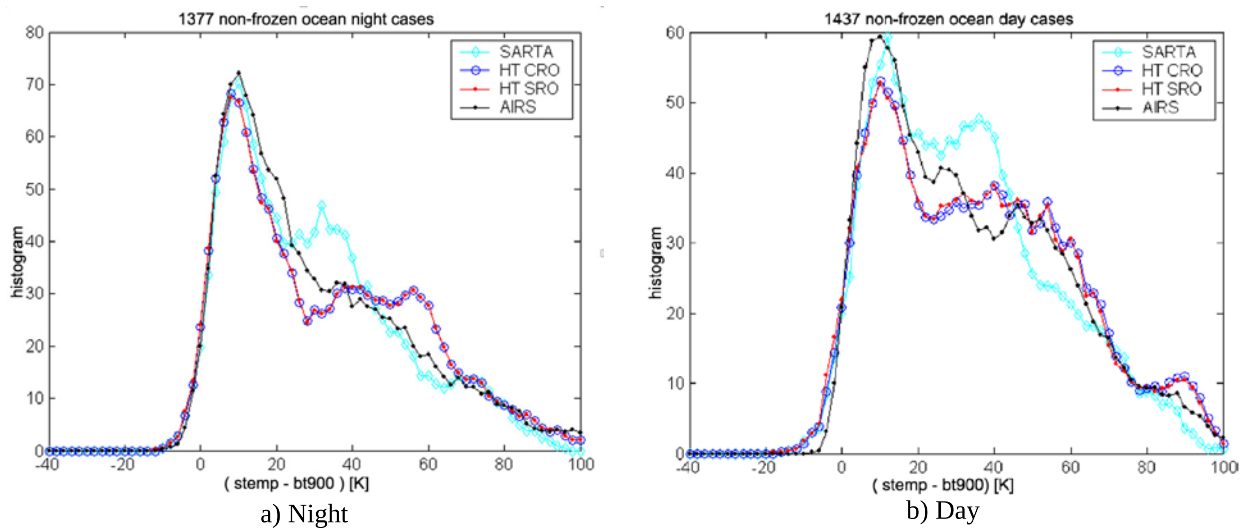
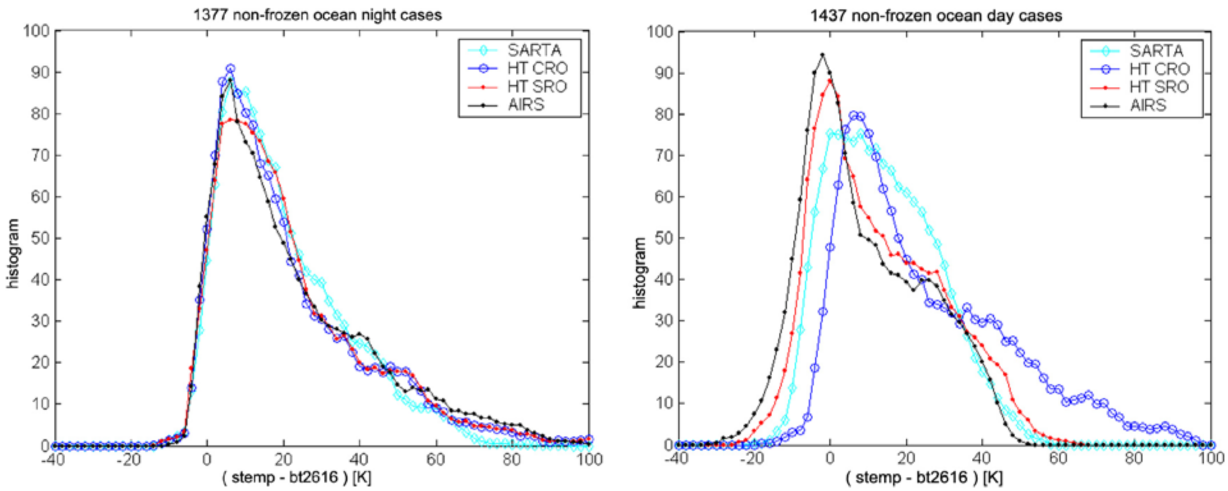


Figure 2. Histograms for (stemp-bt900) observed by AIRS and calculated by three representative RTMs



a) Night b) Day

Figure 3. Histograms for stemp-bt2616 observed by AIRS and calculated by three representative RTMs

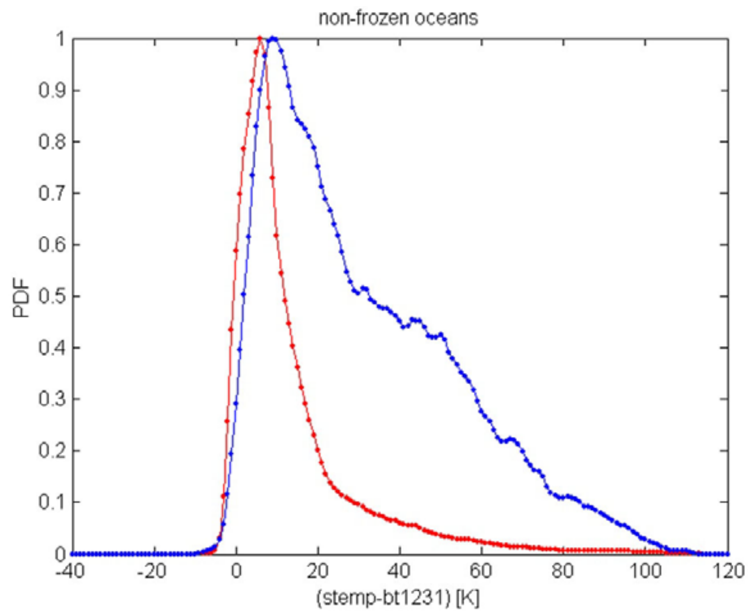


Figure A1. The red trace shows the distribution of the cloud effect (stemp-bt1231) for an area-representative random sample of the non-frozen oceans. The blue trace is the distribution of the stratified sample.