A Process-Based Climatological Evaluation of AIRS Level 3 Tropospheric Thermodynamics over the High-Latitude Arctic

JOSEPH SEDLAR

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/ESRL/Global Monitoring Division, Boulder, Colorado

MICHAEL TJERNSTRÖM

Department of Meteorology, Stockholm University, Stockholm, Sweden

(Manuscript received 28 November 2018, in final form 4 June 2019)

ABSTRACT

Measurements from spaceborne sensors have the unique capacity to fill spatial and temporal gaps in ground-based atmospheric observing systems, especially over the Arctic, where long-term observing stations are limited to pan-Arctic landmasses and infrequent field campaigns. The AIRS level 3 (L3) daily averaged thermodynamic profile product is widely used for process understanding across the sparsely observed Arctic atmosphere. However, detailed investigations into the accuracy of the AIRS L3 thermodynamic profiles product using in situ observations over the high-latitude Arctic are lacking. To address this void, we compiled a wealth of radiosounding profiles from long-term Arctic land stations and included soundings from intensive icebreaker-based field campaigns. These are used to evaluate daily mean thermodynamic profiles from the AIRS L3 product so that the community can understand to what extent such data records can be applied in scientific studies. Results indicate that, while the mid- to upper-troposphere temperature and specific humidity are captured relatively well by AIRS, the lower troposphere is susceptible to specific seasonal, and even monthly, biases. These differences have a critical influence on the lower-tropospheric stability structure. The relatively coarse vertical resolution of the AIRS L3 product, together with infrared radiation through persistent low Arctic cloud layers, leads to artificial thermodynamic structures that fail to accurately represent the lower Arctic atmosphere. These thermodynamic errors are likely to introduce artificial errors in the boundary layer structure and analysis of associated physical processes.

1. Introduction

Atmospheric observations provide the backbone for fundamental knowledge of how thermodynamic processes in the atmosphere operate. Radiosoundings released operationally from a global network of sounding stations provide crucial in situ measurement profiles of the atmospheric state. While this network is globally distributed, it remains sparse in temporal and spatial coverage and is mainly limited to land-based locations. In the Arctic, where the effects of climate change are amplified (e.g., Screen and Simmonds 2010; Pithan and Mauritsen 2014; Huang et al. 2017), operational profiling from radiosoundings is limited to on the order of 10 land-based stations, distributed on the pan-Arctic coastline (e.g., Nygård et al. 2014). The number of radiosounding stations inland from the Arctic coastline increases as one moves farther south, but these observations also contain less of an Arctic Ocean footprint. As a result, significant gaps remain in the fundamental process-level understanding of atmospheric thermodynamics, thermodynamic interaction with the ocean and sea ice boundaries, and persistent cloudiness across the much of the central Arctic.

Measurements from space have the unique ability to fill the gaps in sparse, surface-based observational networks, both spatially and temporally. From a thermodynamic

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

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAMC-D-18-0306.s1.

Corresponding author: Joseph Sedlar, joseph.sedlar@colorado.edu

DOI: 10.1175/JAMC-D-18-0306.1

profiling perspective, the Atmospheric Infrared Sounder (AIRS) is one such measurement platform. AIRS measures spectral infrared radiation emitted from Earth's surface and the atmosphere (Chahine et al. 2006). Thermodynamic state variables are retrieved from the measured infrared spectra combined with microwave radiances from AMSU, providing sounding profiles across the atmosphere. AIRS was launched on board the *Aqua* satellite as part of the A-Train constellation of polar-orbiting satellites, operational since late 2002 (Parkinson 2003).

AIRS-observed radiances provide a critical asset to numerical weather prediction models. National meteorological centers incorporate AIRS measurements as part of their data assimilation cycle, improving the initialization state of the atmosphere and ultimately contributing to improved weather prediction (Singh et al. 2011; Jones and Stensrud 2012; Reale et al. 2008, 2018). Over the data-sparse Arctic, AIRS has the potential to provide an unprecedented thermodynamic profiling data record. Recently, these profiles have been utilized to improve the understanding of atmospheric processes over the Arctic Ocean. AIRS level 3 (L3) profiles have been exploited to understand the frequency of commonly occurring temperature inversions across the lower Arctic troposphere and the relationship between temperature and water vapor inversion structures (Devasthale et al. 2010, 2011, 2016). AIRS measurements have provided insight into the role of atmospheric transport of heat and moisture into the Arctic and its radiative impact at the surface and the top of the atmosphere (Sedlar and Devasthale 2012; Devasthale et al. 2013; Sedlar and Tjernström 2017). Surface latent heat fluxes have been derived from L3 measurements of near-surface static stability and water vapor over the Arctic Ocean by Boisvert et al. (2013) and Boisvert and Stroeve (2015), linking changes in evaporation with Arctic warming.

Attempts to characterize the global climatological performance of AIRS thermodynamics have shown that AIRS uncertainties are dominated by instrumental sampling biases, while temporal sampling biases are smaller but important within the atmospheric boundary layer (Hearty et al. 2014). These biases were relatively small across the Arctic Ocean and largest in the subtropics where large diurnal variations in thermodynamics are observed. However, the evaluation of Hearty et al. (2014) was made relative to the MERRA atmospheric reanalysis, which may not accurately capture the true physical behavior of the Arctic atmosphere both because of inadequate model physics and that the constraining observations are sparse.

From the above it is clear that the AIRS L3 product has become one of the most widely used datasets to infer important processes in the Arctic atmosphere. Despite the use of AIRS measurements for many process-related Arctic studies, a strict evaluation of AIRS thermodynamic profiles with in situ observations across the Arctic is lacking. Therefore, in this paper we assess the usefulness of the AIRS L3 product to characterize the thermodynamic state of the Arctic atmosphere. In particular, we explore the ability of the AIRS L3 sampling to adequately represent vertical structures previously independently identified by radiosoundings and believed to result from important local processes in the Arctic troposphere, and especially the boundary layer. The aim is to identify features that may lead to a misinterpretation of processes and to elaborate on both strengths and weaknesses in the AIRS L3 product for Arctic research.

For this, a wealth of Arctic radiosoundings are compiled both from long-term land-based stations and from shorter, but intensive, Arctic Ocean field campaigns. These are used as the observational truth to evaluate daily mean collocated thermodynamic profiles from AIRS. The aim is to characterize the vertical distribution of AIRS temperature T and water vapor mixing ratio Q relative to in situ sounding profiles. Section 2 describes the analyzed AIRS and radiosounding data records and includes a description of the method for comparison. Section 3 quantifies the capacity of AIRS in resembling the vertical thermodynamic structures and relates the mean bias error (MBE) and root-mean-square error (RMSE) statistics to process-related phenomena frequently observed across the high-latitude Arctic. A discussion of and conclusions from the results are presented in section 4.

2. Observations and method

a. *AIRS observations*

AIRS passively observes infrared radiances across 2378 hyperspectral channels covering wavelengths from 3.7 to 15.4 μm. Vertical weighting kernels are used to convert the radiances to physical profiles of T and Q through an extensive retrieval process (e.g., Chahine et al. 2006). Stated accuracies for the L3 thermodynamic profiles are relatively robust: 1 K km^{-1} for T and 15% (2 km)$^{-1}$ for Q (http://airs.jpl.nasa.gov/data/product_accuracies/). AIRS thermodynamic profiles came online in mid-2002, and the AIRS instrument remains operative. The analysis in this study focuses on AIRS data from January 2003 through August 2016.

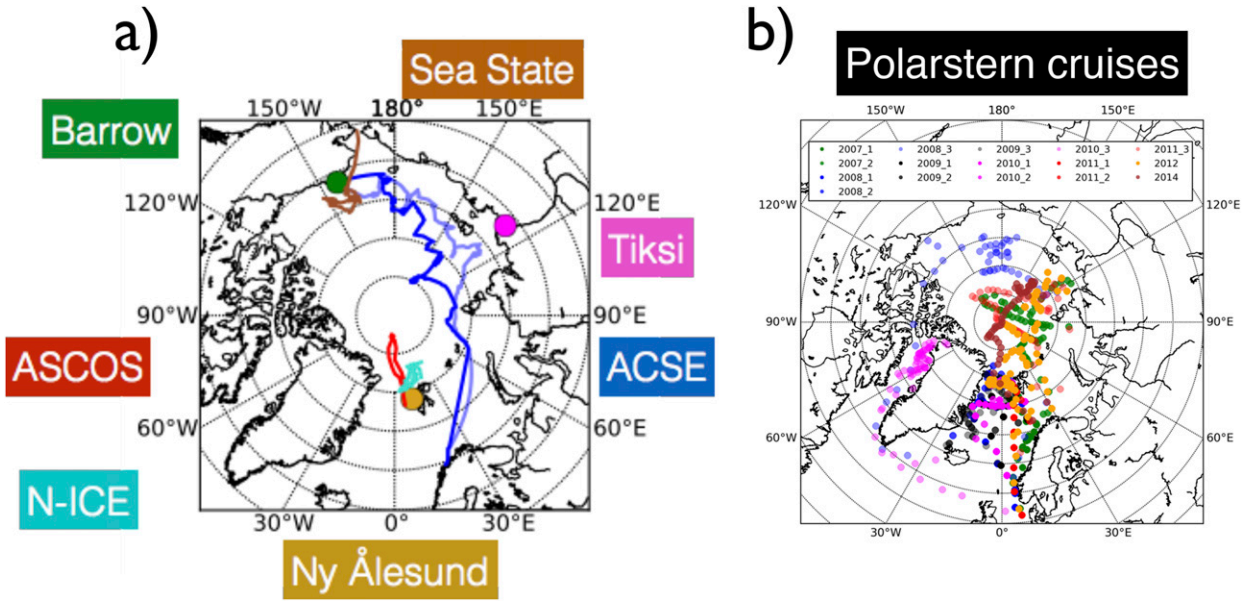

FIG. 1. (a) Locations of the radiosounding launches from the land-based stations (filled circles) and the field campaign tracks (solid lines). (b) *Polarstern* cruises: because of the volume of nearly annual Arctic field campaigns on board the *Polarstern*, the circles represent the individual sounding launch locations; the colors correspond to the year of the field campaign, with different legs of the field campaign within a particular year displayed as a shade of that color.

Housed on board a polar-orbiting satellite, AIRS provides L3 thermodynamic profiles twice daily at approximately 1330 (ascending node) and 0130 (descending node) local time each day. To achieve the L3 products, each local AIRS level 2 (L2) data swath from the cross-track scanning operation measures hyperspectral radiances on a 13.5-km spatial resolution and retrieves standard AIRS L2 profiles. AIRS L3 processing samples all collected L2 data that pass a quality control onto a 1° × 1° equal-angle grid at standard pressure levels, for the ascending and descending orbits separately (Chahine et al. 2006). It is important to realize that the sampling and averaging in going from L2 to L3 data reduces the resolution, and is sensitive to sampling errors primarily due to the sensors' sensitivity to clouds (Yue et al. 2013; Wong et al. 2015). We analyze the latest AIRS/*Aqua* L3 Daily Standard Physical Retrieval (AIRS+AMSU) 1° × 1° V006 data products (https://doi.org/10.5067/Aqua/AIRS/DATA301), as this is the most widely used L3 product from AIRS.

The AIRS instrument is housed on board NASA's Earth Observing System polar-orbiting *Aqua* satellite. Because *Aqua* is a polar-orbiting satellite, AIRS cross-track data swaths from consecutive overpasses overlap at high latitudes. Consequently, multiple overpasses contribute through increasing the sampling on the L3 grid. For analysis in this study, we take the average of the gridded L3 ascending and descending overpasses, resulting in daily mean profiles of L3 T and Q. AIRS L3 thermodynamics are retrieved on standard pressure levels. We focus on the eight lowest resolved tropospheric pressure levels: 1000, 925, 850, 700, 600, 500, 400, and 300 hPa. The vertical analysis of AIRS thermodynamics is performed on the pressure-level data, where the thermodynamics are valid for a particular pressure level and therefore are not layer-averaged quantities.

b. *Radiosounding observations*

A collection of land-based and in situ Arctic Ocean radiosounding profiles have been compiled for the purpose of evaluating the performance of AIRS profiles. Figure 1 shows the locations of the land-based and field campaign radiosounding profiles. The three land-based observatories include 1) Barrow, Alaska (now known as Utqiaġvik); 2) Tiksi, Russia; and 3) Ny-Ålesund, Svalbard; here, radiosounding releases were performed regularly from the same location during the period 2003–16 (Fig. 1a). Soundings were launched at 0000 and/or 1200 UTC at Tiksi and Ny-Ålesund; at Barrow [ARM North Slope of Alaska (NSA) site] they were launched at 0600 and 1800 UTC.

These observatories were chosen based on geographic location, sampling the main pathways of atmospheric transport to the high-latitude Arctic. Barrow and Ny-Ålesund contribute to the Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN; Seidel et al. 2009); Tiksi is a GCOS Upper-Air

TABLE 1. Information related to the radiosounding observations used to evaluate the AIRS thermodynamic profiles. Included are: station or field campaign name, geographic coordinates, time period of analysis, frequency of sounding launches, total number of daily averaged profiles, and radiosounding type.

Site/campaign name	Coordinates	Time period for evaluation	Launch frequency (per day)	No. of daily averaged profiles	Radiosounding manufacturer/model
Barrow	71.3°N, 156.6°W	Jan 2003–Aug 2016	1–2	4992	Vaisala RS92
Tiksi	71.6°N, 128.9°E	Jan 2003–Aug 2016	1–2	4992	Russia-USSR Mars/Russia-USSR MRZ-3A
Ny-Ålesund	78.9°N, 11.9°E	Jan 2003–Aug 2016	1–2	4992	Vaisala RS92
ASCOS	77.9°–87.5°N, 11.1°W–9.6°E	Aug–Sep 2008	4	35	Vaisala RS92
ACSE	71.4°–85.2°N, 25.7°E–178.1°W	Jul 2014–Oct 2014	4	85	Vaisala RS92
N-ICE2015	79.2°–83.3°N, 3.4°–29.8°E	Jan–Jun 2015	2	140	Vaisala RS92
Sea State	65.9°–75.5°N, 148.6°–168.5°W	Oct–Nov 2015	4	35	Vaisala RS92
Polarstern	See Fig. 1b	Jun–Oct for 2007–12 and 2014	1–2	711	Vaisala RS92

Network (GUAN) station; there are no GRUAN stations along the Russian coast to the Arctic. GRUAN is a concerted effort to provide a standard reference network of requirements and uncertainty quantification for upper-air essential climate variables (Bodecker et al. 2016), whereas GUAN is more an agreement on format and procedure, with no rigorous quality control.

Several intensive field campaigns add thermodynamic profile information over the sparsely observed central Arctic sea ice (Figs. 1a,b). A description of the scientific and measurement details of each field campaign is beyond the scope of this study; we direct the reader to the references outlining each field campaign in detail: Arctic Summer Cloud Ocean Study (ASCOS; Tjernström et al. 2014); Arctic Clouds during Summer Experiment (ACSE; Sotiropoulou et al. 2016); Norwegian Young Sea Ice Cruise (N-ICE2015 or N-ICE; Granskog et al. 2016); Sea State and Boundary Layer Physics of the Emerging Arctic (Sea State; Persson et al. 2018); and a series of yearly summer and autumn Alfred Wegener Institute expeditions of the Research Vessel *Polarstern* to the Arctic (*Polarstern*; see König-Langlo 2009).

Table 1 provides details on the radiosounding profile measurement duration, frequency and the type of radiosounding deployed for each measurement site and campaign. As all other measurements, soundings are associated with measurement uncertainty. However, this basic technology has been used for a long time, and radiosoundings are a backbone of the atmospheric observation systems. The Barrow and Ny-Ålesund stations used Vaisala RS92 sondes during this time period. Dirksen et al. (2014) give the uncertainty in temperature to between 0.15 (night) and 0.6 (day) K. The corresponding uncertainty in relative humidity is less than 5%, giving an uncertainty in specific humidity of mostly less than $0.2\,\mathrm{g\,kg^{-1}}$ (temperature sensitive). These values correspond well to those provided by the manufacturer. The main benefit for this study of using GRUAN stations is therefore that it provides long-term data with a consistent uncertainty. Tiksi uses an older sonde technology, MRZ-3A, and the GUAN status refers more to procedure than quality. Finding authoritative uncertainty information for this sonde type is difficult but the body of evidence (e.g., Sun et al. 2010; 2013; Ingleby 2017; Ho et al. 2017) suggests that the temperature uncertainty is larger, maybe by as much as a factor of 2–3, and that uncertainty in humidity is even larger especially at lower temperature and when the sonde has been wetted. All the field campaigns used here deployed Vaisala RS92 sondes; although GRUAN standards are usually not applied to field campaign data, GRUAN and manufacturer uncertainty agree well (Dirksen et al. 2014) and long-term stability is not required for short-term campaigns. Except for N-ICE, all the in situ field campaigns were operational during Arctic summer and autumn (Table 1). Typically one–two sounding profiles were made per day (0000 and/or 1200 UTC); three of the field campaigns made sounding profiles four times daily (0000, 0600, 1200, and 1800 UTC).

Combining the land stations with in situ field campaign radiosoundings provides a large spatial and temporal context of nearly 16 000 daily atmospheric profiles with which to evaluate the performance of collocated AIRS L3 thermodynamics.

c. *AIRS-radiosounding comparison methodology*

All radiosounding releases from the observatories are first converted from UTC to local time based on the longitude of the sounding launch. Available sounding profiles within one day are then temporally averaged to produce daily mean profiles of T and Q. The vertical resolution of radiosounding thermodynamics is far superior compared to AIRS L3 standard pressure levels. The sounding profiles of T and Q were interpolated to the L3 pressure levels using a cubic spline interpolation.

To compare with daily averaged AIRS profiles, the nearest AIRS L3 grid box to the sounding release coordinates is found. For the mobile field campaign observatories, we take the average of all sounding launch coordinates within 1 day in order to reference the AIRS nearest grid point. Daily mean T and Q AIRS profiles are then statistically compared with the corresponding radiosoundings. We found that variability in AIRS thermodynamics from potential errors in the nearest L3 grid matching to sounding coordinates was similar to the monthly temporal variability in AIRS profiles (see Figs. S1 and S2 in the online supplemental material) and therefore is not critical to the interpretation of the results. Further, we acknowledge that AIRS L3 sampling for a particular grid box at high latitudes can contain observations from multiple AIRS orbits; we consider the additional data input to the L3 products advantageous to the statistical analyses performed in this study, effectively acting to reduce the penalty of outliers in the statistics of L3 products.

We do not incorporate any weighting in the temporal averaging of radiosounding or AIRS data in an effort to strictly characterize the ability of the daily mean climatology from AIRS L3 products to represent those from in situ measurements. Since Arctic studies using AIRS data have exclusively relied on analysis of L3 thermodynamics, we believe an evaluation of its thermodynamics using this methodology is warranted. It is noted that the method used in this study will not separate the sampling bias of AIRS L3 products from the retrieval bias.

We analyze data from January 2003 through August 2016, using the AIRS+AMSU L3 V006 data products, and begin by evaluating profiles seasonally in a statistical fashion. These analyses provide a basis to evaluate the strengths and weaknesses of AIRS L3 thermodynamic profiles across the Arctic. We then expand the evaluation focusing on interrelated thermodynamic properties, such as lower-tropospheric static stability structure and the potential bias ensuing from the AIRS L3 product across the frequently cloudy Arctic troposphere.

3. Results

a. Seasonal statistics

The analysis begins by examining the seasonal, vertical distributions of AIRS T and Q in relation to radiosoundings from Barrow, Alaska, for the period 2003–16; similar distributions for the other stations are presented in Figs. S3–S5 of the online supplementary material. Regardless of season, above 600 hPa, T and Q profile-difference relative frequency distributions (RFDs) are consistently narrow and fluctuate around the zero line (Fig. 2); this behavior is expected from water vapor, where mixing ratios decrease dramatically with height (temperature). Even though the absolute differences are small, the difference relative to the observed value can still be large above 600 hPa. The smallest ranges of MBE and RMSE relative to radiosoundings across the mid- to upper troposphere indicate AIRS is suitable in measuring the thermodynamic setting in this part of the Arctic atmosphere. In general, these MBEs above the 600-hPa level, typically around 4 km, are similar to the mean T and Q uncertainties reported by the GRUAN analysis (light blue bounding lines in Fig. 2).

The largest deviations in AIRS thermodynamics from Barrow radiosoundings occur below 600 hPa (Fig. 2). The distributions show an increased spread in AIRS T and Q differences from soundings, consistent with corresponding increases in MBE and RMSE (red lines). During December–February (DJF) and March–May (MAM), AIRS has a relatively large warm and slightly dry bias across the lower troposphere (Figs. 2a,b,e,f). By July–August (JJA) and September–November (SON), a modest cold bias (Figs. 2c,d) and an enhanced dry bias (Figs. 2g,h) extends across a deeper layer through the lower troposphere; the RFD spread of Q differences during JJA is exceptionally large and is reflected in a large RMSE ranging from 0.5 to 1.0 g kg^{-1} (Fig. 2g). However, AIRS T differences during JJA do not reflect a similarly wide distribution spread (Fig. 2c), and the summer RMSE in T is actually lower than during winter and spring (Figs. 2a,b). From these statistics at Barrow (and the other stations, see Figs. S3–S5), a seasonal bias in AIRS lower-tropospheric thermodynamics emerges where 1) AIRS is generally too warm in winter and spring, with large MBE and RMSE for T, and 2) AIRS is too dry with the largest RMSE for Q during summer and autumn, while only modestly biased toward cold.

Figure 3 shows the profiles of MBE and RMSE for AIRS T (Figs. 3a,b) and AIRS Q (Figs. 3c,d) relative to the soundings from Tiksi (red), Ny-Ålesund (blue), and combining all interior Arctic field campaigns (magenta). As for Barrow (Fig. 2), the AIRS thermodynamics differ the least from radiosoundings in the mid- to upper troposphere, except for at Tiksi. Here, AIRS biases are colder and drier than the biases from the stations. Because of the lack of information regarding the sounding accuracy at Tiksi, it is difficult to determine if the larger free-tropospheric biases are a result of sounding sensor technology or an inability of capturing the observed variability by the AIRS L3 products.

Like the distribution shapes observed in Fig. 2, RMSE and the absolute MBE for T and Q both tend to increase

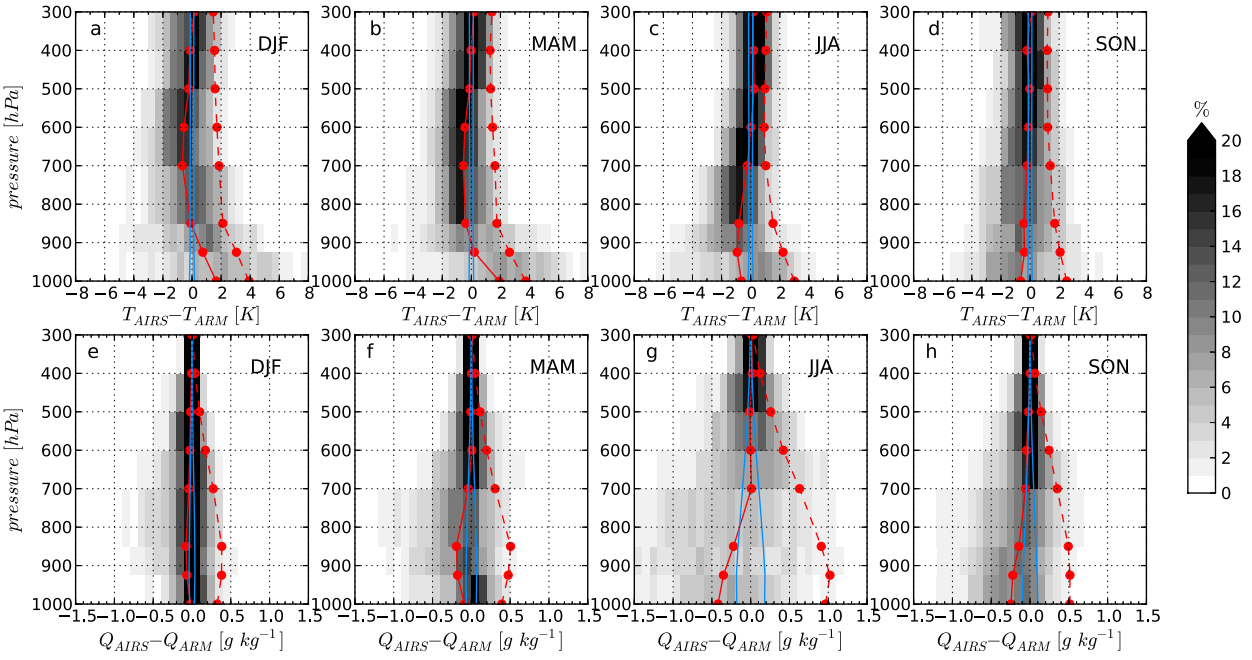

FIG. 2. Seasonal relative frequency distributions (shading; %) of the differences in daily averaged (top) T (K) and (bottom) Q (g kg^{-1}), defined as AIRS minus the radiosounding at Barrow, as a function of pressure (hPa) for (a),(e) DJF, (b),(f) MAM, (c),(g) JJA, and (d),(h) SON. Overlying the distribution contours are the seasonal MBE (solid red) and RMSE (dashed red). Light-blue lines represent the spread in seasonal T and Q uncertainty as a function of pressure level reported in the GRUAN sounding database (Seidel et al. 2009) from Barrow and Ny-Ålesund.

with pressure for all stations. For winter and spring, a warm bias in lower-level AIRS T is present at all observatories (Fig. 3a); the warm bias is not completely confined below 850 hPa at Ny-Ålesund as it is at Barrow and Tiksi. Similarly, the shift to a cold bias during JJA and SON in lower-tropospheric T is common for all observatories. The AIRS cold T bias at Tiksi and Ny-Ålesund during JJA for the lowest atmospheric levels is over 3 times as large as at Barrow and from the in situ field campaigns. The vertical distribution of AIRS T statistics at the field campaign locations are similar in magnitude to the land observatories, especially during JJA; note that these statistical profiles are only shown in the JJA and SON season panels because these were the predominant seasons for these chosen field campaigns (Table 1). For Q, we find that AIRS generally underestimates the water vapor mixing ratio, with the magnitude of the bias dependent upon the season (Fig. 3c). As a consequence, RMSEs are often larger than 0.5 g kg^{-1} across the lower troposphere, regardless of geographic location. Ny-Ålesund was the only observatory with a positive Q bias during DJF and SON for the lowest atmospheric levels (Fig. 3c). It was largest during DJF and corresponded with a small, positive MBE in T, while the MBE in T during SON was negative.

Since these sounding launches are scattered geographically around the Arctic Ocean, a certain degree of variation in the thermodynamic differences between locations is expected. It is likely that a portion of the MBE profile behavior difference at Ny-Ålesund compared with the other sites can be explained by different atmospheric regimes or local land–ocean interactions. The enhanced cold and dry temperature and moisture biases at Tiksi across the mid- to upper troposphere may also be a consequence of extreme seasonal temperature and humidity variations (e.g., Nygård et al. 2014), or even radiosounding sensors of without well-documented quality measures (e.g., Ingleby 2017). However, a robust consensus of AIRS L3 thermodynamics being least (most) similar to radiosoundings across the lower (upper) Arctic troposphere is readily apparent from this statistics.

The extent of the variability in monthly AIRS thermodynamics relative to the soundings is shown in Figs. 4 and 5. For temperature, Fig. 4 shows how the AIRS-sounding differences are distributed when weighted by the respective monthly radiosounding-observed standard deviations for the four highest pressure levels (lower troposphere); the three land stations are combined to form the monthly distributions in Fig. 4, while the distributions combining the summer and autumn

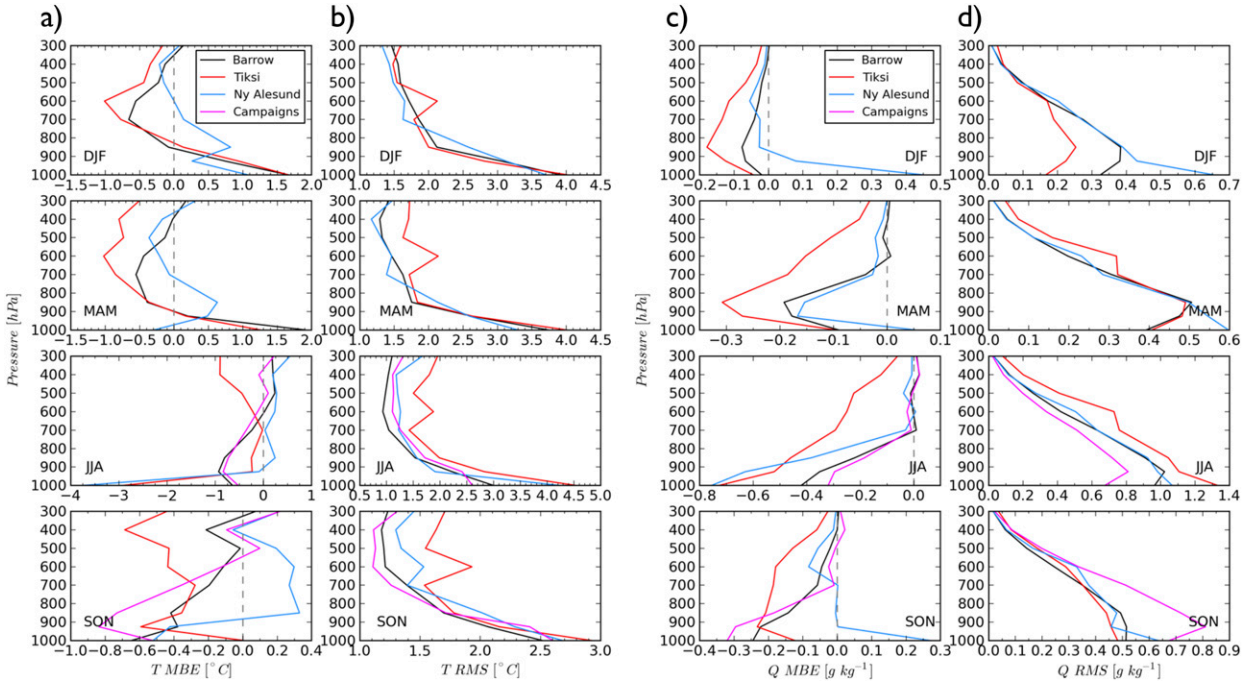

FIG. 3. Seasonal profiles of (a),(c) MBE and (b),(d) RMSE for (left) T and (right) Q for AIRS relative to radiosoundings at Barrow (black), Tiksi (red), Ny-Ålesund (blue), and combining all central Arctic field campaigns (magenta) for the (top) DJF, (top middle) MAM, (bottom middle) JJA, and (bottom) SON seasons.

field campaigns are shown in Fig. S6 of the online supplementary material.

Although the absolute errors can be large (e.g., Figs. 2 and 3), the error distributions are most frequently within ±0.5 of the observed monthly temperature variability. Being that the statistical error distributions are smaller than the observed monthly variability, we conclude that AIRS L3 temperatures have the capacity to provide useful climatological values. However, the error variability distributions at the 1000-hPa level are often broader, especially for JJA (Fig. 4l) and for May and September, preceding and succeeding the summer season. At the same time, the observed monthly standard deviations were also relatively low, around 4 K, during summer. The error variability distribution patterns were similar for the field campaigns in the central Arctic sea ice (Fig. S6), with the lower-tropospheric variability distributions being broader, but also the observed standard deviations were relatively small.

Variability in water vapor has been calculated as the AIRS minus observed difference, divided by the observed value (Fig. 5). Similar to the temperature distributions above, relative error distributions in Q are nearly consistent for all seasons, generally below 50% with relative error distribution peaks often ranging between 0% and 30%. The AIRS stated Q uncertainty is approximately 15% $(2\,\text{km})^{-1}$ (http://airs.jpl.nasa.gov/data/product_accuracies/); below 700 hPa, the peak distributions of AIRS relative error in Q tend to generally occur within this stated uncertainty. Even during JJA, when relative error variability in T below 850 hPa spans a large range (Fig. 4l), the relative error ranges in Q show more moderately defined distribution peaks at smaller relative errors.

b. *Influence of large-scale atmospheric flow on thermodynamic differences*

Expecting fetch to be a factor in the observed thermodynamic variability, seasonal biases in AIRS thermodynamics relative to radiosoundings at Barrow and Tiksi were analyzed with respect to predominant wind direction. Daily averaged atmospheric wind directions at the 700-hPa level were analyzed from ERA-Interim reanalysis (Dee et al. 2011) over a 2° × 2° box around each observatory. Days with a 700-hPa predominant wind direction in the sector between 315° over zero to 45° were considered northerly wind days; winds within the sector from 135° to 225° were considered southerly wind days. These wind sectors were observed for 35%–39% of all seasonal profiles at Barrow, and 40%–46% at Tiksi except for JJA, in which northerly or southerly occurred during 33% of the season. At both sites, the southerly flow regime dominated the northerly regime, typically 3 or more times as frequent (see counts in Figs. 6 and 7). These polarizing wind sectors were

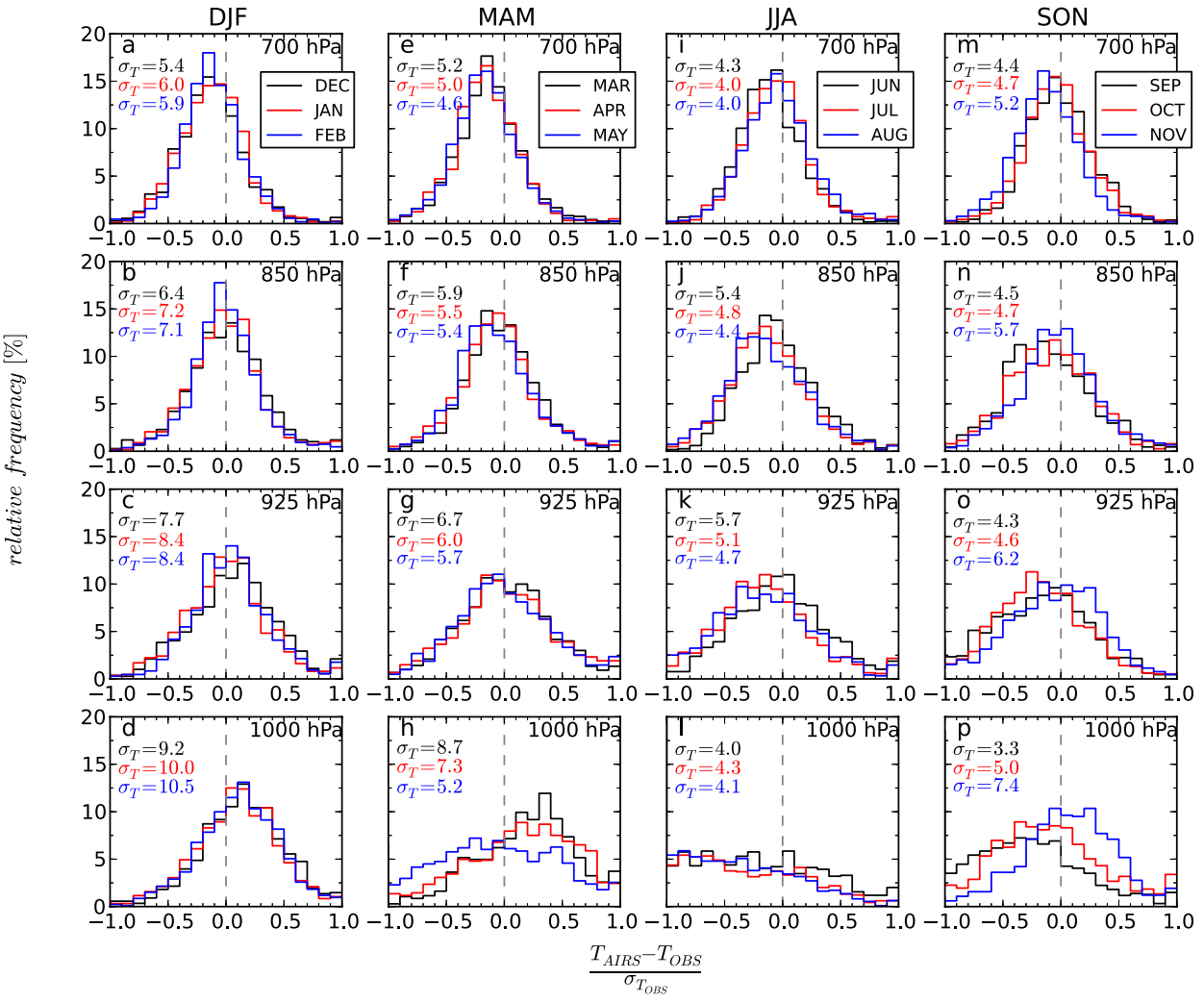

FIG. 4. Monthly relative frequency distributions (%) of the temperature error, AIRS − the observations (OBS), weighted by the observed, monthly standard deviation in temperature as a function of pressure level. The RFDs are created by combining all radiosoundings from the land stations Barrow, Tiksi, and Ny-Ålesund. The rows are for the lowest four standard pressure levels, from 700 down to 1000 hPa. The observed monthly standard deviations used to weight the temperature errors at each level are provided in the top left of each panel.

chosen to examine whether off-ice flow (northerly winds) versus off-land flow (southerly winds) had any impact on T and Q differences between AIRS and radiosoundings at Barrow and Tiksi.

Seasonal error distributions of AIRS minus radiosounding T are shown as box and whisker profiles, with overlying MBE and RMSE profiles, for the 2 wind direction sectors are shown in Fig. 6. At Barrow (Figs. 6a–d), the interquartile and 10th–90th percentile ranges are more widespread for the southerly flow pattern compared to northerly flow; the RMSE profiles also indicate a larger error in T across the lower troposphere when the flow is from the south. The air mass propagating from these polarizing sectors will be influenced by different surface thermodynamic properties, varying between open ocean, sea ice–covered ocean, or land, depending on northerly or southerly flow. The fact that AIRS-sounding T differences exhibit different distributions between the flow regimes suggests the observed variability associated a particular flow regime is not fully represented by the AIRS L3 product.

Airmass advection from the south must pass across a land surface; enhanced daily temperature variability over land compared with that over sea ice appears to signal a flow regime dependence at Barrow. When atmospheric flow is from the south, there is tendency for AIRS to enhance the already positive T bias range at the lowest level above the surface during winter (Fig. 6a), as

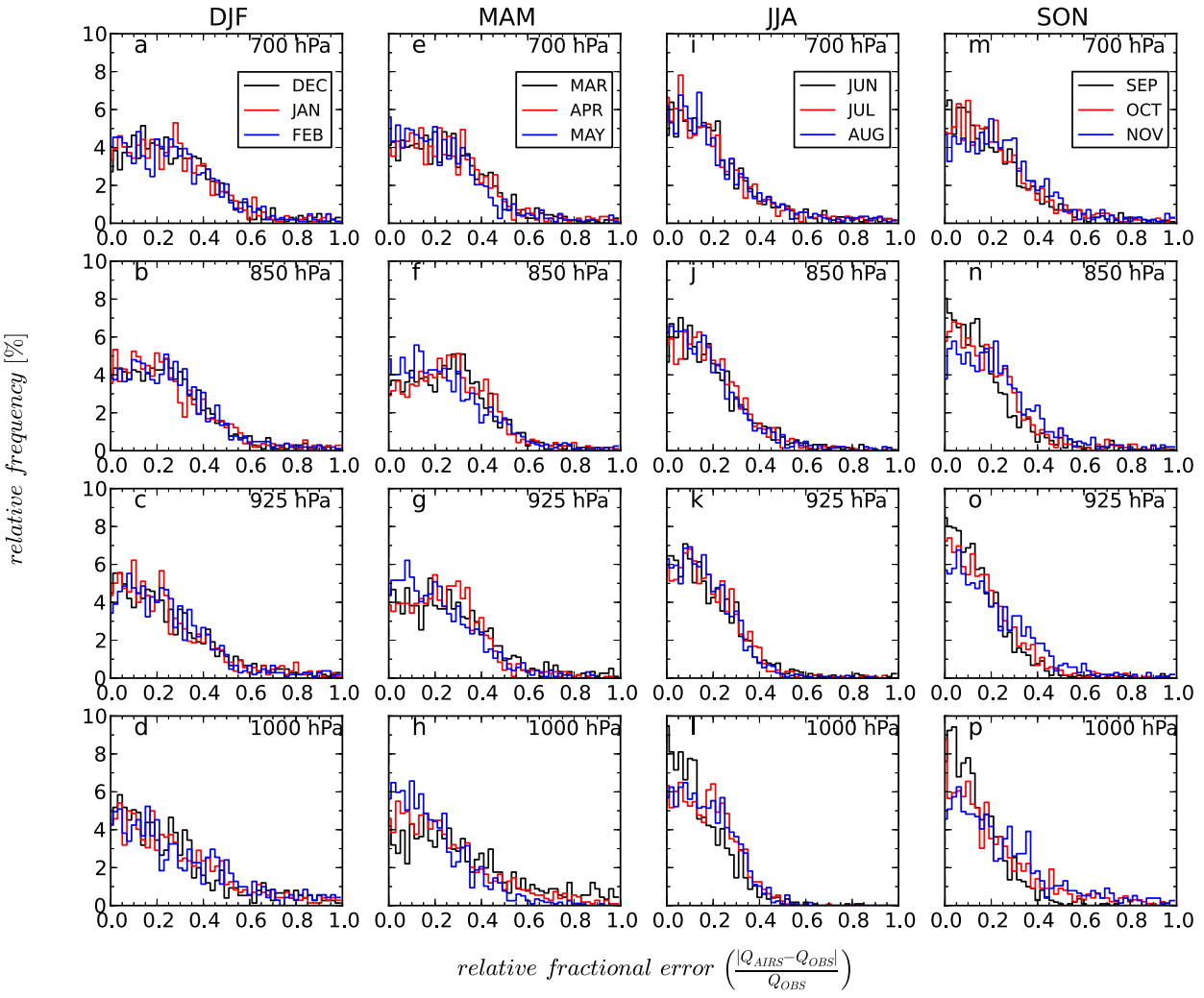

FIG. 5. Monthly relative frequency distributions (%) of the relative fractional error $[(|Q_{\mathrm{AIRS}} - Q_{\mathrm{OBS}}|)/Q_{\mathrm{OBS}}]$ in daily averaged Q, combining all available radiosoundings. The monthly distributions are combined into seasons: (a)–(d) DJF, (e)–(h) MAM, (i)–(l) JJA, and (m)–(p) SON. The rows are for the lowest four standard pressure levels, from 700 down to 1000 hPa. Relative fractional error is divided by 100 such that a fractional error of unity = 100% error relative to the radiosounding observation.

well as increase the negative T bias range across the lower troposphere during summer and autumn (Figs. 6c,d). At Tiksi, the box and whisker T error distribution ranges for northerly and southerly flow are less separable, except at the lowest pressure level, 1000 hPa (Figs. 6e–h). Even though the MBE characteristics show differences between flow origin, the RMSE profiles are very similar above this level. These results suggest that the local surface modification to the air mass near Tiksi does not influence the temperature from the AIRS L3 product capabilities as much as it did over Barrow.

For atmospheric moisture, the dry bias across much of the troposphere identified in Fig. 2 remains, regardless of northerly or southerly atmospheric flow (Fig. 7). Similar to the T difference distributions at Barrow, the range in AIRS minus radiosounding Q is more constrained and subsequently with smaller RMSEs for northerly, off ocean/ice flow compared to southerly off-land flow (Figs. 7a–d). The flow-dependent Q difference ranges are very similar at Tiksi (Figs. 7e–h), with a tendency for lower MBEs and RMSEs across the lower troposphere under northerly flow compared to southerly flow.

c. *Lower-tropospheric process relationships in AIRS and radiosoundings*

Statistical characterizations of AIRS thermodynamics shown in the preceding sections highlight the suboptimal performance of the AIRS L3 product relative to radiosondes across the lower Arctic troposphere. In this section,

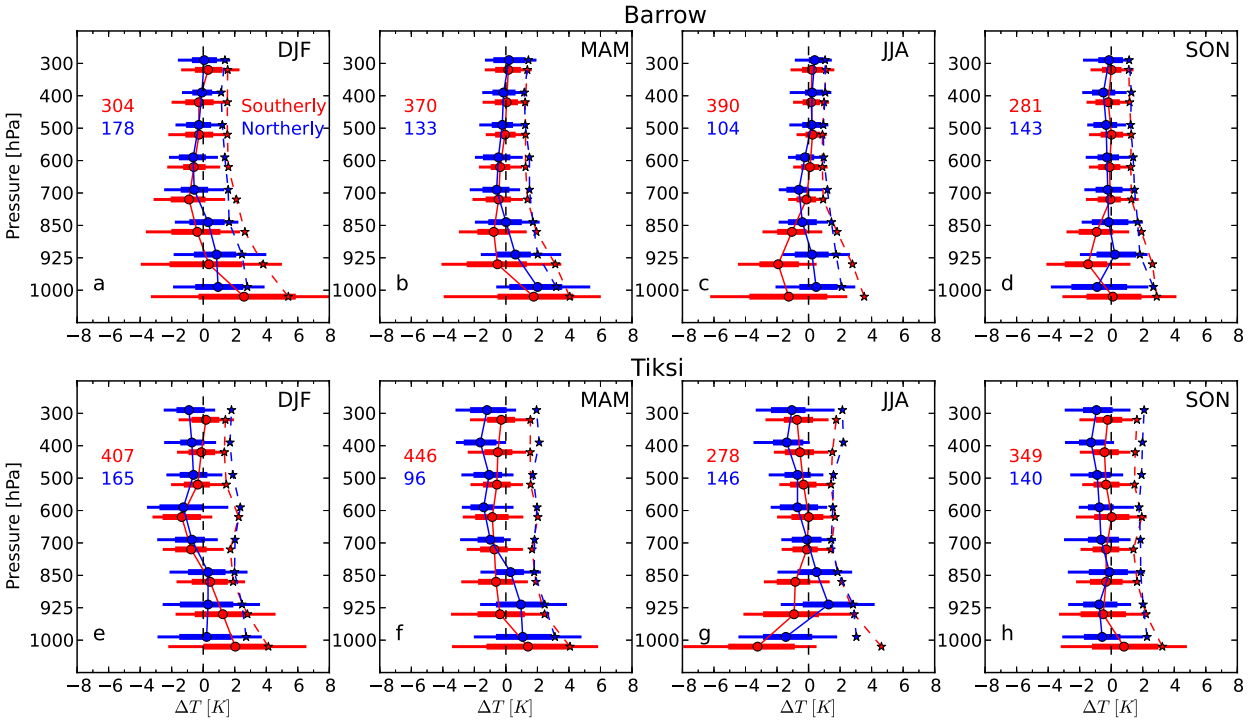

FIG. 6. Seasonal box-and-whisker error distributions (10th–90th and 25th–75th percentiles), MBE (solid lines), and RMSE (dashed lines) of AIRS T relative to radiosoundings at (a)–(d) Barrow and (e)–(h) Tiksi. Distribution and error-statistic profiles for atmospheric flow characterized as southerly at each station are shown in red, and flow characterized as northerly is shown in blue; the number of flow subsamples is included in each panel.

focus shifts to identifying the capacity of the AIRS L3 product in capturing thermodynamic features often observed within the lower Arctic troposphere.

1) GENERAL T AND Q CLIMATOLOGY AT 925 HPA

Time series of daily averaged T and Q from radiosoundings and AIRS at the 925-hPa level for stations with a continuous sounding release are shown in Fig. 8. The 925-hPa level, nominally ranging in height from 500 to 900 m, is chosen because this level is traditionally close to Arctic boundary layer processes (Shupe et al. 2013; Sedlar 2014; Sedlar and Shupe 2014; Sotiropoulou et al. 2014), but we also found that this level also has considerable MBEs and RMSEs (sections 3a and 3b). AIRS T and Q at the multiyear land-based sounding stations exhibit skill in capturing the large seasonal variability from the radiosoundings (Figs. 8a–c). Correlation coefficients (included in each panel) reveal AIRS is capable of reproducing over 88% the variance in T observed by soundings, and over 77% in Q. Climatologically, a relatively strong agreement between AIRS and radiosoundings at all pressure levels exists for all of the locations examined (Table 2). At all 3 land stations, AIRS underestimates the seasonal extremes in T and Q, most notably the winter season cold snaps at Ny-Ålesund (Fig. 8c). Similar conclusions regarding seasonal extreme values were found for the surface air temperature from AIRS L3 over the Greenland ice sheet (Hearty et al. 2018). Nonetheless, AIRS thermodynamics are successful in reproducing the observed climatological records.

Daily mean AIRS T and Q at 925 hPa during ACSE also show a robust likeness to radiosounding data (Fig. 8e). AIRS successfully captures the transition between the end of melt season and the onset of seasonal freeze in late August (e.g., Sotiropoulou et al. 2016). The strong warm and moist air advection episode during the first week of August 2014, described in detail in Tjernström et al. (2015), is also evident; even though AIRS and radiosoundings differ slightly on the timing and magnitude of the event, it is clearly an anomalous event in the AIRS data record matched to the ACSE field campaign cruise track.

The relationship between AIRS and radiosounding 925-hPa T and Q at ASCOS and Sea State (Figs. 8d,f) is less than ideal, particularly at ASCOS. The general trend in T and Q prior to 15 August 2008 at ASCOS is adequate, but large differences emerge for the remainder of the campaign (Fig. 8d). For nearly 1 week in late August, AIRS T is biased 3°–6°C warmer than the

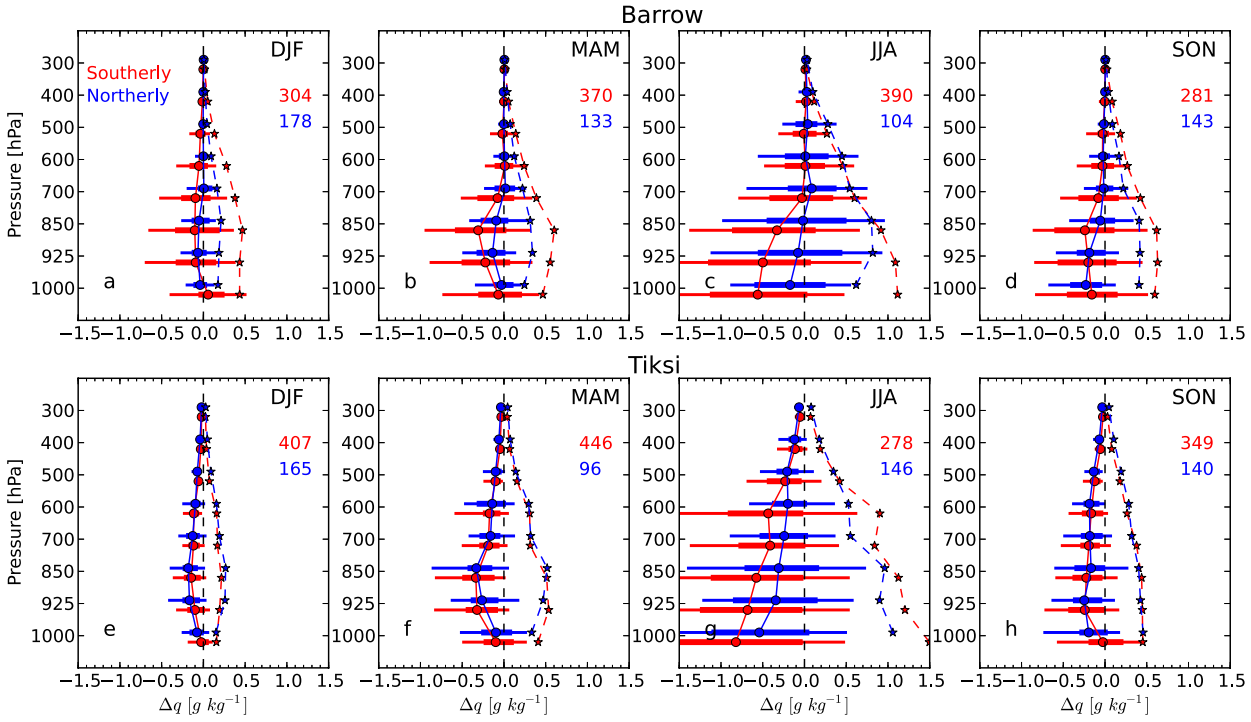

FIG. 7. As in Fig. 6, but for Q.

sounding temperatures. Sedlar et al. (2011) examined the surface energy budget during this week in detail, finding that an initial reduction in low-level mixed-phase clouds and lower-tropospheric cooling between 22 and 23 August 2008 was critical in modifying the sea ice to transition from melt season to freeze-up. AIRS T captures this brief period but fails to replicate the following week. This week is characterized by a near-adiabatic, persistently cloud-capped boundary layer with a strong capping inversion; a situation that AIRS seems to have a problem capturing. Both T and Q records are in better agreement for Sea State, although there are also days where large discrepancies between AIRS and sounding temperatures are present (Fig. 8f).

2) LOW-LEVEL TEMPERATURE AND MOISTURE STRUCTURE

Figure 9 shows the seasonal relationship between AIRS and radiosounding low-level T (Figs. 9a–d) and Q (Figs. 9e–h) differences between the 925- and 1000-hPa levels as number density distributions; all datasets have been combined for this analysis. The difference in T between these two levels is useful in determining the temperature structure in the lower troposphere, which is critically important for characterizing the static stability of the lower troposphere. Generally, the seasonal transition in lower-tropospheric stability observed in the radiosounding data is well captured by AIRS. Positive stability dominates in winter (Fig. 9a), with a transition to more frequent negative T differences across the lower troposphere in spring, continuing through summer and autumn (Figs. 9b–d).

Despite AIRS' ability to reflect the observed seasonal lower-tropospheric stability, seasonal biases are readily identifiable. The strength of positive stability values observed during spring tends to be underestimated by AIRS (Fig. 9b), with substantial scatter also observed during winter (Fig. 9a). The positive MBE in AIRS T at 1000 hPa observed during these seasons (Figs. 2a,b and 3a) is likely the reason for the underestimation of T925 minus T1000. Likewise, the magnitude of observed negative stability between 925 and 1000 hPa during summer and autumn is underestimated by AIRS (Figs. 9c,d). The underestimation is most dramatic during summer, when AIRS T differences are frequently positive, while the radiosoundings reveal negative temperature differences near the moist adiabatic lapse rate. Decreased correlation coefficients between AIRS and soundings, from 0.73 during DJF and MAM to 0.64 during JJA and SON, reflect these seasonal transition differences.

The relationship between AIRS and sounding differences in Q are less correlated than for T, with coefficient r values falling between 0.3 and 0.4 (Figs. 9e–h). During winter and spring, the peak distributions in Q differences occur at slightly positive values for both AIRS and

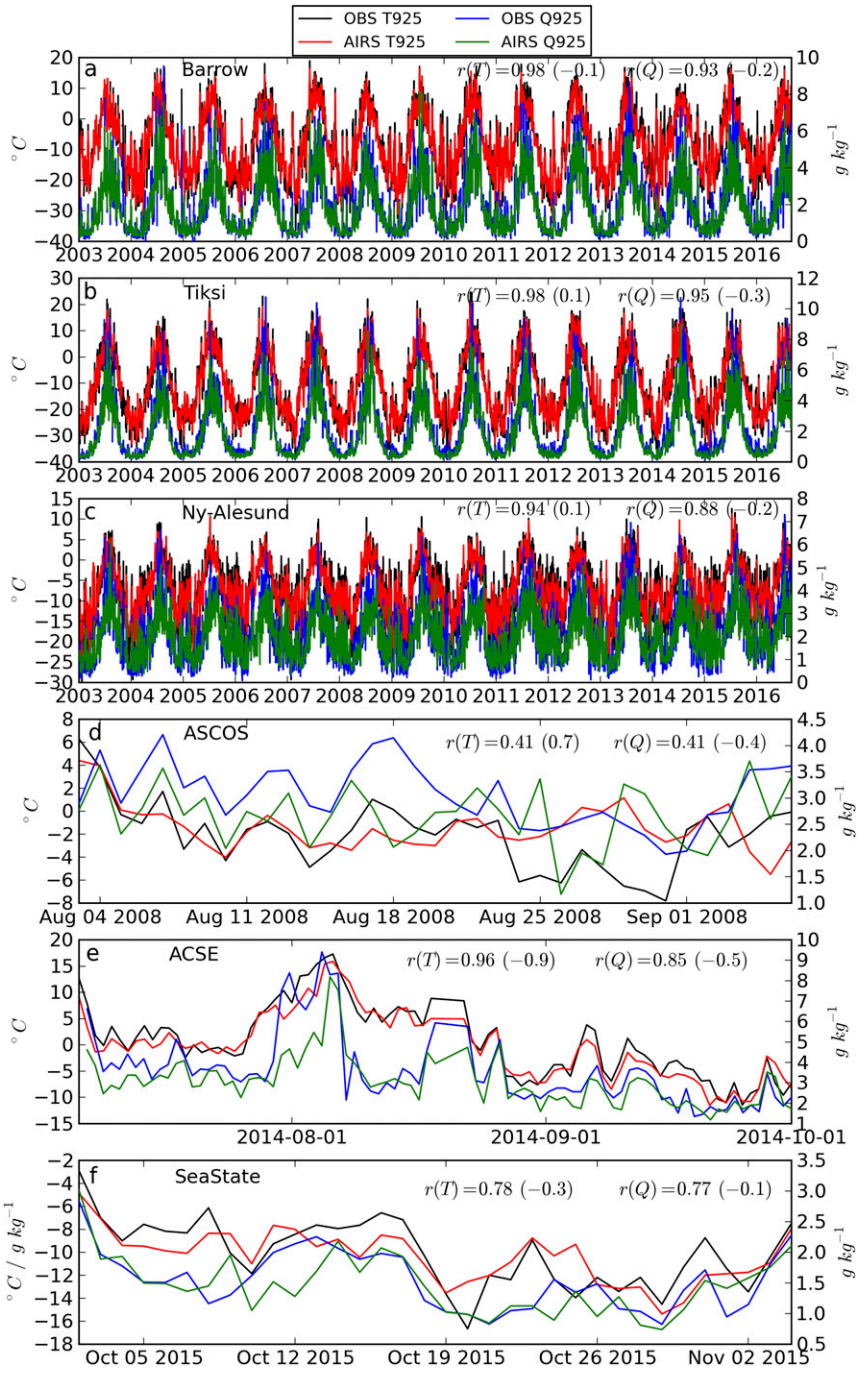

FIG. 8. Time series of daily averaged T (K) at 925 hPa from radiosoundings (black) and from AIRS (red) (y axis on left-hand side) and daily averaged Q ($g\,kg^{-1}$) at 925 hPa from radiosoundings (blue) and from AIRS (green) (y axis on right-hand side) for (a) Barrow, (b) Tiksi, (c) Ny-Ålesund, (d) ASCOS, (e) ACSE; and (f) Sea State. Correlation coefficients and MBEs (in parentheses) between radiosoundings and AIRS thermodynamics are provided in each panel.

soundings (Figs. 9e,f), although AIRS does have a tendency to underestimate the magnitude. When negative Q differences are observed, AIRS often largely overestimates the decrease in moisture. Observed summer and autumn Q differences are most frequently negative, and AIRS manages to represent this shift (Figs. 9g,h). However, the spread in distributions is large during summer, suggesting that the two datasets are occasionally

TABLE 2. Correlation coefficients for T/Q between AIRS and radiosoundings at the three standard pressure levels (1000 = 1000 hPa, etc.) for each individual sounding station or field campaign analyzed.

	Barrow	Tiksi	Ny-Ålesund	ASCOS	ACSE	N-ICE	Sea State	*Polarstern*
T1000/Q1000	0.96/0.95	0.97/0.95	0.92/0.85	0.56/0.48	0.78/0.77	0.93/0.92	0.90/0.85	0.88/0.84
T925/Q925	0.98/0.93	0.98/0.95	0.94/0.88	0.41/0.41	0.96/0.85	0.94/0.90	0.78/0.77	0.91/0.76
T850/Q850	0.99/0.92	0.99/0.94	0.96/0.87	0.83/0.45	0.98/0.86	0.97/0.92	0.89/0.80	0.95/0.79
T700/Q700	0.99/0.90	0.99/0.92	0.98/0.87	0.84/0.69	0.98/0.88	0.98/0.89	0.95/0.89	0.97/0.74
T600/Q600	0.99/0.89	0.98/0.82	0.98/0.77	0.91/0.78	0.99/0.88	0.98/0.86	0.97/0.91	0.97/0.77
T500/Q500	0.99/0.88	0.99/0.87	0.98/0.83	0.88/0.64	0.99/0.82	0.99/0.79	0.96/0.85	0.97/0.76
T400/Q400	0.98/0.87	0.98/0.88	0.98/0.84	0.94/0.86	0.99/0.86	0.98/0.86	0.96/0.81	0.97/0.75
T300/Q300	0.97/0.90	0.97/0.80	0.96/0.86	0.85/0.83	0.97/0.88	0.98/0.84	0.86/0.85	0.94/0.76

anticorrelated. An AIRS overestimation in negative Q between 925 and 1000 hPa reemerges during autumn (Fig. 9h).

3) LIMITED VERTICAL RESOLUTION AND TEMPORAL VARIABILITY

Figure 10 shows an example of radiosounding and corresponding matched AIRS L3 thermodynamic profiles over the central Arctic sea ice during ASCOS at 87.5°N on 27 August 2008. The 0000 (blue) and 1200 UTC (red) radiosoundings show a significant amount of finescale vertical structure and temporal variability in the thermodynamic profiles. This variability in T is most pronounced in the lower troposphere (Fig. 10a), while in Q, the variability is mostly found at levels above 850 hPa (Fig. 10b). From near the surface to approximately 900 hPa, radiosounding profiles reveal decreasing T with height below a temperature inversion associated with a low cloud layer. Corresponding AIRS L3 profiles at valid for 0130 (descending) and 1330 (ascending) UTC fail in capturing this important thermodynamic structure. Both ascending and descending profiles show an increasing T in a stably stratified structure below 925 hPa, resulting primarily from overly warm temperatures retrieved at the 925-hPa level. These excessively warm temperatures cause AIRS profiles to fail in reproducing the vertical structure in the in situ observed profiles, where an elevated temperature inversion capping a near-adiabatic boundary layer structure was present in both soundings and even resolved in the daily average of the two radiosounding profiles (green in Fig. 10a). This vertical structure is the most common for the summer Arctic Ocean atmospheric boundary layer (Sedlar et al. 2012; Shupe et al. 2013;

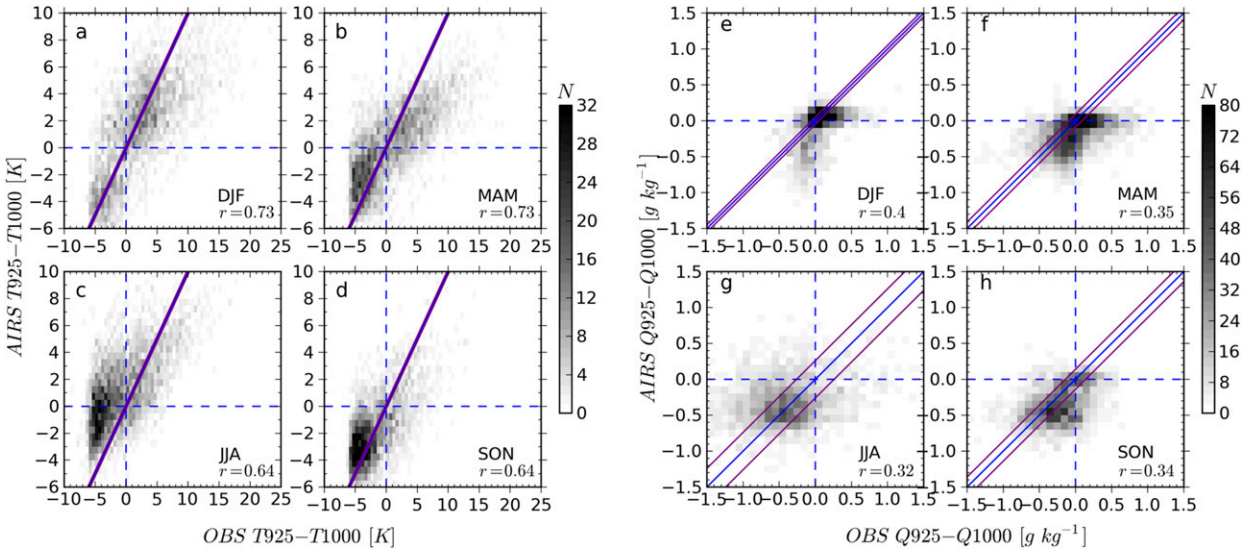

FIG. 9. Seasonal relative frequency distributions (counts N) of (a)–(d) binned T differences (K, with 0.5-K bin width) and (e)–(h) binned Q differences (g kg^{-1}, with 0.1 g kg^{-1} bin width) between the 925- and 1000-hPa levels from daily averaged AIRS vs radiosounding profiles after combining data from all observational sites. The 1:1 line is shown in blue; purple lines span the seasonal uncertainty range in T and Q radiosounding observations provided from the Barrow and Ny-Ålesund GRUAN datasets. Seasonal correlation coefficients are provided in the bottom-right panel corners.

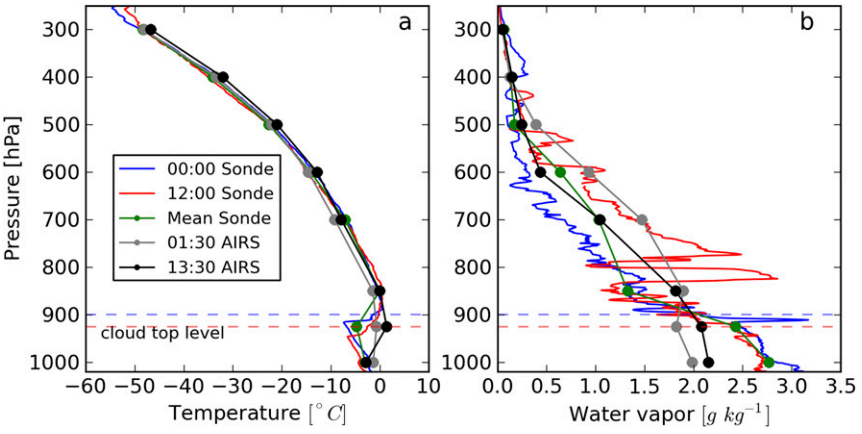

FIG. 10. Local (a) temperature and (b) water vapor mixing ration radiosounding profiles from 0000 UTC (blue) and 1200 UTC (red) on 27 Aug 2008 at approximately 87.5°N, 13°E, during ASCOS; the daily averaged T and Q profiles from these two radiosoundings at standard pressure levels are in green. The gray and black lines are the collocated AIRS L3 T and Q profiles from 0130 UTC (descending node) and 1330 UTC (ascending node) on the same day. Horizontal lines mark the vertically pointing radar-derived cloud-top levels at 0000 (blue) and 1200 (red) UTC.

Sedlar and Shupe 2014; Sotiropoulou et al. 2014) and is also quite common in winter (e.g., Tjernström and Graversen 2009).

A comparison of the 0000 and 1200 UTC Q profiles from radiosoundings and AIRS is even more ambiguous. Across much of the free troposphere, neither of the AIRS overpasses are representative of the temporal evolution of Q found in the radiosoundings (Fig. 10b); the 0130 UTC AIRS Q profile actually better matches the radiosounding from 1200 UTC, while the 1330 UTC AIRS profile is closer to the 0000 UTC sounding. Nearer the surface, AIRS Q is underestimated by at least 0.5 g kg^{-1} across the lower two layers relative to radiosoundings; at 850 hPa, AIRS Q is overestimated by a similar magnitude.

It is important to understand how frequently lower-level temperature inversion structures are actually found in the radiosounding observations, especially when artificial thermal structures like those in Fig. 10a may be created by the AIRS L3 product. Tjernström and Graversen (2009) found such inversions in essentially all of the soundings from the SHEBA experiment. A comparison of the frequency of occurrence of increasing temperatures from 1000 to 925 hPa (Fig. 11a) and from 925 to 850 hPa (Fig. 11b) for the long-term land station soundings matched with AIRS reveals unique spatial and temporal differences in lower-tropospheric temperature structure. At Barrow, AIRS and soundings agree reasonably well with regard to the monthly occurrence of an increasing T inversion structure between the lowest two levels (Fig. 11a, black), although there is a slight overestimate of their occurrence in early summer; the agreement at Tiksi is also adequate, except during summer months when AIRS substantially more often observes a T increase relative to the soundings (blue). Increasing T between 1000 and 925 hPa is surprisingly rare at Ny-Ålesund (red). AIRS reflects this low frequency of occurrence, but again during summer the presence of increasing temperatures is largely overrepresented. Elevated temperature increases between 925 and 850 hPa are present in the radiosoundings during all months (Fig. 11b). However, AIRS underestimates the occurrence of these temperature inversions at all three stations, except during January and February at Tiksi. Between May and October, AIRS almost never indicates an elevated stable layer, while the soundings have a modest occurrence frequency of 10%–25%.

These results and the example profiles of Fig. 10 highlight a few important characteristics of this AIRS L3 thermodynamic products evaluation. Averaging the thermodynamics to daily mean profiles may cause an artificial profile shape that may never be realized, both in the radiosoundings as well as with the AIRS thermodynamics. Since AIRS provides the L3 profiles on both ascending and descending orbits (1330 and 0130 local time), we also performed a more robust comparison by matching these overpasses with the corresponding local radiosoundings, when the local satellite overpass and local sounding time matched within 2 h. Statistical analyses like those in Figs. 2–5 did not differ significantly from those for daily averaged results (not shown).

The limited number of discrete pressure levels also contributes to an erroneous vertical thermodynamic

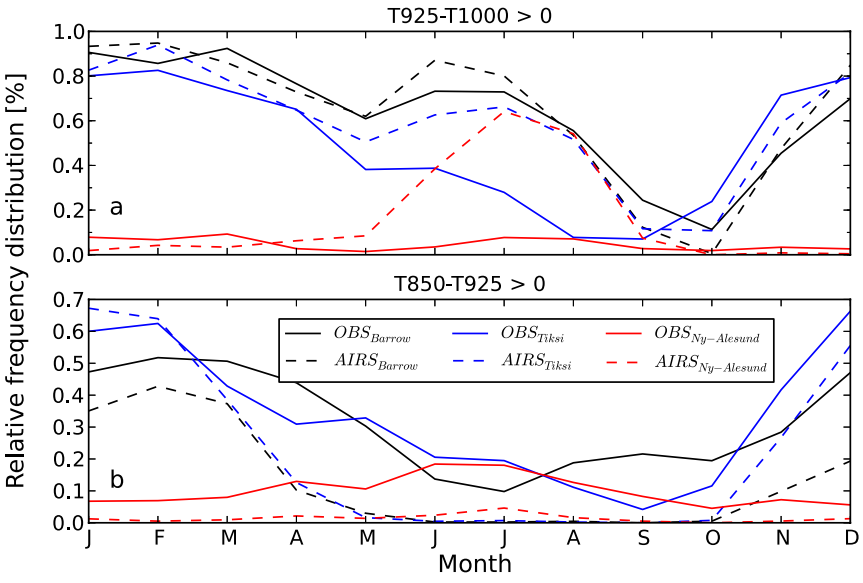

FIG. 11. Relative frequency distributions of positive temperature differences between the (a) 925- and 1000-hPa levels and between the (b) 850- and 925-hPa levels as a function of month. Results are shown for Barrow (black), Tiksi (blue) and Ny-Ålesund (red), where solid lines are temperature differences from radiosounding observations and dashed lines are from AIRS observations.

structure. Over the summer and autumn sea ice, several studies have indicated that the lower troposphere contains a mixed layer from the surface up to ~0.2–1 km, resulting from surface-layer turbulent mixing and cloud-driven buoyant overturning mixing, capped by an inversion (e.g., Tjernström et al. 2004; Sedlar et al. 2011, 2012; Shupe 2011; Shupe et al. 2013; Sedlar and Shupe 2014; Sotiropoulou et al. 2014; Brooks et al. 2017). AIRS L3 thermodynamic profile products are only available at two vertical levels that may characterize this boundary layer depth. If the retrieved T at 925 hPa happens to be within the commonly observed T inversion separating the boundary layer/cloud-mixed layer and the free troposphere, this can introduce an erroneous stably stratified profile between the levels. Likewise, large T biases, particularly at the 1000-hPa level, will further contribute toward a flawed temperature structure; the T profiles in Fig. 10a are an example of producing this mistaken lower-tropospheric stability.

4) THERMODYNAMIC ERRORS CORRESPONDING TO RADIATIVE STATES AT BARROW

The thermodynamic profiles shown in Fig. 10 occurred during a period with low-level cloud made up of a geometrically thin layer of liquid droplets and ice crystals precipitating from this layer (Sedlar et al. 2011; Shupe et al. 2013). Compared to radiosoundings, both AIRS T and Q profiles reveal relatively large differences across the level of cloud top (Fig. 10, dashed lines) and farther down below the top of the cloud. Above 925 hPa, the AIRS T did not differ greatly from the radiosoundings; the retrieved Q profiles still contain biases relative to the soundings, although the general profile shape is more accurate above the cloud-top level. This suggests that the retrieval of temperature and water vapor both within and below the cloud layer may be compromised by the cloud itself. For cloudy scenes, AIRS uses a cloud-clearing technique to estimate the clear-sky radiances that would be observed across the otherwise cloudy layer (Susskind et al. 2006). Considering the relatively large error distributions, MBEs, and RMSEs in AIRS lower-tropospheric thermodynamics, here an attempt to identify the influence of clouds on the T and Q from the AIRS L3 product at Barrow is performed.

Stramler et al. (2011) suggested that the net surface longwave radiation is an important metric to determine the radiative state of the atmosphere. Since clouds greatly enhance the absorption and reemission of longwave radiation to the surface, the net longwave flux (LWN) can be used to identify radiatively opaque (cloudy sky) and radiatively clear (clear sky or radiatively inactive) atmospheric states. Classifying the atmosphere into radiative states has since been used to examine process-level relationships related to clouds and radiation, and airmass transformations (e.g., Morrison et al. 2012; Engström et al. 2014; Pithan et al. 2014; Persson et al. 2017). Cloud boundaries can vary rather dramatically

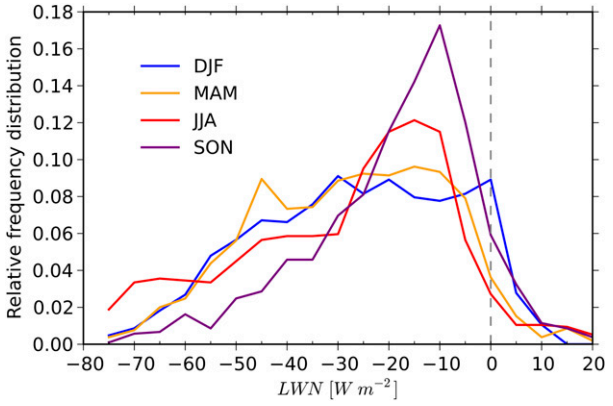

FIG. 12. Seasonal relative frequency distribution of daily averaged surface net longwave radiation (label LWN; $W\,m^{-2}$) at Barrow from 2003 to 2016 for DJF (blue), MAM (yellow), JJA (red), and SON (purple).

within the time period of 1 day, along with the formation and dissipation of multiple cloud layers; as such, a daily average of cloud fraction or cloud boundaries from surface-based remote sensing instrumentation may not be a useful metric to classify the influence of clouds on AIRS retrieved thermodynamics. Therefore, we employ the radiative state methodology proposed by Stramler et al. (2011), using LWN observations at Barrow.

During winter at SHEBA, Stramler et al. (2011) found a distinct bimodal distribution in surface LWN, which made it convenient to separate the atmospheric states radiatively based on LWN. The RFDs of seasonal LWN at Barrow during 2003–16 do not have the same bimodal distribution (Fig. 12) as was found for SHEBA by Stramler et al. (2011). The actual surface LWN in a cloudy state depends on the temperature at the cloud base, which will decrease with height to the cloud. Here, the separation is, somewhat arbitrarily, made at LWN $< -30\,W\,m^{-2}$ for the radiatively clear state and $> -20\,W\,m^{-2}$ for the radiatively opaque state, based on these RFDs. This resulted in a range of 46%–48% of the days as radiatively clear, and 33%–36% as radiatively opaque. For SON, the thresholds are shifted to -25 and $-15\,W\,m^{-2}$, respectively, to reflect the shift in the observed LWN distribution, yielding 33% radiatively clear cases and 42% radiatively opaque.

From these separated radiative states, vertical profiles of error distributions (RFDs), MBEs and RMSEs were computed seasonally for T and Q at Barrow and are shown in Fig. 13, similar to Fig. 2. While the profiles of MBE show the same pattern of biases observed for all sky conditions (Fig. 2), there are differences in the RFDs between radiative states; the spread in the error distributions across the troposphere is larger for the radiatively opaque state (Figs. 13b,d) compared to radiatively clear state (Figs. 13a,c), especially for winter and spring. This increased error distribution is primarily present from near the surface up to ~600 hPa. Additionally, across this lower to midtropospheric layer, the RMSEs for the radiatively clear state also decrease more rapidly with increasing height compared to the radiatively opaque state, especially for the Q profiles.

Figure 14 highlights the relative difference in seasonal RMSE profiles between the radiatively opaque and radiatively clear samples shown in Fig. 13. RMSE relative difference is defined as $[\mathrm{RMSE}_{\mathrm{opaque}}(\mathrm{season})-\mathrm{RMSE}_{\mathrm{clear}}(\mathrm{season})]/\mathrm{RMSE}(\mathrm{season})$, namely, the RMSE difference between opaque and clear cases is weighted by the seasonal RMSE (see Fig. 2). Below 700 hPa, RMSE in AIRS T (Fig. 14a) and Q (Fig. 14b) ranges from a few percent to over 20% larger for radiatively opaque samples compared to the radiatively clear samples; relative differences in RMSE are even larger during spring (yellow). Cloud fractions are frequently high during all seasons across the lower Arctic troposphere (e.g., Shupe et al. 2011). The absolute increase in spread of the error distributions, combined with the increased RMSE in the lower troposphere during the radiatively opaque atmospheric state, suggest that the presence of clouds has a considerable impact on the accuracy of AIRS L3 product thermodynamics. This is especially true across the lower troposphere where Arctic cloud fractions are typically large.

4. Discussion and summary

A rigorous evaluation of daily averaged temperature T and water vapor mixing ratio Q profiles from the AIRS L3 V006 product relative to Arctic in situ radiosoundings has been performed. To mimic how the broader climate science community may utilize AIRS thermodynamic profiles, we chose to analyze the L3 T and Q profiles by averaging *Aqua* overpass nodes to compute the daily mean profiles at standard pressure levels. AIRS profiles are collocated with radiosounding standard pressure levels thermodynamics from nearly 16 000 daily averaged radiosounding profiles following a strict standard pressure level comparison. While comparisons of AIRS thermodynamic structure have been evaluated against soundings over the midlatitude continental United States (Botes et al. 2012) and globally against reanalysis (Hearty et al. 2014), an evaluation across the Arctic using in situ observations as reference has been missing from the literature. This is particularly concerning considering the relatively recent applications of AIRS thermodynamics for process-level studies across the data-sparse Arctic.

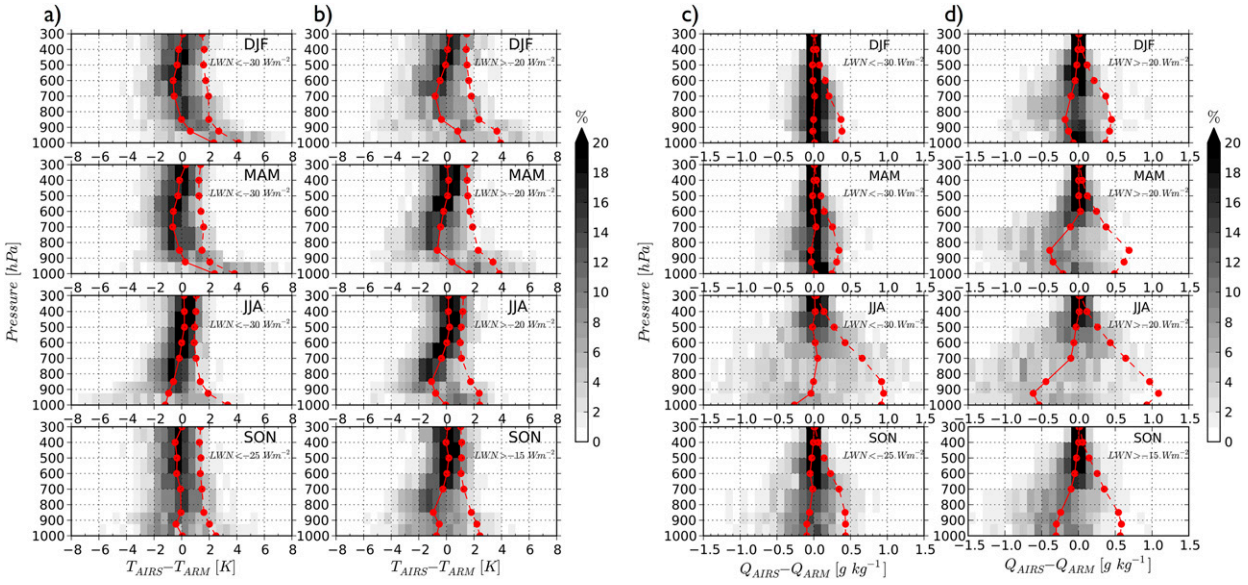

FIG. 13. Seasonal error distributions (contours; %) of (left) T error (K) and (right) Q error at Barrow (defined as AIRS minus radiosoundings); seasonal MBE (solid red) and RMSE (dashed red) profiles are also shown. Shown are the error distributions (a),(c) during the radiatively clear LWN atmospheric state subsamples and (b),(d) for the radiatively opaque LWN atmospheric state subsamples (see the text).

We find that, relative to radiosoundings, AIRS is successful in replicating the climatological records of T and Q at standard pressure levels. Over the long-duration land stations, AIRS shows a strong agreement, capturing seasonal and annual cycles of T and Q, with correlation coefficients frequently larger than 0.9 for T and 0.8 for Q. Additionally, this agreement tends to improve with height across the troposphere. This climatological accuracy of AIRS L3 thermodynamics in capturing the large annual variability is arguably its strongest suit. AIRS ability to represent the seasonal and annual cycle in T and Q on a daily mean basis suggests that AIRS L3 product has the capacity to be useful in climatological and thermodynamic anomaly studies, as in, for example, Sedlar and Devasthale (2012) and Sedlar and Tjernström (2017).

The vertical profiles of bias and RMSE were in general found to be relatively consistent regardless of the location of surface observatory around the Arctic. This is a promising result considering the troposphere

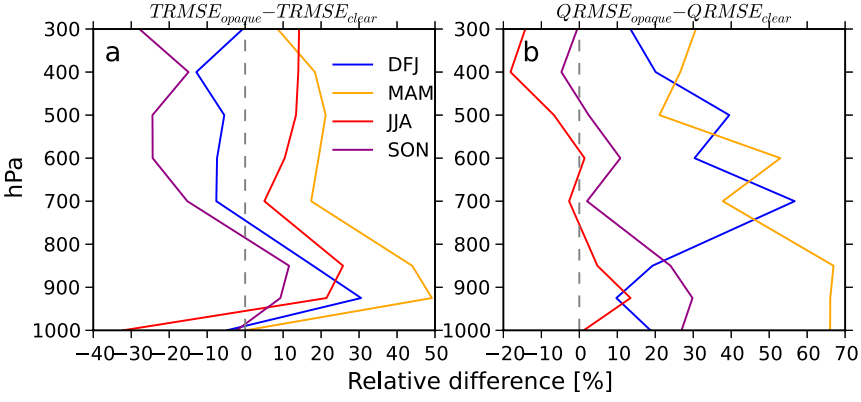

FIG. 14. Relative differences (%) in the vertical profile of (a) T RMSE and (b) Q RMSE between radiatively opaque and radiatively clear atmospheric states. Relative difference is defined as $2 \times (\mathrm{RMSE}_{\mathrm{opaque}} - \mathrm{RMSE}_{\mathrm{clear}})/(\mathrm{RMSE}_{\mathrm{opaque}} + \mathrm{RMSE}_{\mathrm{clear}})$, where the difference in RMSE is weighted by the mean seasonal RMSE as a function of pressure. A positive relative difference indicates a greater RMSE when clouds are present relative to that for clear skies.

overlying the Arctic Ocean is poorly sampled by radiosoundings; here AIRS thermodynamic soundings have the potential to fill a significant gap in data coverage. AIRS is capable of replicating the mid- to upper-troposphere thermodynamic structures at a level of accuracy that is within the level of uncertainty of both the AIRS L3 product and the radiosoundings.

Across the lower troposphere, however, distinct biases are present. AIRS is often too warm during winter and spring, modestly cold during summer and autumn, and generally too dry during all seasons. The RMSEs in T and Q tend to be the largest nearest the surface, and their magnitudes varied the most for moisture. Despite this, the relative errors in the lower troposphere were generally consistent for moisture across the seasons. Relative T errors instead showed the largest seasonal variation. During summer and early autumn, the error variability in temperature errors exceeded ±50% of the observed monthly standard deviations, while the errors were frequently below ±30% for the remaining seasons.

While these results suggest the thermodynamic constraints on T and Q vary dramatically with the season, we find that AIRS L3 product's thermodynamics may be prone to errors due to the presence of persistent, low-level Arctic cloudiness (e.g., Curry et al. 1996; Wang and Key 2005; Shupe et al. 2011). We separated the thermodynamics profiles at Barrow into radiatively clear and radiatively opaque cases (e.g., Stramler et al. 2011) and examined the bias and RMSE statistics for effectively cloud-free and effectively cloudy cases. We found that the error distributions experienced a larger spread and greater RMSE across the lower troposphere below 600 hPa for the effectively cloudy compared to effectively clear samples. Only small changes in the error distributions or statistics were identified across the upper troposphere between the two samples. These results point toward errors and difficulties in AIRS cloud-clearing radiance retrievals (e.g., Susskind et al. 2006) through atmospheric layers that are often cloudy.

The standard pressure level vertical distribution of AIRS L3 thermodynamics manage to provide sampling across the whole troposphere. However, the finite vertical sampling associated with L3 products appears to be insufficient across the lower troposphere, where large variability is commonly observed. We found that artificial stability structures were common in the AIRS profiles, particularly across the lowest 3 vertical levels: 1000, 925 and 850 hPa. While AIRS T structure is able to capture the frequently observed temperature inversions and positive thermal stability over the lower troposphere during cold, dark months, there is a tendency for these stable structures to be more frequently found in AIRS records compared to the radiosoundings during summer. We conclude that these artificial stability structures are primarily a caveat of two concurrent retrieval issues: 1) relatively large T biases at low levels, which are likely associated with persistent cloud cover; and 2) finite vertical levels that coarsely sample the transition between the Arctic cloud-driven mixed layer and the free troposphere; the interface between these two layers is often a strong temperature inversion of varying geometric thickness (Sedlar et al. 2012; Shupe et al. 2013; Sedlar and Shupe 2014; Sotiropoulou et al. 2014). Whether an AIRS L3 product provides values at a standard pressure level below, within, or above an inversion layer will make a crucial difference in the stability structure when using the finite differencing between adjacent layers. The example profiles in Fig. 10 highlight how these errors look in reality. Furthermore, artificially retrieved stability structures are likely to cause problems for applications of AIRS thermodynamics, such as near-surface bulk turbulent heat fluxes (Boisvert et al. 2013; Boisvert and Stroeve 2015), that require accurate knowledge of lower-tropospheric stability stratification.

Acknowledgments. This work was funded by the European Union's Horizon 2020 Research and Innovation Grant Agreement 727890. The authors kindly acknowledge the NASA Jet Propulsion Laboratory for providing freely available AIRS data. Likewise, we acknowledge the Integrated Global Radiosonde Archive (IGRA) and the Atmospheric Radiation Measurement (ARM) program for providing sounding and other datasets at Barrow, Tiksi, and Ny-Ålesund. Last, we are very grateful for the scientists and crew members from each of the Arctic field programs for their efforts in making high-quality in situ measurements and providing us access to these data. Thank you to the four anonymous reviewers for providing important and relevant feedback on this study.

REFERENCES

Bodeker, G. E., and Coauthors, 2016: Reference upper-air observations for climate: From concept to reality. *Bull. Amer. Meteor. Soc.*, **97**, 123–135, https://doi.org/10.1175/BAMS-D-14-00072.1.

Boisvert, L. N., and J. C. Stroeve, 2015: The Arctic is becoming warmer and wetter as revealed by the Atmospheric Infrared Sounder. *Geophys. Res. Lett.*, **42**, 4439–4446, https://doi.org/10.1002/2015GL063775.

——, T. Markus, and T. Vihma, 2013: Moisture flux changes and trends for the entire Arctic in 2003–2011 derived from EOS *Aqua* data. *J. Geophys. Res. Oceans*, **118**, 5829–5843, https://doi.org/10.1002/jgrc.20414.

Botes, D., J. R. Mecikalski, and G. J. Jedlovec, 2012: Atmospheric Infrared Sounder (AIRS) sounding evaluation and analysis of

the pre-convective environment. *J. Geophys. Res.*, **117**, D09205, https://doi.org/10.1029/2011JD016996.

Brooks, I. M., and Coauthors, 2017: The turbulent structure of the Arctic summer boundary layer during the Arctic Summer Cloud-Ocean Study. *J. Geophys. Res. Atmos.*, **122**, 9685–9704, https://doi.org/10.1002/2017JD027234.

Chahine, M. T., and Coauthors, 2006: AIRS: Improving weather forecasting and providing new data on greenhouse gases. *Bull. Amer. Meteor. Soc.*, **87**, 911–926, https://doi.org/10.1175/BAMS-87-7-911.

Curry, J. A., W. B. Rossow, D. Randall, and J. L. Schramm, 1996: Overview of Arctic cloud and radiation characteristics. *J. Climate*, **9**, 1731–1764, https://doi.org/10.1175/1520-0442(1996)009<1731:OOACAR>2.0.CO;2.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data analysis system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Devasthale, A., U. Willén, K.-G. Karlsson, and C. G. Jones, 2010: Quantifying the clear-sky temperature inversion frequency and strength over the Arctic Ocean during summer and winter seasons from AIRS profiles. *Atmos. Chem. Phys.*, **10**, 5565–5572, https://doi.org/10.5194/acp-10-5565-2010.

——, J. Sedlar, and M. Tjernström, 2011: Characteristics of water-vapour inversions observed over the Arctic by Atmospheric Infrared Sounder (AIRS) and radiosondes. *Atmos. Chem. Phys.*, **11**, 9813–9823, https://doi.org/10.5194/acp-11-9813-2011.

——, ——, T. Koenigk, and E. J. Fetzer, 2013: The thermodynamic state of the Arctic atmosphere observed by AIRS: Comparisons during the record minimum sea ice extents of 2007 and 2012. *Atmos. Chem. Phys.*, **13**, 7441–7450, https://doi.org/10.5194/acp-13-7441-2013.

——, ——, B. H. Kahn, M. Tjernström, E. J. Fetzer, B. Tian, J. Teixeira, and T. S. Pagano, 2016: A decade of spaceborne observations of the Arctic atmosphere: Novel insights from NASA's AIRS instrument. *Bull. Amer. Meteor. Soc.*, **97**, 2163–2176, https://doi.org/10.1175/BAMS-D-14-00202.1.

Dirksen, R. J., M. Sommer, F. J. Immler, D. F. Hurst, R. Kivi, and H. Vömel, 2014: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde. *Atmos. Meas. Tech.*, **7**, 4463–4490, https://doi.org/10.5194/amt-7-4463-2014.

Engström, A., J. Karlsson, and G. Svensson, 2014: The importance of representing mixed-phase clouds for simulating distinctive atmospheric states in the Arctic. *J. Climate*, **27**, 265–272, https://doi.org/10.1175/JCLI-D-13-00271.1.

Granskog, M. A., P. Assmy, S. Gerland, G. Spreen, H. Steen, and L. H. Smedsrud, 2016: Arctic research on thin ice: Consequences of Arctic sea ice loss. *Eos, Trans. Amer. Geophys. Union*, **97**, https://doi.org/10.1029/2016EO044097.

Hearty, T. J., and Coauthors, 2014: Estimating sampling biases and measurement uncertainties of AIRS/AMSU-A temperature and water vapor observations using MERRA reanalysis. *J. Geophys. Res. Atmos.*, **119**, 2725–2741, https://doi.org/10.1002/2013JD021205.

Hearty, T. J., III, J. N. Lee, D. L. Wu, R. Cullather, J. M. Blaisdell, J. Susskind, and S. M. J. Nowicki, 2018: Intercomparison of surface temperatures from AIRS, MERRA, and MERRA-2 with NOAA and GC-Net weather stations at Summit, Greenland. *J. Appl. Meteor. Climatol.*, **57**, 1231–1245, https://doi.org/10.1175/JAMC-D-17-0216.1.

Ho, S.-P., L. Peng, and H. Vömel, 2017: Characterization of the long-term radiosonde temperature biases in the upper troposphere and lower stratosphere using COSMIC and Metop-A/GRAS data from 2006 to 2014. *Atmos. Chem. Phys.*, **17**, 4493–4511, https://doi.org/10.5194/acp-17-4493-2017.

Huang, J., and Coauthors, 2017: Recently amplified arctic warming has contributed to a continual global warming trend. *Nat. Climate Change*, **7**, 875–879, https://doi.org/10.1038/s41558-017-0009-5.

Ingleby, B., 2017: An assessment of different radiosonde types 2015/2016. ECMWF Tech. Memo. 807, 70 pp., https://www.ecmwf.int/sites/default/files/elibrary/2017/17551-assessment-different-radiosonde-types-20152016.pdf.

Jones, T. A., and D. J. Stensrud, 2012: Assimilating AIRS temperature and mixing ratio profiles using an ensemble Kalman filter approach for convective-scale forecasts. *Wea. Forecasting*, **27**, 541–564, https://doi.org/10.1175/WAF-D-11-00090.1.

König-Langlo, G., 2009: Upper air soundings during *POLARSTERN* cruise ARK-XXIII/1. Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven, PANGAEA, accessed 27 October 2017, https://doi.org/10.1594/PANGAEA.849232.

Morrison, H., G. de Boer, G. Feingold, J. Harrington, M. D. Shupe, and K. Sulia, 2012: Resilience of persistent Arctic mixed-phase clouds. *Nat. Geosci.*, **5**, 11–17, https://doi.org/10.1038/ngeo1332.

Nygård, T., T. Valkonen, and T. Vihma, 2014: Characteristics of Arctic low-tropospheric humidity inversions based on radio soundings. *Atmos. Chem. Phys.*, **14**, 1959–1971, https://doi.org/10.5194/acp-14-1959-2014.

Parkinson, C. L., 2003: *Aqua*: An Earth-observing satellite mission to examine water and other climate variables. *IEEE Trans. Geosci. Remote Sens.*, **41**, 173–183, https://doi.org/10.1109/TGRS.2002.808319.

Persson, P. O. G., M. D. Shupe, D. Perovich, and A. Solomon, 2017: Linking atmospheric synoptic transport, cloud phase, surface energy fluxes, and sea-ice growth: Observations of midwinter SHEBA conditions. *Climate Dyn.*, **49**, 1341–1364, https://doi.org/10.1007/s00382-016-3383-1.

——, and Coauthors, 2018: Shipboard observations of the meteorology and near-surface environment during autumn freezeup in the Beaufort/Chukchi Seas. *J. Geophys. Res. Oceans*, **123**, 4930–4969, https://doi.org/10.1029/2018JC013786.

Pithan, F., and T. Mauritsen, 2014: Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nat. Geosci.*, **7**, 181–184, https://doi.org/10.1038/ngeo2071.

——, B. Medeiros, and T. Mauritsen, 2014: Mixed-phase clouds cause climate model biases in Arctic wintertime temperature inversions. *Climate Dyn.*, **43**, 289–303, https://doi.org/10.1007/s00382-013-1964-9.

Reale, O., J. Susskind, R. Rosenberg, E. Brin, E. Liu, L. Riishojgaard, J. Terry, and J. C. Jusem, 2008: Improving forecast skill by assimilation of quality-controlled AIRS temperature retrievals under partially cloudy conditions. *Geophys. Res. Lett.*, **35**, L08809, https://doi.org/10.1029/2007GL033002.

——, E. L. McGrath-Spangler, W. McCarty, D. Holdaway, and R. Gelaro, 2018: Impact of adaptively thinned AIRS cloud-cleared radiances on tropical cyclone representation in a global data assimilation and forecast system. *Wea. Forecasting*, **33**, 909–931, https://doi.org/10.1175/WAF-D-17-0175.1.

Screen, J. A., and I. Simmonds, 2010: The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, **464**, 1334–1337, https://doi.org/10.1038/nature09051.

Sedlar, J., 2014: Implications of limited liquid water path on static mixing within Arctic low-level clouds. *J. Appl. Meteor. Climatol.*, **53**, 2775–2789, https://doi.org/10.1175/JAMC-D-14-0065.1.

——, and A. Devasthale, 2012: Clear-sky thermodynamic and radiative anomalies over a sea ice sensitive region of the Arctic. *J. Geophys. Res.*, **117**, D19111, https://doi.org/10.1029/2012JD017754.

——, and M. D. Shupe, 2014: Characteristic nature of vertical motions observed in Arctic mixed-phase stratocumulus. *Atmos. Chem. Phys.*, **14**, 3461–3478, https://doi.org/10.5194/acp-14-3461-2014.

——, and M. Tjernström, 2017: Clouds, warm air, and a climate cooling signal over the summer Arctic, 2017. *Geophys. Res. Lett.*, **44**, 1095–1103, https://doi.org/10.1002/2016GL071959.

——, and Coauthors, 2011: A transitioning Arctic surface energy budget: The impacts of solar zenith angle, surface albedo and cloud radiative forcing. *Climate Dyn.*, **37**, 1643–1660, https://doi.org/10.1007/s00382-010-0937-5.

——, M. D. Shupe, and M. Tjernström, 2012: On the relationship between thermodynamic structure and cloud top, and its climate significance in the Arctic. *J. Climate*, **25**, 2374–2393, https://doi.org/10.1175/JCLI-D-11-00186.1.

Seidel, D. J., and Coauthors, 2009: Reference upper-air observations for climate: Rationale, progress, and plans. *Bull. Amer. Meteor. Soc.*, **90**, 361–369, https://doi.org/10.1175/2008BAMS2540.1.

Shupe, M. D., 2011: Clouds at Arctic atmospheric observatories. Part II: Thermodynamic phase characteristics. *J. Appl. Meteor. Climatol.*, **50**, 645–661, https://doi.org/10.1175/2010JAMC2468.1.

——, V. P. Walden, E. Eloranta, T. Uttal, J. R. Campbell, S. M. Starkweather, and M. Shiobara, 2011: Clouds at Arctic atmospheric observatories. Part I: Occurrence and macrophysical properties. *J. Appl. Meteor. Climatol.*, **50**, 626–644, https://doi.org/10.1175/2010JAMC2467.1.

——, P. O. G. Persson, I. M. Brooks, M. Tjernström, J. Sedlar, T. Mauritsen, S. Sjogren, and C. Leck, 2013: Cloud and boundary layer interactions over the Arctic sea ice in late summer. *Atmos. Chem. Phys.*, **13**, 9379–9400, https://doi.org/10.5194/acp-13-9379-2013.

Singh, R., C. M. Kishtawal, and P. K. Pal, 2011: Use of Atmospheric Infrared Sounder clear-sky and cloud-cleared radiances in the Weather Research and Forecasting 3DVAR assimilation system for mesoscale weather predictions over the Indian region. *J. Geophys. Res.*, **116**, D22116, https://doi.org/10.1029/2011JD016379.

Sotiropoulou, G., J. Sedlar, M. Tjernström, M. D. Shupe, I. M. Brooks, and P. O. G. Persson, 2014: The thermodynamic structure of summer Arctic stratocumulus and the dynamic coupling to the surface. *Atmos. Chem. Phys.*, **14**, 12 573–12 592, https://doi.org/10.5194/acp-14-12573-2014.

——, and Coauthors, 2016: Atmospheric conditions during the Arctic Clouds in Summer Experiment (ACSE): Contrasting open water and sea ice surfaces during melt and freeze-up seasons. *J. Climate*, **29**, 8721–8744, https://doi.org/10.1175/JCLI-D-16-0211.1.

Stramler, K., A. D. Del Genio, and W. B. Rossow, 2011: Synoptically driven Arctic winter states. *J. Climate*, **24**, 1747–1762, https://doi.org/10.1175/2010JCLI3817.1.

Sun, B., A. Reale, D. J. Seidel, and D. C. Hunt, 2010: Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics. *J. Geophys. Res.*, **115**, D23104, https://doi.org/10.1029/2010JD014457.

——, ——, S. Schroeder, D. J. Seidel, and B. Ballish, 2013: Toward improved corrections for radiation-induced biases in radiosonde temperature observations. *J. Geophys. Res. Atmos.*, **118**, 4231–4243, https://doi.org/10.1002/jgrd.50369.

Susskind, J., C. Barnet, J. Blaisdell, L. Iredell, F. Keita, L. Kouvaris, G. Molnar, and M. Chahine, 2006: Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover. *J. Geophys. Res.*, **111**, D09S17, https://doi.org/10.1029/2005JD006272.

Tjernström, M., and R. G. Graversen, 2009: The vertical structure of the lower Arctic troposphere analysed from observations and the ERA-40 reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 431–443, https://doi.org/10.1002/qj.380.

——, C. Leck, P. O. G. Persson, M. L. Jensen, S. P. Oncley, and A. Targino, 2004: The summertime Arctic atmosphere: Meteorological measurements during the Arctic Ocean Experiment 2001. *Bull. Amer. Meteor. Soc.*, **85**, 1305–1321, https://doi.org/10.1175/BAMS-85-9-1305.

——, and Coauthors, 2014: The Arctic Summer Cloud Ocean Study (ASCOS): Overview and experimental design. *Atmos. Chem. Phys.*, **14**, 2823–2869, https://doi.org/10.5194/acp-14-2823-2014.

——, and Coauthors, 2015: Warm-air advection, air mass transformation and fog causes rapid ice melt. *Geophys. Res. Lett.*, **42**, 5594–5602, https://doi.org/10.1002/2015GL064373.

Wang, X., and J. R. Key, 2005: Arctic surface, cloud and radiation properties based on the AVHRR Polar Pathfinder dataset. Part I: Spatial and temporal characteristics. *J. Climate*, **18**, 2558–2574, https://doi.org/10.1175/JCLI3438.1.

Wong, S., E. J. Fetzer, M. Schreier, G. Manipon, E. F. Fishbein, B. H. Kahn, Q. Yue, and F. W. Irion, 2015: Cloud-induced uncertainties in AIRS and ECMWF temperature and specific humidity. *J. Geophys. Res. Atmos.*, **120**, 1880–1901, https://doi.org/10.1002/2014JD022440.

Yue, Q., E. J. Fetzer, B. H. Kahn, S. Wong, G. Manipon, A. Guillaume, and B. Wilson, 2013: Cloud-state-dependent sampling in airs observations based on *CloudSat* cloud classification. *J. Climate*, **26**, 8357–8377, https://doi.org/10.1175/JCLI-D-13-00065.1.