Surface Air Temperature and Humidity from Intersatellite-Calibrated HIRS Measurements in High Latitudes

LEI SHI
NOAA/National Climatic Data Center, Asheville, North Carolina

GE PENG
NOAA/National Climatic Data Center, and STG, Inc., Asheville, North Carolina

JOHN J. BATES
NOAA/National Climatic Data Center, Asheville, North Carolina

(Manuscript received 3 January 2011, in final form 21 July 2011)

ABSTRACT

High-latitude ocean surface air temperature and humidity derived from intersatellite-calibrated High-Resolution Infrared Radiation Sounder (HIRS) measurements are examined. A neural network approach is used to develop retrieval algorithms. HIRS simultaneous nadir overpass observations from high latitudes are used to intercalibrate observations from different satellites. Investigation shows that if HIRS observations were not intercalibrated, then it could lead to intersatellite biases of \(1^\circ\text{C}\) in the air temperature and 1–2 g kg\(^{-1}\) in the specific humidity for high-latitude ocean surface retrievals. Using a full year of measurements from a high-latitude moored buoy site as ground truth, the instantaneous (matched within a half-hour) root-mean-square (RMS) errors of HIRS retrievals are 1.50 \(^\circ\text{C}\) for air temperature and 0.86 g kg\(^{-1}\) for specific humidity. Compared to a large set of operational moored and drifting buoys in both northern and southern oceans greater than 50\(^\circ\) latitude, the retrieval instantaneous RMS errors are within 2.6 \(^\circ\text{C}\) for air temperature and 1.4 g kg\(^{-1}\) for specific humidity. Compared to 5 yr of International Maritime Meteorological Archive in situ data, the HIRS specific humidity retrievals show less than 0.5 g kg\(^{-1}\) of differences over the majority of northern high-latitude open oceans.

1. Introduction

The High-Resolution Infrared Radiation Sounder (HIRS) has been on board the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellite series for more than 30 yr. There are 20 channels in the HIRS instrument, providing measurement in a spatial resolution of approximately 20 km at nadir. Among these channels, channels 7, 8, 10, and 11 are designed to measure the surface and lower-atmosphere temperature and humidity. Central wavenumbers for these channels are approximately 750 cm\(^{-1}\) for channel 7, 900 cm\(^{-1}\) for channel 8, 1225 or 800 cm\(^{-1}\) (varied for different satellites) for channel 10, and 1365 cm\(^{-1}\) for channel 11 [see Shi et al. (2008) for variations of central wavenumbers from satellite to satellite]. Channel 7 observes near-surface air temperatures. Channel 8 is a surface window channel. Channel 10 is in the water vapor continuum band for near-surface observation. Channel 11 is a lower-atmosphere water vapor channel. This study examines the high-latitude sea surface air temperature (\(T_a\)) and surface specific humidity (\(Q_a\)) derived from HIRS measurements.

The surface temperature and humidity are key components in computing surface turbulent heat fluxes. Past studies (Curry et al. 2004; Jackson et al. 2006) showed that a significant portion of errors for current air–sea heat flux datasets is due to uncertainties in retrieving \(T_a\) and \(Q_a\). Liu and Curry (2006) also showed that the discrepancies of the interannual variability and decadal trend of latent heat flux among satellite-derived products and reanalyses are primarily caused by the difference in \(Q_a\).
These studies illustrated the need to improve $T_a$ and $Q_a$ retrievals for their application in deriving sensible heat flux (SHF) and latent heat flux (LHF).

In recent years, a number of studies have used satellite microwave sensors to derive sea surface air temperature and humidity. For example, Jackson et al. (2006) examined the retrieval of surface air temperature and specific humidity using multisensor microwave observations. The sensors used included the Advanced Microwave Sounding Unit-A (AMSU-A), Special Sensor Microwave Imager (SSM/I), and Special Sensor Microwave Temperature Sounder (SSM/T-2). The use of multiple sensors led to a reduction in root-mean-square (RMS) error to 0.96 g kg$^{-1}$ from previous values of 1.33–1.49 g kg$^{-1}$ in specific humidity, and a reduction in RMS error to 1.96°C from previous values of 3.28°C–4.60°C in air temperature. The retrievals were later further improved through refinements to the regression formula, training dataset, collocation procedure, and height adjustment to 10 m (Jackson et al. 2009). Zong et al. (2007) used the Advanced Microwave Scanning Radiometer for Earth Observing System (EOS) (AMSR-E) to derive near-surface specific humidity. Roberts et al. (2010) developed retrieval algorithms based on the neural network technique to derive sea surface temperature, air temperature, specific humidity, and wind speed. Jackson and Wick (2010) derived sea surface temperature and surface air temperature from AMSU-A measurements using the multivariable linear regression approach. These studies advanced our knowledge of satellite retrieval of surface temperature and humidity. Some of these studies used intersatellite-calibrated data (e.g., Roberts et al. 2010). However, many of these past studies did not examine intersatellite biases. Because of the independence in the calibration based on an individual instrument’s channel spectral response function, along with other factors, biases exist from satellite to satellite. Without intersatellite calibration, developed algorithms cannot be applied to other satellites for climatological studies.

In the present study we perform an intersatellite calibration based on simultaneous nadir overpass (SNO) observations from HIRS. HIRS measurements in the NOAA series of polar-orbiting satellites are calibrated to a baseline satellite to form a temporally homogeneous database. A neural network technique is used to develop $T_a$ and $Q_a$ retrieval schemes.

2. Data processing

a. Intersatellite calibration

HIRS has been used to derive temperature and humidity profiles since the early days of its observation (e.g., Chedin et al. 1985). There have been more than a dozen satellites carrying the HIRS instruments in the NOAA polar-orbiting satellite series. To minimize intersatellite biases, measurements from HIRS SNOs are examined (Cao et al. 2005; Shi et al. 2008). These measurements are taken at the orbital intersections of each pair of satellites viewing the same earth target within a few seconds. The satellite intersections are found in the regions from $+70^\circ$ to $+80^\circ$, and from $-70^\circ$ to $-80^\circ$ latitude zones once every 8–9 days. These data are ideally suited for intersatellite calibration over high latitudes.

The intersatellite biases derived from SNO for channels 7, 8, 10, and 11 are shown in Fig. 1. Channel 7 is located at the sharp transmission line of the infrared spectrum for sensing the near-surface air temperature. For about half of the satellite pairs (NOAA-7 and -8, NOAA-14 and -15, NOAA-15 and -16, and NOAA-16 and -17), the variations of channel 7 bias across the observed temperature ranges are larger than 1 K. For other pairs of satellites with bias changes of less than 1 K, the variations are also notable. The bias values for channel 8 are generally within $\pm 0.2$ K for most of the satellite pairs, except for the NOAA-6 and -7 pair, which has biases of around 0.4 K. Very large bias values are observed for several satellite pairs of channel 10. As discussed in Shi et al. (2008), one of the contributing factors for the large biases are due to the spectral changes of channel 10 in HIRS instruments on board different NOAA polar satellites. The center frequency of channel 10 was near 1225 cm$^{-1}$ for NOAA-6 to -10 and for NOAA-12, but changed to 796 cm$^{-1}$ for NOAA-11 and -14, and near 802 cm$^{-1}$ for NOAA-15 and after. The figure also shows that the channel 10 biases tend to be larger at low brightness temperatures than at high brightness temperatures. Large differences reflect the level of sensitivity each frequency has to the presence of water vapor and temperature near the surface. The biases for channel 11 are mostly within $\pm 0.6$ K. More than half of the satellite pairs have brightness temperature–dependent bias variations of larger than 0.5 K. Based on the intersatellite bias dataset, the HIRS channel brightness temperature data from individual satellites are adjusted to a base satellite. In this study the base satellite is chosen to be NOAA-12. This satellite is in the middle of the 30-yr HIRS time series, and it is the same base satellite used in an earlier study by Shi and Bates (2011). The data processing follows an established procedure by Shi and Bates (2011) in that the adjustments are dependent on the observing brightness temperatures.

b. Neural network retrieval

A neural network approach is used to develop temperature and humidity retrieval algorithms. The training
dataset is constructed by using a radiative transfer model, the Radiative Transfer for Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder version 9 (RTTOV-9), to simulate a diverse sample of reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) system (Chevallier 2001). One year of surface temperature and specific humidity data from ship measurements are also incorporated as a calibration dataset. The ship measurements include data from voluntary observing ships covering a large area of northern high-latitude and a smaller portion of southern high-latitude oceans. Among the many variables, the ECMWF data include surface skin temperature, air temperature, and specific humidity. The global model reanalysis data are selected instead of the surface in situ measurement mainly for two reasons. First, the global in situ surface observations are unevenly distributed. The surface observations over many of the remote areas are either lacking or insufficient. It may lead to a bias of the retrieval scheme toward the environment where denser observation sites are located. And, second, the HIRS pixel resolution is 20 km at nadir and close to 30 km at the sides of a scan line. Many of the land-based and near-coast surface observations are not spatially homogeneous enough to represent the conditions over such large areas. To reduce potential bias from radiative transfer model simulation, ship measurements adjusted to 10 m in height are collocated with retrievals. The differences between the two datasets are derived from a temperature- and humidity-dependent calibration dataset. The calibration is then applied to all retrievals. The RTTOV model is a broadband model in that the integration over the channel response is simulated directly. A description of RTTOV can be found in Saunders et al. (1999). RTTOV was one of the models in an intercomparison of HIRS and AMSU channel radiance computation (Garand et al. 2001). When a selection of six HIRS channel simulations is compared to a line-by-line model, the standard deviation of a HIRS near-surface channel (channel 10) in the RTTOV simulation is 0.25 K, while the standard deviations of RTTOV simulations for mid- and upper-atmosphere channels vary from 0.09 to 0.55 K. In general, the RTTOV model was
found to provide good performance among the participating models. In the current study, the radiative transfer model is run with 36 vertical levels from pressure levels of 1010–0.1 hPa, along with surface parameters. Because the HIRS dataset has been calibrated to a base satellite, the training dataset is constructed using radiative model simulation of the base satellite.

Fully connected, feed-forward, back-propagation neural networks following the approach used by Shi (2001) are applied in developing the retrieval scheme. A similar neural network structure was also discussed in Roberts et al. (2010). In the present study different architectures with different numbers of layers and transfer functions are examined for the temperature and humidity retrieval.

A three-layer network, including one input layer, one hidden layer, and one output layer, is applied. The use of a hyperbolic tangent function to propagate to the hidden layer and a logistic transfer function to propagate to the output layers was found to give the optimum network performance for the type of data studied. For the hidden layer, when more neurons are added to a back-propagation neural network, more degrees of freedom are obtained. The network is usually able to store more complex patterns. However, an increased number of neurons that produce a tighter fit to the training set can cause poor generalization for new cases. For the Ta and Qa retrievals, experiments are done for a large range of hidden neuron numbers. Good network performances are found when the numbers of hidden neurons are in the range of 40–55 neurons. Among the total of 24 808 patterns in the training set, 20% of data are randomly chosen and set aside for testing during neural network training to select the neural network structure that has optimal performance. The back-propagation network is trained automatically based on specified training criteria. When the predetermined convergence criteria are met, the network uses the input variables in the testing dataset and retrieves the output elements. Then, the retrieved output variables are compared with the output variables in the testing set to obtain errors. After numerous iterations, the network parameters are saved when no improvement is found for a specified number of test trials. The input set for the retrievals includes the HIRS surface and lower-atmosphere temperature and water vapor channels 7, 8, 10, and 11, and the carbon dioxide concentration. Because channel 7 is in the carbon dioxide band, the change of carbon dioxide concentration can have an effect on channel 7 measurement, and therefore the carbon dioxide concentration is included in the input to account for this effect. Though the retrieved variables include surface skin temperature, Ta, and Qa over both land and ocean surfaces, the current stage of our study has been focused on examining Ta and Qa retrievals over high-latitude ocean surfaces.

c. Cloud removal

The HIRS data are first processed to remove cloudy pixels, and limb correction is applied. The cloud-clearing procedure follows the method detailed in Jackson et al. (2003). The process is accomplished using a simplified method based on the International Satellite Cloud Climatology Project (ISCCP) cloud detection approach (Rossow and Garder 1993). This approach combines spatial and temporal variations in the brightness temperature and applies thresholds to these variations to detect clouds. To minimize any possible cloud contamination, the thresholds were chosen to remove clouds at the expense of removing some clear-sky observations. Wylie et al. (2007) compiled cloud cover derived from HIRS and cloud cover derived from the Geoscience Laser Altimeter System (GLAS) lidar on the Ice, Cloud, and Land Elevation Satellite (ICESat) spacecraft. It was found that the geographical distributions of cloud cover show general agreement between both GLAS and HIRS. The largest difference between GLAS and HIRS was found at the poles north and south of 70° latitude. In the Arctic Ocean, HIRS reported 0.1 more cloud frequencies than GLAS.

In the high latitudes strong temperature inversions often exist in the lower atmosphere. Clouds with tops in the inversions appearing warmer than the surface and surrounding areas can be mistakenly labeled as clear pixels. In an effort to remove these remaining cloudy pixels, the characteristics of the HIRS surface and lower-atmosphere sounding channels in lower-atmospheric temperature inversion layers are analyzed by using simulations from RTTOV-9. Among the sampled ECMWF profiles discussed previously, the profiles with temperature inversions are selected. These profiles are divided into two groups, one with clear sky only and the other with clouds in the lower-atmospheric inversion layer. The differences between channels 7 and 8 brightness temperatures are obtained from the RTTOV-9 simulation. These are the two temperature channels sensing the lowest portion of the atmosphere and the surface. For each of the two groups of profiles, differences between surface skin temperatures and temperatures at 700 and 850 hPa are calculated, respectively. Scatterplots of the HIRS channel differences and the surface skin and air temperature differences are presented in Fig. 2. In the figure, Ts stands for surface skin temperature, the “+” denotes the clear-sky atmosphere, and the “o” denotes the atmosphere with clouds in the temperature inversion layer. Figure 2 shows that when there are clouds in the inversion layer, channel 8 brightness temperatures are generally larger than channel 7 brightness temperatures, even though for some clear-sky profiles that have temperature
inversions the brightness temperatures of channel 8 can also be larger than those of channel 7. However, the plots show that in the region where channel 8 brightness temperatures are less than channel 7 brightness temperatures, the profiles are almost all composed of clear-sky profiles. On the basis of this result, criteria are developed to further remove the possible cloudy HIRS pixels in temperature inversions that were labeled with clear sky during the initial cloud removal process. This new method removes an addition of ~5% of pixels from the original all-sky data. The HIRS clear-sky dataset used in this study is very conservative with likely fewer clear-sky observations than reality. However these cloud-removing procedures are necessary in minimizing contamination from clouds.

3. Comparison with in situ data

One full year of buoy data from International Comprehensive Ocean–Atmosphere Dataset (ICOADS) are used to examine ocean surface retrievals. The buoy data were not used in developing the retrieval algorithm and are considered as an independent validation dataset. The measurements from buoy data have been adjusted to 10-m height to achieve uniformity. A description of the dataset is provided by Worley et al. (2005). Among all the moored buoys, we selected a site from the highest latitude location that provides a full year of hourly observations. There is no land surface in the vicinity of the buoy; therefore, the buoy measurement is expected to be comparable to the HIRS field-of-view observation. HIRS records from 2002 are collocated to the site at 61.2°N, 1.1°E. The temporal evolution of buoy (solid gray line) and HIRS Ta records derived from NOAA-14 to -17 measurements are displayed in the top panel of Fig. 3. The panel shows that the HIRS Ta values follow the variation of buoy temperatures very closely. The HIRS data are able to capture the seasonal variability of air temperature as observed by the buoy data. The bottom panel of Fig. 3 displays the differences of time-matched
(within a half-hour) HIRS retrieval and buoy measurement. Almost all of the difference values are within 3°C, with a standard deviation of the difference values being 1.47°C. The instantaneous RMS error for the Ta retrieval is 1.50°C. Similarly, Fig. 4 displays the matched surface specific humidity between HIRS retrieval and buoy observation for the same buoy site. The upper panel shows collocated HIRS retrievals and the hourly buoy measurement. The lower panel shows the differences of time-matched (within a half-hour) HIRS retrievals and buoy observations of Qa. The HIRS-derived Qa values aligned with the surface observations closely. Throughout the year, the differences between the HIRS retrievals and buoy measurements are mostly within 2 g kg⁻¹. The standard deviation of the difference values is 0.77 g kg⁻¹, and the RMS error is 0.86 g kg⁻¹.

The HIRS retrievals were based on intersatellite-calibrated channel brightness temperature measurements. To examine the effect of the intersatellite calibration, we also plotted the collocated HIRS retrievals in Figs. 5–6 using non-intersatellite-calibrated HIRS measurements. Compared to Fig. 3, the upper panel of Fig. 5 shows more HIRS-retrieved Ta points that are lower than buoy observations, and the lower panel shows more scattered Ta differences. The standard deviation of differences increased from 1.47° to 1.60°C, and the RMS error increased from 1.50° to 1.64°C when nonintercalibrated HIRS brightness temperature inputs are used. Because the intersatellite biases of channel 8 are generally small and within 0.2 K, the differences between Figs. 5 and 3 mostly come from the differences between the intercalibrated and nonintercalibrated channel 7 brightness temperatures. The Qa retrievals in Fig. 6 using nonintercalibrated HIRS channel measurements are also significantly lower than those in Fig. 4. Figure 6 shows that the mean bias between HIRS retrievals and buoy observations increased from −0.38 to −1.07 g kg⁻¹ in Fig. 4 to −1.07 g kg⁻¹ in Fig. 6, the standard deviation of differences increased from 0.77 to 1.00 g kg⁻¹, and the RMS errors increased from 0.86 to 1.46 g kg⁻¹ when the nonintcalibrated inputs are used.

Figure 7 compares the HIRS-retrieved Ta and Qa values between using intercalibrated and nonintercalibrated HIRS channel brightness temperature inputs. The upper panel shows that the nonintercalibrated Ta retrievals from NOAA-15 and -17 are consistently lower than the
intercalibrated retrievals. Referring back to Fig. 1, channel 7 has large positive biases between NOAA-14 and -15 (i.e., NOAA-15 brightness temperatures are lower than NOAA-14 brightness temperatures) and between NOAA-16 and -17. The lower brightness temperature measurements from NOAA-15 and -17 resulted in smaller values of \( T_a \) retrievals. There are negative brightness temperature biases between NOAA-15 and -16. The decreased brightness temperatures in NOAA-15 (cf. NOAA-14) and increased brightness temperatures in NOAA-16 (cf. NOAA-15) mostly cancel out each other’s intersatellite biases; thus, the NOAA-16 \( T_a \) retrievals are not affected as much as the NOAA-15 and -17 \( T_a \) retrievals, though the nonintercalibration effect on NOAA-16 \( T_a \) retrievals is still notable. To illustrate the intersatellite biases in a more straightforward way, we calculated the averages of channel corrections in relation to the NOAA-12 base satellite for the HIRS measurements at the buoy site, and the results for channels 7, 8, 10, and 11 are listed in Table 1. The table shows that, on average, large corrections of greater than 1 K need to be applied to channel 7 on NOAA-15 and -17. This explains why the HIRS \( T_a \) and \( Q_a \) retrievals from NOAA-15 and -17 are consistently lower when inputs without the corrections are used. Channel 8 is a stable channel, and the intersatellite corrections for all satellites are small and within 0.2 K. Because of intersatellite biases, the \( \sim 0.5^\circ \)C decrease of retrievals for NOAA-16 in the warmer \( T_a \) range and the \( \sim 1^\circ \)C decrease of retrievals for NOAA-15 and -17 in the entire \( T_a \) range are nonnegligible discontinuities for long-term climate studies.

![Fig. 6](image1.jpg)

*Fig. 6.* As in Fig. 4, but showing the \( Q_a \) retrievals using nonintercalibrated HIRS input.

![Fig. 7](image2.jpg)

*Fig. 7.* Comparisons of HIRS retrievals between the ones derived from intercalibrated input and the ones without intercalibration. The (a) \( T_a \) and (b) \( Q_a \) comparisons.

<table>
<thead>
<tr>
<th>Channel 7</th>
<th>Channel 8</th>
<th>Channel 10</th>
<th>Channel 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOAA-14</td>
<td>0.47</td>
<td>0.10</td>
<td>0.18</td>
</tr>
<tr>
<td>NOAA-15</td>
<td>1.36</td>
<td>0.13</td>
<td>0.41</td>
</tr>
<tr>
<td>NOAA-16</td>
<td>-0.27</td>
<td>0.17</td>
<td>-0.81</td>
</tr>
<tr>
<td>NOAA-17</td>
<td>1.50</td>
<td>0.07</td>
<td>-0.35</td>
</tr>
</tbody>
</table>

**Table 1.** Averaged bias corrections (K) of channels 7, 8, 10, and 11 for NOAA-14 to -17 satellites in relation to the base satellite NOAA-12.
In the lower panel of Fig. 7, which displays the Qa comparison, NOAA-15 and -17 retrievals using non-intercalibrated HIRS inputs are much lower than the retrievals using intercalibrated HIRS inputs. In the warmer Qa retrieval range of around $10 \text{ g kg}^{-1}$, the biases resulted from using nonintercalibrated HIRS brightness temperatures can be $0.5 \text{ g kg}^{-1}$ for NOAA-14 and -16 and 1–2 g kg$^{-1}$ for NOAA-15 and -17. These intersatellite discontinuities are caused by the combined effect from multiple channels, primarily from channels 10, 11, and 7. The Qa observations in high-latitude oceans are generally in the range of 4–11 g kg$^{-1}$. The 0.5–2 g kg$^{-1}$ of intersatellite discontinuities can pose significant errors in long-term analyses.

The Ta and Qa are two of the components in the bulk aerodynamic formulas that are often used in calculating sensible and latent heat fluxes. For SHF, the formula takes the form $\text{SHF} = \rho_a C_h C_p U (\text{SST} - \text{Ta})$, and the LHF formula is in the form of $\text{LHF} = \rho_a C_e L_v U (\text{Qs} - \text{Qa})$, where $\rho_a$ is air density, $C_h$ and $C_e$ are the bulk exchange coefficients for heat and humidity, respectively, $C_p$ is the specific heat capacity of air, $U$ is the near-surface wind speed, SST is for sea surface temperature, $L_v$ is the latent heat of vaporization, and $\text{Qs}$ is the saturated specific humidity. A number of past studies (e.g., Curry et al. 2004; Jackson et al. 2006; Smith et al. 2011) showed that a significant portion of errors in heat flux computations using the bulk formulas is due to uncertainties in retrieved Qa and Ta. The impact of Ta and Qa retrieval biases on SHF and LHF resulting from intersatellite differences depends on other factors such as $U$ and SST. Because SST – Ta values are often less than $3^\circ \text{C}$ over sea surfaces, a bias of $1^\circ \text{C}$ in Ta can easily lead to 20%–30% error in the estimate of SHF. Similarly, a bias of 1 g kg$^{-1}$ in Qa can also lead to over 20% error in the estimate of LHF.

To further examine the retrievals over high-latitude oceans, we collocated (within a half-hour and in the HIRS field of view) the HIRS Ta and Qa retrievals with operational moored buoys in northern oceans and drifting buoys in both northern and southern oceans greater than $50^\circ \text{N}$ latitude in 2002. Comparison plots between HIRS retrievals and buoy observations are provided in Figs. 8 and 9 for Ta and Qa, respectively. Figure 8 shows that HIRS-retrieved Ta values are generally consistent with buoy measurements. The differences between the two datasets are mostly clustered within $3^\circ \text{C}$ of the unity correlation line. The instantaneous RMS errors of HIRS-derived Ta are 2.50$^\circ \text{C}$ compared to drifting buoys and 2.59$^\circ \text{C}$ compared to moored buoys. The biases between

![Fig. 8](image_url)
HIRS Ta retrievals and in situ measurements are 0.18°C compared to drifting buoys and −0.51°C compared to moored buoys. For the Qa comparison in Fig. 9, the scatterplots also show consistency between HIRS retrievals and buoy measurements. There are no Qa measurements from drifting buoys; therefore, only comparisons with moored buoys are available in Fig. 9. The RMS error between the HIRS Qa retrievals and moored buoy observations is 1.44 g kg\(^{-1}\) and the bias is −0.11 g kg\(^{-1}\). The RMS errors from comparison to all available high-latitude buoys shown in Figs. 8 and 9 are larger than those from comparison to a single buoy site shown in Figs. 3 and 4. In general, when compared to a larger amount of observations made by different types of instruments, the uncertainties become larger. The uncertainties can come from different instrumentations within moored (or within drifting) buoys, or from potentially larger differences between moored and drifting buoys. This is indicated by the different patterns of comparisons between retrieved Ta to drifting (lower left panel in Fig. 8) and moored (lower right panel in Fig. 8) buoys. Different designs in drifting and moored buoy instrumentations could explain, at least partially, the different patterns in the plots. Past studies have found that different instrumentations produced notably different measurements (e.g., Kent and Taylor 1995).

To examine the effect of intersatellite biases on the retrievals, we also calculated the RMS errors and biases for retrievals that are matched to the same drifting and moored buoys using nonintersatellite-calibrated HIRS brightness temperatures as input. The output resulted in increases of RMS errors for both Ta and Qa. The RMS errors for Ta using nonintercalibrated input increased to 2.60°C compared to drifting buoys and increased to 2.71°C compared to moored buoys. For Qa derived from nonintercalibrated input, the RMS error increased to 1.56 g kg\(^{-1}\). The biases also increased to −0.24°C compared to drifting buoys and −1.14°C compared to moored buoys for Ta, and they increased to −0.714 g kg\(^{-1}\) for Qa. These results further demonstrate the importance of performing intersatellite calibration.

ICOADS International Maritime Meteorological Archive (IMMA) version 2.4 data were used in the Jackson et al. (2009) and Jackson and Wick (2010) studies as a validation dataset of satellite-retrieved Qa. To examine the spatial differences between HIRS-retrieved Qa and in situ measurements of Qa, we collocated HIRS-retrieved Qa with a recent version of IMMA data, version 2.5. As described in Jackson et al. (2009), the quality control procedure includes removing derived Qa values below 0 and above 30 g kg\(^{-1}\), and removing observations exceeding 2.8\(\sigma\) from the climatology based on National Climatic Data Center trimming flags. HIRS retrievals are matched with IMMA data within 3 h and within 0.25° latitude–longitude. The differences between matched HIRS retrievals and IMMA data are then mapped to 2.5° latitude–longitude grids. Because in situ observations in the southern high latitudes are very sparse, only comparisons for northern high latitudes are plotted in Fig. 10. Figure 10 shows that the differences between HIRS-derived Qa and in situ data are mostly within ±0.5 g kg\(^{-1}\) over the open oceans, especially in the northern Pacific Ocean. However, scattered differences of larger than 1 g kg\(^{-1}\) are found in higher latitudes and near a few coastal areas, such as the Greenland coast. One potential source of the larger difference near coast lines could come from the uncertainty in the heights of ship measurements. The ships operating near coastal areas are likely very different from transocean vessels, and therefore

![Fig. 9](image-url)
have different platform heights. Furthermore, there are much more frequent samplings over the relatively lower latitudes of 50°–60°, and the samplings in latitudes greater than 65° become increasingly sparse toward the North Pole. In a study of sampling errors from voluntary observing ship observations, Gulev et al. (2007) showed that in poorly sampled regions random sampling errors can amount to 2–2.5 g kg\(^{-1}\) for specific humidity. A few values of larger than 1.5 g kg\(^{-1}\) on coastal lines in Fig. 10 are likely due to the fact that these HIRS pixels contain partial land surfaces, while their retrievals are compared to ship observations over water. Most of these coastal line matches occurred during daytime hours of warm seasons, and the land surface air temperatures were larger than water surface temperatures.

4. Summary and discussion

The SNO observations enabled intersatellite calibration of HIRS channel measurement in high latitudes. The SNO data show intersatellite biases of as large as 1.3 K for channel 7 and 0.6 K for channel 11. The intersatellite bias can be as large as 8 K for channel 10. Examinations based on NOAA-14 to -17 data show that if the nonintercalibrated HIRS channel brightness temperatures were used for Ta and Qa retrievals, then it could lead to Ta intersatellite biases as large as 1°C and Qa intersatellite biases of 1–2 g kg\(^{-1}\). If HIRS channel brightness temperature intersatellite biases were not corrected, then they could lead to significant discontinuities in long-term time series.

For the application of using a satellite infrared sensor to derive surface properties, it is important to remove all of the pixels that contain clouds. The cloud removal process can be complicated by the frequent temperature inversion layers in high latitudes. Clouds with tops in the inversions appearing warmer than the surrounding surface can be mislabeled as clear pixels. In this study RTTOV-9 simulations are used to analyze the relationship of the channel 8 and 7 difference in lower-atmospheric temperature inversions. The simulations show that when there are clouds in the inversion layer channel 7 brightness temperatures are lower than channel 8 brightness temperatures. This criterion is used to further remove clouds in an inversion layer that could be mistakenly labeled as clear in the first step of cloud removal process.

Based on intersatellite-calibrated HIRS measurements, Ta and Qa retrieval algorithms are developed based on a neural network approach, and the retrievals over high-latitude oceans are examined. It is shown that Ta and Qa variations are generally consistent with high-latitude buoy observations. Compared to a full year of buoy observations at a northern high-latitude site, the instantaneous RMS errors of HIRS retrievals are 1.50°C for Ta and 0.86 g kg\(^{-1}\) for Qa. Compared to a year of moored and drifting buoys in both northern and southern high latitudes, the retrieval instantaneous RMS errors are 2.5°–2.6°C for Ta and 1.4 g kg\(^{-1}\) for Qa. Compared to several years of IMMA in situ data, the HIRS-retrieved Qa show less than 0.5 g kg\(^{-1}\) of differences for the majority of northern high-latitude open oceans. Because only clear-sky data from an infrared sounder are used in this study, the retrievals cover only cloud-free areas on a certain day. On a long-term basis if this is the only dataset in the analysis, then it has the potential of producing clear-sky-related bias (e.g., dry bias). Estimation of the potential bias would require long-term, global measurement under all-sky conditions. This would be a valuable work, but is beyond the scope of the current study. Retrievals from microwave sensors can penetrate the majority of clouds and provide a much larger spatial coverage (but not all-sky conditions because of the removal of rainy pixels) over the sea surfaces. However, HIRS measurements can be used to derive temperature and humidity not only over the ocean but also over land surfaces, while microwave sounder retrievals of Ta and Qa are limited to water surfaces. Both infrared and microwave sensors have their advantages and limitations, and their applications should be used integrally along with in situ observations.

In addition to intersatellite calibration work being done over high latitudes, preliminary work has been carried out to extend HIRS intersatellite calibration to mid- and low latitudes (Cao et al. 2009; Lindfors et al. 2011). Further work will allow the extension of HIRS long-term
temperature and humidity retrievals to all latitudes. The current study generates high-latitude ocean Ta and Qa from HIRS measurements as two of the key components in deriving surface sensible and latent heat fluxes. Complementing other satellite and surface in situ measurements, the HIRS-derived Ta and Qa add to the growing database and knowledge of the surface flux research community.

Acknowledgments. We are grateful to Darren Jackson of Cooperative Institute for Research in Environmental Sciences, University of Colorado for providing HIRS cloud-clearing and limb correction codes and gridded in situ data. We also wish to thank Changyong Cao and Likun Wang of NOAA/STAR for providing the HIRS SNO datasets, and we thank Shawn Smith of The Florida State University for providing ship and buoy measurements. We also thank the anonymous reviewers for their useful comments.

REFERENCES


