| A Regime-Dependent Retrieval Algorithm for                                      |
|---|
| Near-Surface Air Temperature and Specific Humidity from Multi-Microwave Sensors |
|   |
| Lisan $Yu^*$ and Xiangze Jin  |
| Woods Hole Oceanographic Institution, Woods Hole, MA                            |
|   |
|   |
|   |
|   |
|   |
|   |
|   |
|   |
|   |
|   |
| Remote Sensing of Environment   |
| May 2018  |
|   |
|   |
|   |
|   |

# 25 Key words:

- 26 Near-surface air temperature and specific humidity, moored surface buoys, retrieval algorithm,
- 27 microwave passive sensors, multisensor retrieval.

#### Abstract

29

30 Near-surface specific humidity  $(q_a)$  and air temperature  $(T_a)$  over the global ocean are important 31 meteorological variables, but they cannot be retrieved directly from remote sensing. Many efforts 32 have been made to develop algorithms that derive  $q_a$  and  $T_a$  from multisensor microwave and/or 33 infrared observations using in situ measurements as training datasets. However, uncertainty 34 remains large in the resultant  $q_a$  and  $T_a$  retrievals. In this study, 147 moored surface buoys are 35 used to examine how buoy measured  $q_a$  and  $T_a$  are related to satellite microwave brightness 36 temperature (Tb) on the spatial scale from the warm/humid tropics to the cold/dry high latitudes. 37 It is found that the Tb –  $q_a$  and Tb –  $T_a$  relations are structured along two distinct, near-linear 38 bands, with the primary band in the warm/humid regime and a secondary (weaker) band in the 39 cold/dry regime. The step-like transition (or separation) between the two regimes occurs at 8 – 10 g kg<sup>-1</sup> for  $q_a$  and 14 – 17°C for  $T_a$ . The evidence suggests that one algorithm may not be 40 41 sufficient to extract  $q_a$  and  $T_a$  from Tb in all regimes. Therefore, a high-latitude enhancement is 42 added to the global algorithm so that the  $q_a$  and  $T_a$  retrievals in the dry/cold regime can be 43 specifically addressed. The new algorithms are applied to 11 microwave sensors, including 44 SSM/I, SSMIS, and AMSU-A, from 1988 to 2016. Based on the 475,717 buoy collocations during the 29-year period, the retrieved  $q_a$  and  $T_a$  have root-mean-square differences of 0.82 g 45 kg<sup>-1</sup> and 0.51°C, respectively. 46

- 47
- 48
- 49
- 50

#### 51 **1. Introduction**

52 Near-surface air temperature  $(T_a)$  and specific humidity  $(q_a)$  over the global ocean are 53 important meteorological variables. They set one of the driving conditions for turbulent 54 exchanges of heat and water vapor at the air-sea interface (Stephens 1990; Yu and Weller 2007; 55 Lorenz et al. 2010), and influence many ocean and atmospheric processes that are central to the 56 Earth's weather and climate (Bretherton et al. 2004; Sherwood et al. 2010). Water vapor is a 57 nature greenhouse gas. With the rise of atmospheric temperature in response to the addition of 58 anthropogenic gases, there are more evaporation of water vapor from the ocean (Yu 2007) and 59 more absorption of water waypor in the air (Held and Soden 2006) which further increase 60 atmospheric temperature. This positive water vapor feedback is regarded as a key driver of the 61 amplification of any warming caused by changes in atmospheric CO2 (Manabe and Wetherald, 62 1967; Hansen et al. 1984). Thus, there is a pressing need for well-calibrated, consistent, and 63 continuous long-term data records for  $T_a$  and  $q_a$  to keep tracking variability and long-term 64 change of the near-surface thermal conditions and to improve the quantification and modeling of 65 the water vapor feedback.

Measurements of  $T_a$  and  $q_a$  are conventionally obtained at sea by moored buoys and ships 66 67 of opportunity, but these observations are sparse in both space and time (e.g. Kent and Taylor 68 1996; Dai 2006; Willett et al. 2008; Berry and Kent 2009). Space-borne sensors that provide 69 continuous global coverage would be a desirable observing platform. However, retrieving  $T_a$  and 70  $q_a$  at a level of a few meters above the surface proves difficult for the space-borne technology, 71 because the measured radiation is emitted from relatively thick atmospheric layers rather than from single levels (Schulz et al. 1993). Making use of satellite measured total column of water 72 73 vapor (or total precipitable water (PW)) to estimate  $q_a$  and  $T_a$  at the near surface has been

74 actively investigated ever since the pioneer work by Liu and Niiler (1984), Liu (1986; 1988), and 75 Liu et al. (1991). The Liu studies showed that most of the spatial and temporal variability of 76 water vapor is confined in the lower part of the atmospheric column. The decoupling of the 77 atmospheric boundary layer from the higher atmosphere enables the vertical coherence of the humidity variability, leading to a high correlation between the surface-level  $q_a$  and the total PW 78 79 observed by radiosondes. Liu (1988) indicated that PW measured by scanning multichannel 80 microwave radiometer (SMMR) on Seasat is similar in form to the PW observed by radiosondes 81 (Liu and Niiler 1984; Liu 1986), both of which can be used to estimate  $q_a$  on monthly timescales with a root-mean-square (rms) error of about 0.4 g kg<sup>-1</sup> in the tropics and 0.8 g kg<sup>-1</sup> over the 82 global ocean. Hsu and Blanchard (1989) examined the PW- $q_a$  relation in the context of 13 field 83 84 experiments over the oceans, and suggested that the relation should also work for instantaneous 85 surface humidity retrievals. Subsequent algorithm developments have been particularly boosted 86 by the availability of a series of Special Sensor Microwave/Imager (SSM/I) onboard the Defense 87 Meteorological Satellite Program (DMSP) satellites since July 1987, and by the launch of the 88 Advanced Microwave Sounding Unit-A (AMSU-A) on the National Ocean and Atmospheric 89 Administration (NOAA) series of polar orbiting meteorological satellites in May 1998. The 90 SSM/I is a conically scanning microwave radiometer with seven frequency channels. The 91 AMSU-A is a multi-channel microwave radiometer that performs atmospheric sounding of 92 temperature and moisture levels by passively recording atmospheric microwave radiation in 93 multiple wavelengths. A survey of the algorithms that have been developed in the past 30 years 94 is provided in Table 1.

As shown in Table 1, the algorithm development can be loosely categorized into three
stages. The first stage is represented by the work of Schulz et al. (1993) and Chou et al. (1995),

who used the SSM/I measurements to first determine the water vapor in the lowermost 500 m of 97 98 the atmosphere and then predict  $q_a$ . This approach underscores the fact that for most oceanic 99 situations, the water vapor in the homogenous atmospheric boundary layer can be related to  $q_a$  by 100 a simple linear regression (Taylor 1982). The second stage is ushered in by Schlüssel et al. 101 (1995), who proposed a direct  $q_a$  retrieval from SSM/I Tb in 5 channels (19v, 22v, 37v, 19h, and 102 37h GHz, with h and v denote horizontal and vertical polarization, respectively). This approach 103 seems to be able to reduce the error propagation associated with the two-step procedure of Schulz et al. (1993), leading to an improved  $q_a$  retrieval. The approach was subsequently 104 105 extended to Tb measurements from other platforms including the Advanced Microwave 106 Scanning Radiometer for EOS (AMSR-E), the WindSat Polarimetric Radiometer, and the 107 Special Sensor Microwave Imager/Sounder (SSMIS) (Jones et al. 1999; Bentamy et al. 2003; 108 Jackson et al. 2006; 2009; Kubota and Hihara 2008; Roberts et al. 2010). The third stage features 109 the use of atmospheric soundings from instruments like the AMSU, the High-resolution Infrared 110 Radiation Sounder (HIRS), and the Special Sensor Microwave Water Vapor Profiler (SSM/T and 111 SSM/T-2). The sounders do not directly provide shallow surface measurements, but the detailed 112 profile information observed by the sounders can help to remove the variability in the total 113 column measurements that are not associated with the surface. Utilization of microwave and infrared sounders, with or without microwave radiometers, to estimate  $q_a$  and  $T_a$  has been 114 115 explored by Shi (2001), Shi et al. (2012), Jackson et al. (2006; 2009), and Jackson and Wick 116 (2010).

117 Ironically, despite the efforts and the progress that have been made,  $q_a$  and  $T_a$  remain the 118 leading error source for satellite-based air–sea heat flux datasets (Curry et al. 2004; Jackson et al. 119 2006; Prytherch et al. 2015). Jin et al. (2015) conducted a buoy-based evaluation of two satellite-

120 derived products: one is the  $q_a$  product by the Goddard Satellite-based Surface Turbulent Fluxes 121 (GSSTF) v3 (Chou et al. 1995; Shie et al. 2012) and the other is the  $q_a$  and  $T_a$  products from the 122 multi-instrument microwave regression (MIMR) by Jackson et al. (2006; 2009) and Jackson and 123 Wick (2010). Evaluation with the 137 air-sea buoys over the global scale showed that the biases in  $q_a$  and  $T_a$  are strongly regime dependent, featuring warm and wet biases in the tropical 124 125 warm/humid region and cold and dry biases in the extratropical cold/dry region. For instance, the mean difference between MIMR and buoy  $q_a$  changes from a wet bias of 0.8 g kg<sup>-1</sup> in the tropical 126 latitudes between  $10^{\circ}$ S –  $10^{\circ}$ N to a dry bias of –0.8 g kg<sup>-1</sup> at high latitudes poleward of 45°. Jin et 127 128 al. (2015) eventually obtained an improved analysis on 0.25° resolution through implementing a 129 buoy-based bias correction to MIMR and GSSTF3 followed by an objective synthesis with the 130 1°-gridded  $q_a$  analysis produced by the Objectively Analyzed air-sea Fluxes (OAFlux; Yu and 131 Weller 2007; Yu et al. 2008). However, the extensive use of buoys and OAFlux-1° for bias 132 correction at both regional and global scales is not a sustainable solution for improving the  $q_a$ 133 and  $T_a$  retrievals.

134 Esbensen et al. (1993) raised a cautious note about applying the Liu method to the global scale. When estimating  $q_a$  from SSM/I observations, they obtained systematic discrepancies of 135 over 2 g kg<sup>-1</sup> in the tropics. They postulated that in regions of persistent strong subsidence, most 136 137 of the water vapor is trapped near the surface and so the Liu's method can work reasonably well. 138 In regions of active convection, a significant amount of water vapor can exist aloft where it is 139 acted upon by processes that are less correlated with humidity at the near-surface level. Hence, 140 the Liu method in the tropics and the storm-tracked regions may not work in the same way as 141 that in the middle and high latitudes. If this is the case, the regime-dependent biases in MIMR 142 and GSSTF  $q_a$  (and  $T_a$ ) that are identified by Jin et al. (2015) may well be caused by error

143 compensations when fitting the regimes that have different vertical structures of water vapor. 144 The retrieval algorithms for  $q_a$  and  $T_a$  that have been developed so far are based primarily on two approaches: linear regression (Schlüssel et al 1995; Bentamy et al. 2003; Jackson et al. 2006; 145 146 2009; Kubota and Hihara 2008) and nonlinear artificial neural network (Jones et al. 1999; Shi 147 2001; Shi et al. 2012; Roberts et al. 2010). All these algorithms require in situ observations as 148 reference datasets to train the empirical relationships (or transfer functions) between the input 149 (Tb) and output  $(q_a, T_a)$ . Usually, the relationships are assumed globally valid and applicable to 150  $q_a$  and  $T_a$  in all regimes. If the findings of Esbensen et al. (1993) hold true, a globally applicable 151 algorithm would not be sufficient to represent the full range of the relations between Tb and  $(q_a)$ 152  $T_a$ ).

153 Moored air-sea buoys have been continually deployed in dynamically active regions 154 since the late 1980s, providing high-frequency research-quality time series measurements over 155 146 locations globally (Figure 1) (Send et al. 2010). They serve as anchor points for satellite 156 retrievals (section 2). Therefore, the objective of this study is threefold: (i) to examine the 157 relationships between satellite microwave Tb observations and buoy measurements of  $q_a$  and  $T_a$ 158 on the global scale (section 3), (ii) to explore an enhancement approach to address the challenges 159 induced by regime differences (section 4), and (iii) to explore a synergistic strategy for multisensory retrieval of  $q_a$  and  $T_a$  starting from 1988 when SSM/I became available (also 160 161 section 4). Evaluation with independent buoy measurements is conducted (section 5), and a 162 summary and conclusion is given (section 6).

163

164 **2. Data** 

165 2.1 Satellite sensors

This study uses Tb observations from microwave sensor series on the following satellite platforms: SSM/I sensors (F08, F10, F11, F13, F14, and F15) and SSMIS sensor (F17) on DMSP satellites, and AMSU-A sensors (N15, N16, N17, and N18) on NOAA meteorological satellites. The time line for each of the 11 sensors is shown in Figure 2, and general characteristics of the 11 sensors are summarized in Table 2.

171 SSM/I was first launched in July 1987 on the DMSP satellite F08. The instrument has 172 seven passive channels that sense vertically and horizontally polarized radiation at 19.3, 37.0, 173 and 85.5 GHz and vertically polarized radiation at 22.2 GHz. It has a 1394 km scanning swath 174 with a footprint ranging from 13 km to 69 km, depending on the channel and geographic location 175 (Hollinger et al., 1990). SSMIS is a series of passive microwave conically scanning imagers and 176 sounders onboard the DMSP satellites. It is designed to combine and extend the imaging and 177 sounding capabilities of three previously separate DMSP microwave sensors: the SSM/T-1 178 temperature sounder, the SSMI/T- 2 moisture sounder, and the SSM/I. The SSMIS instrument 179 measures microwave energy at 24 channels, with 21 discrete frequencies from 19 to 183 GHz 180 and a swath width of 1700 km. Although 4 SSMIS instruments have flown subsequently since 181 October 2003, only F17 Tb observations were used. SSMIS F16 has radiance anomalies due to 182 calibration issues (www.rmess.com), and F18 and F19 datasets were not available when the work 183 commenced.

The AMSU-A is a cross-track scanner with 15 channels (Goldberg, 1999), with 11 channels (4 through 14) located in the 50-60 GHz oxygen absorption band that is capable of sampling the temperature of the atmosphere in various layers. It has approximately 50 km spatial resolution at nadir and a swath width of 2343 km. The Tb observations from SSM/I and SSMIS are version 7 products processed by Wentz (2013) and those from AMSU-A by Zou and Wang(2013).

190

191 2.2 Moored surface buoys

192 Of all 147 moored surface buoys assembled (Figure 1), 68 buoys are from the Tropical 193 Atmosphere Ocean/TRIangle Trans-Ocean buoy Network (TAO/TRITON) in the tropical Pacific 194 Ocean (McPhaden et al., 1998), 21 are from the Pilot Research Moored Array in the Atlantic 195 (PIRATA) (Bourlès et al., 2008), and 28 are from the Research Moored Array for Africa-Asian-196 Australian Monsoon Analysis and Prediction (RAMA) in the tropical Indian Ocean (McPhaden 197 et al., 2009). The remaining 30 buoys that are not associated with the tropical moored arrays are 198 listed in Table 3. At each buoy location, measurements of air-sea variables (e.g. near-surface 199 temperature, humidity, wind, barotropic pressure, downwelling short- and/or long-wave, sea-200 surface temperature (SST), etc.) are provided at every 10- to 15-min intervals. Collocated satellite Tb retrievals and buoy measurements of  $q_a$  and  $T_a$  were compiled for 201 202 2010 to allow examination of the relationships between these variables. The year 2010 was 203 chosen because it was the time that the South Ocean Flux Station (SOFS) started to operate 204 (Schulz et al. 2012) and the buoy provides a valuable reference for the southern high latitudes. 205 There were 125 buoy time series available in 2010, with 117 from the three tropical moored 206 arrays. The 8 other buoys include the surface mooring at the Kuroshio Extension Observatory 207 (KEO) (Cronin et al., 2008), the Ocean Station PAPA (OSP) in the Gulf of Alaska (Kamphaus et 208 al. 2008), the SOFS (Schulz et al. 2012), the Woods Hole Oceanographic Institution (WHOI) 209 Hawaii Ocean Timeseries Site (WHOTS), the WHOI Northwest Tropical Atlantic Station 210 (NTAS), the WHOI Stratus Deck regions of the eastern tropical Pacific (STRATUS), and 2

buoys from the National Data Buoy Center (NDBC) at Stations 46075 (160.817°W, 53.983°N)
and 46085 (142.492°W, 55.868°N).

The development of the retrieval algorithms used 125 buoys in 2010. There are additional 214 22 buoys that were deployed either before or after 2010, and these buoys plus the 125 buoys 215 outside of 2010 were taken as independent validation data in the study.

216 The three tropical arrays carry the Automated Temperature Line Acquisition System 217 (ATLAS) buoys or TRITON buoys (McPhaden et al., 1998). The WHOI buoys are equipped 218 with the Improved METeorology (IMET) or Air Sea Interaction–METeorology (ASIMET) 219 systems (Weller and Anderson, 1996; Moyer and Weller, 1997). A comparison of the ATLAS, 220 TRITON, and IMET mooring meteorological sensors using a land-based cross-validation 221 approach showed that the three systems measure to equivalent standards of accuracy (Payne et 222 al., 2002). The relative humidity (RH) sensors are accurate to  $\pm 1\%$  RH and air temperature 223 sensors are better than 0.1°C (Colbo and Weller, 2009).

224

## **3.** Global relationships between satellite Tb and buoy measured $q_a$ and $T_a$

The Tb observations from four vertically polarized channels were used to relate buoy 226 227 measured  $q_a$  and  $T_a$ , with three frequencies (19, 22, and 37 GHz) from SSM/I and SSMIS, and 228 one (52 GHz) from AMSU-A. The selection is in line with the findings of Jackson et al. (2006) 229 that these channels are most sensitive to water vapor. Microwave spectroscopy has a weak water 230 vapor absorption line centered on the 22 GHz (Liu 1988; Wentz 1997), which enables the signal 231 from the water vapor to be better separated from the signal from the cloud (i.e., dew condensed 232 water). The 19 and 37 GHz bands are away from the spectral peak and are more absorbent to 233 water vapor. The AMSU 52 GHz channel is near the characteristic absorption peak of oxygen

and contains information about the physical temperature of the atmospheric molecules and its
 perspective profiles. A combination of the two spectra, water vapor and oxygen, can provide
 needed information of water vapor and temperature.

237 Buoy  $q_a$  and  $T_a$  are usually measured at a height of 2 - 3 m above the water (McPhaden 238 et al., 1998; Colbo and Weller, 2009). These  $q_a$  and  $T_a$  measurements were height adjusted to the 239 standard 2m using the Coupled Ocean-Atmosphere Response Experiment (COARE) version 3.5 algorithm (Fairall et al. 2003). Accordingly, the reference height for the  $q_a$  and  $T_a$  retrievals is set 240 241 at 2m in this study. This is different from most satellite retrieval algorithms that use ship-based 242 measurements as reference and obtain  $q_a$  and  $T_a$  at a height of 10 m (Bentamy et al. 2003; 243 Jackson et al. 2006; 2009; Kubota and Hihara 2008; Roberts et al. 2010; Shi et al. 2012). The 2 244 m  $q_a$  and  $T_a$  have the advantage of making a direct comparison with the atmospheric reanalysis, 245 as atmospheric models output the two variables at the 2-m height. 246 Collocated satellite Tb retrievals and buoy measurements of  $q_a$  and  $T_a$  were compiled for 247 the year of 2010. Three sensors were in operation in 2010 (Figure 2), SSMIS F17, AMSU-A 248 N15 and N18. To illustrate the spatial distribution of Tb variability, daily Tb fields from F17 19, 22, 37, and N15 52 GHz are shown on January 1<sup>st</sup>, 2010 (Figure 3). These daily fields were 249 250 constructed from ascending and descending passes associated with each channel. Our calculation 251 shows that SSMIS provides 84% of daily coverage over the global ocean and AMSU-A provides 252 97%. The spatial patterns of Tb from the three SSMIS channels all look similar, but the 253 magnitude of Tb is largest at 22 GHz because the weak water vapor absorption line near the 22 254 GHz is the best place for observing emission from the lower atmosphere. The AMSU measures

the atmospheric temperature profiles at channels near 60 GHz, which is an oxygen absorption

256 line. The ASMU-A Tb at 52 GHz is known for its sensitivity to temperature near the surface (e.g. 257 Jackson et al. 2006).

258 To collocate satellite Tb measurements with buoy  $q_a$  and  $T_a$  measurements, we used a 259 time and space window of 90 mins and 50 km. The space-window of 50 km is selected on the 260 basis of spatial resolution of the sensors. The spatial resolution is 50 km for AMSU, and 28 - 43261 km for SSM/I and SSMIS (Zou and Wang 2013; Wentz 2013). The collocation times were based 262 on the local overpass time of the sensor. The DMSP polar-orbiting satellite F17 has local 263 ascending overpass times at 6:25 pm. The NOAA polar-orbiting satellites fly Sun synchronous 264 orbits, with local ascending overpass times of ~2 pm for N18 and ~7 pm for N15 in 2010. F17 265 and N15 can be collocated within the time window of 90 mins, while the 2 pm N18 does not 266 collocate with either of them. To ensure the coherence in observed variability between sensors, 267 only F17 and N15 were taken as reference data. A total of 31,793 buoy-SSMIS collocations and 268 36,291 buoy-AMSU-A collocations were obtained. The triple collocations between buoy, SSMIS, 269 and AMSU-A totaled 17,975. Note that the search period for the satellite-buoy collocations were limited to the year of 2010 at all buoy sites except for SOFS. As the buoy at the latter site was 271 deployed in March 2010, the search period was extended to February 2011 to make it a full year. The scatter diagrams of the Tb– $q_a$  collocations (Figure 4) are used to examine how the 272 two variables correlate on the global scale. Strikingly, two bands of the Tb- $q_a$  relation are 273 274 displayed at all four frequencies, with a primary band in the humid regime (i.e., buoy  $q_a$  higher than 10 g kg<sup>-1</sup>) and a secondary (weaker) band in the dry regime (i.e. buoy  $q_a$  lower than 8 g kg<sup>-1</sup>). For each regime, the Tb– $q_a$  relation is near linear at 22 and 52 GHz channels, showing that Tb is 277 most sensitive to the near-surface humidity at the spectral peaks of water vapor and oxygen. At 278 19 and 37 GHz channels, the relation is near linear in the dry regime but highly nonlinear in the

270

275

highly humid regime. The increase of Tb with  $q_a$  hits a plateau when  $q_a$  reaches ~ 20 g kg<sup>-1</sup>, and  $q_a$  remains little changed with further increase in Tb.

281 The step-like regime shift in the Tb– $q_a$  relation is equally featured in the scatter diagrams of the Tb– $T_a$  collocations (Figure 5), with one band in the warm regime (i.e. buoy  $T_a$  higher than 282 283 17°C) and the other in the cold regime (i.e. buoy  $T_a$  lower than 14°C). The transition between the 284 warm and cold regimes occurs for  $T_a$  in the range of 14–17°C, which is more distinct and step-285 like than the transition between the humid and dry regimes (Figure 4). In the cold regime, the 286 relationship between the two variables is generally linear though the scatter is wider. On the other hand, the outstanding features that are identified in the Tb- $q_a$  scatter diagrams are also 287 observed. For instance, the 22 and 52 GHz channels produce a more linear Tb– $T_a$  relation than 288 289 the 19 and 37 GHz channels. The Tb– $T_a$  relation has a plateau when  $T_a$  exceeds 26°C, where 290 further increase in Tb is not accompanied by an increase in  $T_a$ .

These patterns of the Tb– $q_a$  and Tb– $T_a$  relations (Figures 4-5) are in good agreement with the patterns derived from ship reports that were constructed by Jackson et al. (2006). However, the step-like shift in the Tb– $q_a$  and Tb– $T_a$  relations between the tropical warm/humid and the higher-latitude cold/dry zones are more vividly captured by buoy measurements in this study. The regime-dependent shift in the relation indicates that a global model is not sufficient to represent the characteristics of the relation in each regime.

297

### **4. A global retrieval algorithm with high-latitude enhancement**

299 4.1 The retrieval strategy

In the literature (see Table 1), two general methods are used to develop the Tb-based
retrieval algorithms: a multivariate linear regression (e.g., Schlüssel et al. 1995; Bentamy et al.

| 302 | 2003; Kubota and Hihara 2008; Jackson et al. 2006; 2009) and ANNs (e.g., Jones et al. 1999; Shi      |
|-----|--|
| 303 | 2001; Roberts et al. 2010). ANNs are nonlinear models (or transfer functions) that are trained to    |
| 304 | extract hidden structures and relationships in data, have gained broad applications to retrieving    |
| 305 | various geophysical parameters (Thiria et al. 1993; Krasnopolsky et al. 1995). The types of          |
| 306 | model most often used by ANNs are the logistic sigmoid function and the hyperbolic tangent           |
| 307 | function (a tanh function), both of which feature a characteristic S-shaped curve.                   |
| 308 | The Tb– $q_a$ and Tb– $T_a$ relations (Figures 4-5) are dependent of the thermodynamical             |
| 309 | regimes of the near-surface properties (Esbensen et al. 1993). The relations in the warm/humid       |
| 310 | regime not only depart from the relations in the cold/dry regime but also differ in linearity and    |
| 311 | sensitivity. In this case, an S-shaped mathematical function is not able to capture the complexity   |
| 312 | of the relationships. Special treatment needs to be given for the cold/dry regime at higher          |
| 313 | latitudes.   |
| 314 |  |
| 315 | 4.2 A high-latitude enhanced global retrieval algorithm  |
| 316 | To better address the regime-shift in the Tb– $q_a$ and Tb– $T_a$ relations, we propose a three-     |
| 317 | step retrieval approach. The first step is to construct a global model, the second step to construct |
| 318 | a model for the high-latitude cold/dry regime, and the third step to obtain a high-latitude          |
| 319 | enhanced global retrieval algorithm through merging the two models. The retrieval model is           |
| 320 | based on the second-order polynomial regression and the formulation can be expressed as              |

321 follows:

322 
$$q_a = \sum_{m}^{4} \left( a_{1,0} + a_{1,m} T b_m + b_{1,m} (T b_m)^2 \right)$$
(1)

323 
$$T_a = \sum_{m}^{4} \left( a_{2,0} + a_{2,m} T b_m + b_{2,m} (T b_m)^2 \right)$$
(2)

where m=1,..4 denotes the four channels (19v, 22v, 37v, and 52v GHz), and *a* and *b* are the polynomial coefficients for regressing Tb from the four channels to  $q_a$  and  $T_a$ , respectively. All coefficients were determined by the 125 buoy time series in 2010, except for the SOFS buoy where the period of March 2010 to February 2011 was used (Table 3).

The three-step procedure is executed as follows. Equations (1)-(2) are used for both the global model (hereafter GM; Step 1) and the high-latitude model (hereafter HLM; Step 2), and the only difference between the two models is the training of the polynomial coefficients. For the GM, the 125 buoy time series in 2010 were all used, while for the HLM, only four buoy time series were taken. The four buoys are SOFS, OSP, NDBC 46075 and 46085 (Table 3) that are located poleward of 40° latitude. The coefficients *a* and *b* that were trained for the two sets of models are listed in Table 4.

Once the two models are obtained, the next step (Step 3) is to produce two sets of  $q_a$  and  $T_a$  daily mean fields that correspond to the respective GM and HLM. The two sets of daily fields are then merged over the transition zones depicted by the scatter diagrams (Figures 4-5). The transition zone occurs around 8-10 g kg<sup>-1</sup> for  $q_a$ , and 14-17°C for  $T_a$ . Hence, the three-step approach can be designed as follows.

340 
$$x = \begin{cases} x(GM), & \text{if } x \ge xl \\ x(HLM), & \text{if } x \le xu \\ x(GM) \times \alpha + x(HLM) \times (1-\alpha), & \text{if } xl \le x \le xu \end{cases}$$
(3)

341 where *x* denotes  $q_a$  or  $T_a$ . We define  $\alpha = (\bar{x} - xl)/(xu - xl)$ , where  $\bar{x}$  is the average of *x*(GM) and 342 *x*(HLM), and *xu*(*xl*) is the upper (lower) bound of the transition zone. A flowchart of the three-343 step approach is shown (Figure 6).

| 344 | Production of the daily-mean $q_a$ and $T_a$ fields on 01 January 2010 is illustrated (Figures                      |
|-----|---|
| 345 | 7a-f) to provide some insights on how the three-step approach is used. On each day, two sets of                     |
| 346 | $q_a$ and $T_a$ fields are produced on each day in correspondence to the GM and HLM, respectively,                  |
| 347 | but neither field is used on the global scale. The GM-based $q_a$ and $T_a$ are cut off at the lower                |
| 348 | bound of the transition zone, which is set at 8 g kg <sup>-1</sup> for $q_a$ and 14°C for $T_a$ (Figures 7a,d). The |
| 349 | HLM-based estimates are cut off at the upper bound of the transition zone, which is set at 10 g                     |
| 350 | kg <sup>-1</sup> for $q_a$ and 17°C for $T_a$ (Figures 7b,e). Using Eq.(3), the two sets of estimates are combined  |
| 351 | over the transition zone, leading to the merged global daily-mean $q_a$ and $T_a$ fields (Figures 7c,f).            |
| 352 |   |

# 4.3 Assessment of the $q_a$ and $T_a$ estimates in the extratropical oceans

Our use of the three-step approach aims to improve the estimates of  $q_a$  and  $T_a$  in the less 354 355 humid and less warm extratropical region, where one Tb retrieval can fit into both GM and HLM 356 (Figures 4-5). To assess the improvement made by the three-step approach, we evaluated the  $q_a$ and  $T_a$  retrievals produced at two buoy sites, OSP (145°W, 50°N) and KEO (144.6°E, 32.4°N) 357 (Figures 8–11). At each site, buoy measurements were a reference for three sets of  $q_a$  and  $T_a$  time 358 359 series: those derived from the GM, the HLM, and the merge of the two. Station OSP is located in the higher latitude cold/dry regime, where buoy measured  $q_a$  was mostly below 10 g kg<sup>-1</sup>, which 360 361 is the upper bound of the transition zone (Figure 8). It can be seen that the GM does not work 362 well at this site, where the  $q_a(GM)$  time series is weakly underestimated in the drier regime ( $q_a <$ 8 g kg<sup>-1</sup>) between Days 1 and 120 and largely overestimated in the more humid regime ( $q_a > 8$  g 363 kg<sup>-1</sup>) between Days 150 and 250. By comparison, the HLM works better. The  $q_a$ (HLM) time 364 365 series (Figure 8b) has an overall good performance against buoy time series except for the period (Days 150–250) when  $q_a$  gets into the transition zone (i.e.  $q_a > 8 \text{ g kg}^{-1}$ ). Following the three-step 366

approach outlined in Eq.(3), we used  $q_a$ (HLM), not  $q_a$ (GM), for  $q_a < 8$  g kg<sup>-1</sup>, and combined the two sets of estimates for  $q_a \ge 8$  g kg<sup>-1</sup>. The resultant  $q_a$  time series (Figure 8c) represents an improvement over both  $q_a$ (HLM) and  $q_a$ (GM).

370 The  $T_a$  time series at Station OSP was produced in a similar manner:  $T_a$ (GM) (Figure 9a) is not used except for  $T_a > 14^{\circ}$ C, where it is combined with  $T_a$ (HLM) (Figure 9b) to produce an 371 improved  $T_a$  in the transition zone (Figure 9c). Clearly, the GM and HLM perform well in each 372 designated regime, but neither of them are desirable in the transition zone. When the two models 373 are combined, they do lead to improved  $q_a$  and  $T_a$  estimates over the full range of values. 374 The climate at the KEO buoy site is different, as the buoy is located in a western 375 376 boundary current region. In 2010, buoy measurements were available from Days 100 to 280, during which  $q_a$  was predominantly greater than 8 g kg<sup>-1</sup> (Figure 10a). In this humid regime, 377  $q_a$ (HLM) is poorly produced while  $q_a$ (GM) performs well (Figures 10a–b), and the final  $q_a$  is 378 estimated primarily by the GM according to Eq.(3). It is seen that HLM is nonresponsive for  $q_a$ 379 greater than 10 g kg<sup>-1</sup> (Figure 10b), because the HLM is trained by four buoys located in the 380 381 cold/dry regime. Hence, the HLM works only in the dry/cold regime, and has no skill in the 382 warm/humid regime.

383

### 4.4 Treatments for SSM/I-based retrievals for the pre-AMSU period

AMSU-A was available only after October 1998 (Table 2), and so the retrieval of  $q_a$  and  $T_a$  for the pre-AMSU period needs to be obtained from the SSM/I sensors (Figure 2). The second-order polynomial fitting algorithms in Eqs. (1)-(2) are then revised as follows:

388 
$$q_a = \sum_{m}^{3} \left( a_{1,0} + a_{1,m} T b_m + b_{1,m} (T b_m)^2 \right)$$
(4)

389 
$$T_a = \sum_{m}^{3} \left( a_{2,0} + a_{2,m}Tb_m + b_{2,m}(Tb_m)^2 \right) + a_{2,SST}SST + b_{2,SST}(SST)^2 \right)$$
(5)

390 where m=1,2,3 denotes the three SSM/I channels (19v, 22v, and 37v). Compared to Eqs. (1)-(2), 391 the main difference is the inclusion of SST in the  $T_a$  retrieval model (Eq. (5)). The use of SST 392 serves two purposes. One is to increase the daily coverage of SSM/I sensors and the other is to 393 improve the accuracy of the SSM/I-based  $T_a$  retrievals (e.g. Jackson and Wick 2010). The 394 algorithm was trained by the same set of buoy time series in 2010 but excluding AMSU-A Tb in 395 the algorithm (Figure 6). The new set of coefficients is provided in Table 5. 396 The retrievals derived from SSM/I only sensors might differ from those derived from the 397 combined use of SSM/I and AMSU-A. To address this issue, two sets of retrievals were 398 examined, with one produced using the SSM/I+AMSU-A and the other using SSM/I only 399 algorithms. The mean differences between the two sets of retrievals for the three years 2000 -400 2002 were shown (Figure 11). For  $q_a$  (Figures 11a,b,c), the differences between the two mean 401 fields are distributed in zonal bands with alternating signs. Compared to  $q_a(SSMI)$ , 402  $q_a$ (SSMI+AMSU-A) tends to be lower (i.e. drier) in the mid-latitudes (25–45° North and South) 403 and higher (i.e. wetter) in the tropical and higher latitudes. In addition, in the coastal regions off 404 California, Peru, and Benguela, the differences between  $q_a$ (SSMI+AMSU-A) and  $q_a$ (SSMI) are particularly pronounced, exceeding 0.5 g kg<sup>-1</sup>. For  $T_a$  (Figures 11e,f,g), the pattern of the 405 406 differences between  $T_a(SSMI+AMSU-A)$  and  $T_a(SSMI)$  shows similar spatial structure. 407 For both  $q_a$  and  $T_a$ , the spatial pattern of the differences has only a slight year-to-year 408 variation. The time series comparison (Figure 12) shows that the gap between the SSM/I only 409 and the SSM/I+AMSU-A retrievals is essentially near-constant for the overlapping 2000-2002 410 period. This suggests that the mean difference pattern averaged over the three-year period 411 (Figures 11d,h) could be used as a reference to level the mean difference of the pre-AMSU

412 retrievals. The adjustment enables the merge of the pre-AMSU period to produce an 11-sensor 413 based time series for  $q_a$  and  $T_a$  retrievals from 1988 to 2016. The 29-year averaged global distribution of the retrieved  $q_a$  and  $T_a$  in January, July, and annual mean (Figure 13) shows the 414 415 seasonal contrast of the warm/humid regime in the tropics and the cold/dry regime at higher 416 latitudes.

417

418

#### 5. Error statistics based on 147 buoys

419 The buoy time series measurements served two tasks in developing the three-step 420 retrieval approach: training the empirical relationships prescribed in the retrieval algorithms 421 (Eqs.(1)-(5)), and validating the  $q_a$  and  $T_a$  retrievals. The first task used 125 buoy time series 422 collected in 2010, and the second task used 147 independent buoys outside 2010 (Figure 1). To 423 provide a complete assessment of the  $q_a$  and  $T_a$  daily global fields retrieved from 11 sensors over 424 the 29-year period, the 147 buoy time series in all available periods are used to evaluate the 425 statistical properties of the two retrieved products. Summary plots of the buoy-based statistical 426 properties of the retrieved  $q_a$  and  $T_a$  daily products were constructed (Figures 14-15).

427 The mean retrieved-minus-buoy differences, which are a measure of bias in products, 428 were constructed at the 147 buoy locations by averaging over the available measurement periods at each location (Figures 14a-b). The bias in  $q_a$  is generally small, within ±0.5 g kg<sup>-1</sup> at most sites, 429 430 including the eastern equatorial Pacific, the south Indian Ocean, the North Atlantic, and the 431 higher Northern and Southern latitudes. Large mean differences, with magnitude exceeding  $\pm 0.5$ g kg<sup>-1</sup>, occur in regions of the equatorial central Pacific and Atlantic Oceans, and also the Bay of 432 433 Bengal. These large differences are associated with a highly humid regime where  $q_a$  is usually greater than 18 g kg<sup>-1</sup> (Figure 13a). By comparison, the bias in  $T_a$  is small, within ±0.3°C, at 434

almost all locations. The standard deviations (STD) of daily differences between retrieved and measured values are generally in the range of 0.5–1.0 g kg<sup>-1</sup> for  $q_a$  and 0.3–0.5°C for  $T_a$  (Figures 14c-d). The correlation coefficients between retrieved and measured daily  $q_a$  time series are high, greater than 0.8 at all locations. The correlation is even higher for  $T_a$ , exceeding 0.9 at most sites (Figures 14e-f).

The overall statistical properties of the retrieved  $q_a$  and  $T_a$  daily values at all buoy sites are summarized in two scatter plots that were constructed from a total of 475,717 collocations (Figures 15a-b). The mean differences (biases) of the retrieved  $q_a$  and  $T_a$  are -0.29 g kg<sup>-1</sup> and -0.13°C, respectively. The STD differences of the retrieved  $q_a$  and  $T_a$  are 0.77 g kg<sup>-1</sup> and 0.50°C, respectively. The root-mean-square (rms) differences are 0.82 g kg<sup>-1</sup> for the retrieved  $q_a$  and 0.51°C for the retrieved  $T_a$ . The respective correlation coefficients with the buoy counterparts are high, 0.94 for  $q_a$  and 0.98 for  $T_a$ .

447

#### 448 **6. Summary and conclusions**

This study uses buoy measurements to show that the  $Tb-q_a$  and  $Tb-T_a$  relations differ 449 450 between the warm/humid tropics and the cold/dry high latitudes and to develop new retrieval algorithms to address the regime dependence. It is found that the  $Tb-q_a$  and  $Tb-T_a$  relations are 451 452 structured along two distinct, near-linear bands, with a primary band in the warm/humid regime 453 and a secondary (weaker) band in the cold/dry regime. The step-like transition (or separation) between the two regimes occurs at 8–10 g kg<sup>-1</sup> for  $q_a$  and 14–17°C for  $T_a$ . The evidence suggests 454 that one algorithm may not be sufficient to extract  $q_a$  and  $T_a$  from Tb in all regimes. Therefore, 455 456 we propose to add a high-latitude enhancement to the global algorithm so that the  $q_a$  and  $T_a$ 457 retrievals in the dry/cold regime can be specifically addressed. This leads to the development of a 458 three-step retrieval approach. The first step is to construct a global model, the second step to 459 construct a model for the high-latitude cold/dry regime, and the third step to merge the two 460 models to create a high-latitude enhanced global retrieval algorithm. Both the global and the 461 high-latitude retrieval models are based on the second-order polynomial regression, and the 462 difference between the two models is how the regression coefficients are trained. Buoy 463 measurements in 2010 are used as the training datasets. The global model is trained with 125 464 buoy time series that are available in 2010, and the high-latitude model is trained with 4 buoys 465 that are located poleward of 45° latitude north and south. The merge of the two models is 466 conducted in the respective transition zone for  $q_a$  and  $T_a$ .

The new algorithms are applied to 11 microwave sensors from SSM/I, SSMIS, and AMSU-A between 1988 and 2016. Since AMSU-A was not available before 1998, the  $q_a$  and  $T_a$ retrievals based on SSM/I only are slightly higher than those retrievals derived from SSM/I+AMSU. Our experiments show that the difference pattern has little change from year to year, and hence an ensemble mean difference pattern is applied to level down the  $q_a$  and  $T_a$ retrievals before 1998.

The buoy time series measurements served two tasks in developing the three-step retrieval approach: training the empirical relationships prescribed in the retrieval algorithms and validating the  $q_a$  and  $T_a$  retrievals. The first task used 125 buoy time series collected in 2010, and the second task used 22 independent buoys that were deployed either before or after 2010. To provide a fuller assessment of the  $q_a$  and  $T_a$  daily global fields retrieved from 11 sensors over the 29-year period, the 147 buoy time series in all available periods are used to evaluate the statistical properties of the two retrieved products. On the basis of the 475,717 buoy collocations 480 during the 29-year period, the retrieved  $q_a$  and  $T_a$  have root-mean-square differences of 0.82 g 481 kg<sup>-1</sup> and 0.51°C, respectively and a mean bias of -0.29 g kg<sup>-1</sup> and -0.13°C, respectively.

482 As expected, the global algorithm with the high-latitude enhancement shows improved  $q_a$ 483 and  $T_a$  retrievals at high latitudes. Nevertheless, the Tb- $q_a$  and Tb- $T_a$  relations are more complex than we could cover in this study. The Tb- $q_a$  and Tb- $T_a$  plateaus at very warm ( $T_a > 26^{\circ}$ C) and 484 very humid ( $q_a > 20 \text{ g kg}^{-1}$ ) conditions need to be further investigated and better represented in 485 486 the retrieval algorithm. However, these conditions are generally associated with tropical 487 convective activities. As pointed out by Esbensen et al. (1993), the Liu method (1986) in the 488 tropics and the storm-tracked regions may not work in the same way as in the middle and high 489 latitudes because in regions of active convection, significant amounts of water vapor can exist 490 aloft where it is acted upon by processes that are less correlated with the humidity at the near-491 surface level. Whether there are ways to improve the  $q_a$  and  $T_a$  retrievals in the convective 492 regime are yet to be fully determined.

493

#### 494 Acknowledgements

495 The project is sponsored by the NOAA Ocean Observation and Monitoring Division (OOMD) 496 under grant NA14OAR4320158. We thank the following data sources for providing the data 497 used in this study. SSMI and SSMIS were downloaded from Remote Sensing Systems 498 (http://www.ssmi.com/) The AMSU-A was acquired from NOAA's National Climatic Data 499 Center (http://www.ncdc.noaa.gov). Buoy measurements from TAO, PIRATA, and RAMA 500 programs were downloaded from http://pmel.noaa.gov, and those from ARC, KEO and OSP 501 were from http://www.pmel.noaa.gov/ocs. The WHOI buoys (Arabian Sea, ASREX91, 502 ASREX93, Bay of Bengal, CLIMODE, CMO, COARE, MLML91, NTAS, PACS, SESMOOR,

- 503 SMILE, Stratus, Subduction, WHOTS) were from http://uop.whoi.edu. The SOFS mooring is
- 504 part of the Australian Integrated Marine Observing System (IMOS) with data available from
- 505 <u>https://portal.aodn.org.au</u>. The OOI SO and Argentine buoys were from
- 506 <u>https://ooinet.oceanobservatories.org</u>. We thank Bob Weller and Nan Galbraith for providing the
- 507 OOI Irminger Sea buoy measurements.

#### 508 **References**

- 509 Berry, D. I., and E. C. Kent (2009), A new air-sea interaction dataset from ICOADS with
- 510 uncertainty estimates. *Bull. Am. Meteorol. Soc.*, **90**(5), 645–656.
- 511 Bentamy, A., K. B. Katsaros, A. M. Mestas-Nunoz, W. M. Drennan, E. B. Forde, and H. Roquet
- 512 (2003), Satellite estimates of wind speed and latent heat flux over the global oceans. J. Clim.,

**16**, 637–656.

- 514 Bourlès, B., R. Lumpkin, M.J. McPhaden, F. Hernandez, P. Nobre, E. Campos, L. Yu, S. Planton,
- 515 A. Busalacchi, A.D. Moura, J. Servain, and J. Trotte (2008), The PIRATA Program: History,
- accomplishments, and future directions. *Bull. Am. Meteorol. Soc.*, **89**, 1111-1125.
- 517 Bretherton, C. S., M. E. Peters, and L. E. Back (2004), Relationships between water vapor path
- and precipitation over the tropical oceans, J. Clim., 17, 1517–1528.
- Chou S.-H., R. M. Atlas, C-L. Shie, and J. Ardizzone (1995), Estimates of surface humidity and
  latent heat fluxes over oceans from SSM/I Data, *Mon. Weather Rev.*, 123, 2405-2425.
- 521 Colbo, K. and R. A. Weller (2009), The accuracy of the IMET sensor package in the subtropics.
- 522 J. Atmos. Ocean. Technol., **26**(9), 1867-1890.
- 523 Cronin, M. F., C. Meinig, C. L. Sabine, H. Ichikawa, and H. Tomita (2008), Surface mooring
- network in the Kuroshio Extension. *IEEE Systems Special Issue on GEOSS*, **2**(3), 424-430.
- 525 Curry, J. A., A. Bentamy, M.A. Bourassa, D. Bourras, E.F. Bradley, M. Brunke, S. Castro, S.H.
- 526 Chou, C.A. Clayson, W.J. Emery, L. Eymard, C.W. Fairall, M. Kubota, B. Lin, W. Perrie, R.A.
- 527 Reeder, I.A. Renfrew, W.B. Rossow, J. Schulz, S.R. Smith, P.J. Webster, G.A. Wick, and X.
- 528 Zeng (2004), SEAFLUX. Bull. Am. Meteor. Soc., 85, 409-424.
- 529 Dai, A. (2006), Recent climatology, variability, and trends in global surface humidity. J.
- 530 *Clim.*, **19**, 3589–3606.

- 531 Esbensen, S. K., D. B. Chelton, D. Vockers, and J. Sun (1993), An analysis of errors in Special
- Sensor Microwave Imager evaporation estimates over the global oceans. *J. Geophys. Res.*, 98,
  7081-7101.
- 534 Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson (2003), Bulk
- 535 parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. J.

536 *Clim.*,**16**, 571–591.

- 537 Goldberg, M. D.(1999), Generation of retrieval products from AMSU-A: Methodology and
- validation. *10th International TOVS Study Conference*, Boulder, Colorado, 27 January 2
  February.
- 540 Hansen, J., A. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, R. Ruedy, and J. Lemer (1984),
- 541 *Climate sensitivity: Analysis of feedback mechanisms.* In: Climate Processes and Climate
- Sensitivity, J. E. Hansen and T. Takahashi, eds., Washington, D.C., American Geophysical
  Union. pp. 130-163.
- 544 Held, I.M. and B.J. Soden, 2006: Robust Responses of the Hydrological Cycle to Global
- 545 <u>Warming.</u> J. Climate, 19, 5686–5699, <u>https://doi.org/10.1175/JCLI3990.1</u>
- 546 Hollinger, J.P., J. L. Peirce, and G. A. Poe (1990), SSM/I Instrument Evaluation. *IEEE Trans.*
- 547 *Geosci. Remote Sens.*, **28**(5), 781–790.
- 548 Hsu, S. A., and B. W. Blanchard (1989), The relationship between total precipitable water and
- 549 surface-level humidity over the sea surface: A further evaluation. J. *Geophys. Res.*, **94**, 14,539-
- 550 14,545.
- Jackson, D.L., G.A. Wick, and J.J. Bates (2006), Near-surface retrieval of air temperature and
- 552 specific humidity using multi-sensor microwave satellite observations. J. Geophys. Res. 111,
- 553 D10306, doi:10.1029/2005JD006431.

- Jackson, D. L., G. A. Wick, and F. R. Robertson (2009), Improved multisensor approach to
- satellite-retrieved nearsurface specific humidity observations. J. Geophys. Res., 114, 1-13,
- 556 doi:10.1029/2008JD011341.
- 557 Jackson, D.L. and G. A. Wick (2010), Near-surface air temperature retrieval derived from
- 558 AMSU-A and sea surface temperature observations. J. Atmos. Ocean. Technol., 27, 1769-1776.
- Jin, X., L. Yu, D. L. Jackson, and G. A. Wick, (2015), An improved near-surface specific
- 560 humidity and air temperature climatology for the SSM/I satellite period. J. Atmos. Oceanic
- 561 *Technol.*, **32**, 412–433, doi: http://dx.doi.org/10.1175/JTECH-D-14-00080.1.
- 562 Jones, C., P. Peterson, and C. Gautier (1999), A new method for deriving ocean surface specific
- humidity and air temperature: An artificial neural network approach. J. Appl. Meteorol., 38,
  1229–1245.
- 565 Kamphaus, R., M. Cronin, C. Sabine, S. Emerson, C. Meinig, and M. Robert (2008), New
- surface mooring at Station PAPA monitors climate. *PICES Press*, **16**(2), 26-27.
- 567 Kent, E.; Taylor, P.K. (1996), Accuracy of humidity measurements on ships: Consideration of
- 568 solar radiation effects. J. Atmos. Ocean. Technol., 13 (6). 1317-1321.10.1175/1520-
- 569 0426(1996)013<1317:AOHMOS>2.0.CO;2
- 570 Krasnopolsky, V. M., L. C. Breaker, and W. H. Gemmill (1995), A neural network as a nonlinear
- transfer function for retrieving surface wind speeds from the special sensor microwave imager.
- 572 J. Geophys. Res., 100, 11 033–11 045.
- 573 Kubota, M., and T. Hihara (2008), Retrieval of surface air specific humidity over the ocean using
- 574 AMSR-E measurements. *Sensors*, **8**, 8016–8026, doi:10.3390/s8128016.
- 575 Liu, W. T. (1986), Statistical relation between monthly precipitable water and surface-level
- humidity over global oceans. *Mon. Weather Rev.*, **114**, 1591–1602.

- 577 Liu, W. T. (1988), Moisture and latent heat flux variabilities in the tropical Pacific derived from
  578 satellite data. *J. Geophys. Res.*, 93, 6749–6760.
- Liu, W.T., and P.P. Niiler (1984), Determination of monthly mean humidity in the atmospheric
  surface layer over oceans from satellite data. *J. Phys. Oceanogr.*, 14, 1451-1457.
- 581 Liu, W.T., W. Tang, and P.P. Niiler (1991), Humidity profiles over oceans. J. Clim., 4, 1023582 1034.
- 583 Lorenz, D. J., E. T. DeWeaver, and D. J. Vimont (2010), Evaporation change and global
- warming: The role of net radiation and relative humidity, J. Geophys. Res., 115, D20118,
- 585 doi:10.1029/2010JD013949.
- Manabe, S., and R. T. Wetherald (1967), Thermal equilibrium of atmosphere with a given
  distribution of relative humidity, *J. Atmos. Sci.*, 24, 241–259.
- 588 McPhaden, M. J., and coauthors (1998), The Tropical Ocean-Global Atmosphere observing
- 589 system: A decade of progress, *J. Geophys. Res.*, **103**(C7), 14,169–14,240,
- 590 doi:10.1029/97JC02906.
- 591 McPhaden, M. J., G. Meyers, K. Ando, Y. Masumoto, V. S. N. Murty, M. Ravichandran, F.
- 592 Syamsudin, J. Vialard, L. Yu, and W. Yu (2009), RAMA: The Research Moored Array for
- African-Asian-Australian Monsoon Analysis and Prediction. *Bull. Am. Meteor. Soc.*, **90**, 459480.
- 595 Moyer, K. A. and R. A. Weller (1997), Observations of surface forcing from the Subduction
- 596 Experiment: A comparison with global model products and climatological data sets. *J. Clim.* 10,
  597 2725-2742.
- 598 Payne, R. E., K. Huang, R. Weller, H.P. Freitag, M. F. Cronin, M.J. McPhaden, C. Meinig, Y.
- 599 Kuroda, N. Ushijima, and R.M. Reynolds (2002), A comparison of buoy meteorological

- systems. WHOI Tech. Rep. WHOI-2002-10, 67 pp., Woods Hole Oceanogr. Inst., Woods Hole,Mass.
- 602 Prytherch, J., E. C. Kent, S. Fangohr, and D. I. Berry (2014), A comparison of SSM/I-derived
- global marine surface-specific humidity datasets. *Int. J. Climatol.*, **35**, 2359–2381.
- 604 doi:https://doi.org/10.1002/joc.4150.
- Roberts, J. B., C. A. Clayson, F. R. Robertson, and D. L. Jackson (2010), Predicting near-surface
- atmospheric variables from Special Sensor Microwave/Imager using neural networks with a
- 607 first-guess approach. J. Geophys. Res., 115, D19113, doi:10.1029/2009JD013099.
- 608 Schlüssel, P., L. Schanz, and G. Englisch (1995), Retrieval of latent heat flux and longwave
- 609 irradiance at the sea surface from SSM/I and AVHRR measurements. *Adv. Space Res.*, **16**(10),
- 610 107–116.
- 611 Schulz, J., P. Schluessel, and H. Grassl (1993), Water vapour in the atmospheric boundary layer
- 612 over oceans from SSM/I measurements. *Int. J. Remote Sens.*, **14**, 2773 2789.
- 613 Schulz, E. W., S. A. Josey, and R. Verein (2012), First air-sea flux mooring measurements in the
- 614 Southern Ocean, *Geophys. Res. Lett.*, **39**, L16606, doi:10.1029/2012GL052290.
- 615 Send, U., Weller, R., Wallace, D., Chavez, F., Lampitt, R., Dickey, T., Honda, M., Nittis, K.,
- 616 Lukas, R., McPhaden, M. and Feely, R. (2010) "OceanSITES" in *Proc. of OceanObs'09:*
- 617 Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21-25
- 618 September 2009, Hall, J., Harrison, D.E. and Stammer, D., Eds., ESA Publication WPP-306,
- 619 doi:10.5270/OceanObs09.cwp.79
- 620 Sherwood, S. C., R. Roca, T. M. Weckwerth, and N. G. Andronova (2010), Tropospheric water
- 621 vapor, convection and climate. *Rev. Geophys.*, **48**, RG2001, doi:10.1029/2009RG000301.

- 622 Shi, L. (2001), Retrieval of atmospheric temperature profiles from AMSU-A measurement using
- 623 a neural network approach. J. Atmos. Ocean. Technol., 18, 340–347.
- 624 Shi, L., G. Peng, and J. J. Bates (2012) Surface air temperature and humidity from intersatellite-
- 625 calibrated HIRS measurements in high latitudes. J. Atmos. Ocean. Technol., 29, 3-13,
- 626 doi:10.1175/JTECH-D-11-00024.1.
- 627 Shie, C.-L., K. Hilburn, L. S. Chiu, R. Adler, I-I Lin, E. Nelkin, J. Ardizzone, and S. Gao (2012),
- 628 Goddard Satellite-Based Surface Turbulent Fluxes, Daily Grid, version 3. Edited by Andrey
- 629 Savtchenko., Greenbelt, MD, USA: Goddard Earth Science Data and Information Services
- 630 Center (GES DISC), Accessed Enter User Data Access Date at
- 631 doi:10.5067/MEASURES/GSSTF/DATA301.
- 632 Stephens, G. L. (1990), On the relationship between water vapor over the oceans and sea surface
  633 temperature. *J. Clim.*, **3**, 634–645.
- Taylor, P. K. (1982), Remote sensing of atmospheric water content and sea surface latent heat
- 635 flux. Proc. Annual Technical Conf. on Remote Sensing and the Atmosphere, Reading, United
- 636 Kingdom, Remote Sensing Society, 265–272.
- 637 Thiria, S., C. Mejia, F. Badran, and M. Crepon (1993), A neural network approach for modelling
- 638 nonlinear transfer functions: Application for wind retrieval from spaceborne scatterometer data.
- 639 J. Geophys. Res., 98, 22 827–22 842.
- 640 Weller, R. A. and S. P. Anderson (1996), Surface meteorology and air-sea fluxes in the western
- 641 equatorial Pacific warm pool during the TOGA Coupled Ocean-Atmosphere Response
- 642 Experiment. J. Clim., 9(8), 1959-1990.
- 643 Wentz, F.J. (1997), A well-calibrated ocean algorithm for SSM/I. J. Geophys. Res., 102, C4,
- 644 8703**-**8718.

- 645 Wentz, F. J. (2013), SSM/I Version-7 Calibration Report. Report number 011012, Remote
  646 Sensing Systems, Santa Rosa, CA, 46 pp.
- 647 Willett, K. M., P. D. Jones, N. P. Gillett, and P. W. Thorne (2008), Recent changes in surface
- humidity: Development of the HadCRUH dataset, J. Clim., 21, 5364–5383,
- 649 doi:10.1175/2008JCLI2274.1.
- 450 Yu, L., and R. A. Weller (2007), Objectively analyzed air-sea heat fluxes for the global ice-free
- 651 oceans (1981-2005). Bull. Am. Meteor. Soc., **88**(4), 527-539
- 452 Yu, L., 2007. Global variations in oceanic evaporation (1958-2005): The role of the changing
- 653 wind speed. J. Climate, 20(21), 5376-5390.
- 4654 Yu, L., X. Jin, and R. Weller (2008), Multidecade Global Flux Datasets from the Objectively
- Analyzed Air-sea Fluxes (OAFlux) Project: Latent and Sensible Heat Fluxes, Ocean
- 656 Evaporation, and Related Surface Meteorological Variables, OAFlux Project Tech. Rep. OA-
- 657 2008-01, 64 pp.
- 658 Zou, C.-Z., and W. Wang (2013), Climate algorithm theoretical basis document (C-ATBD) -
- AMSU radiance fundamental climate data record derived from integrated microwave inter-
- 660 calibration approach, *Tech. Rep.*, NOAA, Asheville, N. C.
- 661
- 662

# 663 List of Tables

- Table 1. Summary of the Qa/Ta algorithms developed in the past 30 years
- 665 Table 2. List of satellite sensors used in the study
- Table 3. List of the 30 buoys that are not associated with TRITON/TAO, PIRATA, and RAMA
- 667 Table 4. Regression coefficients for Equations (1)-(2)
- 668 Table 5. Regression coefficients for Equations (4)-(5)

| 670 | Figure captions   |
|-----|---|
| 671 | Figure 1. Locations of 147 moored surface buoys. The 121 red squares denote that the              |
| 672 | measurements made by those buoys in 2010 were used as training datasets for the global model,     |
| 673 | while the 4 red squares with black edges located at high latitudes were used as training datasets |
| 674 | for the high latitude model. The 22 blue squares denote that the measurements at these            |
| 675 | locations were made either before or after 2010 so that they can be used as independent           |
| 676 | validation datasets.  |
| 677 | Figure 2. Timeline and duration of the 11 sensors used in the study.                              |
| 678 | Figure 3. Daily-mean fields of brightness temperature (Tb) on 01 January 2010 observed by         |
| 679 | SSMIS F17 channels at (a) 19 GHz, (b) 22 GHz, (c) 37 GHz, and AMSU-A channel at (d) 52            |
| 680 | GHz. The daily mean fields are the average of the ascending and descending passes.                |
| 681 | Figure 4. Scatter diagrams of buoy measured $q_a$ versus Tb from SSMIS F17 channels at (a) 19     |
| 682 | GHz, (b) 22 GHz, (c) 37 GHz, and AMSU-A channel at (d) 52 GHz. The diagrams are based             |
| 683 | on a total of 31,793 tri-collocations between buoy, SSMIS, and AMSU-A that are constructed        |
| 684 | in 2010.  |
| 685 | Figure 5. Same as Figure 4 but for buoy measured $T_a$ .  |
| 686 | Figure 6. A flowchart of the three-step procedure that incorporates a high-latitude enhancement   |
| 687 | into the global retrieval algorithm.  |
| 688 | Figure 7. An illustration of how the three-step procedure works using the data generation on 01   |
| 689 | January 2010 as an example. The $q_a$ fields produced by (a) the global model (GM), (b) the       |
| 690 | high-latitude model (HLM), and (c) merge of the GM and HLM. The $T_a$ fields produced by (d)      |

691 the global model (GM), (e) the high-latitude model (HLM), and (f) merge of the GM and HLM.

- Figure 8. Examining the performance of the GM, HLM, and the merged approach for  $q_a$
- 693 retrievals using buoy measurements at Ocean Station PAPA 145°W, 50.0°N. Time series
- 694 comparison for (a) GM, (c) HLM, and (e) the merged, and the scatter diagrams of buoy versus
- 695 (b) GM, (d) HLM, and (f) the merged.
- 696 Figure 9. Same as Figure 8 but for  $T_a$ .
- Figure 10. Same as Figure 8 but for  $q_a$  at the KEO site 144.6°E, 32.4°N.
- Figure 11. Using SSM/I sensors to retrieve  $q_a$  (left) and  $T_a$  (right) with and without AMSU-A.
- 699 The mean differences between the two  $q_a$  patterns in (a) 2000, (b) 2001, (c) 2002, and (d) the
- three-year mean. The mean differences between the two  $T_a$  patterns in (e) 2000, (f) 2001, (g)
- 701 2002, and (h) the three-year mean.
- 702 Figure 12. Annual-mean time series of SSM/I-only retrievals (red line), SSM/I+SSMIS+AMSU
- retrievals (black), and the mean-adjusted SSM/I-only retrievals (grey) for (a)  $q_a$  and (b)  $T_{a.}$
- Figure 13. Time-mean fields of the retrieved  $q_a$  in (a) January, (b) July, and (c) annual mean, and
- the retrieved  $T_a$  in (d) January, (e) July, and (f) annual mean. The mean fields are constructed for the period 1988-2016.
- Figure 14. Statistical properties of the retrieved  $q_a$  and  $T_a$  daily products evaluated at the 147
- buoy locations. Retrieved-minus-buoy differences averaged over the available measurement
- periods for (a)  $q_a$  and (b)  $T_a$ , the standard deviations of daily retrieved-minus-buoy differences
- for (c)  $q_a$  and (d)  $T_a$ , and the correlation coefficient between retrieved and measured daily time
- 711 series for (e)  $q_a$  and (f)  $T_a$ .
- Figure 15. Scatter plots of the retrieved versus observed parameters for (a) specific humidity at
- 713  $2m(q_a)$  and (b) air temperature  $(T_a)$  based on buoy time series measurements at 147 locations
- 714 for the period 1988 2016.

| Algorithm                      | Sensors                        | Retrieval Methodology for $q_a$   | Training<br>data                 | Retrieval<br>Methodology for <i>T<sub>a</sub></i>  |
|--------------------------------|--------------------------------|---|----------------------------------|--|
| Liu (1986)                     |                                | First algorithm to relate $q_a$ to IWV  | Radiosonde<br>soundings          | N/A  |
| Liu (1988)                     | SSM/I                          | N/A   | Radiosonde<br>soundings          | Obtained from surface<br>mixing ratio assuming a<br>constant relative<br>humidity of 80% |
| Schulz et al.<br>(1993)        | SSM/I                          | Linear regression using IWV<br>in the lower 500m boundary<br>layer.                 | Radiosonde<br>soundings          | N/A  |
| Chou et al.<br>(1995)          | SSM/I                          | EOF-based approach using<br>IWV in the lower boundary<br>layer.                     | Radiosonde<br>soundings          | N/A  |
| Schluessel et<br>al.<br>(1995) | SSM/I                          | Linear regression using Tb<br>from five channels (19v, 22v,<br>37v, 19h, and 37h).  | Radiosonde<br>soundings          | N/A  |
| Jones et al.<br>(1999)         | SSM/I                          | First neural network approach<br>using SSMI/I IWV and NCEP<br>SST.                  | OISST<br>(NCEP<br>version)       | Simultaneously<br>retrieved  |
| Bentamy et al.<br>(2003)       | SSM/I                          | Similar to Schluessel et al.<br>(1995) but using a 4-channel<br>regression (no 37v) | Ship<br>observations             | N/A  |
| Kubota and<br>Hihara (2008)    | AMSR-E                         | Similar to Schluessel et al.<br>(1995) but using a 12-channel<br>regression         | NCEP and<br>Ship<br>observations | N/A  |
| Shi (2001)                     | AMSU-<br>A                     | N/A   | NCEP                             | Neural network<br>approach   |
| Jackson et al.<br>(2006; 2009) | AMSU-<br>A<br>SSM/I<br>SSM/T-2 | Multivariate linear regression  | Ship<br>Observations             | Simultaneously<br>retrieved  |
| Shi et al.<br>(2012)           | HIRS                           | Neural network approach   | ECMWF<br>and ships               | Simultaneously<br>retrieved  |
| Roberts et al.<br>(2010; 2012) | SSM/I                          | Neural network approach   | Research<br>vessels and<br>buoys | Simultaneously<br>retrieved  |
| This Paper                     | AMSU-<br>A<br>SSM/I<br>SSMIS   | A 4-Channel nonlinear<br>regression with high-latitude<br>enhancement               | Buoys                            | Simultaneously<br>retrieved  |

Table 1. Summary of the Qa/Ta algorithms developed in the past 30 years

Table 2. List of satellite sensors used in the study

| Sensor     | Period<br>in use       | Channel<br>in use | Swath<br>Width | Daily<br>coverage |  |
|------------|------------------------|-------------------|----------------|-------------------|--|
| SSM/I F08  | 01/01/88 –<br>12/31/91 |                   |                |                   |  |
| SSM/I F10  | 12/31/90 –<br>08/31/97 |                   | 1394 km        |                   |  |
| SSM/I F11  | 12/01/91 –<br>03/01/00 | 19v<br>22-r       |                | 74.50/            |  |
| SSM/I F13  | 05/01/95 –<br>11/01/09 | 22v<br>37v        |                | 74.3%             |  |
| SSM/I F14  | 05/01/97 –<br>12/01/05 |                   |                |                   |  |
| SSM/I F15  | 12/01/99 –<br>12/31/05 |                   |                |                   |  |
| SSMIS F17  | 01/01/07 –<br>12/31/16 | 19v, 22v, 37v     | 1700 km        | 83.5%             |  |
| AMSU-A N15 | 10/01/98 –<br>12/31/16 |                   |                |                   |  |
| AMSU-A N16 | 01/01/01 –<br>12/31/06 | 521               | 00.40.1        | 06.0%             |  |
| AMSU-A N17 | 06/01/02 –<br>12/31/03 | 324               | 2343 KIII      | 90.970            |  |
| AMSU-A N18 | 05/02/05 –<br>12/31/16 |                   |                |                   |  |

| Buoy          | Location        | Period used in study |
|---------------|-----------------|----------------------|
| Arabian Sea   | 61.5°E, 15.5°N  | 10/16/94 - 10/19/95  |
| ARC           | 30.0°E, 38.5°S  | 12/01/10 - 03/08/11  |
| ASREX91       | 132.0°W, 49.2°N | 11/01/91 - 01/06/92  |
| ASREX93       | 69.7°W, 33.9°N  | 12/15/93 - 03/23/94  |
| Bay of Bengal | 89.5°E, 18.0°N  | 12/08/14 - 01/29/16  |
| CLIMODE       | 65.0°W, 38.0°N  | 11/14/05 - 02/08/07  |
| СМО           | 70.5°W, 40.5°N  | 07/31/96 - 06/12/97  |
| TOGA COARE    | 156.0°E, 1.8°S  | 10/22/92 - 03/03/93  |
| KEO           | 144.6°E, 32.4°N | 06/16/04 - 12/31/16  |
| MLML91        | 20.8°W, 59.5°N  | 04/30/91 - 09/05/91  |
| NTAS          | 51.0°W, 14.8°N  | 03/31/01 - 12/31/16  |
| PACS(n)       | 125.4°W, 9.9°N  | 04/30/97 - 09/13/98  |
| PACS(s)       | 124.6°W, 2.8°S  | 04/21/97 - 09/19/98  |
| РАРА          | 145.0°W, 50.0°N | 06/08/07 - 12/31/16  |
| SESMOOR       | 61.2°W, 42.5°N  | 10/18/88 - 03/07/89  |
| SMILE         | 123.5°W, 39.6°N | 11/15/88 - 05/14/89  |
| SOFS          | 142.0°E, 46.8°S | 03/17/10 - 12/24/12  |
| SPURS         | 38.0°W, 24.6°N  | 09/14/12 - 09/30/13  |
| Stratus       | 85.0°W, 20.0°S  | 10/08/00 - 12/31/16  |
| WHOTS         | 158.0°W, 22.8°N | 09/15/04 - 12/31/16  |
| Subduction    | 34.0°W, 33.0°N  | 06/19/91 - 06/20/93  |
| Subduction    | 22.0°W, 33.0°N  | 06/19/91 - 06/20/93  |
| Subduction    | 29.0°W, 25.5°N  | 06/19/91 - 06/20/93  |
| Subduction    | 22.0°W, 18.0°N  | 06/19/91 - 06/20/93  |
| Subduction    | 34.0°W, 18.0°N  | 06/19/91 - 06/20/93  |
| OOI SO        | 89.2°W, 54.4°S  | 12/17/15 - 11/28/16  |
| OOI Argentine | 42.6°W, 42.9°S  | 11/14/15 - 11/08/16  |

Table 3. List of the 30 buoys that are not associated with TRITON/TAO, PIRATA, and RAMA 727

| OOI Irminger Sea | 39.5°W, 59.9°N  | 09/13/14 - 01/25/16 |
|------------------|-----------------|---------------------|
| NDBC 46075       | 160.8°W, 53.9°N | 05/08/04 - 12/31/16 |
| NDBC 46085       | 142.5°W, 55.9°N | 05/05/07 - 12/31/16 |

- 732 Table 4. Regression coefficients for Equations (1)-(2)

Add one more digit for the highlighted values (i.e. need 5 digits after the decimal point if the

735 integer digit is 0)

| Model                | A <sub>0</sub> | A <sub>19v</sub> | A <sub>22v</sub> | A <sub>37v</sub> | A <sub>52v</sub> | В <sub>19v</sub> | B <sub>22v</sub> | B <sub>37v</sub> | B <sub>52v</sub> |
|----------------------|----------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| q <sub>a</sub> (GM)  | 1423.34        | 0.46967          | 0.43401          | -0.92292         | -11.494          | -0.00071         | -0.00072         | 0.00155          | 0.02336          |
| q <sub>a</sub> (HLM) | -127.10        | -0.21113         | 0.71712          | -0.78268         | 1.1918           | 0.00062          | -0.00139         | 0.00153          | -0.00222         |
| T <sub>a</sub> (GM)  | -389.96        | 0.48775          | 0.75817          | -1.3532          | 2.8253           | -0.00093         | -0.00156         | 0.00265          | -0.00463         |
| T <sub>a</sub> (HLM) | -624.73        | -1.7726          | 1.4135           | -1.4004          | 6.1955           | 0.00459          | -0.00299         | 0.00268          | -0.01220         |

Table 5. Regression coefficients for Equations (4)-(5) 740

| /41 | 7 | 4 | 1 |
|-----|---|---|---|
|-----|---|---|---|

| Model                | A <sub>0</sub> | A <sub>19v</sub> | A <sub>22v</sub> | A <sub>37v</sub> | B <sub>19v</sub> | B <sub>22v</sub> | <b>B</b> <sub>37v</sub> | A <sub>SST</sub> | B <sub>SST</sub> |
|----------------------|----------------|------------------|------------------|------------------|------------------|------------------|-------------------------|------------------|------------------|
| q <sub>a</sub> (GM)  | 20.773         | 1.0004           | 0.68113          | 1.0004           | -0.00166         | -0.00100         | 0.00306                 | N/A              | N/A              |
| q <sub>a</sub> (HLM) | 36.335         | -0.48568         | 0.93744          | -0.4857          | 0.00123          | -0.00180         | 0.00157                 | N/A              | N/A              |
| T <sub>a</sub> (GM)  | -14.114        | 0.06469          | 1.4800           | -1.4004          | 0.00003          | -0.00293         | 0.00271                 | 0.36543          | 0.00989          |
| T <sub>a</sub> (HLM) | 170.11         | -2.6251          | 2.0226           | -1.2309          | 0.00660          | -0.00433         | 0.00261                 | 1.2728           | _<br>0.02626     |









/34

755 Figure 2





### 757 Figure 3



Figure 4 



Figure 5





770 Figure 7



Figure 8



774 Figure 9



Figure 10





778 Figure 11









782 Figure 13



784 785 Figure 14



