

A comparison of SSM/I-derived global marine surface-specific humidity datasets

John Prytherch,^{a*} Elizabeth C. Kent,^a Susanne Fangohr^b and David I. Berry^a

^a National Oceanography Centre, Southampton, UK

^b School of Ocean and Earth Science, University of Southampton, UK

ABSTRACT: Satellite-based microwave sensors have, since the 1980s, provided a means to retrieve near-surface marine specific humidity (q_a), accurate estimation of which is necessary for climate and air–sea interaction applications. Seven satellite measurement-derived monthly mean humidity datasets are compared with one another and with a dataset constructed from *in situ* measurements. The means, spatial and temporal structures of the datasets are shown to be markedly different, with a range of yearly, global mean q_a of $\sim 1 \text{ g kg}^{-1}$. Comparison of the datasets derived using the same satellite measurements of brightness temperature reveals differences in q_a that depend on the source of satellite data; the processing and quality control applied to the data; and the algorithm used to derive q_a from the satellite measurements of brightness temperature. Regional differences between satellite-derived q_a due to the choice of input data, quality control and retrieval algorithm can all exceed the accuracy requirements for surface flux calculation of $\sim 0.3 \text{ g kg}^{-1}$ and in some cases can be several g kg^{-1} in monthly means for some periods and regions.

KEY WORDS marine climatology; satellite observations; surface humidity; air–sea interaction; specific humidity; SSM/I

Received 14 May 2013; Revised 22 July 2014; Accepted 9 August 2014

1. Introduction

Satellite-based microwave measurements have become an established data source for determining several important climatic variables in the marine system, including columnar atmospheric water vapour concentration, w (kg m^{-2}) and near-surface properties, such as wind speed and specific humidity, q_a (g kg^{-1}). Global coverage of the world's oceans at a time interval of a few days by orbiting microwave sensors has been available since the 1980s and several dataset products have been constructed from these measurements. These provide researchers with both the meteorological variables and air–sea fluxes obtained from them using bulk formulae. Humidity of the near-surface atmosphere is designated an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS, 2010).

Air–sea heat and freshwater fluxes are key components of the Earth's climate system. Near-surface humidity contrasts drive evaporation and exert an important influence on the global energy and hydrological cycles. Comparisons of global marine latent heat flux products identify variation in the near-surface atmospheric humidity fields as the largest source of flux uncertainty (Bourras, 2006; Tomita and Kubota, 2006; Iwasaki *et al.*, 2010; Andersson *et al.*, 2011; Smith *et al.*, 2011).

Satellite sensors provide data at a higher spatial resolution than are available *in situ* (Bourras *et al.*, 2010).

However, they do not directly measure near-surface q_a , which is therefore derived using empirical relationships with *in situ* observations. Over the ocean, a variety of sources of *in situ* data have been used to derive and validate the near-surface atmospheric humidity retrievals. These sources include observations from research vessels, buoys, radiosonde ascents and routine weather measurements from the Voluntary Observing Ships (VOS) programme, for example those collated in the International Comprehensive Ocean-Atmosphere Data Set (ICOADS, Woodruff *et al.*, 2011).

The Special Sensor Microwave/Imager (SSM/I) instruments carried on board the Defense Meteorological Satellites Programme (DMSP) series of sun-synchronous, near-polar orbiting satellites have been the most commonly used passive microwave sensors for near-surface marine humidity retrievals. The sequence began with satellite F08 (launched 20 June 1987) followed by F10, F11, F13, F14 and F15 (launched 12 December 1999). Successive satellites operated for overlapping time periods to allow for cross calibration. The SSM/I sensor was replaced by the 24 channel Special Sensor Microwave Imager Sounder (SSMIS) on DMSP satellites F16, F17 and F18, which became operational from November 2005 onwards. The primary channels in the SSMIS are similar to those in SSM/I but to date SSMIS measurements have not been used to derive q_a due to problems encountered with calibration (Kunkee *et al.*, 2008).

This analysis compares several prominently available SSM/I-derived near-surface humidity datasets with one

* Correspondence to: J. Prytherch, Department of Earth and Environment, University of Leeds, UK. E-mail: johnprytherch@gmail.com

another and with a dataset constructed from *in situ* measurements. The satellite humidity products compared are: the National Aeronautics and Space Administration (NASA) Goddard Space Science Laboratory Satellite-Based Surface Turbulent Fluxes (GSSTF2b, GSSTF2c and GSSTF3, collectively the Goddard datasets); Hamburg Ocean Atmosphere Parameters and fluxes from Satellite data version 3.2 (HOAPS3.2); French Research Institute for Exploration of the Sea merged flux dataset (IFREMER); Japanese Ocean Flux Data sets with Use of Remote Sensing Observations (J-OFURO2); SeaFlux Turbulent Flux Dataset V1.0 (SeaFluxV1.0). Our analysis extends a study comparing available *in situ* and atmospheric reanalysis humidity products (Kent *et al.*, 2014).

Smith *et al.* (2011) compare nine air–sea flux products including IFREMER and older versions of datasets used in this study: HOAPS2 and GSSTF2. They found large differences between q_a in the different datasets. In particular q_a from IFREMER and HOAPS was underestimated under El Niño conditions and the satellite-derived products had large biases off the coast of Peru which were attributed to the presence of persistent stratocumulus clouds.

Brunke *et al.* (2011) compared 11 datasets, including GSSTF2, GSSTF2b, J-OFURO2 and HOAPS3. They found differences in q_a from measurements on 12 research cruises to vary in magnitude and sign among the satellite datasets, with little consistency among the differences for either a particular cruise or for a particular region.

Bourras (2006) compared five satellite-derived humidity products, including J-OFURO2, HOAPS2 and GSSTF2 with observations from moored buoys in the Tropics (66 buoys), and off the coasts of United States (9 buoys) and Western Europe (4 buoys). He concluded that HOAPS q_a was more accurate than that from GSSTF2.

Iwasaki *et al.* (2010) compared q_a from J-OFURO2 with observations on 15 research cruises. They found an overall overestimate of q_a from J-OFURO2 to be 1.1 g kg^{-1} with RMS difference of 1.2 g kg^{-1} . Tomita *et al.* (2010) compare J-OFURO2 q_a with that measured by two buoys in the Kuroshio extension. In winter J-OFURO2 overestimated q_a relative to one of the buoys by $1.5 \pm 1.1 \text{ g kg}^{-1}$ and in summer underestimated with a difference of $-0.3 \pm 1.4 \text{ g kg}^{-1}$. For the common period of measurement of the two buoys an overestimate by J-OFURO2 of 1.75 g kg^{-1} was observed by one buoy and 0.85 g kg^{-1} at the other.

Santorelli *et al.* (2011) found a mean offset of $-0.3 \pm 0.8 \text{ g kg}^{-1}$ between q_a from IFREMER and buoys in the Tropical Atlantic. In the Mediterranean IFREMER q_a was higher than buoy measurements ($0.5 \pm 1.0 \text{ g kg}^{-1}$) and in the Gulf Stream region IFREMER q_a was on average too low ($-1.9 \pm 2.7 \text{ g kg}^{-1}$).

The focus of this study is on the magnitude and causes of differences between the SSM/I products. The differences are shown to result from choices in input data, data processing route, calibration, quality control and retrieval algorithm and are substantial relative to desired surface flux accuracy.

The instrumentation, retrieval methodology and the datasets used in this comparison are described in Section 2 along with details of gridding and masking used in the analysis. Section 3 presents the results of a comparison of published datasets which are discussed in Section 4 where some conclusions are drawn.

2. Surface humidity measurement and datasets

2.1. SSM/I: satellite programme and data sources

SSM/I's seven microwave channels are located in four frequency bands, sensing vertically and horizontally polarized radiation at 19.35 GHz (hereafter 19V and 19H respectively), 37.0 GHz (37V, 37H) and 85.5 GHz (85V, 85H), and vertically polarized radiation at 22.235 GHz (22V). The swath width is approximately 1394 km. Due to the sensor's large footprint, land can affect ocean brightness temperatures up to 3–4 pixels away from the coastline (Cavalieri *et al.*, 1997, 1999). Most areas of the globe receive two measurements per day and complete global coverage is obtained every 3 days. The transmission and processing routes for SSM/I data are described by Ritchie *et al.* (1998).

The near-surface q_a datasets derive their observations either from the National Environmental Satellite, Data, and Information Service (NESDIS) Temperature Data Records (TDR) or from Remote Sensing Systems (RSS). Both the NESDIS TDR and RSS archives originate from the US Navy's Fleet Numerical Meteorology and Oceanography Center (FNMOC) which receives the SSM/I observations directly from the ground receiving stations. Calibration and initial quality control (e.g. flagging of out of bounds thermistor readings) is carried out by the processing centres, and may differ between them. Figure 1 summarizes the data flow and processing for SSM/I data and Table 1 summarizes the data sources and processing for the datasets compared in this paper.

Ritchie *et al.* (1998) compared the antenna temperatures provided by FNMOC and RSS and found a mean difference of approximately zero and a standard deviation of 0.15 K for each channel. The comparisons presented by Ritchie *et al.* (1998) do not include all the potential sources of differences due to input data for the datasets compared here, in particular inter-satellite calibration

In addition to the humidity datasets, in this study we have used antenna temperature measurements from the RSS version 6 (RSSv6) archive for the F11 satellite only, as distributed by the National Climatic Data Centre (NCDC). Data are provided as antenna temperature, from which brightness temperature $T_b X$, (K; where X is the channel the brightness temperature is obtained from, e.g. 19V) can be obtained via an established method (Wentz, 1991).

2.2. Inter-satellite calibration

Input data to the humidity products compared in this analysis have inter-satellite calibration applied at three different centres: RSS; HOAPS and Colorado State University (CSU) (Figure 1). Additionally one product (IFREMER)

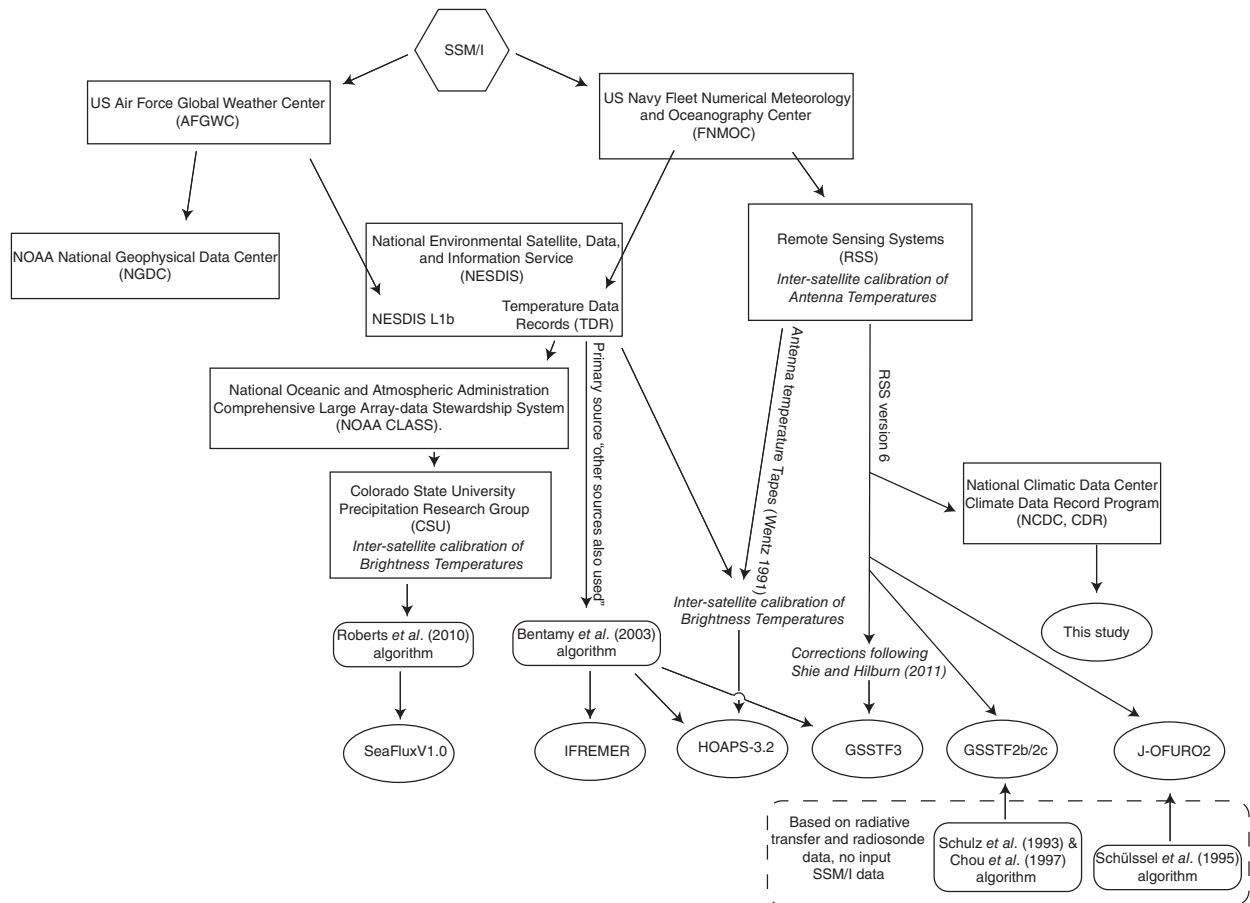


Figure 1. Flow diagram of processing routes for SSM/I measurements for the q_a datasets compared.

uses data without adjustment. Brightness temperature biases between the SSM/I instruments on different satellites are typically less than ± 1 K (Jackson *et al.*, 2006; Wentz, 2010) but can be as large as ± 4 K. Differences are typically larger for the horizontally polarized channels, particularly 85H (Andersson *et al.*, 2010). In addition to inter-sensor differences, the inter-satellite calibration procedure is intended to correct any variation resulting from the differing orbital characteristics of the various DMSP satellites (e.g. Andersson *et al.*, 2010; Wentz, 2010).

Jackson *et al.* (2006) concluded, based on a comparison of estimates from SSM/I and measurements on eight research cruises in 1999 and 2001, that any effect of the inhomogeneities between satellites on q_a was small. However this assessment was based on root mean square (RMS) differences and they noted biases between satellite brightness temperatures 'usually below 1°C '. Much larger inter-satellite differences are seen at other times in the record (Semunegus, 2011).

The RSS inter-satellite calibration is applied to antenna temperatures (Wentz, 2010). Variability of the antenna temperatures with a period of 122 days was thought to be caused by incorrect adjustment for the Earth incidence angle (EIA) and adjusted as such in version 6. However, it was later discovered that the F10 oscillations were a side effect of the modelling of the EIA dependence of the wind-induced component of sea surface emissivity. For

version 7.0, RSS use a radiative transfer model (Meissner and Wentz, 2012) to correct the effects resulting from varying EIA and the 122-day oscillation is not seen (K. Hilburn, 2012, pers. comm.; Wentz, 2013).

The HOAPS inter-sensor calibration as described in Andersson *et al.* (2010) is referenced to the SSM/I sensor on DMSP F11 due to its reliable long-term stability and the large number of overlaps with the other SSM/I sensors. In the case of DMSP F10 an allowance was made for orbit changes (Wentz, 2010) and calibration coefficients were linearly interpolated between values obtained for 1992 and 1996. The HOAPS brightness temperature dataset and calibration information has recently been published with optional inter-sensor calibration and EIA corrections (Fennig *et al.*, 2013).

The approach taken by CSU is to match all of the brightness temperatures to those from the Tropical Rainfall Measuring Mission's (TRMM) Microwave Imager (TMI) (described on http://mrain.atmos.colostate.edu/LEVEL1C/level1c_devssmi.html, accessed 15 July 2014). Differences between the TMI and SSM/I sensors due to the channel frequencies, EIA and spatial resolution were accounted for using radiative transfer modelling.

2.3. Near-surface marine humidity retrieval algorithms

The dependencies on environmental parameters of the brightness temperatures measured by the SSM/I are

Table 1. Characteristics of satellite-derived q_a datasets.

Dataset	Period	Grid resolution	q_a retrieval method	Input data (see also Figure 1)	QC
GSSTF3; Shie <i>et al.</i> (2012); Shie (2012)	July 1987–December 2008	0.25° monthly; 0.25° daily	Bentamy <i>et al.</i> (2003)	RSSv6 SSM/I F08 – F15. + Hilburn and Shie (2011)	Humidity inversion
GSSTF2b; Shie <i>et al.</i> (2010a); Shie (2010) and GSSTF2c; Shie <i>et al.</i> (2011); Shie (2011)	July 1987–December 2008	1° monthly; 1° daily	Chou <i>et al.</i> (1997)	RSSv6 SSM/I F08 – F15. NCEP daily SST	Humidity inversion Large droplet (S93)
HOAPS3.2; Fennig <i>et al.</i> (2012); Andersson <i>et al.</i> (2010)	July 1987–December 2008	0.5° monthly; 0.5° 6-hourly	Bentamy <i>et al.</i> (2003)	SSM/I F08 – F15 from RSS and NESDIS AVHRR SST In house processing, including inter-satellite calibration Fennig <i>et al.</i> (2013)	Large droplet (HOAPS3.2) Humidity inversion
IFREMER; Bentamy <i>et al.</i> (2003)	March 1992–December 2007	1° monthly; 1° weekly	Bentamy <i>et al.</i> (2003)	Primary source NESDIS from FNMOC TDR (no inter-satellite calibration)	Large droplet (S93) Bentamy <i>et al.</i> (1999)
J-OFURO2; Kubota <i>et al.</i> (2002)	January 1988–December 2006	1° monthly; 1° daily	Schlüssel <i>et al.</i> (1995)	RSSv6 SSM/I F08 – F14	Large droplet (S93)
SeaFlux V1.0 C. A. Clayson, pers. comm.	January 1998–December 2007	0.25° monthly; 0.25° 3-hourly	Roberts <i>et al.</i> (2010)	Beta version 3, CSU SSM/I F11 – F15 SST (Reynolds <i>et al.</i> , 2002)	

complex and the appropriate approach to determining near-surface q_a from the measurements is not immediately obvious. This has led to a number of different statistical approaches to retrieving these parameters from SSM/I. Simple linear statistical relationships between SSM/I measurements and marine near-surface q_a have been developed by Schulz *et al.* (1993), Schlüssel *et al.* (1995) and Bentamy *et al.* (2003). The latter two algorithms are used in the datasets analysed and their coefficients are given in Table 2.

The Schlüssel *et al.* (1995; hereafter S95) algorithm was developed using simulated SSM/I data based on radiative transfer calculations using conditions measured during radiosonde ascents (described in detail in Schulz *et al.*, 1993; hereafter S93) and obtains near-surface q_a directly from brightness temperatures, changing from a four channel (as used by S93) to a five-channel regression (addition of the 37H channel; coefficients given in Table 2).

Bentamy *et al.* (2003; hereafter B03) revisited the problem using the four-channel approach of S93 but with a wider global distribution of *in situ* data from the Comprehensive Ocean-Atmosphere Data Set (COADS Release 2; Woodruff *et al.*, 1998). The ship observations were adjusted from an assumed height of 15 m; Kent *et al.*

(2007) subsequently demonstrated that mean humidity measurement heights for this period were about 22 m. B03 adjusted the humidity observations to give an estimate of the 10-m neutral specific humidity (q_{10n} ; W. Drennan, 2012, pers. comm.).

An empirical orthogonal function (EOF)-based approach was developed by Chou *et al.* (1995, 1997). The precipitable water of the bottom 500 m is calculated following S93 and data affected by scattering from large water or ice particles were removed using the same criteria as S93. The EOF parameters, derived from radiosonde soundings, are then used to calculate the expected shape of the humidity profile and hence daily near-surface q_a . A positive bias in the retrieval was noted in the summertime extratropics and Eastern Tropical Pacific and thought to be due to near-surface temperature inversions. To address this, saturation specific humidity at the SST [determined from National Centre for Environmental Prediction (NCEP) daily estimates] was set as an upper bound for near-surface q_a .

Roberts *et al.* (2010) derived a non-linear algorithm using inter-calibrated SSM/I data from the CSU repository for the period 1998–2007 and SST from Reynolds *et al.* (2002) using a neural network technique. The algorithm was trained using research vessel and buoy observations

Table 2. Retrieval algorithm coefficients.

	a_0	a_1	a_2	a_3	a_4	a_5
Bentamy <i>et al.</i> , 2003	-55.9227	0.4035	-0.2944	0.3511	-0.2395	-
Schlüssel <i>et al.</i> , 1995	-80.23	0.6295	-0.1655	0.1495	-0.1553	-0.06695
Schulz <i>et al.</i> , 1993	-116.1763	0.7205	-0.4658	0.3038	-0.0969	-

$$q_a \text{ (g kg}^{-1}\text{)} = a_0 + a_1 T_b 19V + a_2 T_b 19H + a_3 T_b 22V + a_4 T_b 37V + a_5 T_b 37H.$$

The Schulz *et al.* (1993) coefficients are obtained from a combination of their linear algorithms for retrieval of lowest 500 m humidity (w_1 , g cm⁻²), and for retrieval of q_a from w_1 .

for the period 1998–2007 from the SeaFlux database (Curry *et al.*, 2004).

Other available algorithms, such as the linear regression of S93, the neural network methods of Jones *et al.* (1999) and Meng *et al.* (2007), the genetic algorithm of Singh *et al.* (2006) or the multi-sensor retrieval approach of Jackson *et al.* (2006, 2009) are not used in the datasets compared in this paper and so are not discussed here.

A new version of the B03 algorithm has recently been developed, incorporating both SST and air–sea temperature difference data (Bentamy *et al.*, 2013). Bentamy *et al.* (2013) explicitly use SST in their algorithm and show (in their Figure 1) large differences between q_a from the *in situ*-based NOC Surface Flux dataset v2.0 (NOCv2.0; Berry and Kent, 2009, 2011) and B03 that vary with SST. However these large differences are likely an artefact of the method chosen to average and present the data. At the time of writing, the coefficients for the new algorithm are not published and are not used in any available monthly average dataset so it does not feature in the analysis presented here.

2.4. Comparison datasets used

2.4.1. Introduction

A summary of available satellite-derived global datasets is given here and in Table 1. The analysis in this study also uses q_{10} from the *in situ*-based NOCv2.0 (Berry and Kent, 2009, 2011). Further details of this dataset relevant to this study are given in Kent *et al.* (2014).

2.4.2. Goddard Satellite-Based Surface Turbulent Fluxes 2b/2c/3

The GSSTF2b dataset is an update of GSSTF2 (Chou *et al.*, 2003). SSM/I brightness temperatures from RSSv6 are used with the humidity retrieval algorithm of Chou *et al.* (1997). Daily estimates are produced for each of the satellites (F08, F10, F11, F13, F14 and F15) and are weighted equally to give combined daily estimates. Two different versions of GSSTF2b have been produced, set 1 and set 2. Set 1 uses all available satellites; set 2 excludes data for some satellites for some months for the purpose of reducing the negative trend of global q_a with time observed in the GSSTF2b q_a and latent heat fluxes. GSSTF2b set 2 was thought to have a better representation of temporal variability than set 1 and was recommended for applications where temporal stability is important. GSSTF2b is available for the ice-free open

ocean, though the mask used is not described (Shie, 2010).

The difficulties encountered that gave rise to the production of the two different versions of GSSTF2b (Shie *et al.*, 2010a, 2010b) were subsequently attributed in part to changes in the EIA of the SSM/I sensors (Shie, 2011; Shie and Hilburn, 2011). Hilburn and Shie (2011) developed an algorithm to adjust the brightness temperatures for these artefacts. Application of this adjustment algorithm to produce a new brightness temperature dataset reduced the trends seen in the GSSTF2b latent heat fluxes and allowed the inclusion of all SSM/I data in the GSSTF2c dataset construction (GSSTF2c is available from http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldingsMEASURES.pl?PR_OGRAM_List=ChungLinShie_OldVer; more information from <https://climatedataguide.ucar.edu/climate-data/goddard-satellite-based-surface-turbulent-fluxes-version-2c-gsstf-2c>). Accordingly only one version of GSSTF2c is provided (Shie, 2011; Shie *et al.*, 2011). The monthly versions of GSSTF2b (which they denote ‘set 1 combined’) and GSSTF2c (denoted ‘combined’) are used in this study.

In August 2012 GSSTF3 was released (Shie, 2012; Shie *et al.*, 2012) (GSSTF3 is available from http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldingsMEASURES.pl?PR_OGRAM_List=ChungLinShie; more information from <https://climatedataguide.ucar.edu/climate-data/goddard-satellite-based-surface-turbulent-fluxes-version-3-gsstf3>). This version uses the same brightness temperatures as in GSSTF2c but replaces the Chou *et al.* (1997) algorithm for q_a with that of B03.

2.4.3. Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data Version 3.2

The HOAPS3.2 dataset determines q_a from SSM/I following the algorithm of B03. An inter-satellite calibration is applied, with all brightness temperatures adjusted to the measurements of the F11 satellite (Fennig *et al.*, 2012; Andersson *et al.*, 2010) (HOAPS3.2 is available from <http://wui.cmsaf.eu>; more information from <https://climatedataguide.ucar.edu/climate-data/hoaps-hamburg-ocean-atmosphere-parameters-and-fluxes-satellite-data>). A modified version of the S93 large water droplet quality control is applied, removing measurements for which either $T_b 19H > 185$ K, $T_b 37H - T_b 19H > 40$ K or $T_b 37V - T_b 37H < 35$ K (A. Andersson, 2012, pers. comm.). Initial analysis of the B03 algorithm revealed that q_a was biased high in conditions when warmer air moves

over a colder ocean surface, and fog or stratus occurs (A. Andersson, 2012, pers. comm.). Following Chou *et al.* (1997) q_s , determined from AVHRR SST and sampled to the SSM/I grid, is used as an upper bound for q_a to reduce this error (A. Andersson, 2012, pers. comm.). The land mask is that of the Global Land One-km Base Elevation data base (GLOBE Task Team and Others, 1999), extended 50 km into the sea, and with land masses of a diameter less than 5 km removed. A daily sea-ice mask is generated from SSM/I brightness temperatures using the NASA algorithm of Swift *et al.* (1985). After gridding, grid boxes with an average sea-ice fraction above 15% for at least five consecutive days are designated ice covered and removed from the product. The sea-ice margin is then extended 50 km further into the ocean (Andersson *et al.*, 2010). HOAPS3.2 extends the time period covered in the preceding HOAPS-3 dataset. Differences in the global mean q_a of versions 3 and 3.2 are of order 0.01 g kg^{-1} (Andersson *et al.*, 2010).

2.4.4. French Research Institute for Exploration of the Sea merged flux dataset

The IFREMER dataset includes ocean surface fluxes and surface and near-surface parameters (Bentamy *et al.*, 2003). SSM/I brightness temperatures (from the CERSAT archive) are used to derive q_{10m} following the algorithm of B03 and applying the large water droplet quality control exclusion of S93 and the liquid water contamination criteria of Bentamy *et al.* (1999). Daily averages from all available SSM/I measurements are first determined, and then weekly and monthly averages are determined using an objective method that makes use of spatial and temporal structure (A. Bentamy, 2012, pers. comm.). Land and sea-ice areas are removed; with the sea-ice concentration determined on a weekly basis from SSM/I brightness temperatures (Bentamy *et al.*, 2008). (IFREMER is available from <ftp://ftp.ifremer.fr/ifremer/cersat/products/gridded/flux-merged>).

2.4.5. Japanese Ocean Flux Data sets with Use of Remote Sensing Observations

For the J-OFURO2 dataset (release id HF003) q_a was calculated following the algorithm of S95 and the S93 quality control is applied to remove measurements contaminated by large water droplets (Kubota *et al.*, 2002) (http://dtsv.scc.u-tokai.ac.jp/j-ofuro/update_detail_2009_04_01.html, accessed 15 July 2014). DMSP F15 was found to give significantly different values of q_a to e.g. F13 and F14 and was therefore excluded from the dataset. J-OFURO2 q_a is calculated by optimum interpolation using climatological monthly q_a from the Japanese 25-year ReAnalysis (JRA-25, Onogi *et al.*, 2007) as a first guess (M. Kubota, 2012, pers. comm.). (J-OFURO2 is available from <http://dtsv.scc.u-tokai.ac.jp/j-ofuro/>; more information from <https://climatedataguide.ucar.edu/climate-data/japanese-ocean-flux-data-sets-use-remote-sensing-observations-j-ofuro2>).

2.4.6. SeaFlux Turbulent Flux Dataset version 1.0

The SeaFluxV1.0 dataset (C. A. Clayson, pers. comm.) uses inter-calibrated SSM/I brightness temperatures to estimate q_{10m} following the neural network algorithm method of Roberts *et al.* (2010) (SeaFluxV1.0 is available from <http://seafux.org/>). For grid points with missing data, a model-based interpolation, based on the time tendencies of NASA's MERRA reanalysis is applied. The SeaFluxV1.0 dataset uses the land and sea-ice mask provided by NOAA's Optimum Interpolation (OI) Sea Surface Temperature Version 2 (Reynolds *et al.*, 2002).

2.5. Regridding and masking

All of the datasets compared are available from their producers at monthly resolution and these versions of the datasets are used here. In this study, each of the available datasets has been averaged onto the same $1^\circ \times 1^\circ$ grid (that of the NOCv2.0 product) using simple averaging of all non-missing values with latitude weighting. IFREMER, J-OFURO2, GSSTF2b/2c, and NOCv2.0 are already available as $1^\circ \times 1^\circ$ monthly grids, HOAPS3.2 is averaged from its 0.5° grid onto the 1° grid, and SeaFluxV1.0 and GSSTF3 are averaged from their 0.25° grids. In addition, instantaneous SSM/I retrievals from RSSv6 (Wentz, 2010) for the calendar year 1996 were used to generate global humidity q_a maps using the linear algorithms of S95 and B03. RSSv6 has been shown to contain residual biases that depend on EIA (Hilburn and Shie, 2011). Therefore, only measurements from the satellite with the smallest trend in EIA (F11, Hilburn and Shie, 2011; also used as an inter-satellite calibration baseline by Andersson *et al.*, 2010) were included and gridded as per the published datasets.

All the published SSM/I-based datasets, plus NOCv2.0, are made available with land and sea-ice masks applied. The most conservative of these masks (and the one encompassing the entire SSM/I period), that applied to the GSSTF2c dataset (Section 3.1), was applied to all other datasets prior to averaging to 1° resolution.

3. Results

3.1. Data coverage

Figure 2 shows global maps of the fraction of months sampled in each dataset over the duration of the particular dataset. These maps, and those in Figure 3, are at the native spatial resolution, all subsequent maps are at 1° spatial resolution. Also shown is a monthly time series of the fraction of non-land gridboxes containing q_a , again calculated at the native spatial resolution (Figure 2(e)) and zonal mean fractional coverage (Figure 2(f)). The land and sea-ice mask used in the Goddard datasets (GSSTF2c/3) has the largest fraction of missing data as the masking for sea-ice used is invariant over time (Figure 2(a) and (e)). Although GSSTF3 is available at higher spatial resolution than GSSTF2c, the mask applied is similar. The F08 SSM/I was temporarily shut down in December 1987 because

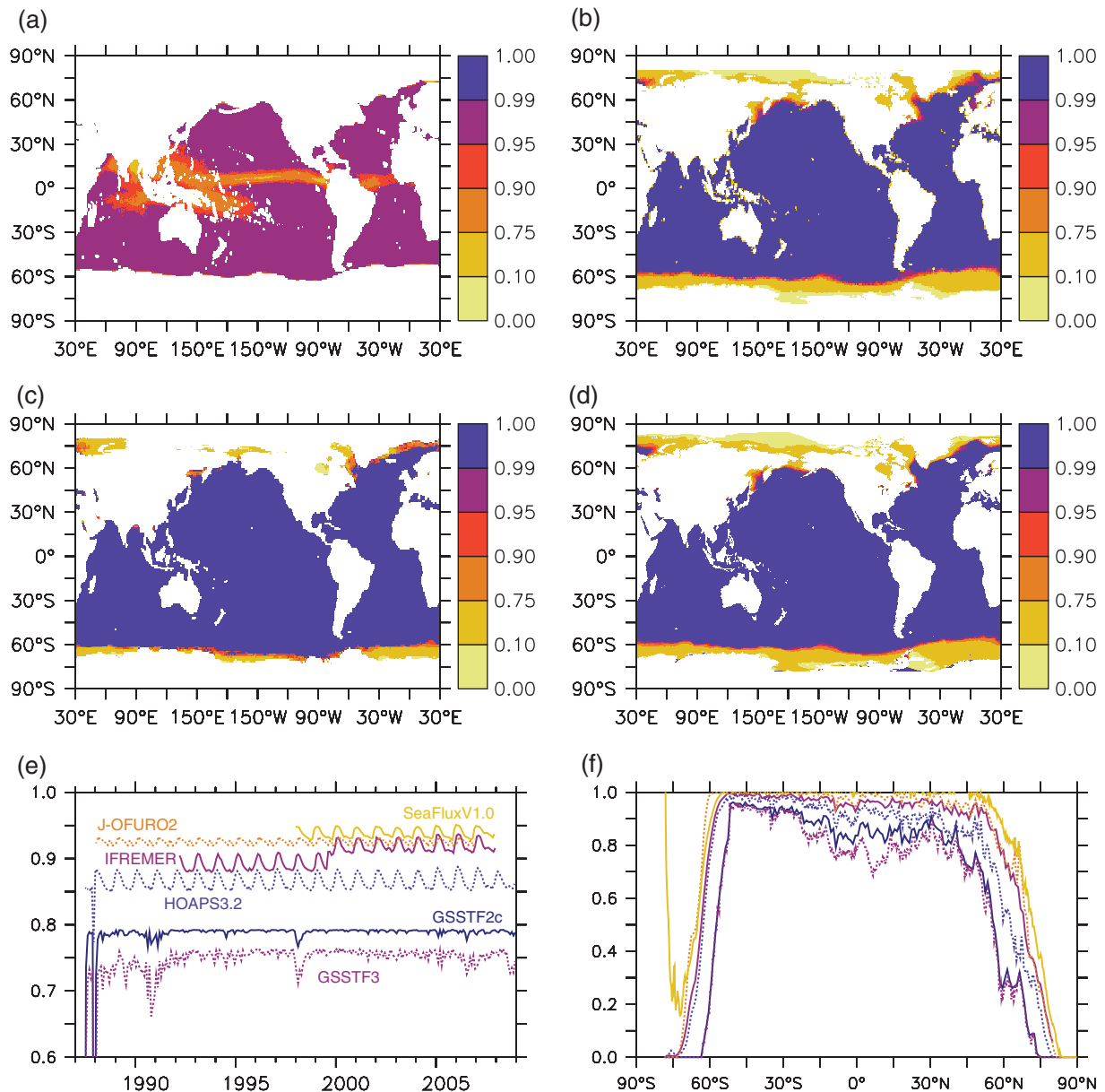


Figure 2. Coverage for each dataset at native resolution. (a) Fraction of gridboxes sampled for GSSTF3 over period of record (as Table 1); (b) as (a) but for IFREMER; (c) as (a) but for J-OFURO2; (d) as (a) but for SeaFluxV1.0; (e) Fractional coverage for 1° ocean-only gridboxes; Dark solid (blue in online): GSSTF2c, Mid dotted (purple in online): GSSTF3, Dark dotted (blue in online): HOAPS3.2, Mid solid (purple): IFREMER, Light dotted (orange): J-OFURO2, Light solid (orange): J-OFURO2 (f) as (e) but for zonal mean fractional coverage. Coverage for panels (e) and (f) has been calculated relative to the surface relief ETOPO5 dataset (NGDC, 1986), averaged to the native grid of each humidity dataset with ocean gridboxes defined as those with negative average surface elevation on the native grid. Latitude weighting has been applied to panel (e).

of thermal problems with the instrument, explaining the missing data from GSSTF2c and lower coverage in HOAPS3.2 for this month. The other datasets have later start dates, for example J-OFURO2 starts in January 1988. GSSTF3 has missing data in tropical regions with very high precipitation (Figure 2(a)); the reason for this missing Tropical data is not clear. J-OFURO2 uses a climatological sea-ice mask (Figure 2(c), (e) and (f)). HOAPS3.2, IFREMER and SeaFluxV1.0 all use a time-varying sea-ice mask based on SSM/I (Figure 2(b) and (d)–(f)). Figure 2(e) shows an increase in the extent of humidity data for the IFREMER dataset over the course

of 1999 that is due to a change in the land mask which occurred in August 1999. The IFREMER dataset also suffers from occasional data loss, particularly in the North Sea (Figure 2(b)), for reasons unknown.

Figure 3 shows monthly mean humidity from each of the datasets in January 2000 at their native spatial resolution over the Northeast Atlantic and Norwegian Sea. The different approaches taken in the different datasets to masking is very clear. The conservative land and ice-masks used in the Goddard datasets (Figure 3(a) and (b)) excludes data north of about 70°N that is unlikely to be affected by sea-ice (compare for example with Figure 3(c)). The HOAPS3.2

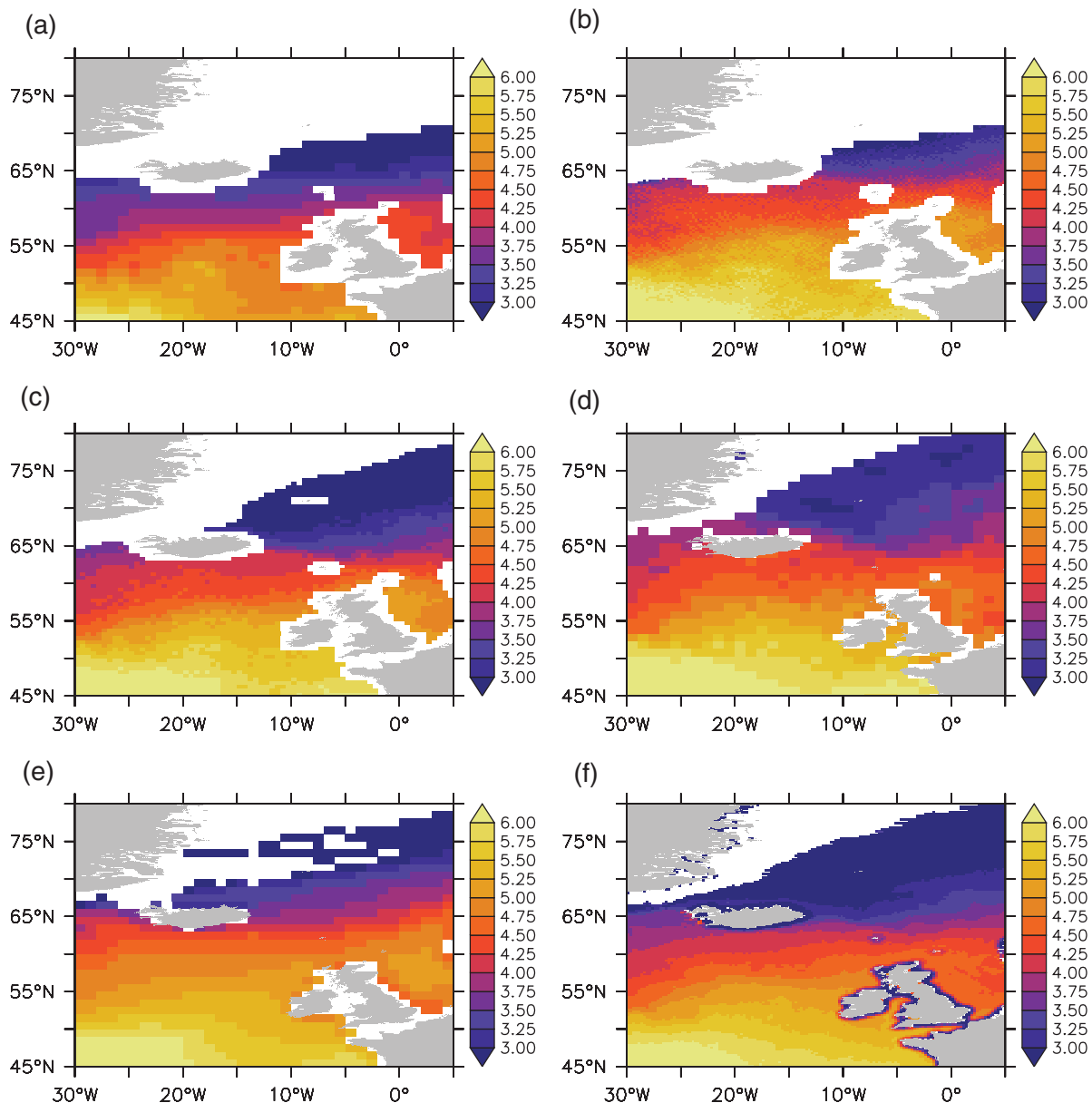


Figure 3. Specific humidity January 2000 (g kg^{-1}). (a) GSSTF2c; (b) GSSTF3; (c) HOAPS3.2; (d) IFREMER; (e) J-OFURO2; (f) SeaFluxV1.0.

land and sea-ice mask (Figure 3(c)) is also relatively conservative, but because it is time-varying excludes less data than the Goddard datasets. The IFREMER data are available closer to the coast and ice-edge than HOAPS3.2, in the latter the masked region is extended by 50 km (Section 2.4.3.). The IFREMER dataset sometimes accepts data in regions between the sea-ice and land, shown in this sample month in Figure 3(d). The J-OFURO2 climatological ice mask is likely to poorly represent the true ice edge in this particular month (Figure 3(e)). The SeaFluxV1.0 dataset contains data much closer to the coast than the other datasets and, like the IFREMER dataset, contains some data near the landward edge of the sea-ice (Figure 3(f)). There is a strip of lower humidity data around the coasts in the SeaFluxV1.0 dataset which, although lower humidity values are expected close to land, may not be reliable due to its proximity to the coast. This is in contrast with

the q_a values from J-OFURO2 which increase close to the eastern coast of the UK (Figure 3(e)).

3.2. Effects of different data sources

The SSM/I data used in the different products derives from different sources (described in Section 2.1) and may have been adjusted to account for differences observed between the measurements from different satellites (described in Section 2.2). The impact of the use of different input data on the calculated humidity can be examined by comparing three datasets which all use the same algorithm (B03): these are IFREMER, HOAPS3.2 and GSSTF3 (Figure 1). We note that different QC has been used in these datasets and the possible effect of differences in QC is examined in Section 3.4. All have been averaged onto a 1° grid and have the conservative GSSTF2c land and ice mask applied in addition to the datasets'

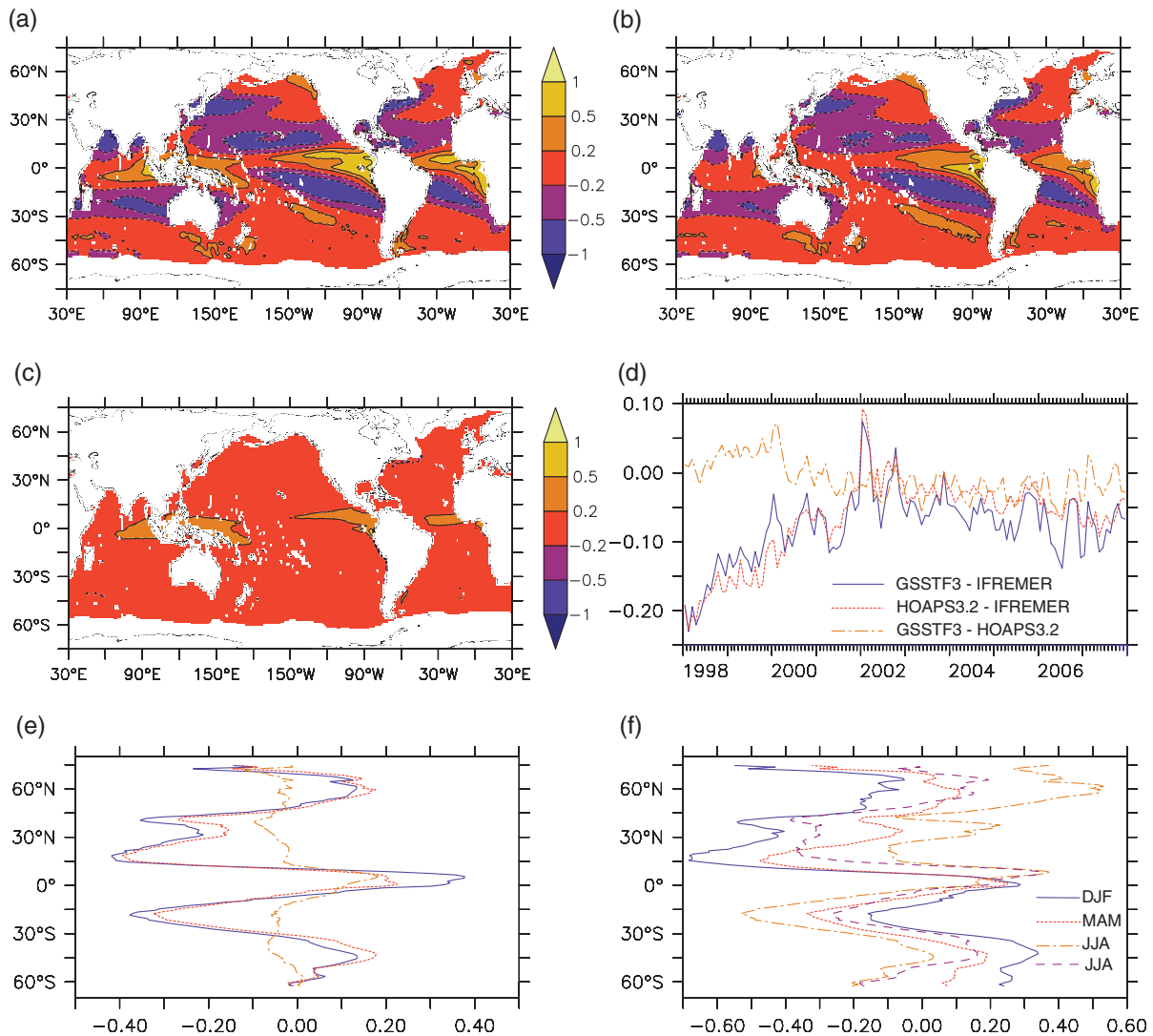


Figure 4. Differences in q_a (g kg^{-1}) for the three humidity products using the Bentamy *et al.* (2003) algorithm, for the 10-year period January 1998–December 2007. (a) 10-year mean, GSSTF3-IFREMER; (b) as (a) but for HOAPS3.2-IFREMER; (c) as (a) but for GSSTF3-HOAPS3.2; (d) global monthly mean GSSTF3-IFREMER (solid line), HOAPS3.2-IFREMER (dotted line), GSSTF3-HOAPS3.2 (dot-dash line); (e) as (d) but for zonal averages; (f) seasonal zonal means for HOAPS3.2-IFREMER, DJF (solid line), MAM (dotted line), JJA (dot-dash line), SON (dashed line).

own masks. Figure 4 compares the humidity fields from each of these products and Table 3 provides some summary statistics of the differences among all of the satellite datasets, calculated using R (R Core Team, 2012) and utilizing the ncdf library (Pierce, 2011). HOAPS3.2 and GSSTF3 both use data following inter-satellite calibration. For GSSTF3 Goddard use RSSv6 data (Wentz, 2010) and apply further processing following Shie and Hilburn (2011) to adjust for the effect of changing EIA. HOAPS apply a simpler form of inter-satellite calibration that does not explicitly correct for EIA as part of their SSM/I data processing stream (Andersson *et al.*, 2010; Fennig *et al.*, 2013). Note that the B03 q_{10n} retrieval algorithm (and the IFREMER dataset) was developed using brightness temperatures without inter-satellite calibration applied. The two datasets with inter-satellite calibration are quite similar (Figure 4(c)–(e)) with a median difference of -0.01 g kg^{-1} ; 50% of the monthly mean 1°

differences, GSSTF3 – HOAPS3.2 fall in the range -0.1 to 0.08 g kg^{-1} (Table 3). Variations in the global mean difference over the 10-year period are small (Figure 4(d)). For the B03 algorithm as applied by HOAPS, any residual biases resulting from imperfect removal of EIA-related error does not appear to affect the global mean. The differences are largest in the Tropics (Figure 4(c) and (e)), the reason for this is unclear but may be related either to EIA or to the QC applied.

Differences of GSSTF3-IFREMER and HOAPS3.2-IFREMER are similar (Figure 4(a), (b), (d) and (e)) and show strong spatial coherence. These patterns remain fairly similar over time, despite the change in mean offset over time (Figure 4(d)). Median differences are $\sim 0.1 \text{ g kg}^{-1}$ (IFREMER slightly higher q_a than either HOAPS3.2 or GSSTF3) and 50% of the differences fall within the approximate range -0.4 to 0.2 g kg^{-1} (Table 3). There are strong seasonal differences between HOAPS3.2

Table 3. Summary statistics for percentiles of monthly mean q_a differences between pairs of satellite datasets calculated over their common period (see Table 1).

Data differences (g kg^{-1})	1%	10%	25%	Median	75%	90%	99%
GSSTF3 – HOAPS3.2	-0.48	-0.20	-0.10	-0.01	0.08	0.20	0.58
GSSTF3 – IFREMER	-1.28	-0.73	-0.41	-0.07	0.23	0.50	1.21
GSSTF3 – J-OFURO2	-1.82	-1.23	-0.78	-0.29	0.06	0.33	0.81
GSSTF3 – SeaFluxV1.0	-1.86	-1.10	-0.57	-0.07	0.28	0.58	1.18
HOAPS3.2 – IFREMER	-1.20	-0.67	-0.37	-0.05	0.22	0.44	0.94
HOAPS3.2 – J-OFURO2	-2.01	-1.35	-0.81	-0.27	0.09	0.37	0.88
HOAPS3.2 – SeaFluxV1.0	-1.69	-1.10	-0.64	-0.07	0.30	0.62	1.22
IFREMER – J-OFURO2	-2.42	-1.40	-0.66	-0.15	0.27	0.64	1.30
IFREMER – SeaFluxV1.0	-1.76	-1.03	-0.51	-0.06	0.30	0.65	1.67
J-OFURO2 – SeaFluxV1.0	-1.13	-0.49	-0.16	0.15	0.45	0.79	1.56

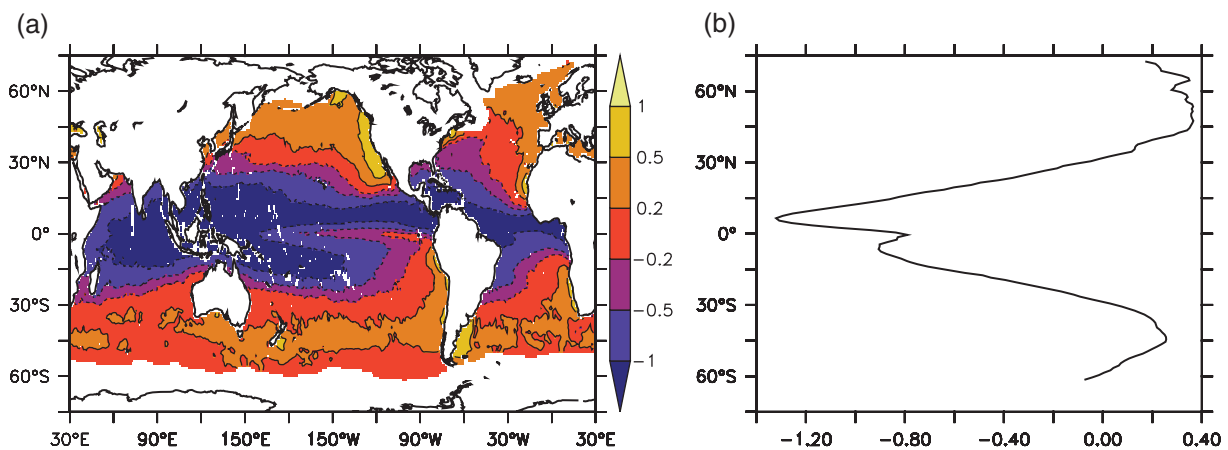


Figure 5. Difference between q_a (g kg^{-1}) calculated from RSSv6 F11 brightness temperatures for 1996 using B03 and S95. Both have large droplet QC applied following S93. (a) Spatial map of mean differences (b) zonal mean differences.

and IFREMER q_a (Figure 4(f)). In the annual zonal mean the IFREMER product is drier near the equator and more humid in a band between 15° and 45° latitude in both hemispheres (Figure 4(e)). Differences poleward of 45° are smaller, as differences change sign with the seasons (Figure 4(f)). In the global mean the effect of the inter-satellite calibration is to reduce the humidity, by about 0.2 g kg^{-1} in 1998, reducing approximately linearly to a small global mean value by the start of 2002 (Figure 4(d)).

3.3. Differences between the algorithms

Four different algorithms are used in the seven different datasets compared (Table 1). For 1996 only we have used RSSv6 data from the F11 satellite to construct humidity estimates using the algorithms of S95 and B03 (Figure 5). The B03 algorithm is drier than S95 over much of the Tropics (except notably the tropical and subtropical South Eastern Pacific), and moister over much of the extratropical regions.

Figure 6 compares the two algorithms in ranges of specific humidity estimates from the NOCv2.0 dataset. The median difference between specific humidities from the S95 and B03 algorithms is -0.21 g kg^{-1} and 50% of the 1° monthly differences fall within the range -1.06 to 0.20 g kg^{-1} . At low specific humidities there is much

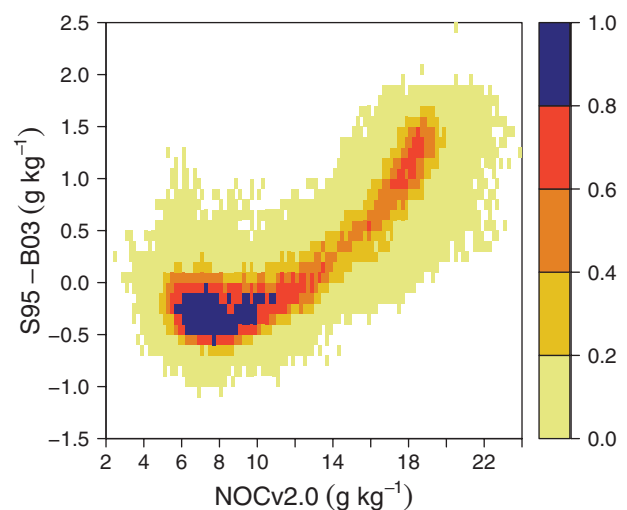


Figure 6. Density plot of difference among 1° area monthly mean q_a (g kg^{-1}) from RSSv6 F11 brightness temperatures for the S95 and B03 algorithms plotted against q_a from NOCv2.0 for 1996 (g kg^{-1}). The q_a algorithms have large droplet QC applied following S93. Each of the five colours represents approximately 20% of the total data points.

better agreement between S95 and B03 than at higher specific humidities where S95 values are typically more than 1 g kg^{-1} higher than B03.

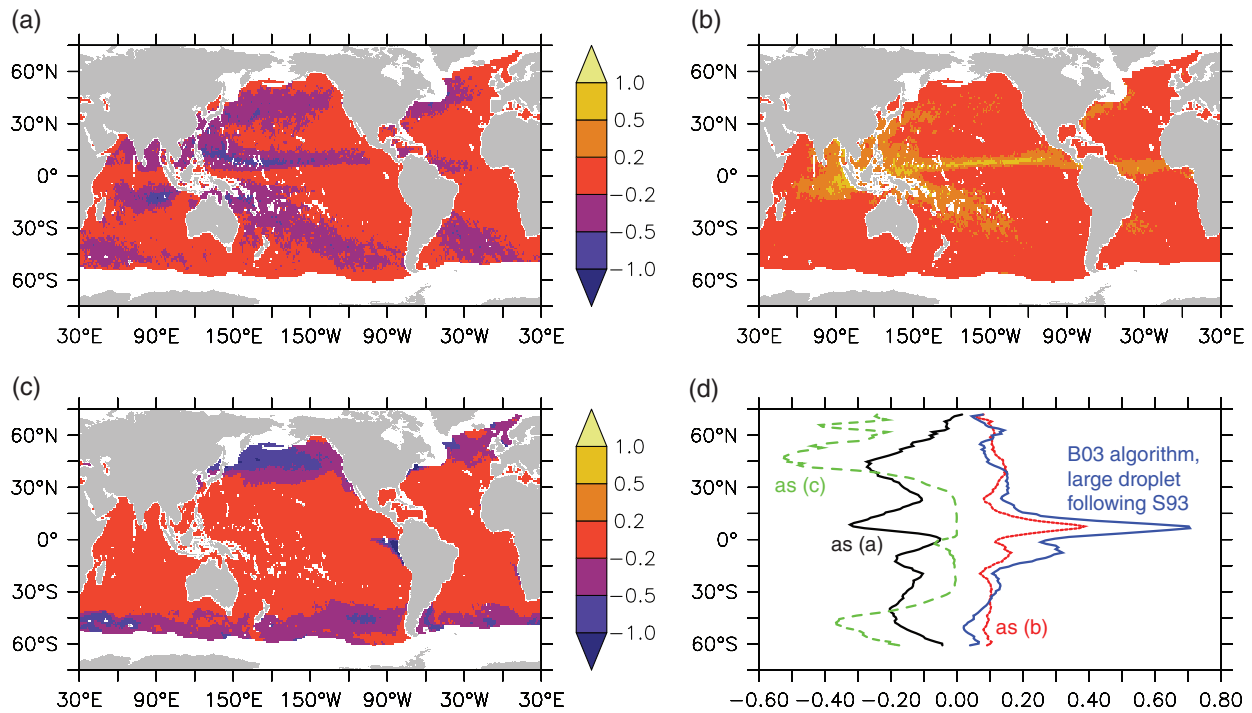


Figure 7. Effect of applying quality control procedures to global q_a estimated from the linear q_a retrieval algorithms. Differences are with QC – without QC, averaged for the year 1996 and derived using RSSv6 F11 brightness temperatures. (a) S95 algorithm, large droplet QC following S93. (b) B03 algorithm, large droplet QC following HOAPS3.2. (c) B03 algorithm, humidity inversion QC. (d) Zonal differences due to QC.

The two algorithms differ in their selection of channels and also differ strongly in the weight given to the different channels (Table 2). This is likely to be a result of multicollinearity among the predictor variables, as there are strong correlations between the brightness temperatures in each pair of channels (r^2 greater than 0.8 in the sample year of F11 data used here). Multicollinearity means that fitted coefficients can vary unpredictably with any changes in the training data used, and results may not be reliable when applied to other data. In the algorithms used here, this feature manifests as the relatively strong dependence of the S95 algorithm on the 19 V channel, compared with the more evenly weighted channels in the B03 algorithm.

3.4. Effect of QC

The different datasets apply different QC procedures to their datasets (Table 1). Which procedures are applied to each dataset is not always apparent from the dataset documentation (Section 2.4). The impact of these procedures can be significant, causing regional variations in the annual mean that can exceed $\pm 1 \text{ g kg}^{-1}$ (Figure 7) and larger variations in individual monthly means (not shown). The effect of application of a particular QC also varies in both strength and sign depending on the underlying algorithm. For example, applying QC to exclude situations with high-liquid water contents and large droplets following S93 (hereafter ‘large droplet QC’) leads to a drying in Tropical and mid-latitude rainbands when applied to the S95 algorithm (Figure 7(a)), but causes a moistening effect in the same regions when applied to the B03

algorithm (HOAPS3.2 version, Figure 7(b)). The reason for the different effect of QC on the different algorithms is unclear but may result from the statistical rather than physical derivation of the algorithms. This means that the effectiveness of any QC applied may influence the sign of any bias under conditions with high liquid water content. The modification of the S93 large droplet QC as applied by HOAPS3.2 reduces the impact of QC (Figure 7(d)). Application of the liquid water contamination QC applied by Bentamy *et al.* (1999) in addition to the large droplet QC of S93 makes only a small difference (not shown). Limiting the maximum value of q_a to be less than q_s , (Chou *et al.*, 1997, here after ‘humidity inversion QC’) applied in the Goddard and HOAPS3.2 datasets has a similar drying effect at higher latitudes on q_a derived from either the B03 (Figure 7(c) and (d)) and S95 algorithms (not shown).

It is expected that excluding periods with high cloud liquid water or rain rates will lead to fair-weather biases in the humidity products relative to datasets produced using other methods, due to the removal of periods more likely to have higher humidities.

3.5. Comparison of GSSTF2b, GSSTF 2c and GSSTF 3 datasets

The GSSTF product has recently gone through several significant revisions. A comparison of the three most recent versions (2b, 2c and 3) is shown in Figure 8. The algorithm change from version 2c to 3 is illustrated in Figure 8(a) (using 2003 as an example) and shows a reduction in q_a in the Tropics in the later version and an increase at higher latitudes. The differences caused by the input

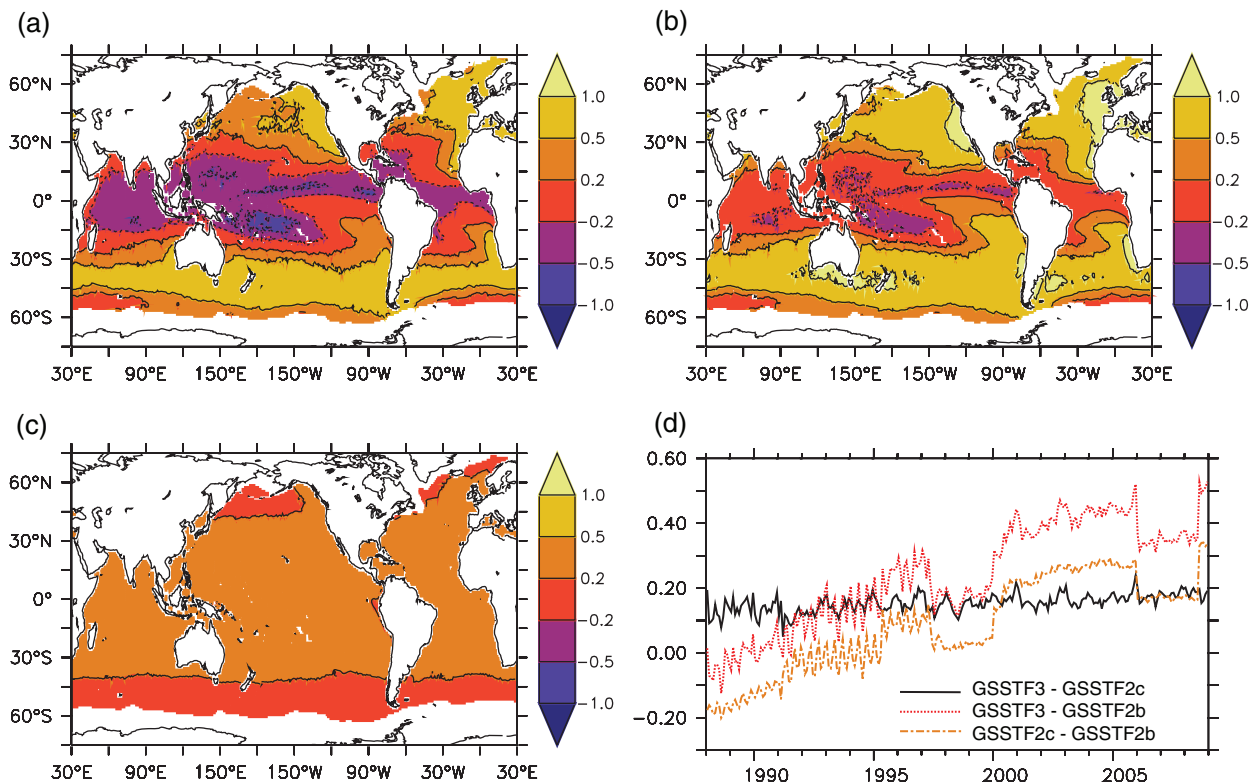


Figure 8. Mean global differences (g kg^{-1}) between iterations of the GSSTF q_a dataset for the calendar year 2003. (a) GSSTF3-GSSTF2c; (b) GSSTF3-GSSTF2b; (c) GSSTF2c-GSSTF2b; (d) global mean: solid line = GSSTF3-GSSTF2c, dotted line = GSSTF3-GSSTF2b, dot-dash line = GSSTF2c-GSSTF2b.

data have a smaller impact in this period (Figure 8(b) and (c)). Figure 8(d) shows clear trends and step changes in the difference between GSSTF2b and GSSTF2c over time, caused by the corrections for changing EIA applied (Hilburn and Shie, 2011). Differences between GSSTF2c and GSSTF3 show smaller variations with time that are due to the algorithm change.

3.6. Comparison of satellite-derived near-surface specific humidity products

Figure 9 shows selected monthly mean differences between the seven satellite-based q_a products plotted against q_a from NOCv2.0 for 2001 and Figure 10 shows global maps of these differences. GSSTF3 and HOAPS3.2 show relatively small differences (Figures 9(a) and 10(a)) as also seen in Figure 4(c). Both datasets use the B03 algorithm and use brightness temperatures that have some form of inter-satellite calibration applied. The differences between IFREMER and HOAPS3.2 are larger (Figures 9(b) and 10(b)) as seen in Figure 4(b). As both datasets use B03 the differences seen are likely to be due to a combination of the effects of inter-satellite calibration on the brightness temperatures and different approaches to QC. For example, HOAPS3.2 applies a humidity inversion QC, whereas IFREMER does not. This may contribute to the higher IFREMER humidities seen in the lower humidity range (Figure 9(b)) although the picture is not simple as IFREMER has lower humidities on average in the Gulf of Alaska (Figure 10(b)).

Figure 9(c) shows the difference between J-OFURO2 and HOAPS3.2. The effect of the humidity inversion QC (applied by HOAPS3.2 but not by J-OFURO2) is clear at q_a below about 12 g kg^{-1} . Aside from this effect, the form of the differences are similar to that expected from the difference between the S95 algorithm (J-OFURO2) and B03 (HOAPS3.2) shown in Figure 6. SeaFluxV1.0, like J-OFURO2, shows higher q_a than HOAPS3.2 at high NOCv2.0 q_a (Figures 9(d) and 10(d)). However the patterns of humidity difference are different (compare Figure 10(c) and (d)). The dryness of SeaFluxV1.0 q_a relative to HOAPS3.2 seen at moderate q_a values (Figure 9(d)) is due to differences in localized regions, e.g. regions of strong negative differences (Figure 10(d)) to the West of Africa, Mexico, Ecuador and off the North West coast of Australia. SeaFluxV1.0 is drier than HOAPS3.2 between NOCv2.0 q_a of $6\text{--}10 \text{ g kg}^{-1}$ (Figure 9(d)), noticeable especially in the Kuroshio and Gulf Stream regions (Figure 10(d)).

GSSTF2c humidities (Chou *et al.*, 1997 algorithm) are lower than those from GSSTF3 (B03) at low humidity values (Figures 9(e) and 10(e)), the reverse is true at higher humidities. The difference between J-OFURO2 and IFREMER (Figures 9(f) and 10(f)) is similar to the J-OFURO2 – HOAPS3.2 differences (Figures 9(c) and 10(c)) but noisier and with a more prominent structure of rainbands. Although the SeaFluxV1.0 and J-OFURO2 differences show relatively little systematic variation in Figure 9(h) (excepting the QC difference seen in e.g.

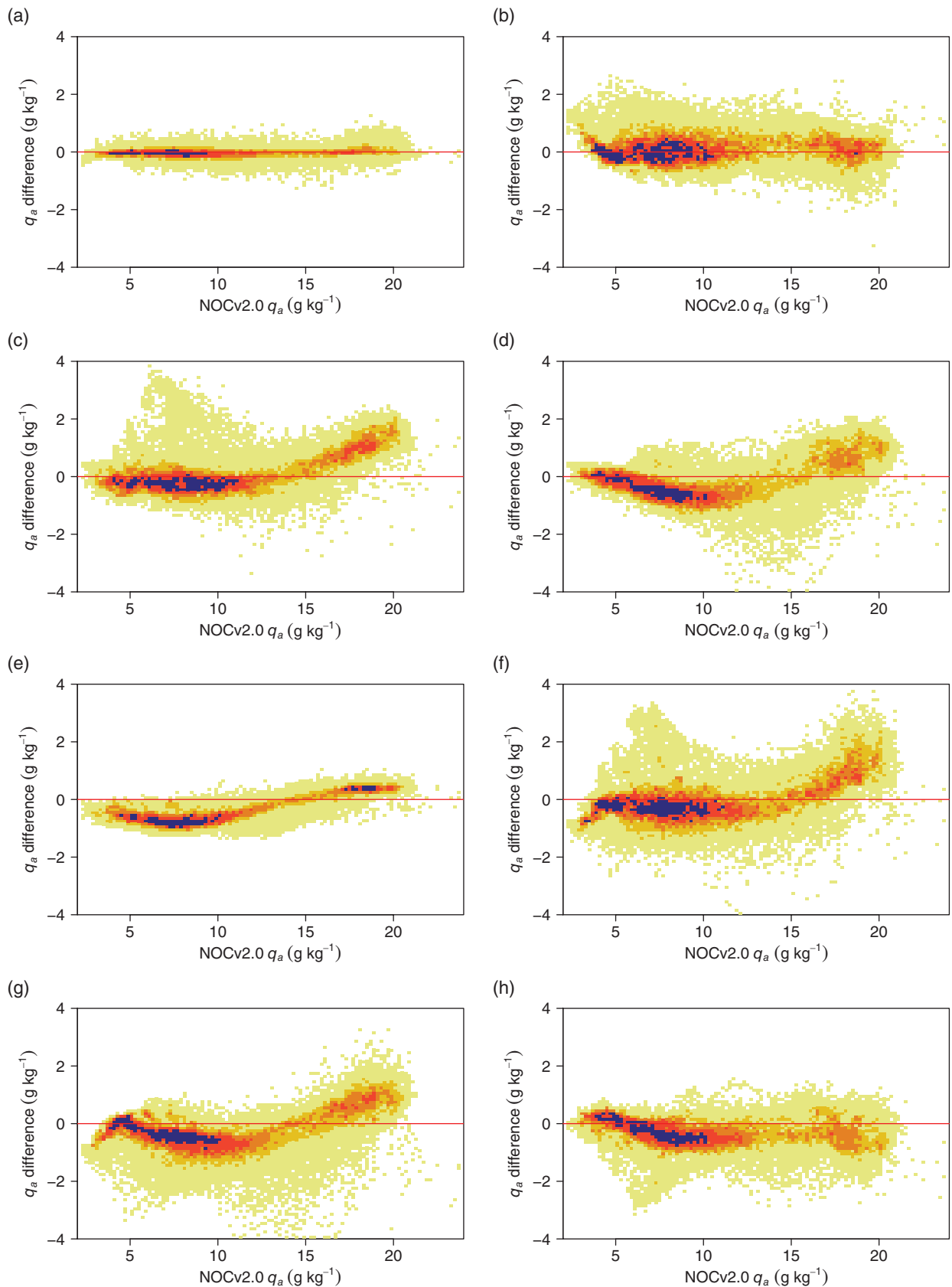


Figure 9. Density plot of differences between q_a from the satellite-derived products, relative to NOCv2.0 (x-axis of all plots). Differences are from 2001 and restricted to data with estimated total error in NOCv2.0 $< 0.3 \text{ g kg}^{-1}$. Colour scale and ranges as for Figure 6. (a) GSSTF3-HOAPS3.2. (b) IFREMER-HOAPS3.2. (c) J-OFURO2-HOAPS3.2. (d) SeaFluxV1.0-HOAPS3.2. (e) GSSTF2c-GSSTF3. (f) J-OFURO2-IFREMER. (g) SeaFluxV1.0-IFREMER. (h) SeaFluxV1.0-J-OFURO2.

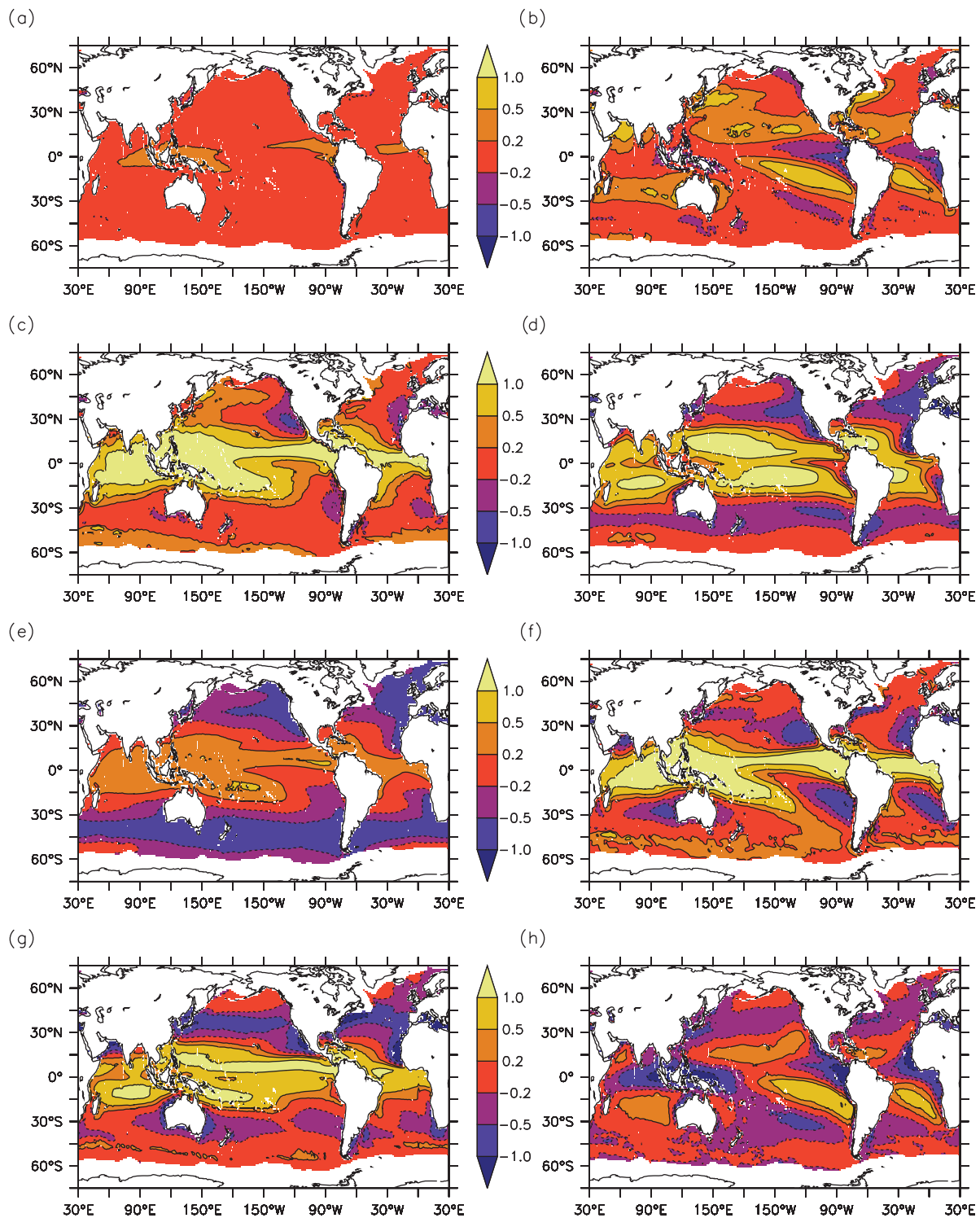


Figure 10. Mean global differences between selected satellite q_a products (g kg^{-1}) averaged over the lifetime of the respective datasets (Table 1). (a) GSSTF3-HOAPS3.2. (b) IFREMER-HOAPS3.2. (c) J-OFURO2-HOAPS3.2. (d) SeaFluxV1.0-HOAPS3.2. (e) GSSTF2c-GSSTF3. (f) J-OFURO2-IFREMER. (g) SeaFluxV1.0-IFREMER. (h) SeaFluxV1.0-J-OFURO2.

Figure 9(c)) the spatial patterns do show significant spatial variability (Figure 10(h)).

In Figure 11 the same differences shown in Figures 9 and 10 are plotted as annual mean differences in three latitude bands (75°N - 25°N , 25°N - 25°S , 25°S - 75°S) and also for 75°N - 75°S . Time series of GSSTF3 and HOAPS3.2 are the most similar of any of the datasets

compared (Figure 11(a)). Differences between IFREMER and HOAPS3.2 change over time, the start of the change seems to be about the same time as the change in the IFREMER land mask noted in Section 3.1 (Figure 11(b)). These differences seem likely to be due to the use of data without inter-satellite calibration by IFREMER. J-OFURO2 dries relative to HOAPS3.2 over time (Figure 11(c)). Although

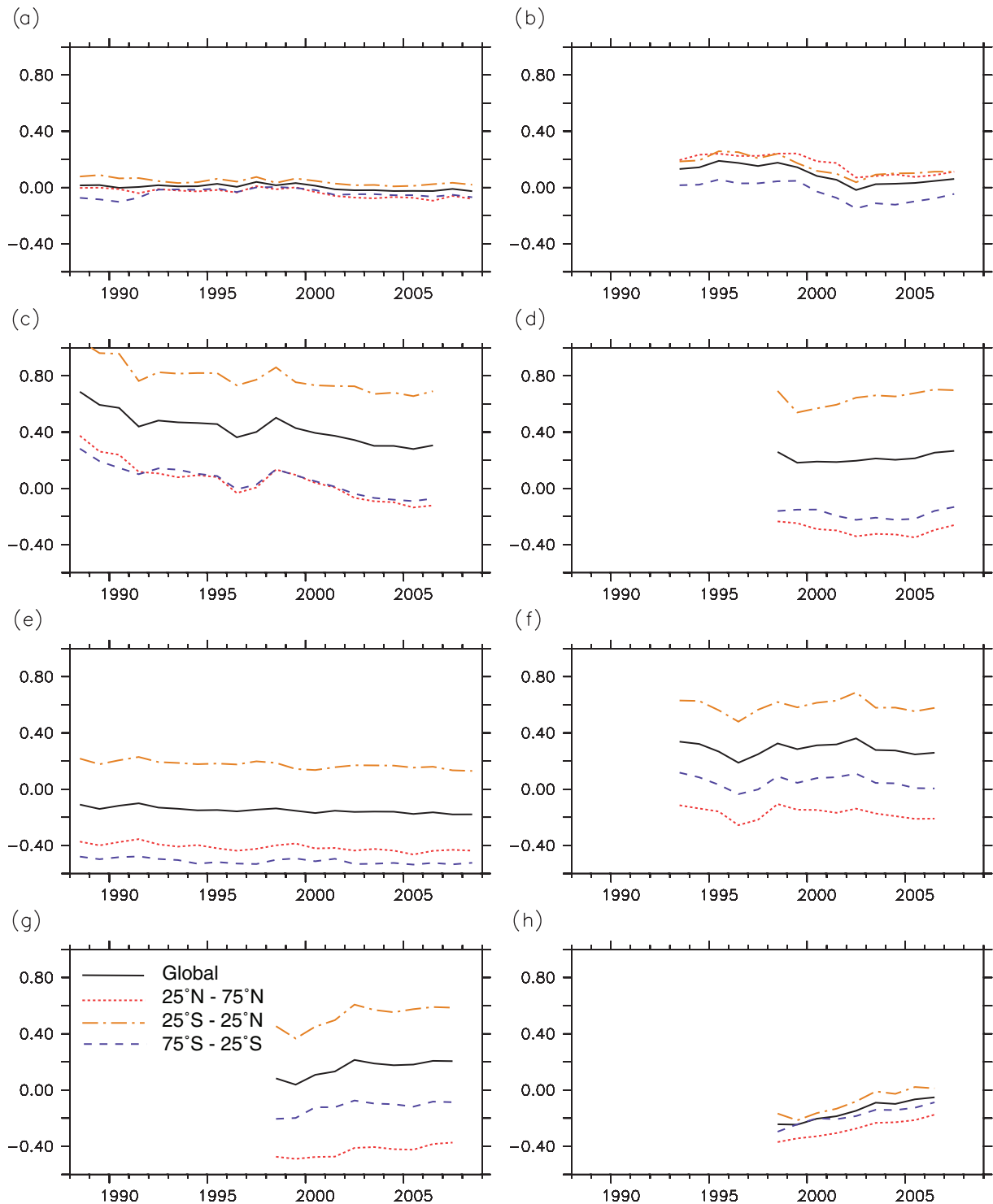


Figure 11. Time series of mean q_a differences (g kg^{-1}) between selected satellite-derived datasets for different regions. Solid line = global; Dotted line = $25^\circ\text{N}-75^\circ\text{N}$; Dot-dash line = $25^\circ\text{S}-25^\circ\text{N}$; Dashed line = $25^\circ\text{S}-75^\circ\text{S}$. (a) GSSTF3-HOAPS3.2. (b) IFREMER-HOAPS3.2. (c) J-OFURO2-HOAPS3.2. (d) SeaFluxV1.0-HOAPS3.2. (e) GSSTF2c-GSSTF3. (f) J-OFURO2-IFREMER. (g) SeaFluxV1.0-IFREMER. (h) SeaFluxV1.0-J-OFURO2.

the differences are not the same magnitude as the differences between GSSTF2b and GSSTF2c (Figure 8) the timings of changes are similar, suggesting that the relative trend between the two datasets is due to problems with RSSv6 identified by Hilburn and Shie (2011).

The most similar pair of datasets is GSSTF3 and HOAPS3.2 (Table 3), which share a common algorithm

and are both constructed from brightness temperatures following inter-satellite calibration (Section 3.2). The largest median and range of differences are seen for comparisons with the J-OFURO2 dataset. J-OFURO2 is the only dataset to use the S95 algorithm, does not apply the humidity inversion QC (see e.g. Figure 9(c)) and will also be affected by biases due to

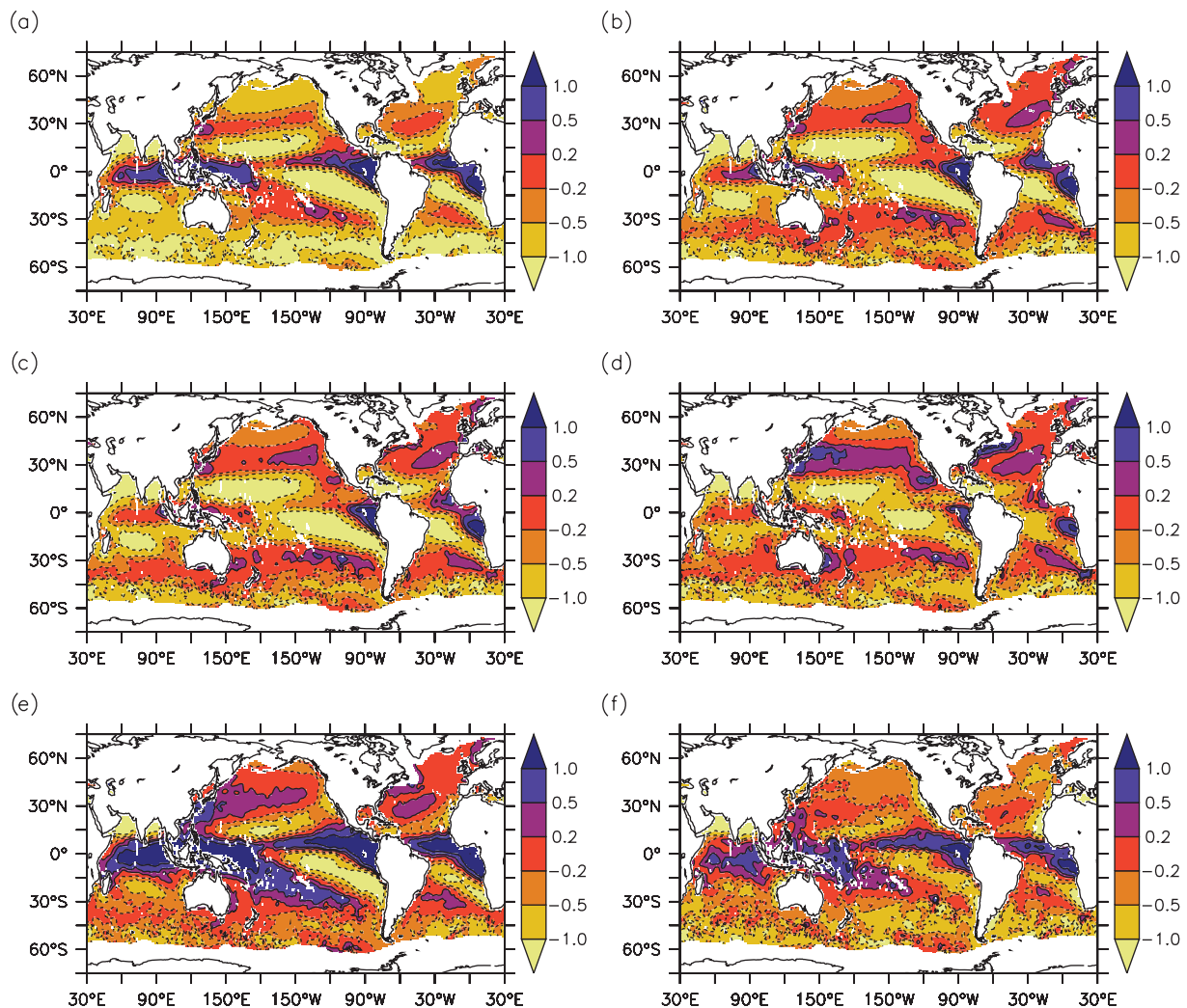


Figure 12. Difference (g kg^{-1}) between satellite products and NOCv2.0 (no uncertainty limit applied), averaged over the lifetime of the respective satellite datasets (Table 1). (a) GSSTF2c-NOCv2.0. (b) GSSTF3-NOCv2.0. (c) HOAPS3.2-NOCv2.0. (d) IFREMER-NOCv2.0. (e) J-OFURO2-NOCv2.0. (f) SeaFluxV1.0-NOCv2.0.

EIA in RSSv6 brightness temperatures (Hilburn and Shie, 2011).

Overall Figures 9–11 show systematic differences between the datasets that depend on source data, QC, the approach taken to gap-filling and interpolation and on algorithmic differences. These cannot be disentangled in all cases, but the range of choices of these three factors made in the generation of the different datasets does allow the identification of the importance of each of them on the characteristics of the final specific humidity dataset.

3.7. Comparison of satellite and *in situ*-derived near-surface q_a products

Figure 12 shows maps of the differences between the satellite-derived specific humidity datasets and estimates from the *in situ* NOCv2.0 dataset. As expected from the previous comparisons there are some large variations among the satellite–*in situ* humidity differences that vary among the satellite humidity datasets. Overall however, the spatial patterns show some consistency. Patterns for GSSTF3, HOAPS3.2 and IFREMER (all using

the B03 algorithm) are similar, as are J-OFURO2 and SeaFluxV1.0 (using the S95 and Roberts *et al.*, 2010 algorithms, respectively). Median climatological monthly mean differences typically show the satellite datasets are drier than NOCv2.0.

Differences will also be expected because NOCv2.0 presents q_{10} whereas the datasets that use either the B03 or Roberts *et al.* (2010) algorithms present q_{10n} (i.e. GSSTF3, HOAPS3.2, IFREMER and SeaFluxV1.0). Conversion of the NOCv2.0 q_{10} to q_{10n} does not improve the comparisons between the *in situ* and satellite humidity estimates while adding substantially to the noise in the *in situ* humidities. Figure 13 illustrates this using climatological monthly mean differences between both SeaFluxV1.0 and HOAPS3.2 and NOCv2.0, the latter with and without adjustment to q_{10n} . For both satellite datasets the mean difference reduces slightly when compared to NOC q_{10n} , but in both cases the distribution of differences actually broadens, most noticeably so for the SeaFluxV1.0 dataset.

Any biases, or noise, in the NOCv2.0 specific humidities will appear across all the panels of Figure 12. However,

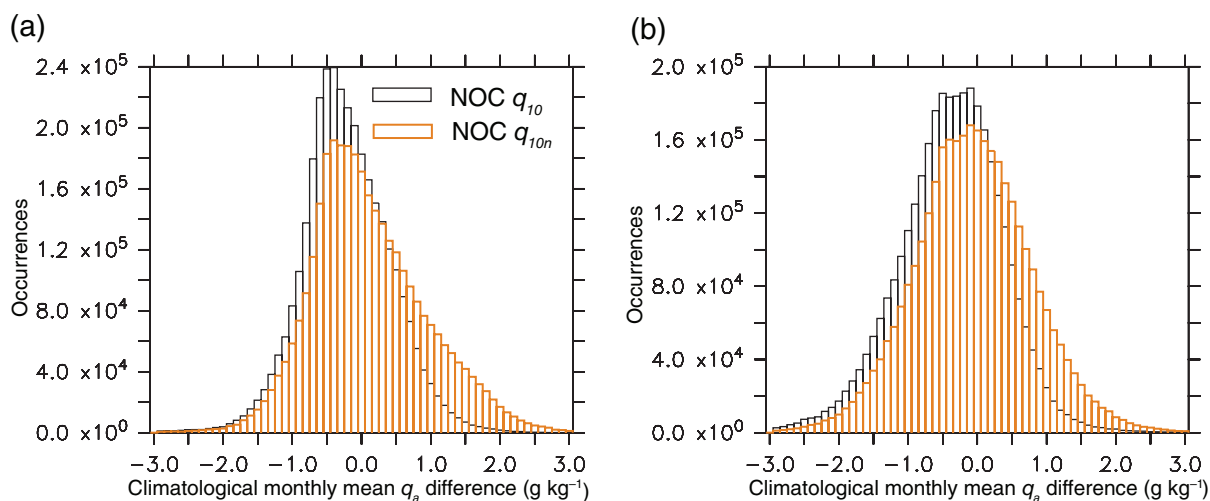


Figure 13. Difference (g kg^{-1}) between climatological monthly mean (1998–2006) satellite products and NOCv2.0 (no uncertainty limit applied) for (a) SeaFluxV1.0 and (b) HOAPS3.2. NOC q_{10n} has been calculated similarly to NOCv2.0 q_{10} (Berry and Kent, 2011).

the patterns of difference do not seem to be related to any likely source of bias in the NOCv2.0 dataset, or to atmospheric stability, but rather to other atmospheric conditions. Biases in NOCv2.0 q_{10} might be expected to show variations with measurement height (Kent *et al.*, 2007); measurement method (Berry and Kent, 2011) or show the characteristic patterns of shipping lanes (Berry and Kent, 2009). The differences shown in Figure 12 (and in Figure 4 of Andersson *et al.* (2010) which also compares HOAPS3 with ERA-Interim (Dee *et al.*, 2011)) are more similar to wind speed biases thought by Atlas *et al.* (2011) to be ‘likely related to persistent atmospheric or oceanic conditions affecting the microwave remote sensing of the ocean surface’. Similar spatial patterns, for the region 38°S to 38°N , are also identified by Liu and Zipser (2013) as having different gradients in the vertical profile of radar reflectivity at 13.4 GHz which they relate to characteristics of the rainfall and convection.

Figure 14 shows the difference between the satellite humidity estimates and NOCv2.0, plotted against NOCv2.0 specific humidities in four sample regions. We have limited the comparison to include only those 1° monthly grid points where the NOCv2.0 uncertainty estimate is $\leq 0.3 \text{ g kg}^{-1}$. This ensures that the grid point is well sampled and any effects of the optimal interpolation procedure should be small. Each region shows a different relationship between satellite and *in situ* q_a that is more similar between the different products than between different regions. This is consistent with the persistent patterns of difference between satellite and *in situ* q_a seen in Figure 12.

Figure 15 shows annual mean q_a from each of the satellite-based datasets for the period 1988–2008 in three wide latitude bands and globally. The estimates show little consistency in terms of mean values, interannual variability or trends. Whereas most of the reanalysis and *in situ*-based datasets compared by Kent *et al.* (2014) showed an increase in q_a over this period, the satellite-based datasets typically decrease or remain similar. The largest

decrease is seen in GSSTF2b (Figure 8(d)) and has already been identified as an artefact (Shie, 2011). Over the short period of these datasets, ranging in length from about 10 to 20 years, it is not possible to draw any firm conclusions about the significance of any trends.

4. Discussion and conclusions

Over the last 30 years there have been numerous attempts to derive relationships between quantities that can be measured from space and near-surface q_a over the global ocean. The motivation is usually to derive estimates of the air–sea flux of latent heat (e.g. Bourassa *et al.* (2010)), and the most commonly used satellite dataset is from the SSM/I instrument flown on the DMSP series of satellites.

SSM/I data is available from several different sources. All of the datasets analysed here derive from data sent directly from the satellites to the FNMOC and based on the results of Ritchie *et al.* (1998) it is usually concluded that differences between this data stream and the same data collected and archived by AFGWC are unimportant. Further processing, and in particular inter-satellite calibration of the SSM/I data, takes place at CSU, HOAPS, RSS and NASA (Goddard Space Science Laboratory). IFREMER also keep an archive of SSM/I data. Currently only CSU and HOAPS provide calibration information for the different satellites and channels that would allow the dataset user to directly quantify the impact of the adjustments.

Any algorithm developed for q_{10} will have characteristics that depend on the satellite and *in situ* data used to derive it, any adjustments that have been applied to one of the data source, and the periods and regions that those data cover. If an algorithm is applied to a brightness temperature dataset with different characteristics to that used to derive it, then the resulting q_a estimates may be biased or noisy compared to q_a derived from the same brightness temperature dataset. SeaFluxV1.0 is the only dataset

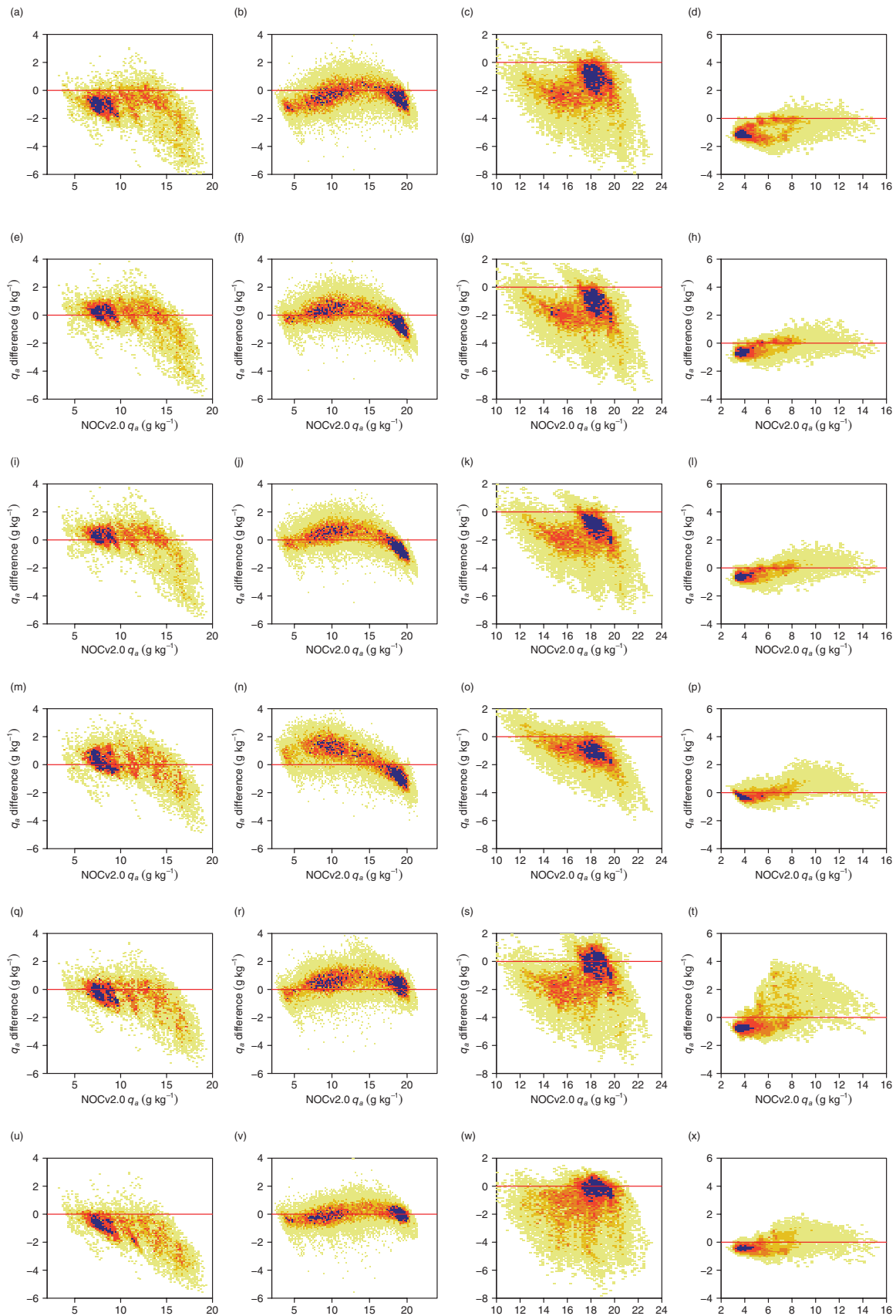


Figure 14. Density plots of regional satellite-NOCv2.0 q_a differences (g kg^{-1} , y-axis) against NOCv2.0 (x -axis) for 1° monthly gridboxes during 1998–2006 and where the total uncertainty estimate for NOCv2.0 $q_a < 0.3 \text{ g kg}^{-1}$. Top row (a–d): GSSTF2c; 2nd row (e–h): GSSTF3; 3rd row (i–l): HOAPS3.2; 4th row (m–p): IFREMER; 5th row (q–t): J-OFURO2; bottom row (u–x): SeaFluxV1.0. Left column: 30°N – 45°N , 0°E – 35°E (Mediterranean); 2nd column: 25°N – 40°N , 120°E – 150°E (Northwest Pacific); 3rd column: 10°N – 30°N – 50°E : 80°E (Arabian Sea); Right column: 40°N – 50°N – 160°E : 180°W (North Pacific). Colour scale and ranges as for Figure 6.

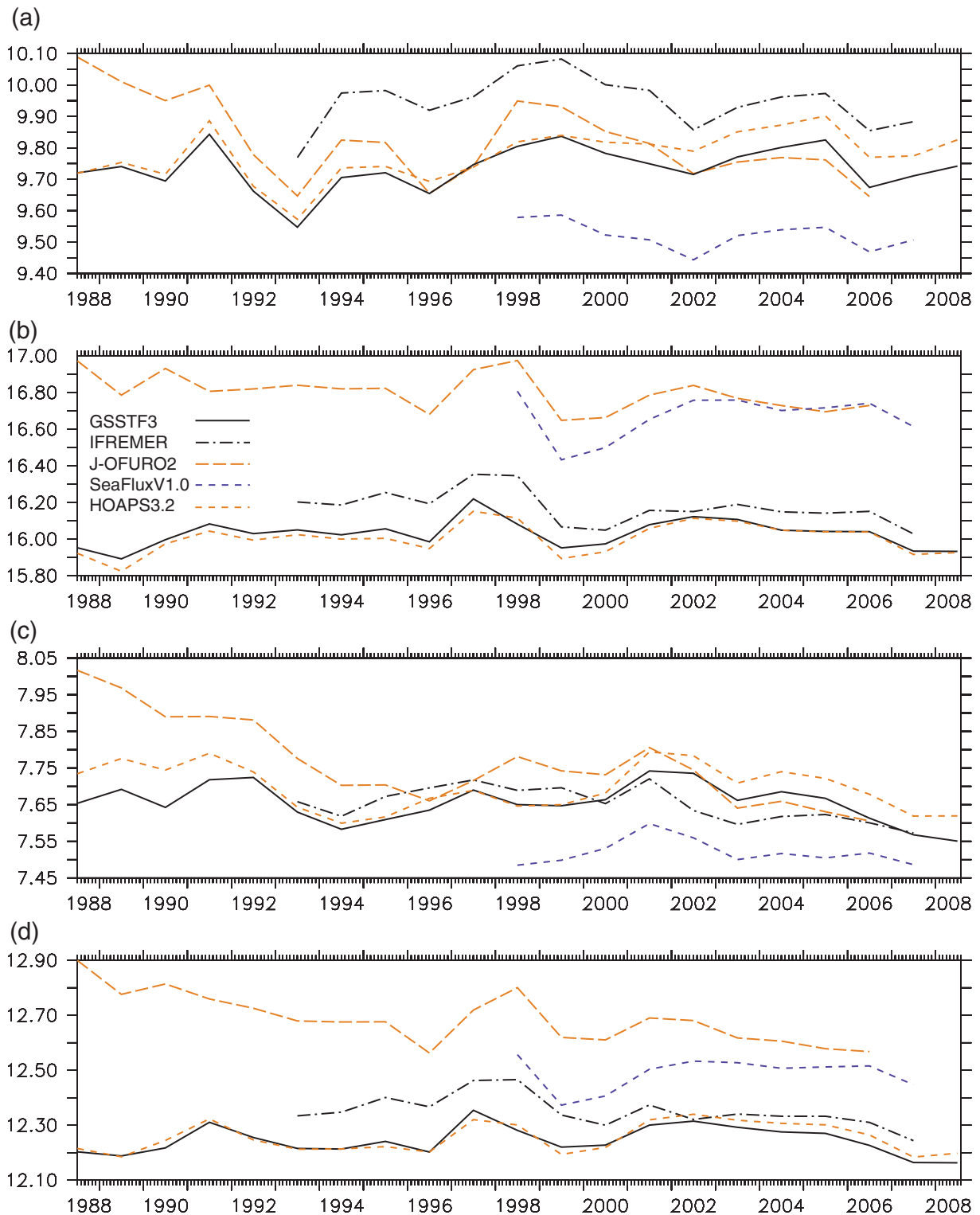


Figure 15. Regional and global time series for the satellite-derived q_a estimates (g kg^{-1}). (a) 25°N - 75°N ; (b) 25°S - 25°N ; (c) 25°S - 75°S ; (d) 75°S - 75°N . Solid black = GSSTF3, Black dot-dash = IFREMER, dot-dash (Orange in online) = J-OFURO2, dashed (Blue in online) = SeaFluxV1.0, dashed (Orange in online) = HOAPS3.2

analysed here that uses an algorithm for q_a that has been developed using inter-satellite calibrated SSM/I data processed in the same way as that used to construct the dataset itself. The B03 algorithm used by IFREMER, HOAPS3.2 and GSSTF3 is derived using data from the IFREMER archive that has not had inter-satellite calibration applied,

and in the cases of HOAPS3.2 and GSSTF3 is applied to SSM/I brightness temperature data derived from different processing routes and that has had inter-satellite calibration applied (Figure 1). The S95 and Chou *et al.* (1997) algorithms used by J-OFURO2 and GSSTF2b/c respectively did not use SSM/I data in their derivation.

Further differences may arise because some of the algorithms are developed using *in situ* estimates of q_{10} (S95, Chou *et al.* (1997)) and some q_{10n} (B03, Roberts *et al.* (2010)). However, the differences between the satellite-derived datasets and between the satellite and *in situ*-derived datasets are not consistent with differences between q_{10} and q_{10n} (Figure 13).

One of the motivations for the development of estimates of q_a from satellite measurements is a lack of *in situ* observations. This lack of *in situ* q_a for calibration and validation is also one of the main challenges, particularly in high latitudes. A variety of different *in situ* data sources have been used and include measurements from radiosonde ascents and q_a from research vessels, merchant vessels and moored buoys. Research vessel measurements of q_a are expected to be of good quality, especially those that have been measured as part of air–sea interaction experiments, but are sparse (Roberts *et al.*, 2010). Measurement of q_a from moored buoys remains challenging (Weller *et al.*, 2008) and is regionally limited to the Tropical arrays and some coastal regions, particularly around United States. VOS humidity measurements are much more numerous than the other *in situ* sources, which helps to overcome their relatively large random uncertainties (Kent and Berry, 2005). Some regions, particularly in the southern hemisphere, remain largely unsampled. The use of VOS humidity measurements requires considerable care. The NOCv2.0 ICOADS-based estimates of q_a have been adjusted to a common reference level of 10 m above sea level and for bias (Berry and Kent, 2011); the effects of these adjustments are presented by Kent *et al.* (2014). Brunke *et al.* (2011) present differences of satellite-derived and research vessel-measured q_a estimates which vary regionally and which show differences between satellite estimates that exceed 1 g kg^{-1} for half of the cruises analysed.

A recent study (Jackson *et al.*, 2009) suggested that the bias adjustment of VOS humidity measurements was not required as they found smaller differences of $\sim 0.05 \text{ g kg}^{-1}$ between q_a from screens and psychrometers than the $\sim 0.3 \text{ g kg}^{-1}$ found by Berry and Kent (2011). Jackson *et al.* (2009) compared ICOADS q_a and that estimated from combined AMSU-A and SSM/I data. They gridded 0.5° and 3-hourly gridded satellite and ICOADS q_a from screens and psychrometers separately. Their approach means that regional biases between the satellite and *in situ* observations (see their Figure 9) will be aliased onto known regional differences in sampling between ships using screens and those using psychrometers. It is also hard to interpret their finding that the use of air temperature estimates biased high by solar heating (Berry *et al.*, 2004) in the calculation of q_a from ICOADS observations (Kent and Taylor, 1996; Dai, 2006) makes only a small difference as no information on any day–night differences in the satellite estimates themselves are presented.

Some of the algorithm training datasets (e.g. B03; Roberts *et al.*, 2010) use subsets of a single large dataset for training and validation of retrieval algorithms. This may introduce unwanted correlations (e.g. from measurements made on board a single vessel). More

generally, biases may be introduced into the algorithms by the non-representativeness of the training data. This is likely to be the cause of regional differences, for example as illustrated in Figure 14.

The different satellite datasets take different approaches to masking of data to remove estimates thought to be affected by the presence of land or of sea-ice. The three Goddard datasets (GSSTF2b/2c/3) take the most conservative approach, using a climatological monthly invariant sea-ice mask, the other datasets allow for monthly variation in sea-ice extent with the use of either climatological monthly (J-OFURO2) or monthly varying masks (HOAPS3.2, IFREMER and SeaFluxV1.0; Figure 2). SeaFluxV1.0 has the largest data coverage, but may suffer from some contamination of q_a in regions close to coasts (Figure 3(f)). Changes to the land mask are planned for the next version of SeaFlux (J. B. Roberts, 2013, pers. comm.)

Despite these factors that complicate the comparison of satellite-derived q_a fields, it is possible to understand some of the characteristics of the datasets in terms of their derivation. Comparison of the three datasets which all use the B03 algorithm (GSSTF3, HOAPS3.2 and IFREMER) shows the effects of the use of different data sources and different approaches to inter-satellite calibration (Figure 4). For these datasets the choice of whether or not to use data that have had inter-satellite calibration applied seems to be the most important.

Two different types of QC are applied to avoid systematic biases in the q_a estimates. The first is to exclude retrievals thought to be adversely affected by large water or ice particles (large droplet QC, implemented in slightly different forms by HOAPS3.2, IFREMER and J-OFURO2). The second is to correct for cases where the estimated q_a is greater than q_s (humidity inversion QC, implemented by GSSTF2b/2c/3 and HOAPS3.2). SeaFluxV1.0 do not apply either type of QC, relying on their algorithm (which also ingests estimates of SST) to avoid these problems. In particular SeaFluxV1.0 uses two different algorithms, the choice depending on the estimate of cloud liquid water (Roberts *et al.*, 2010). The large droplet QC has the biggest effect in regions of high precipitation (Figure 7(a) and (b)) but the impact on the different algorithms is not simple, with the impact changing sign depending on the algorithm used. The effect of the humidity inversion QC appears similar on monthly mean q_a derived from the S95 and B03 algorithms when tested using brightness temperatures. However, of the datasets that do not apply this type of QC, only the J-OFURO2 dataset (using S95) shows a strong effect with the impact on IFREMER (using B03) appearing much smaller (compare Figure 9(b) and (c)).

A limited direct comparison of two different algorithms (S95 and B03) constructed from the same brightness temperature dataset (F11 RSSv6) reveals large and systematic differences in q_a (Figures 5 and 6). These differences, combined with those due to difference in data sources, processing and QC are shown to give substantial differences between monthly mean 1° estimates of q_a in the

available satellite-derived datasets. These differences can reach 4 g kg^{-1} (Figure 9) and show large spatial (Figure 10) and temporal coherence (Figure 11). Comparison with *in situ* q_a estimates shows strong regional variations in the satellite–*in situ* q_a difference (Figures 12 and 14). Mean values, interannual variability and the seasonal cycle all show differences greater than the target accuracies of 0.3 g kg^{-1} and better required for accurate surface flux estimation. Few of the difference distributions are well represented by Gaussian statistics so we present summary statistics of the differences between datasets in the form of percentiles (Table 3).

On the basis of the results presented some general conclusions can be drawn:

1. The choice of the source of the SSM/I brightness temperatures (including processing and inter-satellite calibration), algorithm and QC all impact on the final q_a estimates. These choices therefore need careful consideration and must be fully documented.
2. It follows that any algorithm used to estimate q_a should be developed using satellite and *in situ* measurements from a traceable source, be processed to increase homogeneity and consistency, and all processing and QC should be fully documented. This applies equally to satellite and *in situ* measurements.
3. Algorithms to estimate q_a from satellite observations should be applied to brightness temperatures that have been processed in the same way as those used in the algorithm development. Currently SeaFlux V1.0/Roberts *et al.* (2010) is the only example of an algorithm/dataset combination where this has been done with inter-satellite calibrated data.
4. The large and coherent regional and temporal differences between satellite and *in situ* humidity estimates mean that global statistics on differences can be misleading as areas of positive and negative bias can cancel out.
5. It follows that *in situ* data with wide geographical coverage is needed to develop better algorithms, and independent global *in situ* data is needed to validate the results. This is going to be challenging given the decline of VOS (Kent *et al.*, 2014). Further efforts to aggregate high-quality observations of meteorological variability including humidity from research vessels are likely to prove extremely valuable.
6. *In situ* humidity observations may need to be converted between representations of humidity (e.g. dew-point temperature, relative humidity) and adjusted to a common reference level, usually 10 m over the ocean but estimates at 2 m are typically available from atmospheric reanalyses. The sources of any auxiliary data or metadata used in such conversions or adjustments should be clearly documented, along with clear descriptions of the methods used: including whether the derived estimate of specific humidity is referenced to neutral or ambient atmospheric stability.
7. The development of comprehensive estimates of uncertainty for every grid box for satellite fields of q_a , would

be extremely valuable. At the time of writing none of the satellite-based q_a datasets were accompanied by publicly available estimates of grid box uncertainty.

8. There are large differences among estimates of q_a from modern reanalyses (Kent *et al.*, 2014) and also among the estimates of q_a derived from SSM/I compared in this study. It is not currently possible to state that any of the available estimates of q_a meet the target accuracy for surface flux estimation of better than 0.3 g kg^{-1} on monthly time-scales for large regional averages.

Acknowledgements

This study received funding through the NERC National Centre for Earth Observation (NCEO) and the National Oceanography Centre's National Capability programme.

We are grateful to those who have answered our many questions, including: A. Andersson, A. Bentamy, W. Drennan, K. Fennig, K. Hilburn, D. Jackson, M. Kubota J. B. Roberts and C.-L. Shie. We also gratefully acknowledge the help of A. Andersson, K. Fennig, M. Kubota, J. B. Roberts and C.-L. Shie who all made constructive comments on the draft manuscript. We also thank the National Centre for Atmospheric Research for making the NOCv2.0 data available from the Computational and Information Systems Laboratory Research Data Archive (<http://rda.ucar.edu/datasets/ds260.3/>). The authors wish to acknowledge use of the Ferret programme for analysis and graphics in this paper. Ferret is a product of NOAA's Pacific Marine Environmental Laboratory. (<http://ferret.pmel.noaa.gov/Ferret/>) Finally, the authors wish to thank the two anonymous reviewers and the Editor for their thoughtful and thorough comments on the manuscript.

References

- Andersson A, Fennig K, Klepp C, Bakan S, Grassl H, Schulz J. 2010. The Hamburg ocean atmosphere parameters and fluxes from satellite data – HOAPS-3. *Earth Syst. Sci. Data* **3**: 143–194, doi: 10.5194/essdd-3-143-2010.
- Andersson A, Klepp C, Fennig K, Bakan S, Grassl H, Schulz J. 2011. Evaluation of HOAPS-3 ocean surface freshwater flux components. *J. Appl. Meteorol. Climatol.* **50**: 379–398, doi: 10.1175/2010JAMC2341.1.
- Atlas R, Hoffman RN, Ardizzone J, Leidner SM, Jusem JC, Smith DK, Gombos D. 2011. A cross-calibrated multiplatform ocean surface wind velocity product for meteorological and oceanographic applications. *Bull. Am. Meteorol. Soc.* **92**: 157, doi: 10.1175/2010BAMS2946.1.
- Bentamy A, Queffelec P, Quilfen Y, Katsaros K. 1999. Ocean surface wind fields estimated from satellite active and passive microwave instruments. *IEEE Trans. Geosci. Remote Sens.* **37**: 2469–2486, doi: 10.1109/36.789643.
- Bentamy A, Katsaros KB, Alberto M, Drennan WM, Forde EB, Roquet H. 2003. Satellite estimates of wind speed and latent heat flux over the global oceans. *J. Clim.* **16**: 637–656, doi: 10.1175/1520-0442(2003)016<0637:SEOWSA>2.0.CO;2.
- Bentamy A, Ayina L-H, Drennan WM, Katsaros KB, Mestas-Nunez AM, Pinker RT. 2008. 15 years of ocean surface momentum and heat fluxes from remotely sensed observations. *Flux News* **5**: 14–16.
- Bentamy A, Grodsky SA, Katsaros KB, Mestas-Nunez A, Blanke B, Desbiolles F. 2013. Improvements of air-sea fluxes derived from satellite observations. *Int. J. Remote Sens.* **34**(14): 5243–5261, doi: 10.1080/01431161.2013.787502.

- Berry DI, Kent EC. 2009. A new air-sea interaction gridded dataset from ICOADS with uncertainty estimates. *Bull. Am. Meteorol. Soc.* **90**(5): 645–656, doi: 10.1175/2008BAMS2639.1.
- Berry DI, Kent EC. 2011. Air-sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates. *Int. J. Climatol.* **31**(7): 987–1001, doi: 10.1002/joc.2059.
- Berry DI, Kent EC, Taylor PK. 2004. An analytical model of heating errors in marine air temperatures from ships. *J. Atmos. Oceanic Technol.* **21**(8): 1198–1215, doi: 10.1175/2F1520-0426(2004)021%3C1198:AAMOHE%3E2.0.CO;2.
- Bourassa MA, Gille ST, Jackson DL, Roberts JB, Wick GA. 2010. Ocean winds and turbulent air-sea fluxes inferred from remote sensing. *Oceanography* **23**(4): 36–51, doi: 10.5670/oceanog.2010.04.
- Bourras D. 2006. Comparison of five satellite-derived latent heat flux products to moored buoy data. *J. Clim.* **19**: 6291–6313, doi: 10.1175/JCLI3977.1.
- Brunke MA, Wang Z, Zeng X, Bosilovich M, Shie C-L. 2011. An assessment of the uncertainties in ocean surface turbulent fluxes in 11 reanalysis, satellite-derived, and combined global datasets. *J. Clim.* **24**: 5469–5493, doi: 10.1175/2011JCLI4223.1.
- Cavalieri DJ, Parkinson CL, Gloersen P, Zwally HJ. 1997. Arctic and Antarctic sea ice concentrations from multichannel passive-microwave satellite data sets: October 1978–September 1995 – user's guide. NASA Technical Memorandum 104647, NASA Goddard Space Flight Centre, Greenbelt, MA.
- Cavalieri DJ, Parkinson CL, Gloersen P, Comiso JC, Zwally HJ. 1999. Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets. *J. Geophys. Res.* **104**: 15803–15814, doi: 10.1029/1999JC900081.
- Chou S, Atlas R, Shie C, Ardizzone J. 1995. Estimates of surface humidity and latent-heat fluxes over oceans from SSM/I data. *Mon. Weather Rev.* **123**: 2405–2425, doi: 10.1175/1520-0493(1995)123<2405:EOSHAL>2.0.CO;2.
- Chou S, Shie C, Atlas R, Ardizzone J. 1997. Air-sea fluxes retrieved from special sensor microwave imager data. *J. Geophys. Res. Oceans* **102**: 12705–12726, doi: 10.1029/97JC00978.
- Chou S-H, Nelkin E, Ardizzone J, Atlas RM, Shie C-L. 2003. Surface turbulent heat and momentum fluxes over global oceans based on the Goddard satellite retrievals, Version 2 (GSSTF2). *J. Clim.* **16**: 3256–3273, doi: 10.1175/1520-0442(2003)016<3256:STHAMF>2.0.CO;2.
- Curry J, Bentamy A, Bourassa M, Bourras D, Bradley E, Brunke M, Castro S, Chou S, Clayton C, Emery W. 2004. SeaFlux. *Bull. Am. Meteorol. Soc.* **85**: 409–424, doi: 10.1175/BAMS-85-3-409.
- Dai A. 2006. Recent climatology, variability and trends in global surface humidity. *J. Clim.* **19**: 3589–3606, doi: 10.1175/JCLI3816.1.
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Källberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597, doi: 10.1002/qj.828.
- Fennig K, Andersson A, Bakan S, Klepp C, Schroeder M. 2012. Hamburg ocean atmosphere parameters and fluxes from satellite data – HOAPS 3.2 – monthly means/6-hourly composites. *Satell. Appl. Fac. Clim. Monitor.*, doi: 10.5676/EUM_SAF_CM/HOAPS/V001.
- Fennig K, Andersson A, Bakan S, Schroeder M. 2013. Fundamental climate data record of SSM/I brightness temperatures. *Satell. Appl. Fac. Clim. Monitor.*, doi: 10.5676/EUM_SAF_CM/FCDR_SSMI/V001.
- GCOS. 2010. Implementation plan for the global observing system for climate in support of the UNFCCC. GCOS-138, WMO, Geneva, 180 pp. <http://www.wmo.int/pages/prog/gcos/Publications/gcos-138.pdf> (accessed 9 September 2014).
- GLOBE Task Team and Others, Hastings DA, Dunbar PK, Elphinstone GM, Bootz M, Murakami H, Maruyama H, Masaharu H, Holland P, Payne J, Bryant NA, Logan TL, Muller JP, Schreier G, MacDonald JS (eds). 1999. *The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0*. National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder, CO. Digital Database on <http://www.ngdc.noaa.gov/mgg/topo/globe.html> and CD-ROMs (accessed 9 September 2014).
- Hilburn KA, Shie C-L. 2011. Decadal trends and variability in special sensor microwave/imager (SSM/I) brightness temperatures and Earth incidence angle. Report Number 092811, Remote Sensing Systems, Santa Rosa, CA, 53 pp. http://www.remss.com/papers/tech_reports/hilburn_SSMI_EIA_TB_RSS_Tech_Rpt_092811.pdf (accessed 9 September 2014).
- Iwasaki S, Kubota M, Tomita H. 2010. Evaluation of bulk method for satellite-derived latent heat flux. *J. Geophys. Res.* **115**: C07007, doi: 10.1029/2010JC006175.
- Jackson DL, Wick GA, Bates JJ. 2006. Near-surface retrieval of air temperature and specific humidity using multisensor microwave satellite observations. *J. Geophys. Res.* **111**: D10306, doi: 10.1029/2005JD006431.
- Jackson DL, Wick GA, Robertson FR. 2009. Improved multisensor approach to satellite-retrieved near-surface specific humidity observations. *J. Geophys. Res. Atmos.* **114**: D16303, doi: 10.1029/2008JD011341.
- Jones C, Peterson P, Gautier C. 1999. A new method for deriving ocean surface specific humidity and air temperature: an artificial neural network approach. *J. Appl. Meteorol.* **38**: 1229–1245, doi: 10.1175/1520-0450(1999)038<1229:ANMFDO>2.0.CO;2.
- Kent EC, Berry DI. 2005. Quantifying random measurement errors in voluntary observing ship meteorological observations. *Int. J. Climatol.* **25**(7): 843–856, doi: 10.1002/joc.1167.
- Kent EC, Taylor PK. 1996. Accuracy of humidity measurements on ships: consideration of solar radiation effects. *J. Atmos. Oceanic Technol.* **13**(6): 1317–1321, doi: 10.1175/1520-0426(1996)013<1317:AOHMOS>2.0.CO;2.
- Kent EC, Woodruff SD, Berry DI. 2007. WMO Publication No. 47 metadata and an assessment of voluntary observing ships observation heights in ICOADS. *J. Atmos. Oceanic Technol.* **24**(2): 214–234, doi: 10.1175/JTECH1949.1.
- Kent EC, Berry DI, Prytherch J, Roberts JB. 2014. A comparison of global marine surface specific humidity datasets from in situ observations and atmospheric reanalyses. *Int. J. Climatol.* **34**: 355–376, doi: 10.1002/joc.3691.
- Kubota M, Iwasaka N, Kizu S, Konda M, Kutsuwada K. 2002. Japanese ocean flux datasets with use of remote sensing observations (J-OFURO). *J. Oceanogr.* **58**: 213–225, doi: 10.1023/A:1015845321836.
- Kunkee DB, Poe GA, Boucher DJ, Swadley SD, Hong Y, Wessel JE, Uliana EA. 2008. Design and evaluation of the first special sensor microwave imager/sounder. *IEEE Trans. Geosci. Remote Sens.* **46**(4): 863–883, doi: 10.1109/TGRS.2008.917980.
- Liu C, Zipser EJ. 2013. Why does radar reflectivity tend to increase downward toward the ocean surface, but decrease downward toward the land surface? *J. Geophys. Res.* **118**: 1–14, doi: 10.1029/2012D018134.
- Meissner T, Wentz F. 2012. The emissivity of the ocean surface between 6–90 GHz over a large range of wind speeds and Earth incidence angles. *IEEE Trans. Geosci. Remote Sens.* **50**(8): 3004–3026, doi: 10.1109/TGRS.2011.2179662.
- Meng L, He Y, Chen J, Wu Y. 2007. Neural network retrieval of ocean surface parameters from SSM/I data. *Mon. Weather Rev.* **135**: 586–597, doi: 10.1175/MWR3292.1.
- NGDC. 1986. National Geophysical Data Center/NESDIS/NOAA/U.S. Department of Commerce: ETOPO5 Global Earth Topography, 5-minute.
- Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S, Wada K, Kato K, Oyama R, Ose T, Mannoji N, Taira R. 2007. The JRA-25 reanalysis. *J. Meteorol. Soc. Jpn.* **85**: 369–432, doi: 10.2151/jmsj.85.369.
- Pierce D. 2011. *ncdf: interface to Unidata netCDF data files. R Package Version 1.6.6*. <http://CRAN.R-project.org/package=ncdf> (accessed 9 September 2014).
- R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna. ISBN: 3-900051-07-0. <http://www.R-project.org/> (accessed 9 September 2014).
- Reynolds R, Rayner N, Smith T, Stokes D, Wang W. 2002. An improved in situ and satellite SST analysis for climate. *J. Clim.* **15**: 1609–1625, doi: 10.1175/1520-0442(2002)015<1609:AIIASAS>2.0.CO;2.
- Ritchie AA, Smith MR, Goodman HM, Schudalla RL, Conway DK, LaFontaine FJ, Moss D, Motta B. 1998. Critical analyses of data differences between FNMOC and AFGWC spawned SSM/I datasets. *J. Atmos. Sci.* **55**: 1601–1612, doi: 10.1175/1520-0469(1998)055<1601:CAODDB>2.0.CO;2.
- Roberts JB, Clayton CA, Robertson FR, Jackson DL. 2010. Predicting near-surface atmospheric variables from special sensor

- microwave/imager using neural networks with a first-guess approach. *J. Geophys. Res. Atmos.* **115**: D19113, doi: 10.1029/2009JD013099.
- Santorelli A, Pinker RT, Bentamy A, Katsaros KB, Drennan WM, Mestas-Núñez AM, Carton JA. 2011. Differences between two estimates of air-sea turbulent heat fluxes over the Atlantic Ocean. *J. Geophys. Res.* **116**: C09028, doi: 10.1029/2010JC006927.
- Schlüssel P, Schanz L, Englisch G. 1995. Retrieval of latent heat flux and longwave irradiance at the sea surface from SSM/I and AVHRR measurements. *Adv. Space Res.* **16**(10): 107–116, doi: 10.1016/0273-1177(95)00389-V.
- Schulz J, Schlüssel P, Grassl H. 1993. Water vapor in the atmospheric boundary layer over oceans from SSM/I measurements. *Int. J. Remote Sens.* **14**: 2773–2789, doi: 10.1080/01431169308904308.
- Semunegus H. 2011. Remote Sensing Systems Version – 6 Special Sensor Microwave/Imager Fundamental Climate Data Record, Climate Algorithm Theoretical Basis Document, CDRP–ATBD–0100, Revision 1.0, 26 July 2011, NOAA. <http://www1.ncdc.noaa.gov/pub/data/sds/cdr/docs/rss-v6.0-ssmi-fcdr-source.tar> (accessed 9 September 2014).
- Shie C-L. 2010. Science background for the reprocessing and Goddard satellite-based surface turbulent fluxes (GSSTF2b) data set for global water and energy cycle research. In *NASA GES DISC*, 12 October 2010, NASA Goddard Space Flight Center, Greenbelt, MD, 18 pp. <http://disc.sci.gsfc.nasa.gov/measures/documentation/Science-of-the-data.pdf> (accessed 9 September 2014).
- Shie C-L. 2011. Science background for the reprocessing and Goddard satellite-based surface turbulent fluxes (GSSTF2c) data set for global water and energy cycle research. Science Document for the Distributed GSSTF2c via Goddard Earth Sciences (GES) Data and Information Services Center (DISC), 19 pp. http://disc.sci.gsfc.nasa.gov/measures/documentation/Science_of_the_data.GSSTF2c.pdf (accessed 9 September 2014).
- Shie C-L. 2012. Science background for the reprocessing and Goddard satellite-based surface turbulent fluxes (GSSTF3) data set for global water and energy cycle research. Science Document for the Distributed GSSTF3 via Goddard Earth Sciences (GES) Data and Information Services Center (DISC), 21 pp. http://disc.sci.gsfc.nasa.gov/measures/documentation/Science_of_the_data.GSSTF3.pdf (accessed 9 September 2014).
- Shie C-L, Hilburn K. 2011. A satellite-based global ocean surface turbulent fluxes dataset and the impact of the associated SSM/I brightness temperature. In *Proceedings of the 2011 EUMESAT Meteorological Satellite Conference*, Oslo, 5–9 September, 8 pp. <ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxdocu/gsstf2c/Shie&Hilburn-2011.pdf> (accessed 9 September 2014).
- Shie C-L, Chiu LS, Adler R, Lin I-I, Nelkin E, Ardizzone J. 2010a. The Goddard satellite-based surface turbulent fluxes dataset – Version 2b (GSSTF 2b), global (grid of 1°x1°) daily air-sea surface fluxes from July 1987 to December 2008, distributed via Goddard Earth Sciences (GES) Data and Information Services Center (DISC) in October 2010. <ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxdata> (accessed 9 September 2014).
- Shie C-L, Chiu LS, Adler R, Gao S, Chokngamwong R, Lin I-I, Nelkin E, Ardizzone J, Xie P, Wang F-C. 2010b. A recently revised and produced global air-sea surface turbulent fluxes dataset – GSSTF2b: validations and findings. In *Proceedings of the Joint 2010 CWB Weather Analysis and Forecasting & COAA 5th International Ocean-Atmosphere Conference*, Taipei, 28–30 June, 307–312. <http://disc.sci.gsfc.nasa.gov/measures/documentation/Shie-et-al-COAA2010.pdf> (accessed 9 September 2014).
- Shie C-L, Hilburn K, Chiu LS, Adler R, Lin I-I, Nelkin E, Ardizzone J. 2011. The Goddard satellite-based surface turbulent fluxes dataset – version 2c (GSSTF 2c) [global (grid of 1°x1°) daily air-sea surface fluxes from July 1987 to December 2008] distributed via Goddard Earth Sciences (GES) Data and Information Services Center (DISC) in October 2011. GSSTF2c in HDF-EOS5 format. http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldingsMEASURES.pl?PROGRAM_List=ChungLinShie_OldVer or <ftp://measures.gsfc.nasa.gov/data/s4pa/GSSTF/> (accessed 9 September 2014).
- Shie C-L, Hilburn K, Chiu LS, Adler R, Lin I-I, Nelkin E, Ardizzone J, Gao S. 2012. *Goddard Satellite-Based Surface Turbulent Fluxes, Daily Grid, Version 3*, Savtchenko A (ed). Greenbelt, MD: Goddard Earth Science Data and Information Services Center (GES DISC), doi: 10.5067/MEASURES/GSSTF/DATA301 (accessed 28 September 2012).
- Singh R, Kishtawal C, Pal P, Joshi P. 2006. Surface heat fluxes over global oceans exclusively from satellite observations. *Mon. Weather Rev.* **134**: 965–980, doi: 10.1175/MWR3119.1.
- Smith SR, Hughes PJ, Bourassa MA. 2011. A comparison of nine monthly air-sea flux products. *Int. J. Climatol.* **31**(7): 1002–1027, doi: 10.1002/joc.2225.
- Swift CT, Fedor LS, Ramseier RO. 1985. An algorithm to measure sea ice concentration with microwave radiometers. *J. Geophys. Res. Oceans* **90**: 1087–1099.
- Tomita H, Kubota M. 2006. An analysis of the accuracy of Japanese ocean flux data sets with use of remote sensing observations (J-OFURO) satellite-derived latent heat flux using moored buoy data. *J. Geophys. Res.* **111**: C07007, doi: 10.1029/2005JC003013.
- Tomita H, Kubota M, Cronin MF, Iwasaki S, Konda M, Ichikawa H. 2010. An assessment of surface heat fluxes from J-OFURO2 at the KEO and JKEO sites. *J. Geophys. Res.* **115**: C03018, doi: 10.1029/2009JC005545.
- Weller RA, Bradley EF, Edson JB, Fairall CW, Brooks I, Yelland MJ, Pascal RW. 2008. Sensors for physical fluxes at the sea surface: energy, heat, water, salt. *Ocean Sci.* **4**(4): 247–263, doi: 10.5194/os-4-247-2008.
- Wentz FJ. 1991. *Revision-1 User's Manual SSM/I Antenna Temperature Tapes*. Remote Sensing Systems: Santa Rosa, CA, 70 pp.
- Wentz FJ. 2010. The version-6 calibration of SSM/I. Technical report 102210, Remote Sensing Systems, Santa Rosa, CA, 35 pp. http://images.remss.com/papers/tech_reports/Wentz_SSMI_Version6_Calibration.pdf (accessed 9 September 2014).
- Wentz FJ. 2013. SSM/I version-7 calibration report. Technical Report 011012, Remote Sensing Systems, Santa Rosa, CA, 44 pp. http://images.remss.com/papers/rsstech/2012_011012_Wentz_Version7_SSMI_Calibration.pdf (accessed 9 September 2014).
- Woodruff SD, Diaz HF, Elms JD, Worley SJ. 1998. COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth* **23**: 517–526, doi: 10.1016/S0079-1946(98)00064-0.
- Woodruff SD, Worley SJ, Lubker SJ, Ji Z, Freeman JE, Berry DI, Brohan P, Kent EC, Reynolds RW, Smith SR, Wilkinson C. 2011. ICOADS release 2.5 and data characteristics. *Int. J. Climatol.* **31**(7): 951–967, doi: 10.1002/joc.2103.