Focus Article

Effects of instrumentation changes on sea surface temperature measured *in situ*

Elizabeth C. Kent,[1]* John J. Kennedy,[2] David I. Berry[1] and Robert O. Smith[2]

Measurements of sea surface temperature (SST) are an important climate record, complementing terrestrial air temperature observations, records of marine air temperature, ocean subsurface temperature, and ocean heat content. SST has been measured since the 18th century, although observations are sparse in the early period. Historically, marine observing systems relied on observations made by seafarers and necessary information on measurement methods is often not available. There are many historical descriptions of observing practice and instrumentation, some including quantification of biases between different methods. This documentation has been used, with the available observations, to develop models for the expected biases, which vary according to how the measurements were made, over time and with the environmental conditions. Adjustments have been developed for these biases and some gridded SST datasets adjust for these differences and provide uncertainty estimates, including uncertainties in the bias adjustments. The modern *in situ* SST-observing system continues to evolve and now includes many observations from moored and drifting buoys, which must be characterized relative to earlier observations to provide a consistent record of multi-decadal changes in SST. *WIREs Clim Change* 2010 1 718–728

INTRODUCTION

Measurements of sea surface temperature (SST) have been made for over 200 years. The first data in the International Comprehensive Ocean-Atmosphere Data Set (ICOADS[1]) are from the late 1700s. Observations were extremely sparse until the first international maritime meteorological conference in 1853, which began the data collection program we have today.[2,3] By 1855, the number of SST observations had increased from a few tens of reports per month to a few thousands. Seafarers originally made meteorological observations to aid efficient and safe navigation. More recently, observations have been collected and transmitted in real time for numerical weather prediction and storm warnings as part of the Voluntary Observing Ship (VOS) Scheme.[4] Alongside the ship observations, moored buoy observations of SST become available in 1971, and measurements from surface drifters in 1978. ICOADS contains SST observations from a variety of other sources, including oceanographic and coastal data. Global gridded datasets of SST are available, the longest starting in 1850.[5–9]

Early observations were made by sampling the temperature of water collected in buckets. The recorded SST will vary with the type and size of the bucket and the procedure used to make the measurement. The importance of consistency was recognized and the 1853 conference recommended the use of a common 'abstract log' for recording observations and the use of wooden buckets to measure water temperature.[10] Unfortunately a wide variety of types of bucket, thermometer, and observing practice have been in use.[11,12] The World Meteorological Organization (WMO) recognized the importance of recording methods of measurement and in 1955 introduced their Publication No. 47, a

*Correspondence to: eck@noc.soton.ac.uk

[1]National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

[2]Met Office, Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK

DOI: 10.1002/wcc.55

metadata catalogue for ships participating in the VOS Scheme (known as Pub. 47, see Refs 13,14). Methods of SST measurement can also now be transmitted as part of real-time reports. Comparisons are difficult as each of the measurement methods has different characteristics and biases. To homogenize the SST record requires detailed information on measurement methods and platforms, observation practice, and environmental information. This article reviews the progress that has been made toward gathering the metadata required to identify the methods and practices used, the characterization of biases in data of different types, the development of models of biases, and the construction of bias-adjusted SST datasets that include estimates of uncertainty due to the biases.

METHODS OF MEASUREMENT

What is known as SST is loosely understood as referring to measurements made at a variety of depths in typically the top 20 m of the ocean. Observations of SST will contain biases which depend on the measurement method; these are summarized below. In addition, the surface temperature record will contain variability related to real changes in the surface layer of the ocean. Examples include any diurnal variability that occurs as a result of solar heating when wind speeds are too low to mix the surface waters[15–18] or intense rainfall.[19] Knowledge of the time of day, depth of measurement, and conditions in the preceding hours can therefore help to reconcile temperature observations, although none of the global SST datasets currently available apply such adjustments. The implicit assumption is that the sampling of conditions is regular enough that no regional or time-varying bias is introduced into the datasets by neglecting such effects.

Attempts have been made to hone the definition of ocean surface temperatures by stating the depth of the measurement and introducing the concept of ‘foundation temperature’ as ‘the temperature of the water column free of diurnal temperature variability’ and is the temperature from which the growth of any diurnal thermocline develops each day.[20] Although a useful concept, it requires calculation from existing *in situ* and satellite measurements of temperature using mixed layer models, and few systems can actually measure the foundation temperature.[21]

Ships: bucket measurements

Water samples in buckets are likely to lose heat as the sea surface is typically warmer than the air above it. The exchange of heat between the water sample and the atmosphere will depend on the bucket (e.g., the material it is constructed from, its volume, surface area, and whether it has a lid) and ambient conditions (e.g., the temperature difference between the sample and the surrounding air, relative wind speed over the bucket, atmospheric humidity, and incident solar radiation). The temperature of the final measurement will also depend on how the observation is made (whether the bucket is allowed to equilibrate with the sea temperature—possibly involving taking an initial sample that is discarded—the time taken to haul the bucket, the initial temperature of the bucket, whether the sample is stirred, the time allowed for the thermometer to reach the water temperature, and the type of thermometer used). The depth of the sample may also vary with the size and speed of the ship, sea state, and wind speed.

The earliest SST observations derive from wooden, canvas, and metal buckets[11] (Figure 1). Of these types, wooden buckets are relatively well insulated and tend to have larger volumes leading to smaller temperature changes (typically a reduced temperature as heat loss rather than heat gain is more usual). Models of corrections for wooden and uninsulated canvas buckets show the adjustments to be five to six times greater for the canvas buckets.[11] As the amount of heat loss from uninsulated buckets became recognized, national Meteorological Services began to issue insulated buckets to ships.[22] These were made from insulating materials, such as rubber, or used a double skin construction to allow a layer of sea water to surround the sample being measured.

Comparisons have helped to identify the size of likely biases in observations made using buckets. A comprehensive study[24] compared 16,000 bucket and intake SSTs of various types. The overall intake-bucket difference was $0.3 \pm 1.3^{\circ}$C, but larger differences thought to be related to the bucket measurement were found in winter, at high wind speeds, and for both uninsulated and insulated canvas buckets. Most national-type and specially designed buckets showed smaller differences. Buckets were also relatively cooler compared with intakes when there was precipitation and relatively warmer when there was fog. Buckets were more likely to be relatively cooler before midday and warmer after midday although there is no clear diurnal cycle in the differences. Measurements of water sample temperature in buckets in wind tunnels[25,26] and on ships[23,25,27] and comparisons of bucket and other SST observations[24,28] have shown the impact of environmental conditions and observing practice on bucket observations of SST. Attempts have been made to partition the biases by error source.[29,30] Measurements made for the purpose of comparing

FIGURE 1 | Left to right: wooden bucket, 1891 (courtesy of Scottish Maritime Museum and David Parker); Crawford bucket as described in Ref 23; German metal and leather bucket; and UK Met Office canvas bucket (courtesy of David Parker, Crown Copyright).

measurement methods typically show smaller biases than measurements made as part of routine observing schedules.[31–33]

The evolution of the ocean surface temperature observing system is summarized in Figures 2 and 3, which indicate the platforms and measurement methods contributing to ICOADS over time.

Ships: engine room intake measurements

The earliest steam ships were constructed in the 1830s and by 1914 British steam ships made up well over half of the worlds tonnage.[36] On ships with engines, it was convenient to measure the temperature of the pumped seawater used to cool the engines. Early engine room intake (ERI) measurements may have been taken with a mercury thermometer in a well in the intake pipe or from a dial with temperature intervals of several degrees. The measurement would probably have been taken close to the engines rather than near the inlet.[31,37] Later systems are more likely to use electrical thermometers and some will include a dedicated sensor remote from the engine room giving a measurement closer to the inlet. Biases in ERI measurements are dependent on the details of the system and include fouling, poor exposure of the thermometer in the pipe, air pockets in the thermometer well, flow rate of the pumped system, heat conduction along thermometer and supports, length and insulation of pipe, and the difficulty of reading the scale including parallax errors.[37,38] There are no details of the calibration of ERI thermometers

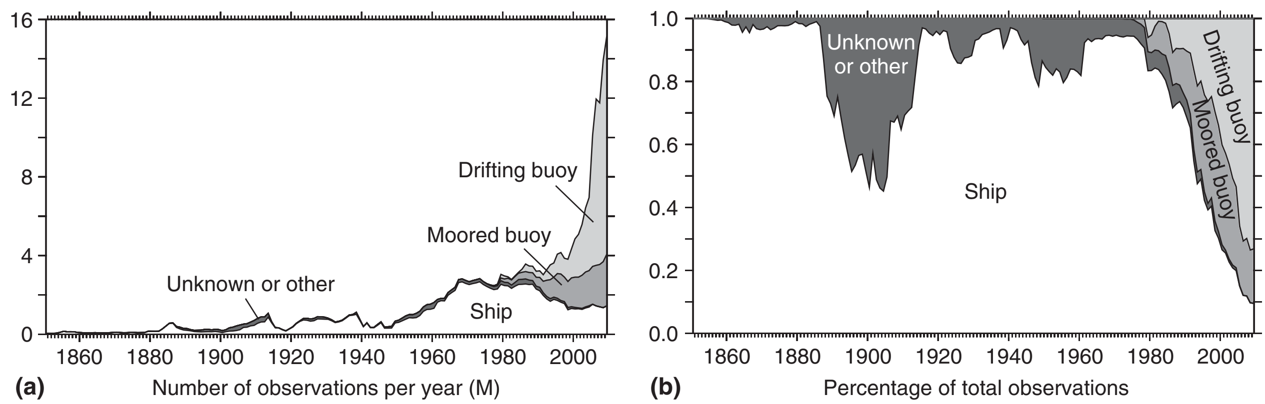

FIGURE 2 | Composition of ICOADSv2.5.[1] (a) Annual number of sea surface temperature observations per year by platform type. (b) As (a) but expressed as a fraction of total number of observations. For the period centered on around 1900, most of the 'unknown or other' observations are of unknown source, and after 1962 almost all have come from oceanographic sources.

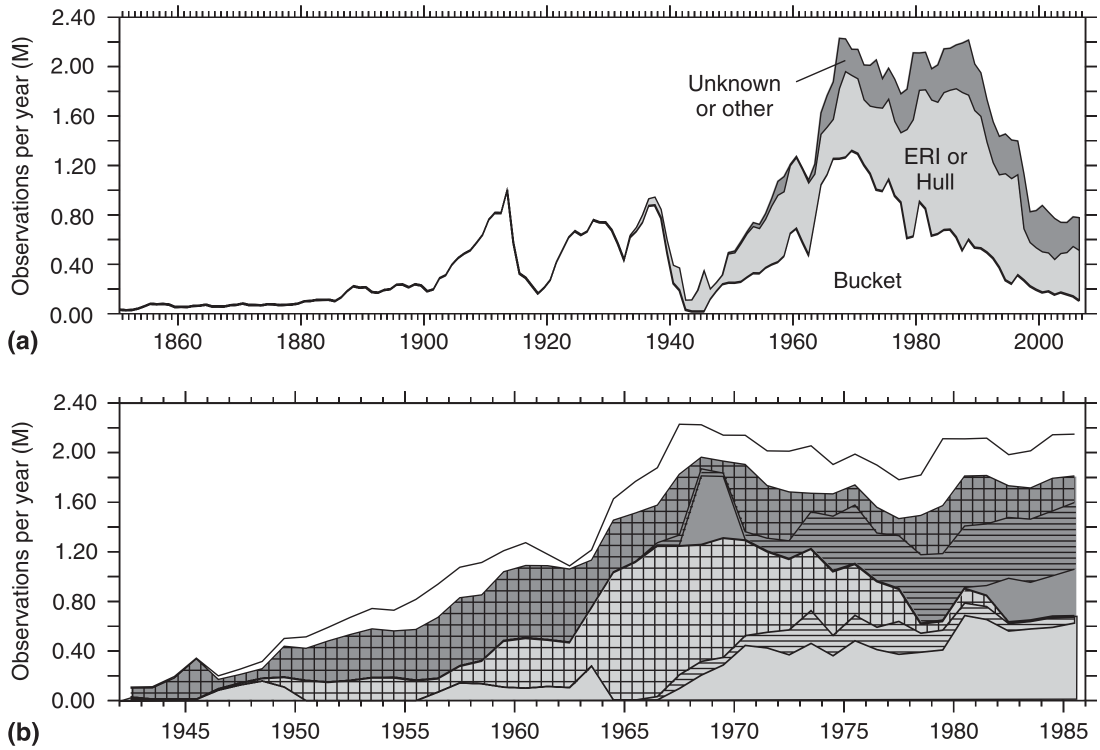

FIGURE 3 | (a) Number of sea surface temperature observations from ICOADSv2.0[34] with measurement methods following Ref 35, excluding data from drifters and buoys. (b) Metadata assignments for period 1942–1986: light gray = bucket, dark gray = engine room intake or hull sensor, white = other or unknown. No pattern = ICOADS metadata, horizontal lines = metadata from Pub. 47, squares = assignment from country preference.

in the literature. The inlet must also be deep enough to be submerged at all loadings and in all weathers. Inlet depth varies by ship type and size: small vessels such as trawlers or research vessels typically measure at 1–3 m, container ships and bulk carriers typically at 7–11 m, and very large ships at even greater depths.[13]

Most studies of biases in ERI measurements show that ERI SSTs are warmer than bucket SSTs.[12,24,37,39,40] Warm biases relative to bucket measurements were smallest for precision thermometers and thermistors, shallow intakes, and thermometers located close to the inlet.[24] Large warm biases were shown for mercury and ‘other’ thermometers, deep inlets, and thermometers positioned far from the inlet. Any cool bias relative to surface measurements that might be expected from measurement at depth is masked by the tendency to see larger warm biases on larger ships with deeper intakes.[24] There are considerable regional variations in measurement depth for those observations in ICOADS between 1995 and 2004 which have Pub. 47 measurement depths.[13] The size of any bias in ERI measurements is therefore dependent on details of the measurement system that are usually unknown and must be estimated from the data themselves.

Ships: other methods

SST measurement method metadata in WMO Pub. 47 includes categories for hull contact and through-hull sensors, trailing thermistor, radiation thermometer, and bait tank thermometer. Of these, only the hull sensors are commonly in use by VOS, and only in recent periods.[13] Hull sensors were found to provide more consistent SST measurements than ERIs and their use was recommended.[28] A further evaluation is now needed as hull sensor measurements now make up nearly a quarter of ship observations. Some underway measurements from research vessels are also incorporated into ICOADS further adding to the diversity of measurement methods.

ICOADS also contains observations of SST from oceanographic profiles from profiling instruments of various kinds and from bottle samples. Near-surface observations from the World Ocean Database[41] are sampled and provide an additional 5–10% to the traditional ship-based observations in the period 1940–2005. ICOADS SSTs from oceanographic and non-oceanographic ships do show some differences[10] and further comparisons are needed.

Moored buoys

The first moored buoy observations of SST become available in ICOADS in 1971. With the exception of the tropical arrays,[42–44] the moorings are largely coastal and most are around the US coasts. Comparison of SST sensors on three pairs of nearby buoys deployed by the US National Data Buoy Centre (NDBC) showed offsets and random variations of a few tenths °C.[45] Recent moored buoy observations show similar random variations.[46]

Moored buoy SST data provide a continuous series of observations at a single location and are potentially of higher quality than other components of the *in situ* observing system because of an annual maintenance and (re-)calibration schedule. However, reliance on data from a single sensor and platform means that identification of any biases or drifts is critical. Changes in buoy hull types, payloads, sensor types, and locations are not currently well documented for the historical record.[47] The availability of such metadata would improve the utility of the moored buoy SST record for climate change detection and research.

Surface drifters

Surface drifter data, first available in 1978, are now the most numerous source of *in situ* SST. It must be noted that despite the large increase in observation numbers, driven partly by the high rate of observations from drifters, the proportion of 1° areas sampled each month is now less than in the period 1965–1995.[4] Drifter design is now standardized, providing a more consistent means of measuring SST *in situ*, but before the early 1990s, a variety of designs were used.[48] Drifting buoy SST has been found on average to be biased cold relative to ship data by between 0.13 and 0.18°C,[49–51] although regional differences are large and vary in sign.[51] Comparisons of SST measurements made with platinum temperature probes installed on 16 drifters with those from the normal type of thermistor showed thermistors outside the 0.1° target accuracy in several deployments and examples of significant drifts.[52] As for moored buoys, improvements in metadata availability for drifters would be valuable.[47]

Other methods

ICOADS also contains observations from coastal and island sites and also from fixed platforms such as oil rigs. The coastal SST data are usually excluded from marine datasets (and from Figure 2) as they are not thought to be representative of ocean values.

ASSIGNING OBSERVING METADATA TO OBSERVATIONS

Observations in ICOADS may have a flag indicating the method of SST measurement. The flag is usually reported with the measurements and may be more up-to-date than, e.g., the annually available Pub. 47 metadata. Problems with flags have, however, been found for certain periods and data sources.[53] Availability of such metadata depends on the data source and many data observations have no accompanying measurement method. Lack of metadata is an important contributor to the uncertainty in SST variations over time. Using metadata from a variety of sources (Figure 3) allows this uncertainty to be modeled.[35] In their analysis, metadata from flags in ICOADS is preferred when available.[35] For observations with no flag, the WMO Pub. 47 metadata was searched for a match. This matching is only possible after 1955 and is dependent on the presence of a valid ship callsign in ICOADS as Pub. 47 metadata is indexed by callsign.[13] When no other information was available, metadata was assigned using information about the recruiting country.[35] Each national fleet is issued with instruments and observing instructions by the national meteorological service and which vary from country to country and over time. For example, Pub. 47 indicates that almost all US-recruited ships report SST from ERIs and all Indian-recruited ships use buckets. The fleets of the United Kingdom, Japan, the Netherlands, and Germany have all increased their use of ERIs over time. SST data with unknown method but known recruiting country can be assigned a measurement method probability that depends on the reporting preferences of the particular country.[35] Uncertainty in bias adjustments were then calculated from ensembles of realizations where observations with unknown metadata were randomly reassigned allowing for uncertainty in the assignment of metadata (e.g., due to ERI measurements being made when bucket observations were impractical[22,32]). Varying schedules for switching from uninsulated canvas buckets to rubber buckets after the Second World War were also included in the ensembles.

Metadata describing hull types, sensor types, and depths for moored and drifting buoys are presently not well integrated with the observational archive.[47] Assignment of metadata for drifting buoys may be complicated by the reissue of identification numbers after a buoy has ceased reporting.

SST BIAS ADJUSTMENT

Although biases in SST observations are well documented, the development of bias adjustments for gridded SST datasets is hampered by the difficulty in ascribing measurement methods described above. Even when measurement method is known, there are further uncertainties in the details of the instrumentation and measurement methodology. An analytical correction scheme for pre-1941 biases in wooden and uninsulated canvas bucket measurements has been developed.[11] The correction, which adjusts observations to the mix of measurement methods over the period 1951–1980, is based on models of heat and moisture transfers from the different types of bucket. Derivation of a correction from these models requires knowledge of the size of the buckets (inner diameter and initial water depth for the canvas bucket model and bucket wall thickness for the wooden bucket model), the time the bucket was exposed on deck, the relative wind speed (which depends on the ship speed, the true wind speed, and the degree of sheltering of the bucket), and the exposure of the bucket to solar radiation, all of which may vary from ship to ship and with time. A range of models was tested with different assumptions and the most important factors were the speed of the ship, which increases systematically over time, and the relative proportions of each bucket type. Information on the types of bucket used was sparse and a linear change in proportions was therefore estimated by optimizing agreement between corrected SST and night-time marine air temperature (NMAT) in the tropics. This suggested that wooden buckets were prevalent in the 1850s, when 70% of buckets were wooden, and went out of use by 1920. The correction varies over time and also geographically due to variations in air–sea temperature difference, humidity, and wind speed (Figure 4). It was noted that any change of the data mix, for example as more historical data are digitized,[54,55] will necessitate a re-evaluation of the corrections. The methodology[11] has been used in several datasets.[5,6,56]

Estimates of bias uncertainty using this methodology have been calculated for the HadSST2-gridded monthly mean dataset.[5] The assumptions made in Ref 11 were reviewed and an ensemble of realizations of bias adjustment covering the likely range of uncertainty in the model assumptions generated. It was concluded that despite the coherence of the bias uncertainties, and the large regional values, the bias adjustments had only a modest effect on the uncertainties in global and hemispheric trends.

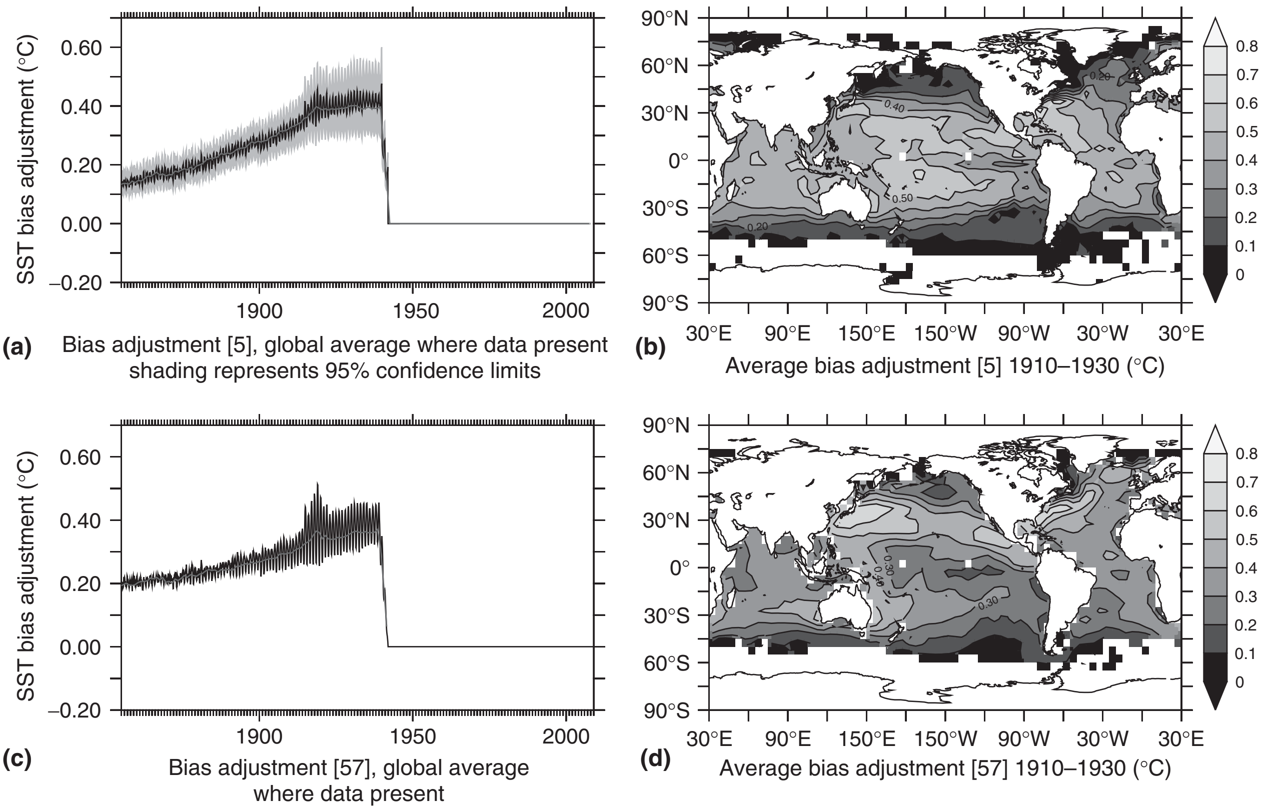

FIGURE 4 | SST corrections (°C). (a) Global average of bias adjustments from Ref 11 as implemented by Ref 5. (b) As (a) but averaged over the period 1910–1930. (c) Global average of bias adjustments from Ref 57. (d) As (c) but averaged over the period 1910–1930.

An alternative methodology for SST bias adjustment for the prewar period is based on the assumption that air–sea interactions on large time- and space-scales were the same over historical periods as for the more recent period 1968–1997.[57] Adjustments were therefore required to account for known biases in NMAT.[56,58] It was assumed that any changes in magnitudes of SST-NMAT differences were due to measurement biases in the SST, and noted that this was an approximation. The average bias adjustments from this scheme are within 0.05°C[11] after 1870 (Figure 4). The adjustment of Ref. 11 tends to follow contours of latent heat flux and that of Ref. 57 the sensible heat flux.[59] This adjustment method[57] has been used in global datasets.[8]

The very data-sparse period around the Second World War has been particularly problematic for SST bias adjustment. Modifications to adjustments were required between 1939 and 1941[5] to account for the incorporation of then newly digitized data from the US merchant marine.[9] Global-average SST from HadSST2 filtered to reduce the impact of short-term variability shows a discontinuity in late 1945, which was the apparent result of uncorrected instrumental biases.[60] Before the step, the majority of observations in ICOADS were made by US ships using the relatively warm ERI method. After the step, the majority were made by British ships using relatively cold buckets. As expected,[11] successive releases of ICOADS[1,9,34,53,61,62] have shown different bias characteristics in this poorly sampled period due to the targeted recovery of data from the archives of different countries.[54,55]

A wider need for bias adjustment for ship SST observations after 1941 was also recognized.[11,12,57,60] Since the 1950s, the relative number of bucket measurements has decreased, but an increasing number of drifting buoy observations since 1978 has mitigated any warm bias that this might have caused (Figures 2 and 3). The approach of Folland and Parker[11] was therefore extended to develop bias adjustments for the full period from 1850 onward,[35] accounting for biases in modern period ship-derived SST[11,12] and a cool bias in SST from drifters relative to that from ships.[49,52] The uncompensated inclusion of large numbers of drifting buoy data in global analyses is thought to have artificially reduced the observed rate of warming in global-average SST over the past 30 years.[35]

COMPARISON WITH OTHER DATA SOURCES

Anomalies in marine air temperatures (MAT) tend to be correlated with anomalies in SST over long space and time scales. Biases present in MAT measurements are largely independent from those in SST. Adjustments must be made to MAT for changes in ship size as larger ships measure relatively higher[56] and solar heating effects are reduced using analytic corrections[63,64] or by use of NMAT.[56] Although every effort has been made to keep SST and NMAT datasets independent, NMAT has been used to correct SST data and verify SST corrections. Localized problems during the 19th century and Second World War require the use of SST or daytime air temperatures to correct NMAT.[58] After 1900, NMAT and SST are independent to a large extent (and completely after 1941) and global and hemispheric averages exhibit similar behavior in this period.[56,65] Despite the SST measurement problems, the uncertainty in global land temperature exceeds that in SST, especially before the first World War,[66] as marine temperatures show much larger spatial and temporal coherence. This means that fewer observations are required to construct large-scale temperature averages. There are indications that since 1980 the land–ocean temperature contrast may be increasing, possibly a consequence of increasing greenhouse gas concentrations, changing atmospheric circulations or biases in the data.[66] Observations of SST have also been helpful in understanding and correcting biases in upper ocean temperature measurements resulting from errors in estimating the fall-rate of expendable probes used to measure ocean temperature profiles.[67,68]

THE FUTURE OBSERVING SYSTEM

Improved methods of measuring SST, combined with a greater understanding of the causes of bias in different observation types, mean that the modern SST-observing system is now of higher quality. *in situ* SST measurements are needed to provide calibration and validation for satellite SST retrievals, and the *in situ* observing system has been designed largely to meet this requirement[69,70] although other requirements such as consistency are also important.[71] The rapid evolution of the observing system, combined with a reduction in numbers of the more traditional ship-based observations[4] (Figure 2), means that improved characterizations of observations from drifting buoys, moored buoys, and ships are needed.[46,51,52] It is expected that more observations of SST will come from expanded use of research vessels,[72] profiling floats,[73] gliders,[74] and marine mammals,[75] which will help reconcile subsurface and near-surface temperature records. Ongoing characterization of observations is necessary to allow the construction of homogeneous datasets.[71] Careful comparisons of satellite and *in situ* SSTs are also needed to account for *in situ* biases and sampling differences that affect the validation of satellite retrievals of SST.[76,77]

CONCLUSION

Biases in the *in situ* SST measurement system are increasingly being understood in terms of the methods used and the environmental conditions at the time of the observation. The importance of retaining observational metadata alongside observations is increasingly being recognized, and newly digitized data should include such information wherever possible. Observational biases have been identified in bucket measurements of SST, which depend on the details of the bucket construction, the ambient conditions, and how the observation is made. Biases tend to be toward cooler measurements and are particularly large when there are large air–sea temperature contrasts, dry conditions, or when the bucket is uninsulated and exposed to strong relative wind speeds. Engine intake measurements have a tendency to be warm, especially in the period before reliable remote reading thermometers could be installed near to the seawater inlet. Models of bias used to adjust observations to create more consistent long-term gridded SST datasets are hampered by lack of metadata and coincident information on environmental conditions. Bias adjustments must be re-evaluated when new or digitized data are added to the record. Despite this, gridded SST datasets that include bias adjustments and estimates of bias uncertainty are available starting from 1850, and uncertainties are small enough to allow determination of climate change.

Improvements to the *in situ* SST record will come from digitization of new data and metadata in data- and metadata-sparse periods and regions. New measurements from satellites and drifting buoys have provided an improved understanding of variability in SST, especially diurnal variability and the impetus for improved models of diurnal variability. Datasets including diurnal variations are likely to be developed. Adjustments for these contributions to differences between SST measurements made in different ways and at different times become more important once the biases discussed in this article have been accounted for.

ACKNOWLEDGEMENTS

The authors would like to thank Nick Rayner and the anonymous reviewers for their help in improving this article. The hard work of the ICOADS team, especially Scott Woodruff, Sandra Lubker, and Steven Worley, has been essential in making the ever-expanding marine meteorological archive easily accessible and freely available. Tom Smith of NOAA kindly provided the data for Figures 4c and 4d. E.C.K. and D.I.B. were supported by the NERC Oceans2025 Theme 10 (SO9). J.J.K. and R.O.S. were supported by the Joint DECC and Defra Integrated Climate Programme—DECC/Defra (GA01101).

REFERENCES

1. Woodruff SD, Worley SJ, Lubker SJ, Ji Z, Freeman JE, et al. ICOADS Release 2.5: extensions and enhancements to the surface marine meteorological archive. *Int J Climatol* 2010. Available at: http://www3.interscience.wiley.com/journal/123210702/abstract. DOI: 10.1002/joc.2103
2. Maury MF. *Explanations and Sailing Directions to Accompany the Wind and Current Charts*, 6th ed. Philadelphia: E.C. and J. Biddle; 1854, 772 pp. + plates. [Note: A report from the 1853 Brussels Conference on pp. 54–60; "Minutes of the Sittings" on pp. 60–88; and "Explanatory notes for keeping the abstract log," and example log forms for men-of-war and merchantmen, on pp. 88–96].
3. Woodruff SD, Diaz HF, Worley SJ, Reynolds RW, Lubker SJ. Early ship observational data and ICOADS. *Clim Change* 2005, 73:169–194.
4. Kent E, Ball G, Berry D, Fletcher J, Hall A, et al. The Voluntary Observing Ship scheme. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society"* Conference, vol. 2, Venice, Italy, 21–25 September 2009: ESA Publication WPP-306; 2010.
5. Rayner NA, Brohan P, Parker DE, Folland CK, Kennedy JJ, et al. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 data set. *J Clim* 2006, 19:446–469.

6. Ishii M, Shouji A, Sugimoto S, Matsumoto T. Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe collection. *Int J Climatol* 2005, 25:865–879.
7. Kaplan A, Cane M, Kushnir Y, Clement A, Blumenthal M, et al. Analyses of global sea surface temperature 1856–1991. *J Geophys Res* 1998, 103:18567–18589.
8. Smith TM, Reynolds RW, Peterson TC, Lawrimore J. Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). *J Clim* 2008, 21:2283–2296.
9. Worley SJ, Woodruff SD, Reynolds RW, Lubker SJ, Lott N. ICOADS release 2.1 data and products. *Int J Climatol* 2005, 25:823–842.
10. Woodruff SD, Diaz HF, Kent EC, Reynolds RW, Worley SJ. The evolving SST record from ICOADS. In: Brönnimann S, Luterbacher J, Ewen T, Diaz HF, Stolarski RS, et al., eds. *Climate Variability and Extremes during the Past 100 Years*, vol. 33, *Advances in Global Change Research*, Netherlands: Springer; 2008, 65–83.
11. Folland CK, Parker DE. Correction of instrumental biases in historical sea surface temperature data. *Q J R Meteor Soc* 1995, 121:319–367.
12. Kent EC, Taylor PK. Toward estimating climatic trends in SST, part 1: methods of measurement. *J Atmos Ocean Technol* 2006, 23:464–475. Doi: 10.1175/JTECH1843.1
13. Kent EC, Woodruff SD, Berry DI. WMO Publication No. 47 Metadata and an assessment of voluntary observing ships observation heights in ICOADS. *J Atmos Ocean Technol* 2007, 24:214–234. DOI: 10.1175/JTECH1949.1
14. WMO, 1955: *International list of selected and supplementary ships*. 1st ed. WMO Report No. 47, 1955, WMO Geneva, 181 pp. [All versions from 1955 to 1972 available online at http://icoads.noaa.gov/metadata/wmo47/cdmp_1955-72/.]
15. Fairall CW, Bradley EF, Godfrey JS, Wick GA, Edson JB, et al. The cool skin and the warm layer in bulk flux calculations. *J Geophys Res* 1996, 101:1295–1308.
16. Bernie DJ, Woolnough SJ, Slingo JM, Guilyardi E. Modeling diurnal and intraseasonal variability of the ocean mixed layer. *J Clim* 2005, 18:1190–1202.
17. Kennedy JJ, Brohan P, Tett SFB. A global climatology of the diurnal variations in sea-surface temperature and implications for MSU temperature trends. *Geophys Res Lett* 2007, 34:L05712. DOI:10.1029/2006GL028920
18. Kawai Y, Wada A. Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: a review. *J Oceanogr* 2007, 63:721–744.
19. Webster PJ, Lukas R. TOGA COARE: the Coupled Ocean—Atmosphere Response Experiment. *Bull Am Meteor Soc* 1992, 73:1377–1416.
20. Donlon C, Robinson I, Casey KS, Vazquez-Cuervo J, Armstrong E, et al. The Global Ocean Data Assimilation Experiment High-Resolution Sea Surface Temperature Pilot Project. *Bull Am Meteor Soc* 2007, 88:1197–1213.
21. Brasnett B. The impact of satellite retrievals in a global sea-surface-temperature analysis. *Q J R Meteor Soc* 2008, 134:1745–1760.
22. Meteorological Office. *Marine Observer's Handbook*, 8th ed.: London: HMSO; 1963, 148pp.
23. Crawford AB. A replacement at last for ... "The Old Oaken Bucket". *Geo Mar Technol* 1966, 2:25–27.
24. James RW, Fox PT. *Comparative sea surface temperature measurements*. World Meteorological Organisation Reports on Marine Science Affairs, Report No 5, WMO-No. 336, 1972, 27 pp.
25. Ashford OM. A new bucket for measurements of sea surface temperature. *Q J R Meteor Soc* 1948, 74:99–104.
26. Roll HU. Water temperature measurements on deck and in the engine room. *Ann Meteor* 1951, 4:439–443.
27. Tauber GM. *The Comparative Measurements of Sea Surface Temperature in the U.S.S.R*, World Meteorological Organisation Technical Note No. 103. Geneva: WMO; 1969, 141–151. Available from: WMO, 7 bis, Avenue de la Paix, CP 2300-1211 Geneva 2, Switzerland.
28. Kent EC, Taylor PK, Truscott BS, Hopkins JS. The accuracy of Voluntary Observing Ship's meteorological observations—results of the VSOP-NA. *J Atmos Ocean Technol* 1993, 10:591–608.
29. Lumby JR. Modification of the surface sampler with a view to the improvement of temperature observation. *Extrait Du Journal Du Conseil International Pour LExploration de ICES J. Mar. Sci* 1928, 3:339–350.
30. Dietrich G. Systematic errors in the observed surface water and air temperatures at sea and their effect on the determination of the heat exchange between the sea and the atmosphere, Translation of Hamburg. *D. Hydrogr. Inst., Dtsch. Hydrogr. Z* 1950, 3:314–324.
31. Brooks CF. Observing water-surface temperatures at sea. *Monthly Weather Rev* 1926, 54:241–253.
32. Amot A. Measurements of sea surface temperature for meteorological purposes, results of observations from ocean weather station M. *Ann Meteor* 1954, 4:1–11.
33. Tabata S. An evaluation of the quality of sea surface temperatures and salinities measured at Station P and Line P in the Northeast Pacific Ocean. *J Phys Ocean* 1978, 8:970–986.
34. Woodruff SD, Diaz HF, Elms JD, Worley SJ. COADS release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys Chem Earth* 1998, 23:517–527.

35. Kennedy JJ, Rayner NA, Smith RO, Saunby M, Parker DE. Reassessing biases and other uncertainties in sea-surface temperature observations since 1850. *J Geophys Res* 2010 (to be submitted).
36. Marshall, MW. Ocean Traders. *From the Portuguese discoveries to the present day*. London, UK: B. T. Batsford Ltd.; 1989, 192pp.
37. Saur JFT. A study of the quality of sea water temperatures reported in the logs of ships' weather observations. *J Appl Meteor* 1963, 2:417–425.
38. Brooks, CF. *Methods of observation at sea, Part 1 - Sea Surface Temperature, Sections 2 and 4*, World Meteorological Organisation, Technical Note No. 2, WMO Report 26, 1954, 35pp.
39. Brooks CF. Reliability of different methods of taking sea-surface temperature measurements. *J Wash Acad Sci* 1928, 18:525–545.
40. Walden H. On water temperature measurements aboard merchant vessels. *Dtsch Hyrdrogr Z* 1966, 19:21–28.
41. Boyer TP, Antonov JI, Garcia H, Johnson DR, Locarnini RA, et al. World Ocean Database 2005. Introduction, Chapter 1. In: Levitus S, ed. *NOAA Atlas NESDIS 60*. Washington, D.C.: U.S. Government Printing Office; 2006, 182 pp.
42. McPhaden MJ, Busalacchi AJ, Cheney R, Donguy J-R, Gage KS, Halpern D, Ji M, Julian P, Meyers G, Mitchum GT, Niiler PP, Picaut J, Reynolds RW, Smith N, and K Takeuchi, The Tropical Ocean-Global Atmosphere (TOGA) observing system: a decade of progress. *J Geophys Res* 1998, 103:14169–14240.
43. Bourlès B, Lumpkin R, McPhaden MJ, Hernandez F, Nobre P, et al. The PIRATA Program: history, accomplishments, and future directions. *Bull Am Meteor Soc* 2008, 89:1111–1125.
44. McPhaden MJ, Meyers G, Ando K, Masumoto Y, Murty VSN, et al. RAMA: The Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction. *Bull Am Meteor Soc* 2009, 90:459–480.
45. Gilhousen DB. A field evaluation of NDBC moored buoy winds. *J Atmos Ocean Technol* 1987, 4:94–104.
46. O'Carroll AG, Eyre JR, Saunders RW. Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations. *J Atmos Ocean Technol* 2008, 25:1197–1207.
47. Woodruff S, Scott N, Berry D, Bourassa M, Gulev S, et al. Surface in situ datasets for marine climatological applications. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: Noordwijk, The Netherlands: ESA Publication WPP-306, 2010.
48. Lumpkin R, Pazos M. Measuring surface currents with Surface Velocity Program drifters: the instrument, its data, and some recent results, Chapter 2. In: Griffa A, Kirwan AD, Mariano AJ, Ozgokmen T, Rossby T, eds. *Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics*. Cambridge: Cambridge University Press, 2006, 500pp.
49. Emery WJ, Baldwin DJ, Schlüssel P, Reynolds RW. Accuracy of in situ sea surface temperatures used to calibrate infrared satellite measurements. *J Geophys Res* 2001, 106:2387–2405.
50. Reynolds RW, Gentemann CL, Corlett GK. Evaluation of AATSR and TMI satellite SST data. *J Clim* 2010, 23:152–165.
51. Kennedy JJ, Rayner NA, Smith RO. Using AATSR data to assess the quality of *in situ* sea surface temperature observations for climate studies. *Remote Sensing Environ* 2010, (submitted).
52. Reverdin G, Boutin J, Martin N, Lourenco A, Bouruet-Aubertot P, et al. Drifters surface temperature measurements. *J Atmos Ocean Technol* 2010 (Accepted subject to minor revision).
53. Slutz, RJ, Lubker SJ, Hiscox JD, Woodruff SD, Jenne RL, Joseph DH, Steurer PM, and Elms JD, 1985: *Comprehensive Ocean-Atmosphere Data Set; Release 1*. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 268 pp. (NTIS PB86-105723).
54. Brohan P, Allan R, Freeman JE, Waple AM, Wheeler D, et al. Marine observations of old weather. *Bull Am Meteor Soc* 2009, 90:219–230.
55. Wilkinson C, Woodruff SD, Brohan P, Claesson S, Freeman JE, et al. RECovery of Logbooks And International Marine Data: The RECLAIM Project. *Int J Climatol* 2010, Available at: http://www3.interscience.wiley.com/journal/123308067/abstract. DOI: 10.1002/joc.2102
56. Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, et al. Global analyses of SST, sea ice and night marine air temperature since the late nineteenth century. *J Geophys Res* 2003, 108:4407. DOI: 10.1029/2002JD002670
57. Smith TM, Reynolds RW. Bias corrections for historical sea surface temperatures based on marine air temperatures. *J Clim* 2002, 15:73–87.
58. Bottomley M, Folland CK, Hsiung J, Newell RE, Parker DE. *Global Ocean Surface Temperature Atlas (GOSTA)*: United Kingdom Meteorological Office and Massachusetts Institute of Technology, Bracknell, Berkshire, UK. 1990, pp 20 and 313 plates.
59. Kent EC, Kaplan A. Toward estimating climatic trends in SST, part 3: systematic biases. *J Atmos Ocean Technol* 2006, 23:487–500. DOI: 10.1175/JTECH1845.1
60. Thompson DWJ, Kennedy JJ, Wallace JM, Jones PD. A large discontinuity in the mid-twentieth century in observed global-mean surface temperature. *Nature* 2008, 453:646–649. DOI:10.1038/nature06982

61. Woodruff SD, Slutz RJ, Jenne RL, Steurer PM. A comprehensive ocean-atmosphere data set. *Bull Am Meteor Soc* 1987, 68:1239–1250.

62. Woodruff SD, Lubker SJ, Wolter K, Worley SJ, Elms JD. Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. *Earth Syst Monitor* 1993, 4:1–8.

63. Berry DI, Kent EC. A new air-sea interaction gridded dataset from ICOADS with uncertainty estimates. *Bull Am Meteor Soc* 2009, 90:645–656. DOI: 10.1175/2008BAMS2639.1

64. Berry DI, Kent EC, Taylor PK. An analytical model of heating errors in marine air temperatures from ships. *J Atmos Ocean Technol* 2004, 21:1198–1215. DOI:10.1175%2F1520-0426(2004)021%3C1198: AA MOHE%3E2.0.CO;2

65. Trenberth KE, Jones PD, Ambenje P, Bojariu, R, Easterling D, et al. Observations: surface and atmospheric climate change. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, et al., eds. *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York: Cambridge University Press; 2007

66. Brohan P, Kennedy JJ, Harris I, Tett SFB, Jones PD. Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. *J Geophys Res* 2006, 111:D12106. DOI:10.1029/2005JD006548

67. Domingues CM, Church JA, White NJ, Gleckler PJ, Wijffels SE, et al. Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature* 2008, 453:1090–1094. DOI:10.1038/nature07080

68. Palmer MD, Antonov J, Ba rker P, Bindoff N, Boyer T, et al. Future observations for monitoring global ocean heat content. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: Noordwijk, The Netherlands: ESA Publication WPP-306; 2010.

69. Reynolds, RW. Specific contributions to the observing system: Sea Surface Temperatures. *Proceedings of Conference on The Ocean Observing System for Climate—OceanObs 99*, St Raphael, France, 25–27 October, 1999.

70. Zhang HM, Reynolds RW, Lumpkin R, Molinari R, Arzayus K, et al. An integrated global observing system for sea surface temperature using satellites and in situ data: research to operations. *Bull Am Meteor Soc* 2009, 90:31–38.

71. Rayner N, Kaplan A, Kent E, Reynolds R, Brohan P, et al. Evaluating climate variability and change from modern and historical SST observations. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: Noordwijk, The Netherlands: ESA Publication WPP-306; 2010.

72. Smith S, Bourassa M, Bradley F, Kent E, Fairall C, et al. Automated underway oceanic and atmospheric measurements from ships. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: ESA Publication WPP-306; 2010.

73. Freeland HJ, Roemmich D, Garzoli SL, Le Traon P-Y, Ravichandran M, et al. Argo—a decade of progress. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: ESA Publication WPP-306; 2010.

74. Testor P, Meyers G, Pattiaratchi C, Bachmayer R, Hayes D, et al. Gliders as a component of future observing systems. In: Hall J, Harrison DE, Stammer D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: ESA Publication WPP-306; 2010.

75. Boehme L, Kovacs K, Lydersen C, Nøst OA, Biuw M, et al. Biologging in the global ocean observing system. In: Hall J, Harrison DE, Stammer, D, eds. *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference*, vol. 2, Venice, Italy, 21–25 September 2009: ESA Publication WPP-306; 2010.

76. Gentemann CL, Wentz FJ, Mears CA, Smith DK. in situ validation of Tropical Rainfall Measuring Mission microwave sea surface temperatures. *J Geophys Res* 2004, 109:C04021. DOI:10.1029/2003JC002092

77. Merchant CJ, Llewellyn-Jones D, Saunders RW, Rayner NA, Kent EC, et al. Deriving a sea surface temperature record suitable for climate change research from the along-track scanning radiometers. *Adv Space Res* 2008, 41:1–11. DOI:10.1016/j.asr.2007.07.041

