Ocean Sci. Discuss., 9, C1659–C1666, 2013
www.ocean-sci-discuss.net/9/C1659/2013/

Ocean Science Discussions

Interactive comment on "Comparing historical and modern methods of Sea Surface Temperature measurement – Part 2: Field comparison in the Central Tropical Pacific" *by* J. B. R. Matthews and J. B. Matthews

J. B. R. Matthews and J. B. Matthews

georobin@uvic.ca

Received and published: 12 March 2013

We thank AR1 for their comments.

General comments >"This paper presents some potentially rather valuable observations made with 3 different types of bucket (wood, canvas and rubber). However the analysis of the differences, and differences with deeper observations is disappointing. The measurements are not placed in their environmental context. A multivariate analysis would be required to understand whether there are any signals, but only single

variable correlations seem to have been performed."

We did not find any evidence of sizeable bucket cooling in our field experiment. Wind tunnel experiments (Ashford, 1949; Roll, 1951) have shown the rate of bucket sample temperature change to be strongly correlated with air-water temperature contrast and wind speed. Indeed Kent and Kaplan (2006) used air-sea temperature difference as the sole environmental variable in their statistical model to derive systematic errors in bucket and intake temperatures. If bucket cooling had occurred in our experiment, we would thus have expected to have observed strong correlations between the inter-bucket temperature differences (e.g. rubber-canvas) and these variables, but none were found. Nor did we find strong correlations with any of our other measured environmental variables (e.g. cloud cover and relative humidity). Given that our aim was to determine whether there was a sizeable change in the temperature of the wood and canvas bucket samples during the exposure period, we thus consider multivariate analysis unnecessary. We wished to directly study the physical aspects of bucket cooling without use of complex statistical relations.

>"It is not made clear what signals are actually expected in the difference between the bucket types, especially given the short exposure time chosen. A plot in Folland (1991, Climate Research Technical Note, No. 14, Sea temperature bucket models used to correct historical SST data in the Meteorological Office) suggests that for the Folland and Parker (1995) wooden bucket model, cooling of around 0.05degC would be expected for a wind speed of 15 m/s, RH of 75%, SST of 30degC, Air Temp of 28degC. By 4 minutes the cooling was predicted to rise to 0.2degC. For the canvas bucket model the difference is 0.3degC rising to around 1degC for 4 minute exposure. The canvas bucket used in the present study is larger than that either of those modelled by Folland and Parker (1995) so the difference might be expected to be less. So for these environmental conditions, at night, the models suggest that the heat loss difference at 1 minute should be around 0.2degC and therefore detectable."

As noted by AR1, the UK Met Office Hadley Centre bucket models simulate rapid cooling under strong positive sea-air temperature contrasts, enhanced at higher wind speeds. However, it is somewhat unclear whether they would simulate sizeable cooling if adapted to represent the buckets we used, typical sample volumes we obtained and more precise exposure conditions experienced. FP95 modelled canvas and wood buckets filled with 3 and 10L of seawater, respectively whereas our versions of these bucket types contained 8 and 5.5L when two-thirds full (a typical sample). FP95 assumed wooden buckets could be readily filled to the brim, an assumption we think unrealistic based on JBRM's practical experience deploying a bucket of this type. Given the same heat losses as for the Folland (1991) wood and canvas bucket results reported by AR1, but for the typical sample volumes we obtained with our versions of these buckets, we'd expect a detectable cooling of ∼0.1C in 1 minute for each. Assuming our rubber bucket samples did not change temperature during the exposure period, we might thus have expected average rubber-wood and rubber-canvas bucket differences of this magnitude. However, given the different diameter to height ratios of our bucket samples compared to those modelled, it is not clear that the heat loss from our samples would actually have been similar to those simulated even for the same exposure conditions and sample volumes. Further, the Mk II appears to have been modelled as though essentially having a lid. We will include brief discussion of how our results compare to those of Folland (1991) in the revised manuscript (after we have obtained this unpublished paper).

>"One of the main conclusions is that bucket measurements of SST do not need correction, this is based on the 1 minute exposure time observations made in the present study, and a selective review of the literature in Part 1. As it stands the study can make no comment on the accuracy, or otherwise, of the Folland and Parker (1995) adjustments. Their choice of a longer exposure time was based on the instructions provided to the observers and evidence from the signals in the data. Other factors such as whether the sample was stirred are also questioned but the conclusions drawn are not based on any new evidence."

To be clear, our exposure time was not chosen per se, rather we estimated that 1 minute was the typical time taken for observers to obtain a temperature reading following sampling. No instructions were provided to the observers with regards the rapidity with which they should conduct the measurements. The length of both the hauling and on-deck periods was likely fairly invariant given the shortness of the bucket haul (∼2.5m) and the fast response time of the thermometers used.

FP95 justify their use of long exposure times by citing instructions for observers that suggest allowing several minutes for thermometer equilibration. However, reports in the literature suggest the actual time allowed was typically only a minute or less. Brooks (1928) notes that bucket temperature readings by the crew aboard the ocean liner SS Finland took 45-60 seconds. This is consistent with an earlier source, Schott (1893) who reports waiting one minute on average to obtain such a reading. These sources are additional to those originally cited in Part 1 and are discussed in the revised version of that paper. They suggest the early liquid-in-glass thermometers used for bucket readings typically approximately stabilised in ∼1 minute. Hauling periods will have varied with deployment height and observer hauling speed. We suggest they were typically ∼1 minute or less for deployments from heights of 10m and under. Thus we expect hauls for deployments from sailing vessels generally only took a few seconds to 10s of seconds.

Our results are most immediately applicable to wood and canvas bucket temperatures obtained underway aboard sailing vessels in the 19th and early 20th century. The volumes of the buckets used for these measurements is poorly known but may have generally been larger than that of the canvas buckets used later aboard motor vessels (∼2-4L, Brooks, 1926; Brooks, 1928). FP95 modelled a wood bucket of 12L capacity and noted that it was similar to a 19th century ship's bucket in the Scottish Maritime Museum. Sampling with such large-volume buckets would have been easier on slower vessels (less drag for a given bucket volume at a slower ship speed). Our results suggest that seawater samples of large volume (5L and above) in wood and canvas buckets do not change temperature appreciably in one minute in the tropics, independent of air-sea temperature contrast and apparent wind speed.

We suggest that, at least in an average sense, errors in bucket temperatures from sample temperature change are sufficiently small (<=0.15C) that 'correction' is unnecessary. We suggest that removal of intake temperatures from historical SST datasets would improve homogeneity by removing spurious short- and long-term trends from changeover between the bucket and intake methods. Note that generation of a new reference period climatology would be required. We further propose that bucket temperatures thought to suffer from large errors due to sample cooling or warming be discarded rather than adjusted (e.g. exclusion of bucket temperatures collected under strong winds using small-volume canvas buckets). Further field studies are required to investigate the extent to which bucket temperatures obtained using buckets of variable type and volume on vessels of variable speed and freeboard may be influenced by such errors. Even still, identification of bucket temperatures in ICOADS that are likely to be in large error due to environmental influence will be difficult given the limited metadata available (e.g. unknown bucket type and volume, unknown vessel type and exposure time).

>"The results on the flow model are interesting (with the James and Shank 1964 study) in demonstrating that heating of the water in the pipe by the warmth of the engine room is unlikely to be important. However Saur (1963) concludes that other factors are likely to have a greater effect, as noted."

>"Overall the conclusions of the study with regard to the climate record are not well-supported by the evidence presented. It is possible that the measurements made could make a useful contribution to understanding the adjustments applied to historical SST observations, but further analysis would be required. The authors should decide what the goal of their analysis should be and ensure that their conclusions are supported by the analyses presented in the context of a more thorough review of the literature."

Historical SST datasets invariably incorporate both bucket and intake temperatures. The literature provides much evidence of the poor quality of intake temperatures. Comparisons between EIT measurements and more accurate temperatures have generally found large systematic errors, with Brooks (1928) reporting EIT as averaging overly-warm by 0.7C on the Finland and Tabata (1978) finding EITs to average 0.3C too warm on a research vessel. Like Tauber (1969), we thus conclude EITs to be unreliable measures of seawater temperature at intake depth (from a scientific standpoint). Our finding that engine room warming is an unlikely contributor to warm bias in EITs provides additional weight to this conclusion since it suggests that the intake thermometers themselves have often not been accurately measuring the temperature of the intake seawater. This could perhaps be due to more direct influence by the warm engine room environment (e.g. contact with warm metal fittings) and/or poor calibration.

Mean EIT errors are highly variable, often shifting widely between cruises on the same ship. Given this variability and that the cause of these errors is largely unknown, we do not think it possible or appropriate to attempt to correct EIT measurements. Further, on our research cruise we found that in situ temperatures at intake depth were consistently cooler than those at the surface both day and night and between different current regimes. Based on these multiple lines of argument, we suggest removal rather than adjustment of intake measurements is most appropriate. Note that the idea of constructing SST records using data from which measurements from certain methods has been removed is not entirely new. Indeed, Kennedy et al. (2011b) present versions of global- and hemispheric-average SST records based solely on bucket measurements.

While the literature suggests that systematic errors can be prevalent amongst both bucket and intake measurements, we cannot directly determine how prevalent such errors are in historical SST datasets. It is therefore important to test the sensitivity of SST records to inclusion of data obtained using these different methods and evaluate potential causes of systematic errors by field and lab experiments and modelling. Towards this aim, we propose an alternative methodological approach for constructing SST datasets in which bucket measurements are left unadjusted and intake measurements are excluded. We think this approach well-justified based on the evidence in the literature that EIT can be in large systematic error (of unknown cause) and that exposure times for bucket measurements may well have been far shorter than previously assumed.

>"Page 2981, line 5: the tropospheric lapse rate is not the appropriate method for adjusting near surface air temperatures. No analysis using the adjusted air temperatures is presented."

No air temperatures were actually adjusted using this technique. Rather, the tropospheric lapse rate was used to suggest that the difference between the air temperature immediately above the sea surface and that measured on deck would have been minimal given the low deck height. The sentence is superfluous given the limited use of our dry bulb temperatures and so will be removed.

>"Page 2981, line 22: accuracy only if recently calibrated. Was it?"

Yes. The thermosalinograph was calibrated at the end of February 2008. This will be noted in the revised manuscript. The manufacturer suggests calibration should be conducted at least once a year.

>"Page 2982, line 25: 70% relative humidity is actually rather low over the ocean"

The mean relative humidity across all hourly measurements was actually higher at $81 \pm 7\%$.

>"Page 2985, lines 19-22: hard to understand what is being said"

These sentences will be reworded.

>"Page 2986, lines 24-26: Figure 8b looks by eye to have a decreasing difference with increasing ship speed."

A linear least squares regression through the data does yield a negative gradient (-0.02) but the r2 value is only 0.06. This will be mentioned in the revised paper.

>"Page 2991, lines 20-24: In Part 1 it is suggested that the contribution of engine intakes may be substantially underestimated, here it is suggested that there aren't very many so excluding them won't make much difference."

Given that the number of intake measurements pre-WWII is assumed small, if in fact twice the number of observations pre-WWII came from intakes than previously assumed (i.e. a substantial underestimate), this would still comprise a small proportion of the record. Regardless, the suggestion that the EIT contribution may have been larger than thought pre-WWII has been removed in the revised version of Part 1.

>"Page 3016, Fig A1: This diagram is different to that given in Part 1 (Figure 1) which has a sea chest."

We are simply modelling the intake as though being a length of pipe containing moving seawater. The temperature of seawater in a sea chest can be assumed equivalent to that of the external seawater immediately beyond the inlet (Tin in Fig. A1). Sea chests as designed to provide a buffer for the effects of changes in vessel speed, ship movement due to wave action and blockage of the pipework upon intake supply.

Interactive comment on Ocean Sci. Discuss., 9, 2975, 2012.