Ocean Sci. Discuss., 9, C1356–C1365, 2012
www.ocean-sci-discuss.net/9/C1356/2012/

Ocean Science Discussions

Interactive comment on "Comparing historical and modern methods of Sea Surface Temperature measurement – Part 1: Review of methods, field comparisons and dataset adjustments" *by* J. B. R. Matthews

R. Matthews

georobin@uvic.ca

Received and published: 19 December 2012

I thank AR2 for their extensive comments. Again there has evidently been some confusion over what is and what is not being suggested. The manuscript will be revised to relay the key ideas with improved clarity and more strongly emphasise important nuances.

General comments

>"Whilst this paper is generally well written and revisits some interesting papers from

the literature that have lost visibility over the years I don't believe it adds anything new to the literature itself."

The suggestion that the paper adds nothing new to the literature does not stand up to much scrutiny. The review examines the evidence for cooling of bucket samples and warming of intake seawater in unparalleled detail, subjecting oft-cited papers to critical analysis for the first time. Many of the papers discussed are not freely available, with several being in the so-called 'grey literature' (e.g. institutional reports). There has thus been a temptation to cite these papers without first giving them a thorough evaluation, a problem I hope the review will counter. The paper synthesises useful historical details about measurement methods from the literature (e.g. intake depths), which I think an important original contribution in its own right.

>"The author draws conclusions that are at odds with those of the papers reviewed and makes a plethora of statements that are based on personal opinion and without any supporting evidence."

Again there has evidently been some confusion as to exactly what the conclusions of the paper are. These have already been clarified in my response to AR1. The reasoning behind statements AR2 feels unsupported is presented below for the individual cases identified.

Major comments

>"Page 2953, line 20 through to 25: This is purely the authors personal opinion based on, I believe, the authors experience in part two of the paper using a pine bucket and not backed up by any of the literature (I'd be happy to be corrected on this point). As such I think it needs to be removed. For example, modern insulated buckets can be heavy, e.g. the bucket issued to German VOS, and have been used from large vessels travelling in excess of 7 knots. It is unclear how this is any less impractical or dangerous than the use of wooden buckets on the early steamships and to me invalidates the authors statement."

OSD

9, C1356–C1365, 2012

Interactive Comment

The key piece of reasoning behind my suggestion that canvas buckets would be favoured over wooden buckets for SST measurement aboard steamships is the different way in which they behave in the water. This is indeed based on my practical experience deploying wood and canvas buckets from a motor-sailing vessel. Our wood bucket bounced along the surface when under-motor, making it difficult to obtain a sample. Once it did dip below the surface, considerable drag was generated on the line due to the bucket's large diameter and inflexible walls. On one occasion when I deployed this bucket under-motor, the drag proved too strong and a lot of line was released, causing the bucket to drift far back towards the stern. I would have been at risk of being pulled overboard had I not been fastened on with a security harness. The canvas and rubber buckets we used did not bounce along the surface when under-motor. However, like the wood bucket, the canvas bucket needed to be towed to obtain a sample, although the drag created was not as large (note that meteorological canvas buckets appear to have typically been smaller than the general-purpose canvas bucket we used). In contrast, our modern rubber meteorological bucket sank near-vertically when deployed, making sample collection quick and easy. The reason for the contrasting behaviour between the bucket types is related to their different surface area to weight ratios. I hypothesize that the German VOS bucket is deliberately made heavy to so that this ratio is small and so the bucket readily sinks without having to be dragged to obtain a sample. I suspect most modern meteorological buckets are designed to do the same.

So I think that in addition to it being impractical to obtain a water sample with a wooden bucket from a steamship, there would also have been a high risk of being pulled overboard and drowned. I thus stand by my point and will add this extended justification to the manuscript.

>"Page 2954, lines 19 – 22: The country of recruitment and nationality of the ships making the observations in ICOADS can be estimated from a combination of the country code, deck and source ID in ICOADS. Using this information the majority of the

observations in ICOADS prior to WWII come from UK, Netherlands and German ships. Whilst they may have been coal burning ships there is no evidence to suggest that they did not use buckets as instructed. The statement that the proportion using EIT may be significantly underestimated is pure conjecture without any further evidence."

Thank you for clarification on why bucket measurements are thought to dominate pre-WWII. My comment arose from a concern that assuming all pre-1941 measurements were by bucket unless otherwise stated (as is done in HadSST3) may greatly underestimate the number of engine intake temperature (EIT) measurements. Presently the metadata is inadequate to investigate this assumption in great detail and so we cannot be certain that the number of intake measurements has not been underestimated. In particular, the difficulty of obtaining bucket samples on motor-vessels and the apparent prevalence of intake thermometers suggest to me this assumption needs further evaluation. Nevertheless, I will remove my comment from the revised version and in its place add your explanation for why bucket measurements are thought to dominate pre-WWII.

>"Page 2959, line 6: The statement “these conclusions cannot not be drawn from averages of noisy bucket – intake temperature differences alone” is not true. As an example I'd ask the author to calculate the t statistic for the James and Fox data and state the statistical significance of the result. Similar statements are also made elsewhere that are equally wrong. In addition to this, the conclusions are not only supported by the data but by an understanding of the physics of the upper ocean and off the processes involved."

Actually it is not being suggested that the average bucket-intake temperature differences quoted are not statistically significant. Indeed, James and Fox (1972) state that their reported average intake-bucket difference of +0.3°C is significant at the 95% level. A t-test calculation shows this to also be significant at the $p<0.01$ level. However, without consideration of the spread in the data and nature of the observations, this statistic can be somewhat misleading. For instance, analysing the histogram pre-

OSD

9, C1356–C1365, 2012

sented by James and Fox (Table III) we find that ∼58% of the differences are positive (intakes warmer than buckets), ∼35% below zero and ∼7% equal to zero. Thus ∼42% of the intake-bucket differences are actually at or below zero, indicating bucket temperatures equal to or warmer than the corresponding intake temperatures. Note that their histogram appears to be missing a few bars, with 94 observations unaccounted for.

What the sentence in question is actually saying is that we cannot determine how close bucket and engine intake temperatures lie to the 'true SST' in the absence of accurate in situ temperature measurements. Here I am taking the 'true SST' to be the actual temperature within the upper few tens of centimetres beneath the surface. We have no assurance that either the bucket or intake temperatures equal the 'true SST' given the potential for errors in each and possible presence of vertical temperature gradients. We cannot separate individual errors in bucket and intake temperatures from relative bucket-intake temperature differences.

>"Page 2960, line 18: Can the author explain the cause of the 0.6C cooling in the bucket measurements relative to the CTD in this study? Could this be explained by drier (i.e. less humid) conditions and hence greater evaporative cooling from the buckets at the time of observation? A quick look at the dew point temperatures from ERA interim suggest that this may be the case for the SURTROPAC 15 values compared to the others."

The finding of a positive average CTD-bucket temperature difference on the SURTROPAC-15 campaign is indeed interesting. The negative average differences found for the two other campaigns can be explained by invoking similar vertical temperature gradients to those reported in Part 2. The different result for SURTROPAC-15 appears to come from the bucket measurements since the average CTD-TSG difference is similar to those for the other campaigns. Since the average CTD-bucket salinity difference is also of anomalous sign for this campaign, the reported average CTD-bucket temperature difference could very well reflect a real environmental signal. Taken together these results would be consistent with a cooling and freshening of the

OSD

9, C1356–C1365, 2012

Interactive Comment

upper few tens of centimetres. Note that both the SURTROPAC-15 and COARE-2 campaigns were conducted during the 1991/92 El Niño event, with COARE-2 being undertaken near its peak intensity. Note also that the CTD-bucket temperature differences are more noisy for the SURTROPAC-15 campaign, with a larger standard deviation reported (0.48 as opposed to 0.34 and 0.22°C).

>" Page 2961, line 7 onwards: Whilst the bucket was not used for collecting the water samples this does not undermine the validity of the experiments. As the author notes, if the bucket were to be used the rates of cooling would be higher due to evaporative cooling through the walls. As a result the cooling rates reported will be a lower limit."

Actually I was implying that the observed cooling rates may have been unrealistically enhanced by suspending the bucket in a windy location and agitating the sample. Without mixing we would expect evaporative cooling to only be apparent near the bucket walls and exposed surface. The experiments do, however, provide the best evidence that seawater samples in small canvas buckets can cool rapidly post-sampling. Note, however, that the majority of this data has never been published outside of the student reports held by the Sea Education Association (SEA). FP95 only used data from a single SEA cruise conducted in February and March 1991.

>"Page 2961, line 15: Why is it unlikely that sailors would have chosen such a location or agitated the water?"

I see no reason why sailors would systematically stand in a windy and/or sun-shaded location. In fact, I would expect most bucket deployments to have been conducted on the leeward side for ease of deployment, as recommended by Brooks (1926). Further, I see no reason why sailors would stir or otherwise agitate the bucket samples unless they were carefully following instructions that recommended doing so. That mariners generally followed such instructions is of itself unlikely. The important question is whether these variants in method actually matter in terms of bucket cooling and this can be readily tested through further field experiments.

>"Page 2961, line 23 to end of paragraph: This is the only direct comparison between buckets described. Measurements using a canvas bucket were found to be 0.3 cooler than coincident measurements using a tin bucket. This increased to 0.6C when made by a member of the ships crew. This is strong evidence that the canvas buckets can undergo significant cooling in contrast to the later conclusion of the author of this paper."

Brooks doesn't attribute the full magnitude of these differences to sample cooling. Of the 0.5°F average difference found between his own canvas and tin bucket measurements, 0.3°F is attributed to evaporative cooling of the water sample and 0.2°F to the initial coolness of the bucket when already wet. He partitions the additional 0.5°F difference found when the quartermasters were making the canvas bucket measurements between further evaporative cooling (0.2°F), cooling by or of the thermometer (0.1°F), reading error (0.1°F) and inexactness in observation time (0.1°F). In reality, I don't think we can invoke physical causation for this additional difference since the quartermasters were only reading their thermometer to 0.5 or 1°F (Brooks was reading his to 0.1°F). Further, the 0.5°F average difference found between Brooks' canvas and tin bucket temperatures is based on only 10 comparisons. Thus I don't think we can take these results to be conclusive evidence for rapid cooling of canvas bucket samples.

>"Page 2965, section 5 (Synthesis and conclusions) onwards: This is the section I have greatest difficulty with. There is very little new presented and the conclusions drawn appear to be at odds with the rest of the paper. As an example of the former, stating that fast response thermometers respond quickly or that accurate measurements can be made when done so carefully are not exactly new ideas or knowledge."

The paper is more concerned with constraining values for descriptive terms like 'fast' and 'accurate' than making the obvious statements suggested. For instance, just how long did historical thermometers take to equilibrate? In what interval were they graduated and to what precision were they read? Further, I think the conclusions of the review (as stated in my response to AR1) are very much consistent with the literature

OSD
9, C1356–C1365, 2012

presented.

>"More seriously, the main conclusion from the author and paper as I read it is that bucket measurements will contain errors of 0.1C due to evaporative cooling at most. This is based on two minor pieces of information contained within the papers reviewed. Firstly, from B26 - the maximum amount of time to take a sample and temperature reading is 1 minute. Secondly, from Folland and Parker (1995; hereafter FP95) the maximum cooling rate experienced by a bucket is 0.1C min-1. Neither of these statements is correct, the sampling time from B26 is likely to be a minimum, not a maximum as evidenced in B26 itself when a member of the ships crew made the measurement rather than Brooks, with average cooling of 0.6C in the canvas bucket sample. As the author of this paper implies, the rates of cooling found by FP95 are also likely to be a minimum due to the buckets not being used to collect the water samples and therefore not experiencing the evaporative cooling through the walls of the bucket. FP95 also not the cooling rates in excess of 0.15C min-1 were found on a number of occasions."

The magnitude of bucket cooling is dependent on cooling rate and exposure time. The latter can be separated into hauling and on-deck periods. Since submitting the paper, my attention has been drawn to several additional estimates of various bucket measurement timings. In Brooks' 1928 paper (recently made available to me) he notes that the mariners aboard the Finland took around 2 minutes to haul up their canvas bucket and suggests they could have reduced this to 1 minute by faster handwork. With regard the on-deck period he notes that readings with the ships' slow-response spherical bulb thermometer usually took around 45-60 seconds. FP95 provide a review of recommended waiting periods for thermometer equilibration. They note that eleven sources published between 1857 and 1925 suggest a waiting period of 2-3 minutes while ten sources published between 1877 and 1981 recommend a waiting period of one minute or less. Personally I think it unlikely that merchant mariners would have waited longer than one minute to get a temperature reading, even if the recommendation was to do so. Even still, there is scope for exposure times to have ranged between tens of seconds and a few minutes. I suspect most would have been around 1-2 minutes.

>"The author of this paper has made no attempt to explain the various differences found between the results of the different papers. For example, B26 note that fairly large positive values have been found in the bucket – EIT over the Grand Banks when from a naïve (i.e. excluding knowledge of the environmental conditions) we would expect a negative value. B26 goes on to state the differences are largely due to limited cooling of the buckets due to high humidity values and stratification of the ocean surface. Trying to understand the differences in the different papers and explain them would have been more useful than just using them to say everything is too uncertain and that no conclusions can be drawn about the EIT – bucket differences."

I actually think a reasonable attempt has been made to explain these differences, where appropriate. For instance, where the papers reviewed suggest reasons for their findings, these have been repeated. Further, there are some conclusions to be made from the reported bucket-intake temperature differences. Firstly, bucket temperatures have generally been found to average cooler than EIT, although invoking physical causes for this is speculative given the noise in the data. Much of this noise likely reflects poor reading of inaccurate, low-precision thermometers rather than actual temperature change of bucket and intake samples.

>"Finally, fair conclusions from the evidence presented in this paper would be: 1) both EIT and bucket measurements of the SST are problematic and need careful consideration when being used;"

I agree. Oceanographers measure temperature to at least 0.1°C, preferably 0.01°C or greater precision. They would not rely on untrained, non-scientist observers for their measurements.

>" 2) that buckets can undergo significant cooling but that this depends on a number of factors, including but not limited to: the environmental conditions (wind humidity, air – sea temperature difference) at time of sampling; the rigorousness with which the

measurements; the type of bucket being used etc;"

Yes, bucket samples can cool rapidly, with the experiments conducted by SEA in the early 1990s providing some of the best evidence. However, several of the field experiments cited in support of this idea (including the original SEA experiments) suffer from significant methodological limitations (e.g. poor observing, historically-unrealistic methods). Further, as reported in Part 2, we found no evidence of bucket cooling in our field experiment. Thus I am in favour of additional field experiments to further test this idea.

>" 3) and EIT measurements can contain significant biases relative to the surface temperature. These are due to heating of the water in the intake pipes (warm bias) and when there is significant stratification of the surface waters (cold bias)."

As demonstrated in Part 2, engine room warming is an unlikely explanation for warm bias in EIT.

>"None of these are new and are already well documented in the literature (including that reviewed by the author)."

As outlined in my response to AR2's general comments, this paper makes several important original contributions to the literature. Indeed, AR1 notes that it would be a useful contribution (after some revision).

Interactive comment on Ocean Sci. Discuss., 9, 2951, 2012.

OSD

9, C1356–C1365, 2012

Interactive Comment