

This discussion paper is/has been under review for the journal Ocean Science (OS).
Please refer to the corresponding final paper in OS if available.

Comparing historical and modern methods of Sea Surface Temperature measurement – Part 1: Review of methods, field comparisons and dataset adjustments

J. B. R. Matthews

School of Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada

Received: 3 August 2012 – Accepted: 27 August 2012 – Published: 20 September 2012

Correspondence to: J. B. R. Matthews (georobin@uvic.ca)

Published by Copernicus Publications on behalf of the European Geosciences Union.

OSD

9, 2951–2974, 2012

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Sea Surface Temperature (SST) measurements have been obtained from a variety of different platforms, instruments and depths over the post-industrial period. Today most measurements come from ships, moored and drifting buoys and satellites. Shipboard methods include temperature measurement of seawater sampled by bucket and in engine cooling water intakes. Engine intake temperatures are generally thought to average a few tenths of a °C warmer than simultaneous bucket temperatures.

Here I review SST measurement methods, studies comparing shipboard methods by field experiment and adjustments applied to SST datasets to account for variable methods. In opposition to contemporary thinking, I find average bucket-intake temperature differences reported from field studies inconclusive. Non-zero average differences often have associated standard deviations that are several times larger than the averages themselves. Further, average differences have been found to vary widely between ships and between cruises on the same ship. The cause of non-zero average differences is typically unclear given the general absence of additional temperature observations to those from buckets and engine intakes.

Shipboard measurements appear of variable quality, highly dependent upon the accuracy and precision of the thermometer used and the care of the observer where manually read. Methods are generally poorly documented, with written instructions not necessarily reflecting actual practices of merchant mariners. Measurements cannot be expected to be of high quality where obtained by untrained sailors using thermometers of low accuracy and precision.

1 Introduction

Sea Surface Temperature (SST) is a fundamental geophysical parameter. SST observations are used in climate change detection, as a boundary condition for atmosphere-only models and to diagnose the phase of the El Niño-Southern Oscillation (ENSO).

OSD

9, 2951–2974, 2012

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The importance of SST to climate science is reflected in its designation as an Essential Climate Variable of the Global Climate Observing System.

Here I review historical and modern methods of SST measurement, field comparisons of measurement methods and adjustments applied to datasets to reduce heterogeneity generated by variable methods. Section 2 describes measurement methods and changes in their prevalence over time. Section 3 reviews studies evaluating shipboard methods by field experiment. The relevance of these studies to homogenisation of SST datasets is assessed by critical analysis of their methodologies and results. Adjustments developed for bucket and engine cooling water intake temperatures are described in Sect. 4. Synthesis and further discussion is provided in Sect. 5.

2 History of SST measurement

Shipboard SST measurements are obtained aboard merchant, navy and scientific vessels. Ships that record SST and other meteorological variables are known as Voluntary Observing Ships (VOS). Today they include container ships, bulk carriers and tankers. International recommendations for SST measurement were first established at the Brussels Maritime Conference of 1853. The conference report proposed that the temperature of surface seawater be measured using wooden buckets (Woodruff et al., 2008). Buckets are thought to have transitioned from predominantly wooden to predominantly canvas between the 1850s and 1920s (Folland and Parker, 1995; referred to as FP95), although other types were used (e.g. tin). I posit that sampling by wooden bucket would have been impractical and dangerous on the steamships that began to replace slower sailing vessels of lower freeboard in the late 19th century. Wooden buckets bounce along the sea surface when suspended from ships travelling at speeds exceeding ~ 7 kt (~ 3.5 ms^{-1}). Canvas buckets are generally lightweight, will not bounce along the surface and can be compacted for storage. They are thought to have remained the dominant bucket type used from the 1920s until their gradual replacement by rubber and other modern “insulated” meteorological buckets in the 1950s

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



and 1960s (Kennedy et al., 2011b). Retrieval of rubber meteorological buckets can be challenging when thrown from the bridge of large modern merchant vessels at say 30 m up and underway at speeds of 20 kt ($\sim 10 \text{ m s}^{-1}$) or more. Hénin and Grelet (1996) note such hauls “can be an arduous and acrobatic process”. Due to supposed differences in their propensities for captured seawater to change temperature following collection, rubber buckets are known as “insulated”, wooden as “partially-insulated” and canvas as “uninsulated” buckets. Evaporation is thought to occur through the walls of canvas buckets but not through the walls of wood or rubber buckets.

A second method of SST measurement evolved with the advent of steamships. To maintain engine temperatures below critical thresholds, large volumes of subsurface seawater were pumped on board for engine cooling. To monitor the efficiency with which the seawater was removing heat from the engine, ships’ engineers began observing seawater temperature in engine cooling water intakes. Meteorologists recognised that intake temperature measured prior to the engine might be representative of seawater temperature at intake depth. Such engine intake temperatures (EIT) are known to have been recorded since at least the 1920s (Brooks, 1926; referred to as B26).

The prevalence of EIT readings in historical SST datasets is largely unknown pre-World War II (WWII) but generally assumed small. However, given that over 80% of the global merchant fleet were coal burning ships at the start of WWII (Bernaerts, 1998), many surely fitted with seawater intake thermometers, their actual contribution could be significantly underestimated. EIT are thought to dominate SST records from 1942–1945 when there was an increase in the proportion of observations coming from US ships, on which this is thought to have been the primary measurement method (Thompson et al., 2008). Furthermore, nighttime bucket sampling would have required a light on deck so was likely avoided during WWII for safety reasons (FP95).

While buckets sample seawater from a fairly consistent depth (the upper few 10s of centimeters), depths sampled by intakes are highly variable. Engine intake inlets are usually close to keel depth to ensure submergence under all sea conditions, with actual

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



sampling depth dependent on ship load. Some ships have dual seawater intakes, one close to keel depth and another 1–2 m higher. The deep intake is used at sea and the upper when in shallow coastal waters or canals. Intake depths reported in the early literature are presented in Table 1. Brooks (1926) reports an intake depth of ~ 7 m on a Canadian Pacific Steamship in the 1920s. James and Shank (1964) estimate intake depths for various US merchant, Navy and Coast Guard observing ships reporting in 1962 and 1963. They defined relations between intake depth and full-load draft for different hull types and categorised observing ships by hull type to estimate their intake depth. Average intake depths for more modern observing ships reporting these between 1995 and 2004 are presented in Table 5 of Kent et al. (2007). Container ships and gas and liquid tankers were found to have intakes at ~ 7–9 m depth while intakes on bulk and livestock carriers often exceeded 10 m. Kent and Taylor (2006) found the average intake depth for observing ships reporting these in 1997 to be 8.4 ± 4.1 m with the deepest being at 26 m. Oceanographic research vessels typically have dedicated seawater intakes for underway scientific measurements, usually sampling at ~ 2–4 m depth. These scientific intakes are distinct from engine intakes in that the intake pipes tend to be of smaller outside diameter (typically 1.25 or 2.5 inches).

With EIT readings traditionally being obtained by ships' engineers for engine monitoring purposes, procedures and instruments have varied from ship to ship and remain unstandardised and poorly documented today. Thermometers are generally mounted within 15 m inboard of the inlet and beyond the seacock (Fig. 1). On modern ships seawater is often piped aboard through a sea chest, a sealed metal box built into the hull with an external grate. In addition to one or more main intake lines, ships can have multiple ancillary lines with temperature measured on several (Saur, 1963).

Two major types of EIT methods can be distinguished; well and faucet. In the well method, a temperature probe is mounted inside a well sunk into the intake pipe to around a third its inside diameter (e.g. Piip, 1974). Wells may be oil-filled and are sometimes referred to as thermometer pockets or thermowells. Rapid conduction across the well casing allows intake temperature to be measured while enabling the probe to be

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



readily removed for maintenance. In some cases the probe may have been directly inserted into the pipe. Both well and direct insertion temperatures are sometimes referred to as injection temperatures. In the faucet method, seawater is sampled from the intake through an attached pipe fitted with a faucet, and its temperature measured externally (e.g. B26; Piip, 1974).

Manually read mercury-in-glass thermometers have been used for both intake (Saur, 1963) and bucket measurements (Collins et al., 1975; Tabata, 1978b). Intake temperatures have also been observed with mercury-in-steel (Collins et al., 1975; Piip, 1974) and platinum resistance thermometers (Tabata, 1978a) and thermistors (Tabata, 1978c). In some cases intake temperatures have been continuously recorded by thermograph (e.g. B26; Piip, 1974). See Hagart-Alexander (2010) for a review of thermometer types and description of thermowells.

In recent decades the number of bucket and engine intake observations has declined, in part due to reduction in the World Meteorological Organization (WMO) VOS fleet from a peak of over 7500 ships around 1985 to under 4000 today (Kennedy et al., 2011b). Shipboard hull contact sensors, that is temperature sensors mounted to the outside or inside of the hull, have increased in prevalence over this period, providing more SST observations than buckets by the late 1990s (Kent et al., 2007). They presently contribute around a quarter of all shipboard measurements (Kent et al., 2010). Other modern dedicated shipboard methods include radiation thermometers, expendable bathythermographs and trailing thermistors.

Since the early 1970s shipboard measurements have been augmented by temperatures from Automated Ocean Acquisition Systems, primarily drifting and moored buoys. Drifting buoys were first introduced in 1978 (Woodruff et al., 2008) and standardised around 1993 (Rayner et al., 2010). While they are purported to measure temperature at a nominal depth of ~25 cm (Kennedy et al., 2007), they oscillate within the wave field such that actual measurement depth can be anywhere within the upper 2 m (Emery et al., 2001). SST has been measured by moored buoys since 1971 (Kent et al., 2010), with observations in the equatorial Pacific since 1984 when the Tropical Atmosphere

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Ocean array was established. Around 70 % of in situ observations were collected by moored and drifting buoys in 2006 (Kennedy et al., 2011b).

Accurate satellite measurements of SST have been available since 1981, following development of the Advanced Very High Resolution Radiometer (Reynolds, 1999), which measures in the thermal infrared. SST has also been measured by satellite-mounted passive microwave radiometers since 1997 (Wentz et al., 2000). These possess an advantage over satellite-mounted infrared sensors in that microwaves penetrate clouds with little attenuation. Satellite observations have greatly improved spatial coverage, particularly in the Southern Ocean where in situ sampling remains sparse.

3 Field evaluation of shipboard methods

3.1 Bucket-intake temperature comparisons

Field evaluations of SST measurement methods have focused on average differences between bucket and engine intake temperatures. Brooks (1926) compared tin bucket and engine intake temperatures collected aboard the Canadian Pacific Steamship *Empress of Britain* on a cruise between New York and the West Indies in February and March 1924. Faucet and injection intake temperatures were found to respectively average 0.1 °F (~0.06 °C) and 0.5 °F (~0.3 °C) warmer than near-contemporaneous tin bucket temperatures. The injection thermometers were mounted at the condenser intake pumps and noted as difficult to read to better than 1 °F (~0.6 °C). The resolution of the fast-response cylindrical bulb thermometer used to measure the tin bucket and faucet temperatures appears to have been 0.1 °F. This was not the thermometer in standard use for bucket measurements aboard the *Empress of Britain*, rather a longer-response spherical bulb thermometer read to 0.5 or 1 °F was used. Brooks suggests the intake seawater may have been warmed prior to measurement and seawater sampled with the tin bucket cooled. Finding an upper difference between faucet and tin bucket temperatures of only 0.25 °F (~0.1 °C), Brooks concluded the upper ocean

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



well-mixed to at least the intake depth (~ 7 m). Brooks, however, noted that sizeable positive average bucket-intake temperature differences were reported in spring and summer in the Grand Banks by a different vessel, attributed to near-surface temperature gradients. Differences across the upper 5 m were found to average 0.6°F ($\sim 0.3^\circ\text{C}$) in daytime and 0.3°F ($\sim 0.2^\circ\text{C}$) at nighttime in this region by the *Tampa* from April to July 1925. Note that 0.6°F was added to the intake readings for supposed parallax error, so the unadjusted measured differences were in fact larger. Similar gradients were found in summertime by James and Shank (1964) for the Western North Atlantic. They found temperature differences between 10 and 30 ft (~ 3 and 9 m) exceeded 0.6°F ($\sim 0.05^\circ\text{C m}^{-1}$) over 15 % of the time in June, July and August while they were $\leq 0.2^\circ\text{F}$ ($\sim 0.02^\circ\text{C m}^{-1}$) over 85 % of the time from September to March. Isothermal conditions were found at least 55 % of the time during the latter period.

Saur (1963) analysed 6826 pairs of bucket and engine intake temperatures obtained aboard 12 US military vessels between May 1959 and January 1962. Three of these vessels were traversing the North Pacific while the remainder were usually stationed ~ 300 miles off the US West Coast. Intake temperatures were found to average $1.2 \pm 1.6^\circ\text{F}$ ($\sim 0.7 \pm 0.9^\circ\text{C}$) warmer than bucket temperatures, although the former were only measured in whole $^\circ\text{F}$. Significant variation in average intake-bucket differences was found between ships, ranging from -0.5 to 3°F (around -0.3 to 1.7°C). Mean offsets also varied considerably between cruises on individual ships, in one case between 0.3 and 1.8°F (around 0.2 and 1°C). Specially-designed buckets and thermometers accurate to at least 0.15°F ($\sim 0.1^\circ\text{C}$) were used for the bucket measurements whereas the intake thermometers were only graduated in intervals of 2 or sometimes 5°F (around 1.1 and 2.8°C). Saur reports that a comparison between intake thermometers from five US Coast Guard Weather Ships to an accurate thermometer found errors between -2 and 3.9°F (around -1.1 and 2.2°C).

The most observation-rich bucket-intake comparison ever conducted was that of James and Fox (1972). They analysed 13 876 pairs of near-simultaneous bucket and intake temperatures obtained aboard WMO VOS ships in the open ocean between

1968 and 1970. Although of global distribution, reports were mainly from the North Atlantic and North Pacific shipping lanes. Intake temperatures averaged $0.3 \pm 0.9^\circ\text{C}$ warmer than bucket readings, with the largest differences exceeding $\pm 2.5^\circ\text{C}$.

With the exception of James and Shank (1964), these studies have variously been cited to support the idea that buckets tend to measure cooler than the “true” sea surface temperature and engine intakes warmer (e.g. Kennedy et al., 2011b). However, these conclusions cannot be reached from averages of noisy bucket-intake temperature differences alone. Such differences being relative cannot directly reveal separate problems with bucket and intake measurements. This leads us to a discussion of terminology. The term “bias” is commonly applied to average bucket-intake temperature differences yet seems inappropriate given that buckets and intakes sample at different depths and both may measure in error. Use of the term “correction” to describe adjustment of bucket temperatures to be more consistent with EIT and vice versa is also unsuitable for these reasons.

Separation of individual problems with bucket and intake temperatures in field comparisons requires supplementary temperature measurements of proven high accuracy and precision at the same sampling depths. Studies by Susumu Tabata published in the late 1970s are the most comprehensive conducted in this regard. Tabata (1978a) compared upper ocean temperatures collected over 1956–1976 by Canadian weather ships at Station P and traversing Line P in the Northeast Pacific. Line P is a ~ 1425 km-long transect extending from the coastal waters of Southwestern Vancouver Island, British Columbia to Station P in the mid-Gulf of Alaska (Crawford et al., 2007). The mean difference between temperatures measured by specially-designed meteorological bucket and accurate reversing thermometer in the upper 1 m was $0.04 \pm 0.13^\circ\text{C}$ over 1969–1976, with bucket temperatures thus concluded accurate to $\pm 0.1^\circ\text{C}$. Like Saur (1963), average bucket-intake temperature differences were found to vary between ships and between cruises on the same ship, although standard deviations were more consistent at ± 0.1 to $\pm 0.2^\circ\text{C}$. Mean cruise intake-bucket temperature differences were $-0.02 \pm 0.12^\circ\text{C}$ and $0.18 \pm 0.20^\circ\text{C}$ on two weather ships over 1962–1967,

Historical and modern SST measurement

J. B. R. Matthews

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

and $-0.05 \pm 0.14^\circ\text{C}$ and $-0.02 \pm 0.15^\circ\text{C}$ for two other weather ships over 1967–1976. Bucket temperatures were measured to precision of at least 0.1°C and intake temperatures to 0.2°C . For all but the single positive offset, these average differences can be considered negligible.

5 Tabata (1978b,d) conducted a similar analysis using measurements obtained on a Canadian oceanographic research vessel in the Northeast Pacific in August and September 1975. Only observations coincident with wind speeds exceeding $\sim 6\text{ m s}^{-1}$ were analysed, conditions under which the upper 10 m was considered isothermal. Engine intake temperatures (inlet at 4 m depth) averaged $0.3 \pm 1.2^\circ\text{C}$ warmer than accurate temperatures from Salinity-Temperature-Depth (STD) meter. Tabata attributed the large standard deviation to reading errors of the intake thermometer by the engine room crew, with the largest differences exceeding $\pm 2^\circ\text{C}$. A minor mean difference of $0.04 \pm 0.15^\circ\text{C}$ was found between rubber bucket and 4 m STD temperature.

10 More recently, Hénin and Grelet (1996) compared meteorological bucket temperatures to Conductivity-Temperature-Depth (CTD) temperatures at 1–2 m depth from research vessels in the Western equatorial Pacific. Bucket temperatures were found to average $0.13 \pm 0.34^\circ\text{C}$ and $0.16 \pm 0.22^\circ\text{C}$ warmer than CTD temperatures on two campaigns and $0.60 \pm 0.48^\circ\text{C}$ cooler on an additional campaign. The warm average differences may be attributable to strong temperature gradients over the upper few meters.

20 3.2 Canvas bucket experiments by the Sea Education Association

The accuracy of canvas bucket temperatures was tested by field experiment in the early 1990s aboard the Sea Education Association (SEA) sailing vessel *Corwith Cramer*. The *Cramer* is the Atlantic sister ship of the *Robert C. Seamans* used in Part 2 of this study, the *Seamans* operating in the Pacific. The experiments, undertaken for the late Reginald Newell of the Massachusetts Institute of Technology, were conducted over several cruises in the North Atlantic and Caribbean. They are described in a series of student project reports in the SEA archives in Falmouth, MA, USA. Underway at between 20 and 40 locations on each cruise, a replica UK Met Office Mk II canvas meteorological

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



bucket was filled with surface seawater and the sample temperature recorded once a minute for 10 min. The Mk II was used to obtain SST measurements in the 1930s and 1940s (FP95). Cooling rates of $\sim 0.05\text{--}0.10^\circ\text{Cmin}^{-1}$ were generally reported, although rates below $0.02^\circ\text{Cmin}^{-1}$ and exceeding $0.15^\circ\text{Cmin}^{-1}$ were sometimes found.

5 FP95 compared measured cooling from one of these cruises to results from their canvas bucket model (discussed in Sect. 4).

These experiments suffer from several historically unrealistic procedural choices. Firstly, the replica canvas bucket itself was often not used for seawater collection, apparently due to concerns this valuable bucket would be damaged. Instead, a plastic bucket was used for sampling and the seawater then poured in to the canvas bucket, thus keeping its walls dry. Second, the canvas bucket was generally hung on a hook in a wind-exposed, sun-shaded location on deck during the 10 min measurement period and agitated every half minute to mix the seawater. Whilst the Mk II must have been held for measurement since its wooden lid and unstiffened canvas construction cause it to collapse when placed on deck, it seems unlikely sailors would have deliberately chosen such a deck location or agitated the seawater sample. Note also that the Mk II is of relatively small volumetric capacity, unrepresentative of larger canvas buckets used historically (discussed further in Part 2).

3.3 Comparisons between bucket types

20 Few comparisons between different bucket types have ever been conducted. James and Fox (1972) report average bucket-intake temperature differences for various bucket types but no direct differences between bucket types. B26 compared canvas and tin bucket temperatures using a canvas bucket that could be placed on deck for measurement. When dropped from the bridge it measured an average of 0.5°F ($\sim 0.3^\circ\text{C}$) cooler than a tin bucket launched from a lower deck, increasing to 1°F ($\sim 0.6^\circ\text{C}$) when the quartermasters took the canvas bucket measurements rather than Brooks himself. The latter result is suggested in part due to the quartermasters' use of the ship's

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



slow-response spherical bulb thermometer rather than the fast-response thermometer used for the other measurements.

An important methodological consideration for bucket measurements is the length of time between sample capture and temperature observation, the so-called “exposure time”. This can be partitioned into a hauling phase and on-deck period. Brooks reports that quick hauls with a 4-quart ($\sim 0.004 \text{ m}^3$) tin bucket took 20 to 30 s from 10 to 20 ft ($\sim 3\text{--}6 \text{ m}$) up on the leeward stern of the *Empress of Britain*, equating to hauling speeds of $\sim 0.15\text{--}0.2 \text{ m s}^{-1}$. Given that his corresponding reported exposure time was $\sim 1 \text{ min}$, it appears the on-deck phase and thus the equilibration period for the fast-response thermometer used was around 30 s. This is consistent with his statement that quick reservoir thermometers become nearly-stationary “within a small fraction of a minute”. The on-deck period has generally been assumed much longer than this (2–3 min or more) based on early written specifications for the length of time liquid-in-glass thermometers should be immersed prior to reading (FP95). However, written instructions do not necessarily equate to the actual practices of merchant mariners. Piip (1974) noted sailors typically received little verbal instruction and that posted written instructions were poorly followed. Further, reporting a discussion, B26 notes “Professor Ward told of the difficulties of getting men to be accurate who are not interested and who do not want to take the observations”. Thus given that bucket temperatures were additional to mariners’ principal activities and without specific recompense, it seems likely they were done in haste and recommended thermometer equilibration periods generally ignored. Written specifications may also have suggested waiting periods longer than strictly necessary for slow-response thermometers to achieve approximate stabilisation. Response time of liquid-in-glass thermometers is almost entirely dependent on bulb diameter (Nicholas and White, 2001), being longer for larger diameter bulbs.

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4 Bucket and engine intake temperature adjustments

FP95 developed physical models for temperature change of seawater samples in wood and canvas buckets following collection. Modelled temperature change is dependent on air-seawater temperature difference, relative humidity and apparent wind speed. Different versions of the models were developed by altering parameters such as ship speed and bucket exposure to solar radiation. Two canvas buckets of different dimension were modelled, one the size of the Mk II and the other around half its diameter at 8 cm. Adjustments were derived for both “fast” and “slow” ships to represent motor and sailing vessels, with ship speed set to 7 and 4 ms⁻¹, respectively.

The FP95 bucket models are particularly sensitive to choice of exposure time. For canvas bucket adjustments in non-equatorial regions with appreciable seasonal SST cycles and sufficient data, exposure time was determined using their finding that cycle amplitudes were generally larger pre-1942 (Folland, 2005). FP95 assumed the larger amplitudes were due to environmental cooling of wood and canvas buckets, the strength of which varies seasonally in their adjustments. Exposure time was altered in 10° latitude bands to find adjustments that would minimize the variance of three pre-1942 30-yr average seasonal cycles relative to the total variance of their complete record. The longest exposure times so derived exceeded 5 min and the shortest were under 2 min. An “optimum integration time” (not reported) was calculated for each model version by averaging over derived times for all 30-yr averages in all latitude bands. The exposure time for the wooden bucket adjustments was set to 4 min everywhere, equating to a 1 min hauling period and 3 min on-deck phase.

To generate final pre-1942 “corrections” the adjustments from different model versions were combined to fit a time-variant ratio of the number of wood to canvas bucket observations and linear increase in ship speed from 4 to 7 ms⁻¹ between 1870 and 1940. The former was set so that the resulting adjustments would minimize the difference between Night Marine Air Temperature (NMAT) and SST anomalies in the tropical Pacific and Southern tropical Indian Ocean between 1856 and 1920. FP95

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



found pre-1942 annual-average Northern and Southern Hemisphere NMAT anomalies were up to 0.5°C larger than the corresponding SST anomalies and attributed this to bucket cooling. It is commonly assumed NMAT and SST anomalies should be similar on seasonal and longer timescales.

5 The FP95 adjustments have been applied with some modifications to pre-1942 bucket temperatures in the UK Met Office Hadley Centre Sea Ice and SST dataset, HadISST (Rayner et al., 2003), and the second and third versions of the Hadley Centre SST dataset, HadSST2 (Rayner et al., 2006) and HadSST3 (Kennedy et al., 2011a,b). Independent bucket adjustments have been applied to the US National Oceanic and Atmospheric Administration's Extended Reconstruction SST version 3, ERSSTv3 (Smith et al., 2008), derived by Smith and Reynolds (2002) using the assumption of similarity between NMAT and SST. Kent et al. (2010) compare bucket adjustments applied to HadSST2 and ERSSTv3. Both generally increase on a global annual-average from the mid-19th century to around 1920 and then plateau through to the late 1930s. In HadSST2 this is due to the FP95 specification of an increased proportion of canvas to wood bucket measurements and "fast" to "slow" ships over this period.

15 As of 2008 in situ observations in historical SST datasets had not been adjusted post-1941. Thompson et al. (2008) suggested a need to apply adjustments to more recent observations, arguing an abrupt 1945 drop of ~0.3°C in global-mean SST from HadSST2 was the result of uncorrected methodological changeover. In HadSST3, adjustments have been applied to measurements from buckets, buoys and engine intakes over the duration of the record (1850–2006). The FP95 "fast ship" adjustments are used post-1941, with their wooden bucket adjustments applied to temperatures from modern "insulated" meteorological buckets. A linear switchover from canvas to the latter is specified over the 1950s and 1960s. As in HadSST2, different realisations of the FP95 adjustments were derived by varying bucket model parameters within their supposed uncertainty ranges.

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Multiple realisations of EIT adjustments were also developed for HadSST3. For measurements obtained in the North Atlantic between 1970 and 1994, adjustments were generated from the EIT errors of Kent and Kaplan (2006). Adjustments for other regions and years were derived by taking the best estimate for the average EIT error from the literature to be 0.2 °C too warm. Note that strictly speaking adjustments are intended to be relative to the mix of observations in the respective dataset reference period (in this case 1961–1990) rather than corrections back towards “true” values.

HadSST3 has been combined with the fourth version of the Climatic Research Unit (CRU) near-surface land air temperature dataset, CRUTEM4 (Jones et al., 2012) to produce a new global instrumental surface temperature record, HadCRUT4 (Morice et al., 2012).

5 Synthesis and conclusions

Various techniques have been used to measure sea surface temperature since the mid-19th century. Methods differ in terms of platform (e.g. ship, buoy or satellite), sampling depth and extent of automation (e.g. manual observation and recording). Shipboard methods include temperature measurement of bucket samples and seawater in engine cooling water intakes. Both are generally poorly documented, with many key methodological details unsatisfactorily known. For bucket measurements these include bucket volumetric capacity, thermometer response time, ship speed, height of throw above the waterline and hauling rate. Detailed analysis of Brooks (1926) revealed that fast-response liquid-in-glass thermometers of the 1920s could approximately equilibrate in tens of seconds.

Details of the engine intake method are particularly poorly constrained. This not being a dedicated scientific method, instruments and procedures vary widely between ships. Important methodological considerations include sampling depth, inlet location, pipe size (inside and outside diameter), engine room air temperature, intake flow velocity and pipe length between inlet and thermometer. Many details of shipboard methods

show general changes over time, although with a great deal of heterogeneity. Observing ships themselves have clearly altered dramatically since the 1850s, with a general increase in average speed, freeboard and the deepest drafts. Intake depths on modern merchant vessels appear typically around 7–10 m, although can exceed 15 m.

5 For all methods, measurement quality is critically dependent on the accuracy and precision of the thermometer used, in addition to the care of the observer where manually read. I posit that accurate temperatures can be obtained with either the bucket or engine intake method provided observation is careful and the thermometer used accurate and precise. Tabata (1978a) found mostly small average bucket-intake temperature differences (0.05 °C or below) for weather ships on which both bucket and intake temperatures were precisely measured. Most historical manual shipboard observations come, however, from merchant ships rather than scientific vessels, on which observation was likely generally less careful.

10 Field experiments evaluating SST measurement methods vary greatly in terms of vessel type(s) (e.g. scientific or merchant; sail or motor), method(s) assessed including their sampling depth(s), the spatial and temporal coverage of measurements (e.g. region(s), season(s) and number of observations) and the thermometer(s) used (type, accuracy and precision). Heterogeneity of individual measurement methods must also be considered when assessing the relevance of field evaluations, for instance, bucket volumetric capacity for bucket temperatures.

15 Bucket-intake comparisons generally do not attempt to separate the contribution of real contrasts in ocean temperature from other factors (e.g. bucket cooling) in derived average bucket-intake offsets. Separation of individual errors in bucket and intake temperatures requires additional measurements of high quality at the same sampling depths. In the general absence of these, for average offsets to be considered representative of real ocean temperature contrasts requires that both methods be inherently accurate and precise. Evidently this is generally not the case, with the intake method appearing particularly prone to error. For instance, Brooks (1926) found injection intake temperatures to average 0.5 °F warmer than tin bucket temperatures,

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



yet the intake thermometer was noted as difficult to read to better than 1 °F. Further, Saur (1963) found EIT averaged 1.2 ± 1.6 °F warmer than bucket temperatures yet the intake thermometers were only graduated in intervals of 2 or 5 °F. In both cases the average differences are of opposite sign to that which would be expected from typical near-surface temperature gradients (cooler with depth).

Reported average bucket-intake offsets cannot be considered generally applicable where their associated standard deviations are large. James and Fox (1972) found intake temperatures averaged 0.3 °C warmer than bucket temperatures, yet the spread of bucket-intake differences was so large (standard deviation of 0.9 °C) that this is not readily apparent from the corresponding histogram. This noise is unsurprising given the heterogeneity of the intake method (e.g. variable sampling depth, instruments and observer care) and the large temporal and spatial coverage of their collated observations. Tabata (1978b,d) found EIT to read 0.3 ± 1.2 °C warmer than accurate STD temperatures, attributing the large standard deviation to thermometer reading errors by the engine room crew.

While contemporary thinking holds that engine intakes generally give warmer temperatures than buckets, the reverse is sometimes found, as would be expected from typical near-surface temperature gradients. For example, Brooks (1926) reports daytime bucket temperatures averaging ~ 0.3 °C warmer than EIT at 5 m in the Grand Banks in spring and summer. The magnitude of average bucket-intake offsets has also been found to vary widely. Saur (1963) found average intake-bucket differences varied between -0.3 and 1.7 °C across 12 US military vessels and from 0.2 to 1 °C between cruises on one ship. I thus agree with the conclusion of James and Fox (1972) that “correction” of intake temperatures is not feasible. Certainly taking the best estimate of EIT error from the literature to be 0.2 °C too warm, as per Kennedy et al. (2011b), will significantly underestimate the true error of many individual EIT observations.

Bucket samples are generally thought to evaporatively cool following collection and intake seawater to warm. While the field experiments conducted by the Sea Education Association in the early 1990s suggested canvas bucket samples of small volume can

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



indeed cool post-sampling, they found the extent of cooling to be highly dependent on the time elapsed between sampling and measurement (the exposure time). Applying the ~ 1 min exposure time reported by Brooks (1926) to the SEA canvas bucket cooling rates yields sample cooling of only ~ 0.1 °C at most. This is at or above the precision to which bucket temperatures have generally been obtained. The bucket adjustments derived by Folland and Parker (1995) are mostly much larger than this (several tenths of a °C), in part due to use of exposure times of several minutes. I posit that merchant mariners largely ignored long thermometer equilibration periods specified in early written instructions such that actual exposure times were short.

On the whole, bucket temperatures appear more reliable than engine intake temperatures. Historical bucket readings are more likely to have been carefully observed and recorded given the greater inherent interest of deck crew in weather observation than that for ships' engineers. Thermometers used also appear more frequently to have been of 0.1 °C or °F resolution, while EIT readings were not required to be of such high precision in their traditional principal role as an engine monitoring tool. Engine intake thermometers may have also more frequently been inaccurate, with Saur (1963) describing a study in which several were found to read in systematic error by between -1 and 2 °C.

In Part 2 an original bucket-intake field comparison in the central tropical Pacific is presented and changes to historical SST datasets proposed based on the results and literature analysis presented here. The likelihood of intake seawater being warmed by engine room air is also investigated by physical modelling.

Acknowledgements. I gratefully acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Training Experience (CREATE) Program in Interdisciplinary Climate Science. I thank Brian Matthews for discussion and Andrew Weaver for comments on an early draft.

References

- Bernaerts, A.: How useful are Atlantic sea-surface temperatures taken during World War II, in: Proceedings Oceanology International '98, Brighton, UK, 10–13 March 1998, Oceanology International, 1998.
- 5 Brooks, C. F.: Observing water-surface temperatures at sea, *Mon. Weather Rev.*, 54, 241–253, doi:10.1175/1520-0493(1926)54<241:OWTAS>2.0.CO;2, 1926.
- Collins, C. A., Giovando, L. F., and Abbott-Smith, K. B.: Comparison of Canadian and Japanese merchant-ship observations of sea-surface temperature in the vicinity of present ocean weather station “P”, 1927–1933, *J. Fish. Res. Bd. Can.*, 32, 253–258, doi:10.1139/f75-023, 10 1975.
- Crawford, W., Galbraith, J., and Bolingbroke, N.: Line P ocean temperature and salinity, 1956–2005, *Prog. Oceanogr.*, 75, 161–178, doi:10.1016/j.pocean.2007.08.017, 2007.
- Emery, W. J., Castro, S., Wick, G. A., Schluessel, P., and Donlon, C.: Estimating sea surface temperature from infrared satellite and in situ temperature data, *B. Am. Meteorol. Soc.*, 82, 15 2773–2785, doi:10.1175/1520-0477(2001)082<2773:ESSTFI>2.3.CO;2, 2001.
- Folland, C. K.: Assessing bias corrections in historical sea surface temperature using a climate model, *Int. J. Climatol.*, 25, 895–911, doi:10.1002/joc.1171, 2005.
- Folland, C. K. and Parker, D. E.: Correction of instrumental biases in historical sea surface temperature data, *Q. J. Roy. Meteorol. Soc.*, 121, 319–367, doi:10.1002/qj.49712152206, 20 1995.
- Hagart-Alexander, C.: Temperature measurement, in: Instrumentation Reference Book, 4th edn., edited by: Boyes, W., Butterworth-Heinemann, Burlington, MA, USA, chapter 21, 269–326, 2010.
- Hénin, C. and Grelet, J.: A merchant ship thermo-salinograph network in the Pacific Ocean, *Deep-Sea Res. Pt. I*, 43, 1833–1855, doi:10.1016/S0967-0637(96)00084-2, 1996.
- 25 James, R. W. and Fox, P. T.: Comparative Sea-Surface Temperature Measurements, WMO, Geneva, Switzerland, 1972.
- James, R. W. and Shank, M. K.: Effect of Variation of Intake Depths on Water Injection Temperatures, Marine Sciences Department, US Naval Oceanographic Office, Washington, DC, USA, 1964.
- 30

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M., and Morice, C. P.: Hemispheric and large-scale land-surface air temperature variations: an extensive revision and an update to 2010, *J. Geophys. Res.*, 117, D05127, doi:10.1029/2011JD017139, 2012.

Kennedy, J. J., Brohan, P., and Tett, S. F. B.: A global climatology of the diurnal variations in sea-surface temperature and implications for MSU Temperature Trends, *Geophys. Res. Lett.*, 34, L05712, doi:10.1029/2006GL028920, 2007.

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., and Saunby, M.: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850, 1. Measurement and sampling uncertainties, *J. Geophys. Res.*, 116, D14103, doi:10.1029/2010JD015218, 2011a.

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., and Saunby, M.: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850, 2. Biases and homogenization, *J. Geophys. Res.*, 116, D14104, doi:10.1029/2010JD015220, 2011b.

Kent, E. C. and Kaplan, A.: Toward estimating climatic trends in SST, Part III: Systematic biases, *J. Atmos. Ocean. Technol.*, 23, 487–500, doi:10.1175/JTECH1845.1, 2006.

Kent, E. C. and Taylor, P. K.: Toward estimating climatic trends in SST, Part I: Methods of measurement, *J. Atmos. Ocean. Technol.*, 23, 464–475, doi:10.1175/JTECH1843.1, 2006.

Kent, E. C., Woodruff, S. D., and Berry, D. I.: Metadata from WMO publication No. 47 and an assessment of voluntary observing ship observation heights in ICOADS, *J. Atmos. Ocean. Technol.*, 24, 214–234, doi:10.1175/JTECH1949.1, 2007.

Kent, E. C., Kennedy, J. J., Berry, D. I., and Smith, R. O.: Effects of instrumentation changes on sea surface temperature measured in situ, *Wiley Interdisciplinary Reviews: Climate Change*, 1, 718–728, doi:10.1002/wcc.55, 2010.

Matthews, J. B. R. and Matthews, J. B.: Comparing historical and modern methods of Sea Surface Temperature measurement – Part 2: Field comparison in the Central Tropical Pacific, *Ocean Sci. Discuss.*, 9, 2975–3019, doi:10.5194/osd-9-2975-2012, 2012.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set, *J. Geophys. Res.*, 117, D08101, doi:10.1029/2011JD017187, 2012.

Nicholas, J. V. and White, D. R.: Traceable Temperatures: An Introduction to Temperature Measurement and Calibration, 2nd edn., Wiley, Chichester, UK, 2001.

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Piip, A. T.: A Critical Description of the CSIRO Sea Surface Temperature and Salinity Sampling Program from Merchant Ships, Marine Laboratory, Cronulla, NSW, Australia, 1974.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, 108, 4407, doi:10.1029/2002JD002670, 2003.
- Rayner, N. A., Brohan, P., Parker, D. E., Folland, C. K., Kennedy, J. J., Vanicek, M., Ansell, T. J., and Tett, S. F. B.: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset, *J. Climate*, 19, 446–469, doi:10.1175/JCLI3637.1, 2006.
- Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., Kennedy, J. J., Woodruff, S. D., Smith, T. M., Donlon, C. J., Breivik, L., Eastwood, S., Ishii, M., and Brandon, T.: Evaluating Climate Variability and Change from Modern and Historical SST Observations, in: *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21–25 September 2009, edited by: Hall, J., Harrison, D. E., and Stammer, D., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.71, 2010.
- Reynolds, R. W.: Specific contributions to the observing system: sea surface temperatures, in: *Proceedings of the International Conference on the Ocean Observing System for Climate – OceanObs'99*, Centre National d'Etudes Spatiales (CNES), St. Raphael, France, 25–27 October 1999, 1999.
- Saur, J. F. T.: A study of the quality of sea water temperatures reported in logs of ships' weather observations, *J. Appl. Meteor.*, 2, 417–425, doi:10.1175/1520-0450(1963)002<0417:ASOTQO>2.0.CO;2, 1963.
- Smith, T. M. and Reynolds, R. W.: Bias corrections for historical sea surface temperatures based on marine air temperatures, *J. Climate*, 15, 73–87, doi:10.1175/1520-0442(2002)015<0073:BCFHSS>2.0.CO;2, 2002.
- Smith, T. M., Reynolds, R. W., Peterson, T. C., and Lawrimore, J.: Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006), *J. Climate*, 21, 2283–2296, doi:10.1175/2007JCLI2100.1, 2008.
- 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. Oceanogr.*, 8, 970–986, doi:10.1175/1520-0485(1978)008<0970:AEOTQO>2.0.CO;2, 1978a.

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Tabata, S.: An examination of the quality of sea-surface temperatures and salinities observed recently in the Northeast Pacific Ocean, Pacific Marine Science Report 78-3, Institute of Ocean Sciences, Sidney, BC, Canada, 1978b.

5 Tabata, S.: Comparison of observations of sea surface temperatures at ocean station P and NOAA Buoy stations and those made by merchant ships traveling in their vicinities, in the Northeast Pacific Ocean, J. Appl. Meteor., 17, 374–385, doi:10.1175/1520-0450(1978)017<0374:COOOSS>2.0.CO;2, 1978c.

10 Tabata, S.: On the accuracy of sea surface temperatures and salinities observed in the Northeast Pacific Ocean, Atmos. Ocean, 16, 237–247, doi:10.1080/07055900.1978.9649032, 1978d.

Thompson, D. W. J., Kennedy, J. J., Wallace, J. M., and Jones, P. D.: A large discontinuity in the mid-twentieth century in observed global-mean surface temperature, Nature, 453, 646–649, doi:10.1038/nature06982, 2008.

15 Wentz, F. J., Gentemann, C., Smith, D., and Chelton, D.: Satellite measurements of sea surface temperature through clouds, Science, 288, 847–850, doi:10.1126/science.288.5467.847, 2000.

20 Woodruff, S. D., Diaz, H. F., Kent, E. C., Reynolds, R. W., and Worley, S. J.: The evolving SST record from ICOADS, in: Climate Variability and Extremes during the Past 100 Years, edited by: Brönnimann, S., Luterbacher, J., Ewen, T., Diaz, H. F., Stolarski, R. S., and Neu, U., Advances in Global Change Research, Vol. 33, Springer, Dordrecht, The Netherlands, 2008.

Historical and modern SST measurement

J. B. R. Matthews

Table 1. Intake depths reported from observing ships of various type in pre-1980 literature. All are for ships contemporary to the publication year except Collins et al. (1975) which are from vessels operating during 1927 to 1933. CSIRO is the Commonwealth Scientific and Industrial Research Organisation of Australia.

Reference	Intake depth(s)	Ship type(s)
Brooks (1926)	22–24 ft 15 ft	Canadian Pacific Steamship (<i>RMS Empress of Britain</i>) 2 US Coast Guard Ice Patrol vessels (<i>Tampa</i> and <i>Modoc</i>)
Saur (1963)	10–22 ft	3 US Military Sea Transport Service ships, 9 US Navy Radar Picket ships
James and Shank (1964)	10–32 ft (average of 21 ft)	US merchant, Navy and Coast Guard vessels
James and Fox (1972)	0–9 m	WMO Voluntary Observing Ships
Piip (1974)	2–6 m	Merchant vessels traversing 0–50° S, 135–180° E and reporting to CSIRO
Collins et al. (1975)	5 m	Canadian merchant vessels

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



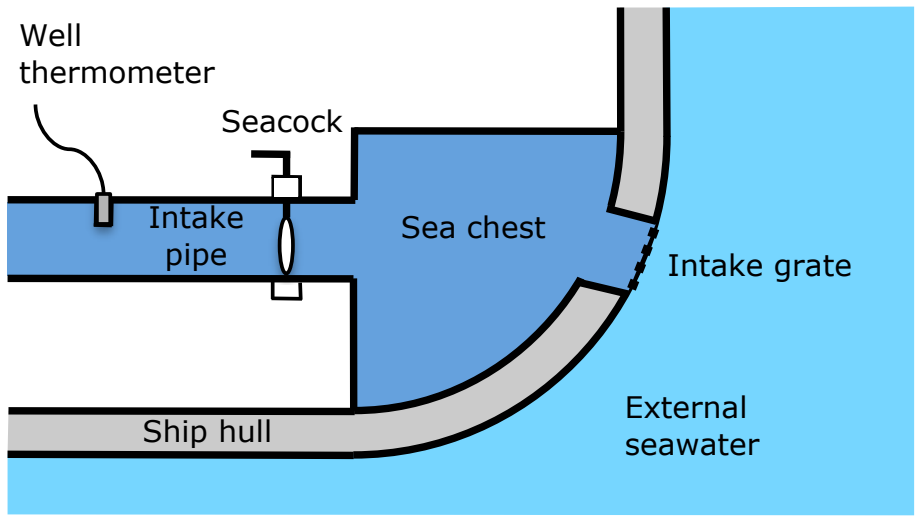


Fig. 1. Schematic of a typical engine cooling water intake system on a modern merchant vessel.

Historical and modern SST measurement

J. B. R. Matthews

Title Page

Abstract Introduction

Conclusions References

Tables Figures

⏪ ⏩

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

