INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **25**: 1007–1022 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.1178

THE EFFECT OF INSTRUMENT EXPOSURE ON MARINE AIR TEMPERATURES: AN ASSESSMENT USING VOSCLIM DATA

DAVID I. BERRY* and ELIZABETH C. KENT
National Oceanography Centre, Southampton, SO14 3ZH, UK

Received 3 May 2004
Revised 25 September 2004
Accepted 11 October 2004

ABSTRACT

Observations of marine air temperature (MAT) by Voluntary Observing Ships (VOS) are known to contain significant biases due to solar heating of the sensor environment. MAT and humidity observations are usually made using wet- and dry-bulb thermometers housed in Stevenson screens, or with psychrometers. These instruments are typically mounted in the bridge wings or on the wheel-house top. If not sited carefully then the instruments can be poorly exposed to the undisturbed environmental conditions and have inadequate ventilation, leading to biased observations of both MAT and humidity.

In this paper we use observations collected as part of the VOS Climate (VOSClim) project to investigate the relationship between instrument exposure and heating errors. The heating errors are estimated as the difference between the observed MAT and the collocated output of a numerical weather prediction model. The instrument exposures are assessed from photographs of the instruments. Currently, photographs of the instruments and sufficient observations exist for 17 VOSClim ships.

Two methods of assessing the instrument exposure using the observations are presented. The first method is based on the skewness of the distribution of estimated heating errors for individual ships. The second method is based on a correction developed to correct the heating errors and uses the ratio of the heating to cooling terms in the correction. When ships are ranked both on the skewness and on the ratio of the heating to cooling terms, there is a statistically significant correspondence between the rankings and the visual assessments of instrument exposure. The skewness of the distribution of estimated errors in MAT is proposed as a simple indicator of instrument exposure. Copyright © 2005 Royal Meteorological Society.

KEY WORDS: VOSClim; ICOADS; ship; air temperature; humidity; exposure; skewness

1. INTRODUCTION

Observations of the marine climate by merchant ships participating in the Voluntary Observing Ships (VOS) programme form an important part of the climate record. These observations have been used in assessments of climate change (e.g. Folland *et al*., 1984; Fu *et al*., 1999; Houghton *et al*., 2001) in the compilation of climatologies (e.g. Rayner *et al*., 2003) and atlases of the surface heat and momentum fluxes (e.g. da Silva *et al*., 1994; Josey *et al*., 1999) and the validation of satellite parameters (e.g. Reynolds and Smith, 1994) and model output (e.g. Josey *et al*., 2001). VOS reports contain a range of surface meteorological parameters, including marine air temperature (MAT), sea-surface temperature (SST), humidity, sea-level pressure and wind speed and direction. VOS meteorological reports are collated, along with observations from moored and drifting buoys and other platforms, in the International Comprehensive Ocean–Atmosphere Data Set (ICOADS; Woodruff *et al*., 1998; Diaz *et al*., 2002). However, in common with all data types, VOS reports are known to contain random and systematic errors and need to be handled with care (e.g. Rayner *et al*., 2003; Kent and Berry, 2005; Kent and Kaplan, in press; Kent and Taylor, in press). The VOS Special Observing

* Correspondence to: David I. Berry, National Oceanography Centre, Southampton, European Way, Southampton SO14 3ZH, UK; e-mail: dyb@noc.soton.ac.uk

Project-North Atlantic (VSOP-NA; Kent and Taylor, 1991; Kent *et al*., 1993a) was therefore set up to assess the error characteristics of observations from the VOS fleet. The VOS Climate (VOSClim) project (JCOMM, 2002) is a successor to the VSOP-NA project and aims to provide a high-quality marine meteorological dataset. One of the main foci for the scientific analysis of the VOSClim dataset is the assessment of biases in VOS weather reports (JCOMM, 2002).

VOS MAT observations are normally made using a dry-bulb thermometer housed in a naturally ventilated Stevenson screen or sling psychrometer. Screens are usually mounted in the bridge wings or on the wheelhouse top. Screens are typically mounted in pairs, one on each side of the ship, so that the windward (better ventilated) screen can be read. Psychrometers are artificially ventilated, either by a mechanical or electric fan or, in the case of the sling psychrometer, manually whirled by the observer. Observers normally make psychrometer observations on the windward bridge wing. The observations are sensitive to the heating of the ship's and sensor environment by solar radiation and typically contain systematic biases due to this heating. These errors have been well documented (Glahn, 1933; Dietrich, 1950; Folland, 1971; Hayashi, 1974; Goerss and Duchon, 1980; Kent *et al*., 1993b; Berry *et al*., 2004). Corrections for these heating errors have been developed (Kent *et al*., 1993b; Berry *et al*., 2004). Naturally ventilated MAT measurements made by buoys also contain heating errors. Anderson and Baumgartner (1998) developed a correction for these errors, which they showed can reach over 3 °C in the tropics under high solar radiation conditions.

The magnitude of the heating errors in VOS MAT observations has been shown to depend on the instrument exposure (e.g. Kent *et al*., 1993b), with poorly exposed instruments having the largest heating errors. Although Kent *et al*. (1993b) conclude that, except for poorly exposed sensors, the errors are independent of exposure, a more careful examination of their results suggests that the heating errors may be up to 30% larger for medium-exposed sensors compared with well-exposed sensors under similar solar radiation conditions. It should be noted that the positioning of the instruments is constrained to locations easily accessible from the bridge; hence, the simple solution of moving the instruments to a better exposed location is not always possible. It is often particularly difficult to site sensors on modern ships with enclosed bridge wings. Although we have information about VOS sensor types and instrument heights from the World Meteorological Organization (WMO) 'List of Selected, Supplementary and Auxiliary Ships' (Publication No. 47; e.g. WMO, 1994), there is no comparable source of information on the exposure of the instruments.

A method to estimate the exposure of air temperature sensors on VOS would allow, for example, the exclusion of MAT reports from ships with poorly exposed sensors from calculations or from gridded datasets. Correction techniques, such as that of Berry *et al*. (2004), can be more effectively applied if it is possible to group together ships which are likely to have similar characteristics, based on likely sensor exposure. Different coefficients can then be used for different VOS groups. Humidity observations are typically made using a wet-bulb thermometer mounted alongside the dry bulb in a screen or psychrometer. Poorly exposed instruments may also give biased estimates of humidity due to inadequate ventilation of the wet bulb (e.g. Met Office, 1981). This would result in overestimation of atmospheric humidity, and a consequent underestimate of evaporative heat transfer from the ocean to the atmosphere. Hence, it would be desirable to develop a method of assessing the instrument exposure to allow quantification of any relationship between instrument exposure and biases in both air temperature and humidity measurements. It should be noted that although air temperature measurements are affected by radiative heating of the sensor environment, the humidity should remain unaffected, as the ship is not usually a source or sink of moisture. If a wet- and dry-bulb psychrometer is used for humidity measurement, then the dew point should be calculated using the measured air temperatures rather than the corrected air temperature (Kent and Taylor, 1996). There may, however, be a correspondence between errors in humidity and temperature, as poorly ventilated sensors are likely to provide elevated temperature estimates and also an underestimate of the wet-bulb depression with a consequent overestimate of atmospheric humidity.

Details of data sources and processing are presented in Section 2. In Section 3.1 the assessment of instrument exposure using photographs of the instrument locations is described. Two methods of assessing the instrument exposure from ship observations are then presented, one based on the skewness of the distribution of heating error estimates for individual ships (Section 3.2) and the other based on a correction for MAT (Berry

et al. 2004; Section 3.3). The three methods of exposure assessment are compared in Section 4. Section 5 is a summary of the results with a discussion of their application.

2. DATA

2.1. Data sources

2.1.1. VOSClim. In this study we use observations from a subset of the VOS fleet taking part in the Joint WMO/Intergovernmental Oceanographic Commission Technical Commission for Oceanography and Marine Meterology (JCOMM) VOSClim project (JCOMM, 2002). The aim of VOSClim is to provide a high quality set of marine meteorological observations merged with the output of a numerical weather prediction (NWP) model with accompanying extensive metadata on the ships, instruments and observing methods. The merged NWP model output allows the quality of the observations to be assessed and, combined with the metadata, should allow the assessment of the error characteristics of individual ships and observing methods.

The metadata have been provided by the participating countries via WMO Publication No. 47. The metadata contains information on *inter alia* the dimensions of the observing ships, methods of observation and instrument heights. In addition, for the VOSClim project, photographs of some ships and instrument locations are available and have been used in this study for the assessment of the exposure of the air temperature sensors.

The NWP model output used as a comparison standard in VOSClim is the background field from the operational global forecast configuration of the UK Met Office Unified Model. The model provides forecast fields every 6 h for the wind speed components, air temperature and humidity at a resolution of 0.83° longitude by 0.56° latitude. The SST is generated once per day independently from the model as a combination of the previous analysis, climatology and 1 day's data. SST data sources used as input include ships, buoys, ocean profiles and advanced very high resolution radiometer (AVHRR) satellite retrievals. These are combined to give an SST thought to be representative of 1 m depth using coefficients derived from comparisons of the satellite retrieval with drifting buoy SST. VOS observations of pressure and wind speed are assimilated into the model, together with radiosonde, satellite and aircraft observations. Some observations will be rejected if they are thought to contain gross errors or fail other quality checks. Temperature and humidity observations from VOS are not assimilated into the global Unified Model; the ship and model air temperatures are thus independent. More details on the model can be found in Met Office (2002), Lorenc *et al*. (1991) and Cullen *et al*. (1997). The forecast fields are interpolated in space and time, and appended to the VOS reports to give collocated model output and VOS observations. It is the difference between each individual VOS MAT report and the collocated NWP forecast MAT that we use in this study as an estimate of the error in the VOS MAT. Both estimates of MAT are adjusted to a common height of 10 m above sea level (see Section 2.3).

We must emphasize that the NWP output is not assumed to be absolutely accurate. Rather, we use the model output to compare statistics of large numbers of observations from widely differing regions and times in order to expose relative differences between ships. This is the same approach used in the VSOP-NA (Kent *et al*., 1993a). There are limitations to this approach and care is needed in the analysis. There is always the concern that errors in the model used as a comparison are misinterpreted as errors in the VOS report. Whilst this can never be completely excluded as a possibility, we have attempted to minimise the likelihood of misinterpretation by ensuring that the ships we are analysing do not report from restricted locations (Table I). This means we are less likely to interpret a regional error in the model as an error in a particular ship. There is a limit to the questions that can be answered using this type of comparison dataset. We cannot, for example, learn about absolute bias. However, if we compare ship and model differences for different types of VOS instrument, then we are likely to learn something about the VOS report.

VOS usually report at 6 h intervals, at 00, 06, 12 and 18 GMT. A smaller number of ships also report at 3 h intervals and some with automatic systems can report hourly. The bulk of the reports in the present study are the 6 h reports; therefore, we have to rely on changes in report longitude, and hence local time, to resolve the diurnal cycle. Again, the possibility of a regional model error being interpreted as a diurnal error is expected to be reduced by the wide geographical variation of the VOSClim ships analysed.

Table I. List of VOSClim ships included in this study. Includes ship callsign, the ship name, the type of ship, the regions in which the ship operates, the period over which observations were made, the height of the temperature sensor and a visual estimate of the air temperature instrument exposure. (E) indicates that the temperature is measured with an electrical resistance thermometer (ERT); other VOSClim ships use mercury thermometers

Callsign	Name	Type	Region	Period (mm/yy)	Height (m)	Exposure
C6KD7	*Chiquita Belgie*	Refrigerated cargo	UK → Caribbean	01/03–08/04	22	Medium
C6LF8	*St Lucia*	Refrigerated cargo	Global	07/02–08/04	25	Good (E)
C6LF9	*Dominica*	Refrigerated cargo	Global	07/02–08/04	25	Good (E)
GBTT	*Queen Elizabeth II*	Passenger ship	Global	07/02–07/04	28	Poor
GLNE	*Discovery*	Research vessel	N. Atlantic	07/02–02/04	17	Good (E)
GVSN	*Oriana*	Passenger ship	N. Atlantic, Mediterranean and South America	01/02–08/04	30	Poor (E)
GXUP	*City of Cape Town*	Container ship	E. Atlantic, UK → S. Africa	07/02–08/04	28	Medium
MHCQ7	*Peninsular Bay*	Container ship	UK → USA via Far East	07/02–08/04	31	Medium (E)
MQEC7	*Newport Bay*	Container ship	N. Hemisphere, global	01/03–08/04	29	Medium (E)
MSTM6	*Providence Bay*	Container ship	N. Hemisphere, global	07/02–08/04	31	Good (E)
MXBC6	*P&O Nedlloyd Southampton*	Container ship	UK → USA via Far East	07/02–08/04	36	Good
MXMM5	*City of London*	Container ship	Global	07/02–08/04	25	Poor
MZGK7	*Glasgow Maersk*	Container ship	N. Hemisphere global	07/02–08/04	28	Poor
ZDLP	*James Clark Ross*	Research vessel	UK → Southern Ocean	09/02–08/04	18	Poor (E)
ZDLS1	*Ernest Shackleton*	Research/supply vessel	UK → Southern Ocean	11/02–05/04	16	Good
ZQAY4	*Grasmere Maersk*	Container ship	UK → Middle/Far East	11/02–08/04	29	Medium
ZQYC5	*P&O Nedlloyd Shackleton*	Container ship	UK → Middle/Far East	07/02–08/04	37	Poor

We should also emphasize that the statistics and results presented in this paper have no utility if we only have the statistics for a single ship. To gain any additional information the statistics for an individual ship need to be compared with the same statistics for other ships, since it is the differences between ships which gives us the additional information. The greater the number of ships we can compare, the greater our confidence will be in the results. Additionally, we expect that absolute values of the exposure-related statistics will vary, depending on the background field used to estimate the MAT errors. However, this will not affect the analysis, since it is the relative differences between ships that we are interested in and not the absolute values.

Table I lists the VOSClim ships used in this study, along with their region of operation, the types of ship, the period over which the observations were made, the height of the instruments and the exposure of the air temperature sensors estimated from photographs of their locations. The ships operate in most regions of the globe and represent a range of types, including container ships, refrigerated cargo ships, research vessels and cruise ships. The air temperature measurements are made at a range of heights between 16 and 37 m, and the exposure of the sensors ranges from good to poor.

All the ships listed in Table I used naturally ventilated screens with a mixture of mercury-in-glass thermometers and remotely read ERTs. As the heating errors in MAT are due to the elevated temperature of

the local sensor environment (Berry *et al*., 2004), we do not expect differences due to the type of thermometer used and expect the characteristics of ERTs and mercury thermometers to be similar. However, the use of ERTs allows the possibility of remote reading of the sensors and the potential for the sensors to be mounted in better exposed conditions further from the bridge (although Table I shows that this is not necessarily the case). We would have liked to examine the relationship between the heating errors in psychrometer observations of MAT and the exposure of the instrument (or measurement location for sling or Assmann-type psychrometers) but no information on instrument exposure was available for VOSClim ships using psychrometers.

2.1.2. VSOP-NA. We have used observations from the VSOP-NA (Kent *et al*., 1993a) to supplement the observations from VOSClim. VSOP-NA was similar to the VOSClim project but limited to the North Atlantic and to the period 1988–90. As with VOSClim, extensive metadata on the ships and observing practices were collected (Kent and Taylor, 1991) and interpolated NWP model output appended to the observations to give collocated model output and VOS observations (Kent *et al*., 1993a). Unlike the forecast model output used in the VOSClim project, the VSOP-NA used the 'analysis' product which was the forecast field with observations assimilated. However, VOS air temperatures were not assimilated (Kent *et al*., 1991), so the ship and model air temperatures in the VSOP-NA can be considered independent. Observations of MAT, including metadata on instrument exposure, were available for analysis from the VSOP-NA data for observations made using both screens and sling psychrometers.

The exposure of the instruments on board the VSOP-NA ships were ranked as part of the project in 10 categories, including information on the fetch over the ship (Table II). We apply the same processing, height correction and data selection as for the VOSClim data. An exception is that, as noted in the next section, no selection for wind direction is required as the VSOP-NA dataset contains direction-dependent exposure ratings.

2.2. Data selection

2.2.1. Selection by relative wind direction. Only observations with a relative wind direction within 45° of either side of the bow have been selected, to reduce any relative wind direction dependence in the results. For example, when the wind is directly over the bow of the *City of Cape Town* (GXUP, Figure 1(a)), the screen will be moderately well exposed. However, for flows from the side there is less obstruction to the air flow and the instrument will be better exposed.

Owing to the relatively high speed of merchant ships, the majority of VOS observations are made with the relative wind within 45° of either side of the bow. For example, before any observations are excluded, 77% of the observations available for the *City of Cape Town* (GXUP) are made with the flow within 45° of the bow. Similar results are found for the other ships used in this study, with the exception of the slower moving

Table II. Exposure index definitions from the VSOP-NA (Kent and Taylor, 1991) and their equivalent exposure rating

Exposure index	Definition	Equivalent exposure
0	Flow fully blocked adjacent to sensor (within 1 m)	Poor
1	Flow fully blocked at medium distance (1 to 4 m)	Poor
2	Flow fully blocked farther away	Poor
3	Flow partially blocked near sensor (within 1 m)	Medium
4	Flow partially blocked at medium distance (1 to 4 m)	Medium
5	Flow partially blocked farther away	Medium
6	Clear flow, long upwind fetch over ship (>30 m)	Good
7	Clear flow, upwind fetch over ship (10 to 30 m)	Good
8	Clear flow, upwind fetch over ship (1 to 10 m)	Good
9	Clear flow, short upwind fetch over ship (<1 m)	Good

Figure 1. (a) Starboard screen on board the *City of Cape Town* (GXUP). The screen has been categorized as having a medium exposure, due to being mounted on open railings, low down and back from the leading edge of the bridge wing. (b) Port screen on board the *Dominica* (C6LF9). The screen has been categorized as well exposed, due to being mounted in an exposed location on the wheel-house top, against open railings and mounted relatively high above the deck. (c) Port screen on board the *Queen Elizabeth II* (GBTT). This screen has been categorized as poorly exposed, due to being close to the deck in a relatively sheltered location and against a solid board. (d) Starboard screen on board the *James Clark Ross* (ZDLP). The screen has been poorly exposed, since it is low down, slightly above a solid metal railing. The screen is also far back from the leading edge of the wheel-house top and situated next to a spotlight

research vessels. The research vessels (the *RRS Discovery* (GLNE), *RRS James Clark Ross* (ZDLP) and the *RRS Ernest Shackleton* (ZDLS1)) report the flow over the bow approximately 50% of the time.

This selection by relative wind direction is not required for the observations from the VSOP-NA dataset since the VSOP-NA dataset contains exposure codes that vary with relative wind direction. Hence, we use all available VSOP-NA observations regardless of the relative wind direction.

2.2.2. Quality control. The VOSClim dataset does not include any quality assurance flags. Outliers have a significant effect on estimates of the shape parameters of distributions, such as skewness, and quality assurance is, therefore, required. After selection by relative wind direction, gross errors were removed by discarding any observation with an absolute ship–model air temperature difference greater than 15 °C, approximately 0.1% of the data (eight observations). The mean and standard deviation of the remaining ship–model differences have then been calculated for each ship individually and the outliers identified as any observation with a

ship–model difference greater than four standard deviations from the mean. Less than 0.5% of the data (32 observations) were identified as outliers and discarded.

The value of four standard deviations was chosen after experimentation with different limits. As the distributions may be strongly skewed, we need to be careful not to truncate the tails of the distribution. The limit of four standard deviations was found to give a good compromise between the removal of outliers and the preservation of the distribution shape. To minimize the effects of any remaining outliers we have used the third L-moment (Hosking, 1990, 2000) as a robust measure of the skewness rather than the more commonly used third central moment (von Storch and Zwiers 1999).

2.3. Height correction

The air temperature observations have been height corrected to remove any bias introduced through the differing model and observation heights. The model output is at a nominal height of 1.5 m and the ship observations are at varying instrument heights (Table I). We have height corrected the data using a neutral temperature profile to correct both the observations and the model to a reference level of 10 m above sea level. There may be additional errors in the air temperature data due to the effects of flow displacement. If there is significant flow distortion, then the air reaching the measurement location may have originated at a different level. Little work has been done, but this displacement could be of the order of 5 m (Yelland *et al*., 2002). If the air temperature adjusts quickly to the displacement, then the effect will be negligible; if not, then the uncertainty is still expected to be an order of magnitude smaller than the heating errors that are the subject of this study.

The wind speed observations (required to calculate the relative wind speed in the correction of Berry *et al*. (2004)) are height corrected using a neutral wind profile by the logging software (TurboWin; Benschop, 1996) used on board the UK VOSClim ships. No additional height correction is therefore required.

3. ESTIMATION OF INSTRUMENT EXPOSURE

3.1. Visual estimation of instrument exposure

Before we can attempt to determine the exposure of each ship's sensors from the observations themselves, we need an independent estimate of the exposure. As part of the VOSClim project, photographs of the ships and their instruments have been collected. Currently, only photographs for some UK and Australian recruited ships are available, all of which use screens. Based on these photographs, the height of the instruments above the deck, physical obstructions and the instrument mountings visible in the photographs, we have classified the exposures as either good, medium or poor (see Table I).

Figure 1(b) gives an example of a screen rated as well exposed on board the *Dominica* (C6LF9). The screen is mounted relatively high above the deck, on open railings and on the outside edges of the wheel-house top. Figure 1(a) gives an example of a medium exposure rated screen on board the *City of Cape Town* (GXUP), which is mounted in the bridge wings. It is mounted fairly low down, close to the deck and quite far from the leading edge of the bridge wing. However, the screen is mounted on open railings and there are no major obstructions nearby. Figure 1(c) gives an example of a poorly exposed screen on board the *Queen Elizabeth II* (GBTT), which is mounted low down, in a sheltered location and against a solid wood board in the bridge wings. Figure 1 shows that that the exposure classification is subjective and depends on the photographs. For example, obstructions may be just out of shot or the instruments may be sited next to a spotlight or vent that is not clearly visible in the photographs.

3.2. Exposure based on skewness

We use the individual height-corrected VOSClim ship–model air temperature differences as an estimate of the errors in the observations. We cannot use the magnitude of the error estimate to assess the instrument exposure, as this value will depend on the region of operation through the effect of solar radiation and the

prevailing winds (Berry *et al*., 2004). If these estimated errors result solely from random errors in either the observation or model output, then the errors should be normally distributed. However, non-random errors, such as the heating bias, could introduce skewness into the error distribution. In the case of MAT radiative heating errors we would expect the distribution to have an increasing positive skewness with increasing heating errors. Since we expect instruments with a poor exposure to have larger heating errors, we would also expect the distribution of observations from a poorly exposed instrument to be more positively skewed than that from a better exposed instrument.

3.3. Exposure of instruments based on the correction of BKT

A correction for the solar radiative heating errors in ship-based observations of the MAT was developed by Berry *et al*. (2004). They use the heat budget of the ship, balancing the solar radiative heating of the sensor with convective and conductive cooling and the storage of heat:

$$\frac{\mathrm{d}(\Delta T)}{\mathrm{d}t} + \overbrace{x_2(x_3 V^{x_4} + x_5)\Delta T}^{\text{convective and conductive cooling}} = \overbrace{x_1 R_{\mathrm{SW}}}^{\text{solar radiative heating}} \tag{1}$$

where ΔT (°C) is the radiative heating error, t(s) is time, V (m s^{-1}) is the relative wind speed and R_{SW} (W m^{-2}) is the incident solar radiation calculated using the Okta model of Dobson and Smith (1988). x_1 through to x_5 are coefficients which are determined empirically and relate to the thermal properties of the ship and the exposure of the sensors. Full details can be found in Berry *et al*. (2004).

The heating errors are modelled as a function of the relative wind speed and the incident solar radiation through the fitted coefficients (Berry *et al*., 2004: equation (1), table II). The values of the coefficients x_2 through to x_5 (Equation (1)) depend on the exposure of the instrument location, whereas the value of the coefficient in the solar radiative heating term x_1 depends on the absorptivity of the sensor environment. The coefficients x_1 and x_2 also depend on the mass and thermal properties of the ship. x_1 and x_2 have the same dependence on the mass and thermal properties of the ship; hence, using the ratio of the heating and cooling terms (Equation (2)), the dependence on the mass and thermal properties (and hence on ship size and type) will cancel. The ratio should then be largely dependent on differences in the exposure of the instruments (coefficients x_3 through to x_5), which we expect to have a large variation (Figure 1(a)–(d)) plus a contribution from the absorptivity of the sensor environment (through x_1). A representative constant relative wind speed of 5 m s^{-1} is used in calculating the ratio for consistency. The contribution of the solar radiation component R_{SW} can be ignored, since a constant value would only act as a scaling factor to the ratio. Our indicator of the instrument exposure is given by

$$\text{Ratio} = \frac{x_1}{x_2(x_3 V^{x_4} + x_5)} \tag{2}$$

4. RESULTS

4.1. Skewness of the error distribution and instrument exposure

4.1.1. VOSClim. Figure 2(a) and (b) shows histograms of the ship–model differences for the *Dominica* (C6LF9) and the *Queen Elizabeth II* (GBTT). The histogram for the *Dominica* (C6LF9) is negatively skewed (Figure 2(a)) and its instruments are well exposed (Figure 1(b) and Table I). The histogram from the *Queen Elizabeth II* (GBTT) is positively skewed (Figure 2(b)) and its instruments are poorly exposed (Figure 1(c) and Table I).

Table III lists the VOSClim ships sorted by the skewness of the error estimates. The error ranges quoted are the bootstrapped confidence intervals calculated using a Monte Carlo simulation with 1000 runs (von Storch and Zwiers, 1999). Also listed in Table III are: the estimated instrument exposures (as in Table I); the

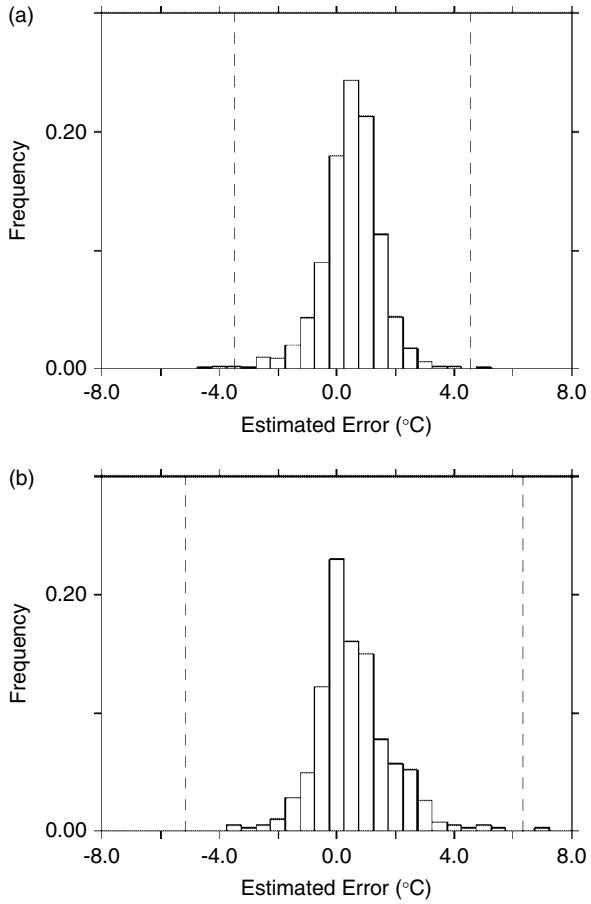

Figure 2. Normalized histograms of estimated MAT errors for individual ships. The quality assurance limits used (four standard deviations from the mean) are indicated by the vertical dashed lines. (a) *Dominica* (C6LF9) using VOSClim ship–model differences as the error estimate; (b) as (a), but for the *Queen Elizabeth II* (GBTT)

mean error estimate, its range and the number of observations. Table III shows a good agreement between the exposure based on the skewness estimates and from the photos. The Spearman rank correlation coefficient (e.g. von Storch and Zwiers 1999) between the ranking based on the skewness and the mid-point ranking based on the photographs gives a value of 0.84, significant at the 99% confidence level. The four ships with the lowest skewness distributions have instruments that are all rated as well exposed; the four ships with the highest skewness are all rated as poorly exposed. All of the sensors judged to be well exposed are in the top half of the table, and those judged to be poorly exposed in the lower half; the medium-exposed sensors tend to be clustered in the centre of the table.

We should remember that the visual estimates of exposure may not be correct, as the classification is fairly subjective. For example, Figure 1(d) shows the starboard screen on board the *James Clark Ross* (ZDLP). We have classified the exposure as poor as it is quite low down, just above a solid metal railing. Additionally, the screen is far back from the leading edge of the wheel-house top and next to a spotlight. However, we could have classified the exposure as medium, since there are relatively few obstructions on the wheel-house top and the flow of air around and through the screen should be relatively unobstructed.

Comparison of Tables I and III shows that the ranking of the ships appears to be independent of both the region in which the ship operates and of the mean heating errors. These results suggest that the skewness of the error estimates is an indicator of the exposure of the instruments and is independent of environmental conditions.

Table III. Range and mean values of the estimated errors (ΔT) for the individual VOSClim ships. Also listed are the skewness of the error estimates, the number of valid observations after preprocessing, and the exposure of the instruments. The errors listed for the mean and skewness are the standard deviation and bootstrapped confidence intervals respectively. The table has been sorted by the skewness

Ship callsign	Range ΔT (°C)	Mean ΔT (°C)	No. of observations	Skewness	Exposure
C6LF9	−3.45 to 3.92	0.51 ± 0.95	871	-0.083 ± 0.019	Good
C6LF8	−5.12 to 5.74	0.48 ± 1.35	726	-0.027 ± 0.025	Good
MSTM6	−4.69 to 5.73	0.58 ± 1.31	715	0.038 ± 0.023	Good
GLNE	−1.91 to 3.60	0.02 ± 0.90	116	0.038 ± 0.062	Good
ZDLS1	−2.93 to 3.80	0.31 ± 1.06	328	0.048 ± 0.028	Good
GXUP	−4.00 to 5.70	0.82 ± 1.21	717	0.048 ± 0.022	Medium
C6KD7	−2.97 to 3.68	0.50 ± 0.90	432	0.049 ± 0.027	Medium
MHCQ7	−3.22 to 4.99	0.48 ± 1.08	765	0.050 ± 0.022	Medium
MXBC6	−2.62 to 4.72	0.80 ± 1.13	586	0.052 ± 0.020	Good
ZQYC5	−2.91 to 4.91	1.01 ± 1.17	587	0.068 ± 0.021	Poor
GBTT	−3.71 to 5.72	0.54 ± 1.26	391	0.095 ± 0.029	Poor
MQEC7	−3.96 to 4.26	0.36 ± 1.23	270	0.096 ± 0.033	Medium
ZQAY4	−4.14 to 5.69	0.72 ± 1.33	297	0.098 ± 0.034	Medium
GVSN	−3.20 to 5.48	0.42 ± 1.24	128	0.134 ± 0.066	Poor
ZDLP	−4.21 to 5.69	0.30 ± 1.24	409	0.142 ± 0.032	Poor
MZGK7	−2.72 to 5.24	0.67 ± 1.25	234	0.145 ± 0.035	Poor
MXMM5	−2.78 to 6.38	1.01 ± 1.49	357	0.166 ± 0.024	Poor

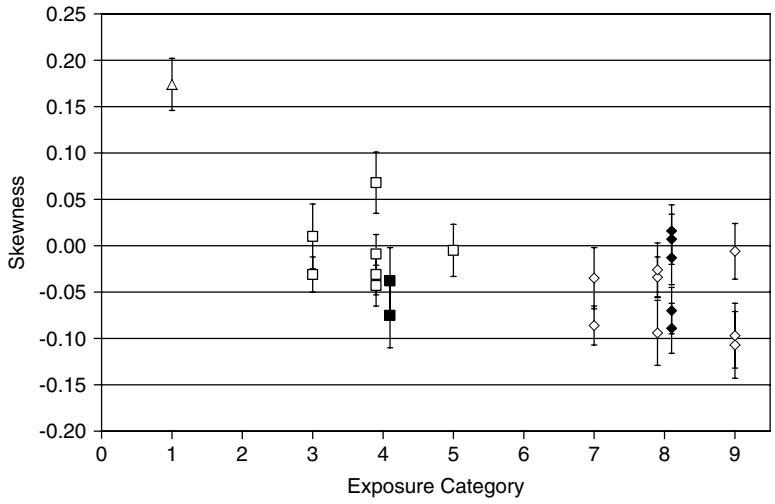

Figure 3. Plot of the skewness against exposure category (0–9) for the VSOP-NA (see Table IV). The open diamonds represent the observations made using screens with a good instrument exposure, the squares a medium exposure and triangles a poor exposure. The filled diamonds represent the observations made using psychrometers with a good exposure and the filled squares a medium exposure. The error bars represent the bootstrapped confidence intervals

4.1.2. VSOP-NA. The skewness has been calculated individually for each different combination of ship and exposure code with more than 200 observations from the VSOP-NA dataset (Table IV). Figure 3 shows the skewness plotted against the exposure. The error bars represent bootstrapped confidence intervals (see Section 4.1.1). The poorly exposed screen (open triangle) has a large positive skewness. The medium-exposed screens (open squares) are scattered around zero and the well-exposed screens (open diamonds) have a negative skewness. There is little difference in the skewness for the different fetch categories (Table II) for medium-

Table IV. Ships from the VSOP-NA, their exposures (see Table II), number of observations and skewness of error estimates. The errors listed for the skewness are the bootstrapped confidence intervals

Ship	Exposure index	No. of observations	Skewness
Screens			
CanMar Ambassador	9	371	-0.107 ± 0.036
Geestbay	9	202	-0.097 ± 0.035
CGM Provence	9	315	-0.006 ± 0.030
Atlantic Cartier[a]	8	303	-0.094 ± 0.035
Edouard LD[a]	8	535	-0.034 ± 0.022
Geestbay	8	505	-0.026 ± 0.029
CGM Provence	7	714	-0.086 ± 0.021
Atlantic Conveyor	7	426	-0.035 ± 0.033
Author	5	341	-0.005 ± 0.028
Geestport	4	622	-0.043 ± 0.022
Geesthaven	4	450	-0.031 ± 0.022
Geestcape	4	622	-0.009 ± 0.021
Nickerie	4	368	0.068 ± 0.033
Margaret Lykes	3	224	0.010 ± 0.035
CanMar Ambassador	3	765	-0.031 ± 0.019
Galveston Bay	1	266	0.174 ± 0.028
Sling psychrometers			
Independent Concept	8	218	-0.089 ± 0.027
Independent Pursuit	8	433	-0.070 ± 0.025
Independent Endeavor	8	263	-0.013 ± 0.029
Nedlloyd Kyoto	8	344	0.007 ± 0.027
Nedlloyd Kingston	8	444	0.016 ± 0.028
Independent Endeavor	4	363	-0.075 ± 0.035
Nedlloyd Neerlandia	4	223	-0.038 ± 0.036

[a] Denotes use of mechanically ventilated POMMAR system, described as a screen in documentation but perhaps more similar to a psychrometer.

(exposures 3, 4 and 5) and well-exposed screens (exposures 7, 8 and 9). This suggests that the fetch has a negligible effect on the heating errors and supports the conclusions of Kent *et al*. (1993b) that the heating errors are independent of ship size and hence fetch. A Spearman rank correlation test between the rank based on the skewness and the mid-point rank based on the VSOP-NA exposure categories shows that they are significantly related at the 99% confidence level with a correlation coefficient of 0.62. Again, this suggests that the skewness shows significant skill at ranking the exposure of the instruments.

The psychrometer observation sites only represented two different exposure categories, codes 4 (medium exposure) and 8 (good exposure). Figure 3 shows that both the medium (filled squares) and good (filled diamonds) exposure-rated psychrometers have a similar, small or negative skewness. The similarity between the medium- and well-exposed psychrometers suggests that the heating errors in psychrometer measurements may be less dependent on the exposure of the instrument than those from screens. There is no significant correlation between the ranking based on the skewness and the VSOP-NA exposure categories for psychrometer observations. However, there are insufficient data to rule out any link between the exposure of psychrometer observations and the skewness of the heating errors, especially as we have no data from poorly exposed psychrometer locations.

4.2. Instrument exposure from the model of Berry et al. (2004)

We have fitted the coefficients in the correction of Berry *et al*. (2004) to the VOSClim data for the ships individually using the Monte Carlo approach described in Berry *et al*. (2004). The fitting routine was run 500

times for each ship using all the observations that pass the data selection (Section 2.2), minimizing the sum of the squared differences between the observed and model MAT values. We need to perform a large number of fits as the solution is strongly non-linear and we want to ensure we find the optimal solution rather than any local minima (Berry *et al*., 2004). The fitted coefficients from the 10 runs giving the largest reduction in the root-mean-square error (RMSE) for each ship (which we call the top 10 runs) have been used to calculate the ratio. The average of these 10 ratio values is then used to assess the instrument exposure. The top 10 runs have been used for each ship since they all give equally good fits, with the exception of the *City of Cape Town* (GXUP, see below), and it is not possible to distinguish between them for the best fit.

Table V lists the values of the average ratio for each ship calculated using a representative relative wind speed of 5 m s^{-1}. Table V also lists the average percentage reduction in RMSE for the top 10 runs, the average residual errors for the top 10 runs, the exposure rating, the skewness (from Table III), and ranking (from Table III), which have all been sorted by the ratio of the heating to cooling terms. The uncertainty values given are the standard deviations of the values for the top 10 runs. As for the skewness estimates of instrument exposure shown in Table III, the ratio estimates of instrument exposure shown in Table V show a reasonable correspondence between the ranking of ships and the exposure of their instruments. Ships with poorly exposed sensors typically appear in the bottom half of the table, and the ships with well-exposed instruments are in the top half. The Spearman rank correlation coefficient between the ranking of the instruments based on the correction of Berry *et al*. (2004) and the photographs is 0.72, statistically significant at the 99% confidence level.

The only ship with a large variation in the percentage reduction in the RMSE is the *City of Cape Town* (GXUP). Using the ratio value from the best fit, i.e. the fit giving the largest reduction in RMSE, decreases the ratio slightly to 2.62×10^{-3}. Using this new value does not significantly change the correlation between the skewness and the ratio. Using the new value to rank the ships improves the ranking slightly, with the *City of Cape Town* (GXUP) moving up the table four places. The correlation between the new ranking and the ranking from the photographs is increased to 0.73.

Table V. Ratio of the heating to cooling terms in the correction of Berry *et al*. (2004) (Equation (2)) from the best fit for each ship using a wind speed of 5 m s^{-1}. Also listed are the percentage reduction in the RMSE and the residual estimated error from the best fit, the skewness and rank from Table III and the exposure of the instruments. The table has been sorted on the ratio of the heating to cooling terms. The errors given for the ratio are the standard deviation of the 10 fits giving the greatest percentage reduction in RMSE for each ship

Callsign	Ratio $\times 10^3$	Reduction in RMSE (%)	Residual error (°C)	Skewness (from Table III)	Rank (from Table III)	Exposure
GLNE	1.72 ± 0.01	3.4 ± 0.2	-0.03 ± 0.00	0.039 ± 0.061	3.5	Good
MHCQ7	1.80 ± 0.02	8.4 ± 0.1	0.13 ± 0.01	0.048 ± 0.021	8	Medium
C6LF9	2.04 ± 0.00	13.1 ± 0.0	0.15 ± 0.00	-0.083 ± 0.019	1	Good
MQEC7	2.35 ± 0.01	11.6 ± 0.0	-0.05 ± 0.01	0.097 ± 0.030	12	Medium
MSTM6	2.35 ± 0.01	7.7 ± 0.1	0.22 ± 0.01	0.038 ± 0.024	3.5	Good
ZDLS1	2.38 ± 0.00	9.1 ± 0.2	0.04 ± 0.01	0.048 ± 0.028	5.5	Good
MXBC6	2.54 ± 0.01	10.8 ± 0.1	0.35 ± 0.00	0.052 ± 0.021	9	Good
C6LF8	2.76 ± 0.04	11.3 ± 0.2	0.08 ± 0.03	-0.027 ± 0.024	2	Good
ZDLP	2.77 ± 0.00	8.2 ± 0.0	0.12 ± 0.00	0.141 ± 0.031	15	Poor
C6KD7	2.79 ± 0.06	22.1 ± 0.0	0.10 ± 0.00	0.047 ± 0.028	7	Medium
GVSN	2.79 ± 0.02	3.1 ± 0.0	0.17 ± 0.00	0.134 ± 0.066	14	Poor
GXUP	2.82 ± 0.02	13.6 ± 1.7	0.37 ± 0.05	0.049 ± 0.022	5.5	Medium
MZGK7	3.22 ± 0.01	28.3 ± 0.1	0.07 ± 0.00	0.144 ± 0.034	16	Poor
ZQAY4	3.38 ± 0.01	22.7 ± 0.1	0.20 ± 0.01	0.097 ± 0.033	13	Medium
ZQYC5	3.54 ± 0.03	24.5 ± 0.3	0.32 ± 0.05	0.066 ± 0.022	10	Poor
MXMM5	3.58 ± 0.02	31.3 ± 0.3	0.13 ± 0.02	0.166 ± 0.024	17	Poor
GBTT	3.61 ± 0.01	17.7 ± 0.1	0.09 ± 0.00	0.097 ± 0.028	11	Poor

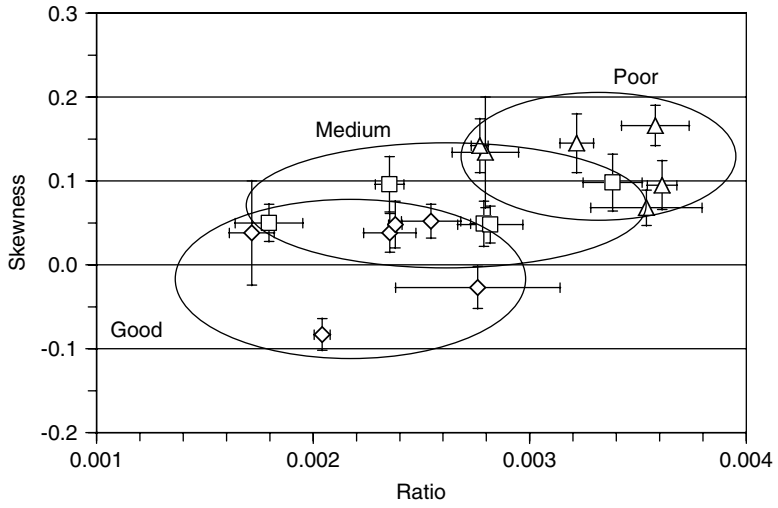

Figure 4. Plot of the skewness (Table III) against the ratio of heating to cooling terms from the correction of Berry *et al*. (2004) (Table V). The diamonds represent the ships with a good instrument exposure, the squares a medium exposure and triangles a poor exposure, based on visual estimates (Table I). The vertical error bars are the bootstrapped confidence intervals for the skewness. The horizontal error bars are the standard deviations of the ratio for the 10 fits giving the greatest reduction in RMSE for each ship. The ovals group together ships in the exposure categories indicated

4.3. Comparison of exposure ratings

Figure 4 shows a plot of the skewness against the ratio of the heating to cooling terms for each of the VOSClim ships. A test of the correlation between the skewness and the ratio values gives a correlation coefficient of 0.55, significant at the 97.5% confidence level. A test of the correlation between the rankings from the two methods gives a Spearman correlation coefficient of 0.61, significant at the 99% confidence level.

5. SUMMARY

The aim of this paper is to find a method of assessing the exposure of air temperature sensors on VOS ships. The exposure of the sensors is expected to have a strong impact on the quality of the measured air temperatures (Kent *et al*., 1993b), but this information is not routinely available from metadata. For two experiments, VOSClim and the VSOP-NA, we have exposure information in the form of photographs of the sensor environments and can make visual assessments of the likely exposure of the instruments. We note, however, that this method, although direct, is fairly subjective and depends on the quality of the photographs.

The first step in assessing the exposure from the MAT observations themselves is to estimate the error in each individual VOS observation. For VOSClim and the VSOP-NA this error is provided by the difference between the observation and collocated model output. It is expected that the statistics of the VOS-model differences can be related to details of how the observations are made (Kent *et al*., 1993a). In the VOSClim dataset the only MAT observations with the required metadata are from ships using screens; in the VSOP-NA dataset, MAT observations from both screens and sling psychrometers were analysed.

Two methods of assessing the exposure of MAT sensors using these error estimates have been presented. The simplest method of assessing the instrument exposure is to relate the exposure to the skewness of the estimated MAT errors for individual ships relative to each other. This method shows statistically significant skill at ranking the exposure of screens on different ships in both the VOSClim and VSOP-NA datasets in an order comparable to visual estimates of instrument exposure made using photographs of instrument sites. The ships with the lowest skewness values, relative to the other ships, have the better instrument exposures. The correlation between the rankings from the skewness and the photographs is significant in both datasets

at the 99% confidence level. For the VSOP-NA screens, the fetch over the ship (Table II) had little effect on the skewness. A correlation between the rankings ignoring the effect of fetch showed the same 99% level of confidence as that including the effects of fetch. However, in the VSOP-NA dataset there was only one poorly exposed screen with enough data for analysis; we cannot exclude the possibility that, with more data, particularly from poorly exposed screens, the effect of fetch could be detected.

No statistically significant difference could be discerned between the skewness estimates for the two medium-exposed measurement sites and the four well-exposed measurement sites for the VSOP-NA psychrometer observations (Figure 3, Table IV). However, lack of data from poorly exposed psychrometer locations means that these results may not give a complete picture.

The second method is more complicated, but relates the exposure directly to the heating errors themselves using the model of Berry *et al*. (2004) and has been applied to the VOSClim data. This method also shows statistically significant skill in identifying ships with instruments in the different exposure categories. The correlation between the rankings based on the model of Berry and the photographs is again significant at the 99% confidence level.

Whilst the correlation between the skewness and ratio exposure estimates (Figure 4) for the individual ships is slightly worse than that between either of the data-based methods and the visual estimates, the correlation is still statistically significant at the 97.5% confidence level. The correlation coefficient between the skewness and the ratio is 0.55. The rank correlation coefficient is 0.61, statistically significant at the 99% confidence level. As the two methods of exposure assessment give comparable results, the simplest method relating instrument exposure to the skewness of the distribution of error estimates is preferred.

It should be emphasized that the results and methods presented in this paper are only valid when we have results from multiple ships. A single skewness (or ratio) value will have no utility on its own, since we will not be able to determine whether the skewness is due to the observations or the comparison field. Only when we compare values from multiple ships operating globally, or in similar regions, will we be able to make any inferences about the exposures.

For other data sources, such as ICOADS, we do not have collocated model output to use in estimating the observational errors. However, preliminary results using differences of individual VOS observations from averages of nearby night-time observations within a window of a few days suggests that errors can also be estimated for ICOADS, but this is work in progress.

Estimates of instrument exposure should have a variety of applications in marine climatology. Knowledge of any relationship between ship characteristics from available metadata, and typical instrument exposure, will guide analysis of MAT and should eventually lead to improved datasets. The VSOP-NA showed that the country of recruitment can be a good indicator of data quality (Kent *et al*., 1993a). The ICOADS MAT observations can be grouped dependent on metadata such as ship length, ship type or country of recruitment, or grouped using estimates of the exposure themselves. This will allow the improved correction of MAT observations for radiation errors (e.g. using the methodology of Berry *et al*. (2004)) for groups of ships judged to have similar instrument characteristics. Alternatively, where there are sufficient observations, an assessment could be made for individual ships. Another application would be to use MAT observations to estimate the exposure of screens and then to analyse humidity observations from instruments in different categories of exposure. Poor ventilation of wet- and dry-bulb thermometers may lead to conditions where the psychrometer coefficient deviates from its assumed value (e.g. Met Office, 1981). An assessment of the characteristics of humidity observations derived from instruments rated by exposure could be used to understand whether such errors lead to detectable bias in flux estimates (e.g. Josey *et al*., 1999).

ACKNOWLEDGEMENTS AND DATA SOURCES

We thank the reviewers of this paper for helping us to make important improvements. This work was funded by grants from the Natural Environment Research Council and the Ministry of Defence Joint Grants Scheme. We would like to thank the participants in the VOSClim project and in particular the Port Meteorological Officers of the UK and Australia for providing photographs of the VOSClim ships and their instruments. The VOSClim data used in this study are freely available from the VOSClim Website

(http://www.ncdc.noaa.gov/VOSClim.html), which is maintained by the US National Climatic Data Center, who act as the VOSClim data collection centre. The UK Met Office provide project management and act as the VOSClim real-time monitoring centre. The WMO provided the metadata. ICOADS was supplied by Steven Worley of the National Center for Atmospheric Research Data Support Section.

We also wish to acknowledge the use of the Ferret program of the NOAA Pacific Marine Environmental Laboratory for analysis and graphics in this paper (http://www.ferret.noaa.gov/). L-moments were calculated with the Fortran L-moments package available from http://lib.stat.cmu.edu/general/lmoments.

REFERENCES

Anderson SP, Baumgartner MF. 1998. Radiative heating errors in naturally ventilated air temperature measurements made from Buoys. *Journal of Atmospheric and Oceanic Technology* **15**: 157–173.

Benschop H. 1996. Windsnelheidsmetingen op zeestations en kuststations: herleiding waarden windsnelheid naar 10-meter niveau (with an English summary). Koninklijk Nederlands Meteorologisch Instituut Technical Report No. 188, KNMI De Bilt.

Berry DI, Kent EC, Taylor PK. 2004. An analytical model of heating errors in marine air temperatures from ships. *Journal of Atmospheric and Oceanic Technology* **21**(8): 1198–1215.

Cullen MJP, Davies T, Mawson MH, James JA, Coulter SC, Malcolm A. 1997. An overview of numerical methods for the next generation U.K. NWP and climate model. In *Numerical Methods in Atmospheric and Ocean Modelling. The Andre J. Robert Memorial Volume*, Lin CA, Laprise R, Ritchie H (eds). NRC Research Press: Ottawa, Canada; 425–444.

Da Silva AM, Young CC, Levitus S. 1994. *Atlas of Surface Marine Data. Vol. 1, Algorithms and Procedures*. *NOAA Atlas Series*. NOAA.

Dietrich G. 1950. Systematic errors in the observed surface water and air temperature at sea and their effect on the determination of heat exchange between the sea and atmosphere. *Deutsche Hydrographische Zettschrift* **3**: 314–324.

Diaz HF, Folland CK, Manabe T, Parker DE, Reynolds RW, Woodruff SD. 2002. Workshop on Advances in the Use of Historical Marine Climate Data (Boulder, Co., USA, 29th Jan–1st Feb 2002). *WMO Bulletin* **51**: 377–380.

Dobson FW, Smith SD. 1988. Bulk models of solar radiation at sea. *Quarterly Journal of the Royal Meteorological Society* **114**: 165–182.

Folland CK. 1971. Daytime temperature measurements on weather ship '*Weather Reporter*'. *Meteorological Magazine* **100**: 6–14.

Folland CK, Parker DE, Yates FE. 1984. World wide marine air temperature fluctuations 1856–1981. *Nature* **310**: 670–673.

Fu C, Diaz HF, Dong D, Fletcher JO. 1999. Changes in atmospheric circulation over Northern Hemisphere oceans associated with the rapid warming of the 1920s. *International Journal of Climatology* **19**: 581–606.

Glahn W. 1933. False measurements of air temperatures on ships. *Der Seewart* **2**: 250–256.

Goerss JS, Duchon CE. 1980. Effect of ship heating on dry-bulb temperature measurements in GATE. *Journal of Physical Oceanography* **10**: 478–479.

Hayashi S. 1974. Some problems in marine meteorological observations, particularly of pressure and temperature. *Journal of Meteorological Research* **26**: 84–87.

Hosking JRM. 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B: Methodological* **52**: 105–124.

Hosking JRM. 2000. Fortran routines for use with the method of L-moments Version 3.03. Research Report RC20525 (90 933), IBM Research Division, T. J. Watson Research Center, Yorktown Heights: New York.

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) 2001. *Climate Change 2001: The Scientific Basis*. Cambridge University Press: Cambridge, UK.

JCOMM. 2002. Voluntary observing ships (VOS) climate subset project (VOSClim) project document, revision 2. WMO/TD-No. 1122. WMO, Geneva.

Josey SA, Kent EC, Taylor PK. 1999. New insights into the ocean heat budget closure problem from analysis of the SOC air–sea flux climatology. *Journal of Climate* **12**: 2856–2880.

Josey SA, Kent EC, Sinha B. 2001. Can a state of the art atmospheric general circulation model reproduce recent NAO related variability at the air–sea interface? *Geophysical Research Letters* **28**: 4543–4546.

Kent EC, Berry DI. 2005. Quantifying random measurement errors in voluntary observing ships' meteorological observations. *International Journal of Climatology* **25**: 843–856; this issue.

Kent EC, Kaplan A. In press. Towards estimating climatic trends in SST, part 3: systematic biases. *Journal of Atmospheric and Oceanic Technology*.

Kent EC, Taylor PK. 1991. Ships observing marine climate: a catalogue of the voluntary observing ships participating in the VSOP-NA. Marine meteorology and related oceanographic activities report no. 25, WMO/TD No. 456. WMO, Geneva.

Kent EC, Taylor PK. 1996. Accuracy and humidity measurements on ships: consideration of solar radiation effects. *Journal of Atmospheric and Oceanic Technology* **13**: 1317–1321.

Kent EC, Taylor PK. In press. Towards estimating climatic trends in SST, part 1: methods of measurement. *Journal of Atmospheric and Oceanic Technology*.

Kent EC, Truscott BS, Hopkins JS, Taylor PK. 1991. The accuracy of ship's meteorological observations, the results of the VSOP-NA. Marine meteorology and related oceanographic activities report no. 26, WMO/TD No. 455. WMO, Geneva.

Kent EC, Taylor PK, Truscott BS, Hopkins JS. 1993a. The accuracy of voluntary observing ships' meteorological observations — results of the VSOP-NA. *Journal of Atmospheric and Oceanic Technology* **10**: 591–608.

Kent EC, Tiddy RJ, Taylor PK. 1993b. Correction of the marine air temperature observations for solar radiation effects. *Journal of Atmospheric and Oceanic Technology* **10**: 900–906.

Lorenc AC, Bell RS, Macpherson B. 1991. The Meteorological Office analysis correction data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society* **117**: 59–89.

Met Office. 1981. *Handbook of Meteorological Instruments, Volume 3, Measurement of Humidity*, 2nd edition. *Met.O. 919c.* HMSO: London.

Met Office. 2002. A new Unified Model. *NWP Gazette*, June. http://www.metoffice.com/research/nwp/publications/nwp_gazette/jun02/index.html [accessed 19 April 2005].

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A. 2003. Global analyses of SST, sea ice and night marine air temperature since the late 19th century. *Journal of Geophysical Research* **108**(D14): 4407. DOI: 10.1029/2002JD002670.

Reynolds RW, Smith TM. 1994. Improved global sea-surface temperature analyses using optimum interpolation. *Journal of Climate* **7**: 929–948.

Von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge, UK.

WMO. 1994. International list of selected, supplementary and auxiliary ships. WMO Report No. 47. WMO, Geneva.

Woodruff SD, Diaz HF, Elms JD, Worley SJ. 1998. COADS release 2 data and metadata enhancements for improvements of marine surface flux fields. *Physics and Chemistry of the Earth* **23**: 517–527.

Yelland MJ, Moat BI, Pascal RW, Berry DI. 2002. CFD model estimates of the airflow distortion over research ships and the impact on momentum flux measurements. *Journal of Atmospheric and Oceanic Technology* **19**: 1477–1499.

