

THE IMPORTANCE OF UNRESOLVED BIASES IN TWENTIETH-CENTURY SEA SURFACE TEMPERATURE OBSERVATIONS

LUKE L. B. DAVIS, DAVID W. J. THOMPSON, JOHN J. KENNEDY, AND ELIZABETH C. KENT

Biases in sea surface temperature observations lead to larger uncertainties in our understanding of mid- to late-twentieth-century climate variability than previously thought.

The surface of the World Ocean warmed by ~ 0.75 K from 1900 to 2016, but the warming did not occur monotonically: temperatures increased during the first half of the twentieth century, decreased slightly during the decades following World War II, and increased rapidly after ~ 1975 (Hartmann et al. 2013). The decreases in ocean temperatures from the 1950s to 1970s are apparent in SSTs averaged over the globe and both the Atlantic and Pacific sectors (Figs. 1k–o, black time series; Figs. 2k–o, black bars).

The absence of warming during the decades following World War II is important because it coincides with steadily increasing concentrations of carbon dioxide over the same period. Several theories have been proposed to explain the absence of warming during this period, including increases in atmospheric sulfate aerosols (Tett et al. 2002; Lamarque et al. 2010; Booth et al. 2012; Myhre et al. 2013; Folland et al. 2018) and decadal variability in the ocean (Delworth and Mann 2000; Baines and Folland 2007; Knight et al. 2006; Semenov et al. 2010). Here we provide novel analyses of SST data separated into the two primary measurement sources to demonstrate that the uncertainty in decadal variability of SST from the 1950s to 1970s is at least as large as the observed decadal variability itself. The results highlight the critical importance of considering uncertainty in SST observations in analyses of observed decadal climate variability.

SST data during the period after 1980 are derived from several in situ and remotely sensed sources (Kent et al. 2010; Kennedy et al. 2011b). But SST data prior to 1980 are derived almost entirely from two in situ sources via “ships of opportunity”: 1) the temperature of seawater in buckets that have been submerged below the ocean surface and then hauled back onto a ship deck (bucket measurements) and 2)

AFFILIATIONS: DAVIS AND THOMPSON—Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado; KENNEDY—Met Office Hadley Centre, Exeter, United Kingdom; KENT—National Oceanography Centre, Southampton, United Kingdom

CORRESPONDING AUTHOR: Luke L. B. Davis, luke.davis@colostate.edu

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-D-18-0104.1

A supplement to this article is available online (10.1175/BAMS-D-18-0104.2).

In final form 25 September 2018

©2019 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

the temperature of the pumped water supply to an engine room [engine-room intake (ERI) measurements] (Kent et al. 2010; Kennedy et al. 2011b). A comparatively small number of hull sensor observations are also included in the ERI category, as the biases in both hull sensor and ERI data are thought to be governed by similar factors (Kennedy et al. 2011b).

Bucket and ERI measurements both exhibit substantial measurement biases (Kent and Kaplan 2006; Rayner et al. 2006; Kent et al. 2010; Kennedy et al. 2011b; Kent et al. 2017; Folland and Parker 1995). ERI measurements are often warm biased because of the transfer of heat from the superstructure of the ship as water passes through pipes, while bucket measurements are often cold biased because of the exchange of latent and sensible heat with the surrounding air. If the mix of measurement types and their relative biases are well understood, then the biases can be adjusted so that they have little effect on the time evolution of spatially averaged temperature data. But if the mix of measurement types is poorly documented, large biases can remain after adjustment, even in widely used climate data sources (Folland and Parker 1995; Thompson et al. 2008; Karl et al. 2015).

In principle, SST data stratified by measurement type provide the opportunity to assess the reproducibility of SST variability in subsets of the data not influenced by changes to instrumentation. With this in mind, the Met Office Hadley Centre developed SST datasets stratified into bucket and ERI measurements in conjunction with the release of their most recent gridded dataset, the Hadley Centre SST dataset (HadSST3; Kennedy et al. 2011a,b). The bucket and ERI data are available over the period 1946–2006 and were developed in the same way as the full HadSST3 dataset (Kennedy et al. 2011a,b), that is, by 1) estimating the measurement types of SST observations in the International Comprehensive Ocean–Atmosphere Data Set, release 2.5 archive (ICOADS2.5; Woodruff et al. 2011); 2) consolidating the observations onto monthly $5^\circ \times 5^\circ$ grids; 3) applying bias adjustment schemes unique to each measurement type; and 4) accounting for parametric uncertainty in the bias adjustment schemes by generating 100 plausible realizations of the adjustments. For each of the bucket-only and ERI-only datasets, observations estimated to be from the other measurement type were ignored, and for this analysis, grid boxes without valid data from

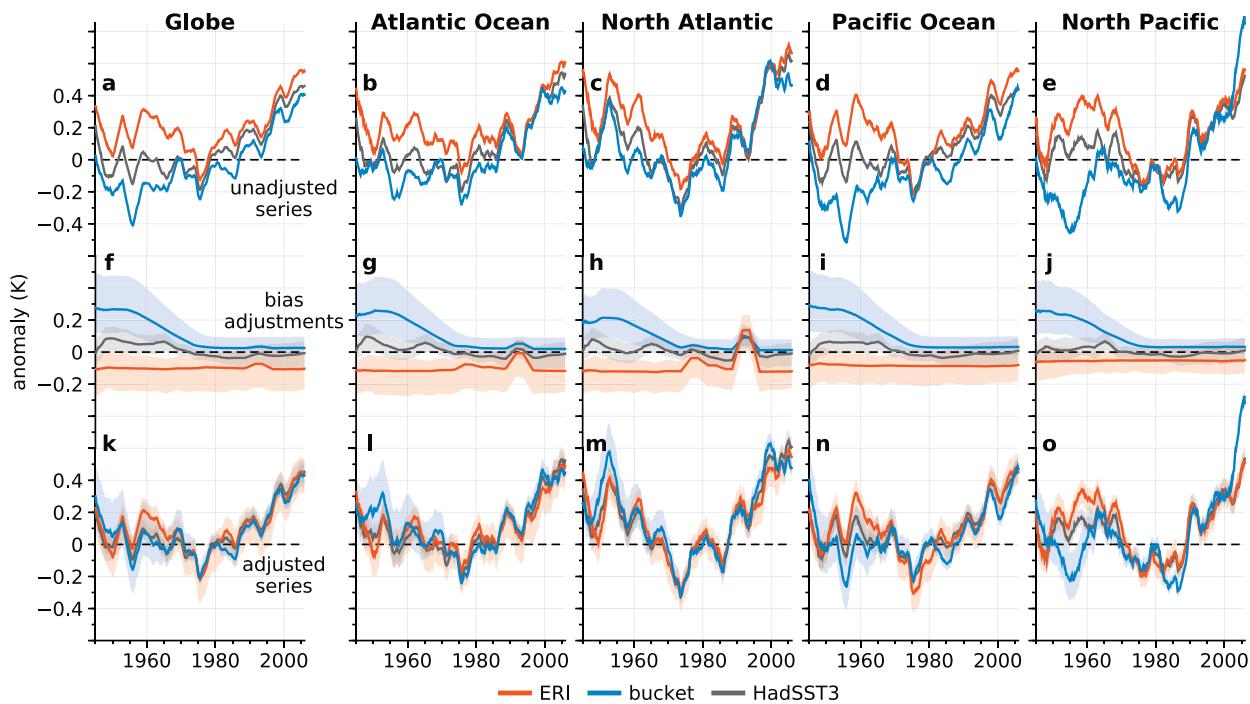


FIG. 1. TIME SERIES OF SEA SURFACE TEMPERATURE ANOMALIES (K; RELATIVE TO 1961–90) FOR THE DATASETS AND REGIONS INDICATED. (a)–(e) Area averages of the unadjusted ERI, bucket, and HadSST3 data; (f)–(j) area averages of the bias adjustments applied to the data; and (k)–(o) area averages of the adjusted data. The North Atlantic and North Pacific time series indicate averages north of 20°N . (middle), (bottom) The shaded regions indicate the range of all 100 realizations of the HadSST3 bias adjustments; the thick lines indicate the average over all 100 realizations. The time series are smoothed by a centered 37-month running mean for display purposes. Note that the ERI, bucket, and HadSST3 data are collocated in space (see appendix).

the other measurement type were excluded. The latter step ensures that the bucket-only and ERI-only data are “collocated,” or have the same spatial coverage through time at the gridbox level. This is critical, as measurement types are often distributed differently across each ocean basin (Kent and Taylor 2006).

The identification of SST methodology is imperfect; in many cases, the ICOADS2.5 metadata do not provide specific information about the measurement method, and hence the measurement type must be estimated from other information, such as country of origin (Kennedy et al. 2011b). Even if the measurement type is indicated by the metadata, the indication is sometimes incorrect (Kent et al. 2007). In other cases, the general type of measurement is known (e.g., bucket), but specific aspects of the measurement (e.g., the construction and insulation of the bucket) are not. Nevertheless, the bucket-

only and ERI-only datasets reflect the best available estimates of mid-twentieth-century SST variability minimally influenced by changes to instrumentation. Together, the two datasets thus provide a unique opportunity to explore uncertainty in observed decadal variability.

THE PROBLEM. The unadjusted bucket and ERI data yield remarkably different renditions of twentieth-century SST variability, particularly prior to ~ 1975 (red and blue time series in Figs. 1a–e; red and blue bars in Figs. 2a–e; Figs. 3a,b). For example, the ERI data exhibit cooling of the Pacific Ocean from 1950 to 1975, whereas the unadjusted bucket data indicate warming (Figs. 1d,e, 2d,e, 3a,b). Likewise, the ERI data exhibit cooling in the global average over the same period, whereas the bucket data indicate warming (Figs. 1a, 2a).

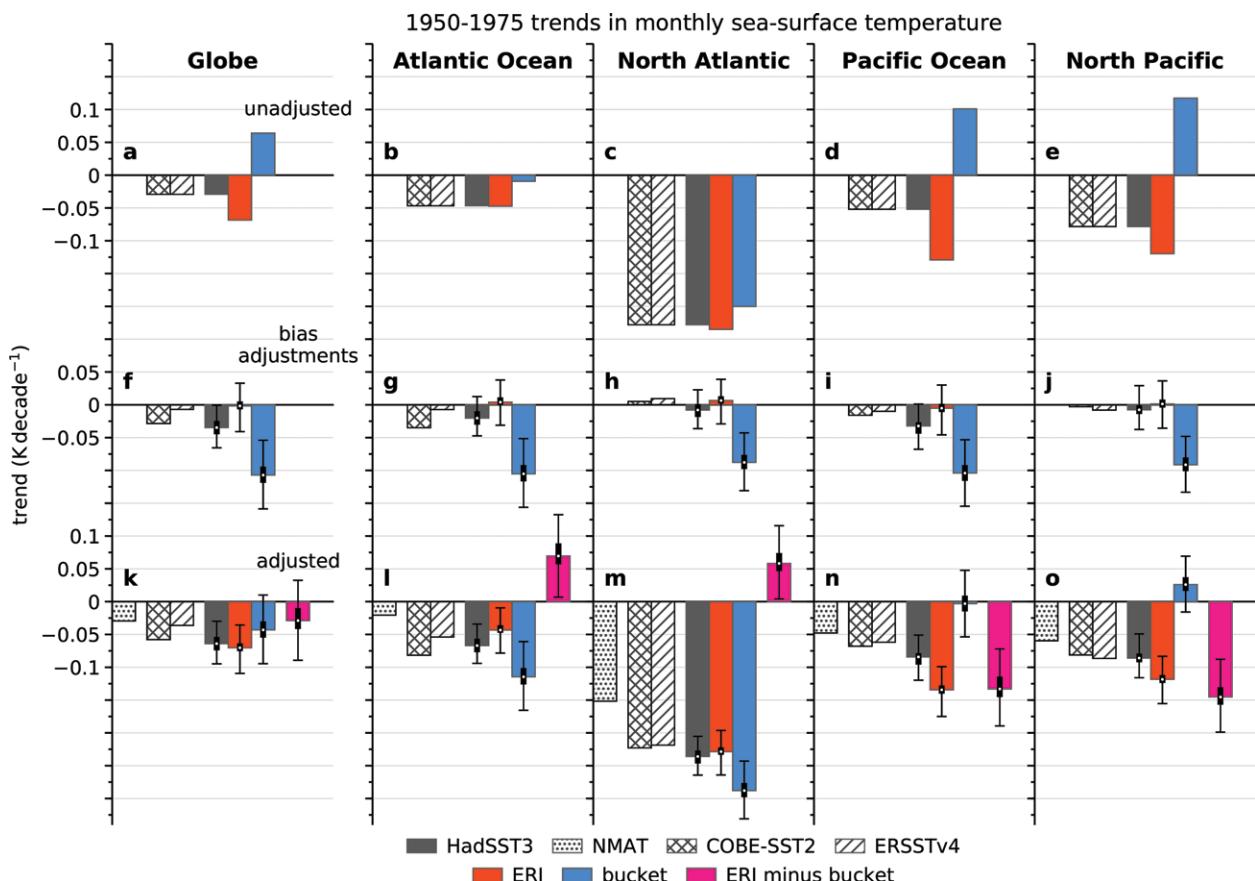


FIG. 2. LINEAR LEAST SQUARES TRENDS IN SEA SURFACE TEMPERATURE ANOMALIES FROM 1950 TO 1975 FOR THE DATASETS AND REGIONS INDICATED. (a)–(e) Trends in the area-averaged unadjusted data, (f)–(j) trends in the area-averaged bias adjustments, and (k)–(o) trends in the area-averaged adjusted data. The North Atlantic and North Pacific trends indicate averages north of 20°N . (middle), (bottom) Thin error bars indicate the range of trends from all 100 realizations of the HadSST3 bias adjustments. The thicker error bars indicate interquartile ranges of the trends. The significance of the mean trends with respect to the internal variability is given in the supplemental material. Note that the COBE-SST2, NMAT, and ERSST4 data are collocated with the ERI, bucket, and HadSST3 data (see appendix). Note also that the NMAT data are only available in their adjusted form.

The adjustments applied to the ERI data using the HadSST3 bias adjustment scheme are mostly stationary in time, with the exception of the short-term bias adjustments applied to the Atlantic sector during the early 1990s (red time series in Figs. 1f–j; see Kent and Kaplan 2006). Hence, they do not notably affect estimates of decadal variability (red bars in Figs. 2f–j; Fig. 3d). In contrast, the adjustments applied to the bucket data introduce a substantial 0.1 K decade^{-1} cooling over the period 1950–75 (blue time series in Figs. 1f–j; blue bars in Figs. 2f–j; Fig. 3e), due to the assumed transition from canvas to rubber buckets (Kent et al. 2010; Kennedy et al. 2011b). The cooling introduced by the adjustments applied to the bucket data ranges from -0.05 to $-0.15\text{ K decade}^{-1}$ across the 100 realizations of the HadSST3 bias adjustments (error bars in Figs. 2f–j). The adjusted bucket temperature data exhibit robust cooling in Atlantic basin averages but not in the global and Pacific basin averages. The Atlantic cooling is apparent in all 100 realizations of the adjusted bucket data (whiskers on blue bars in Figs. 2l,m) and is also statistically significant with respect to the detrended variability in the data

(Figs. ES11,m in the supplemental material; <https://doi.org/10.1175/BAMS-D-18-0104.2>). The adjustments for both ERI and bucket data are roughly stationary in time during the 1976–2006 period (Figs. 4d–f).

The resulting adjusted ERI and bucket data (red and blue time series in Figs. 1k–o; red and blue bars in Figs. 2k–o; Figs. 3g,h) are in closer agreement with each other than their unadjusted counterparts. However, the trends in the SST field over the period 1950–75 are still notably different for the two measurement types. In the global average, the cooling in the adjusted bucket data is roughly half as large as the cooling in the adjusted ERI data (Figs. 1k, 2k). The discrepancies are especially notable in the Pacific sector, where the adjusted ERI data exhibit cooling over the period 1950–75 but the adjusted bucket data exhibit relatively little change in temperature (Figs. 1n,o, 2n,o, 3g,h). In contrast, over the Atlantic sector the adjusted ERI data exhibit significantly weaker cooling than the adjusted bucket data (Figs. 1l,m, 2l,m, 3g,h). These patterns of disagreement are stronger in the North Pacific and North Atlantic during boreal winter (Figs. S2l,m, S3g,h).

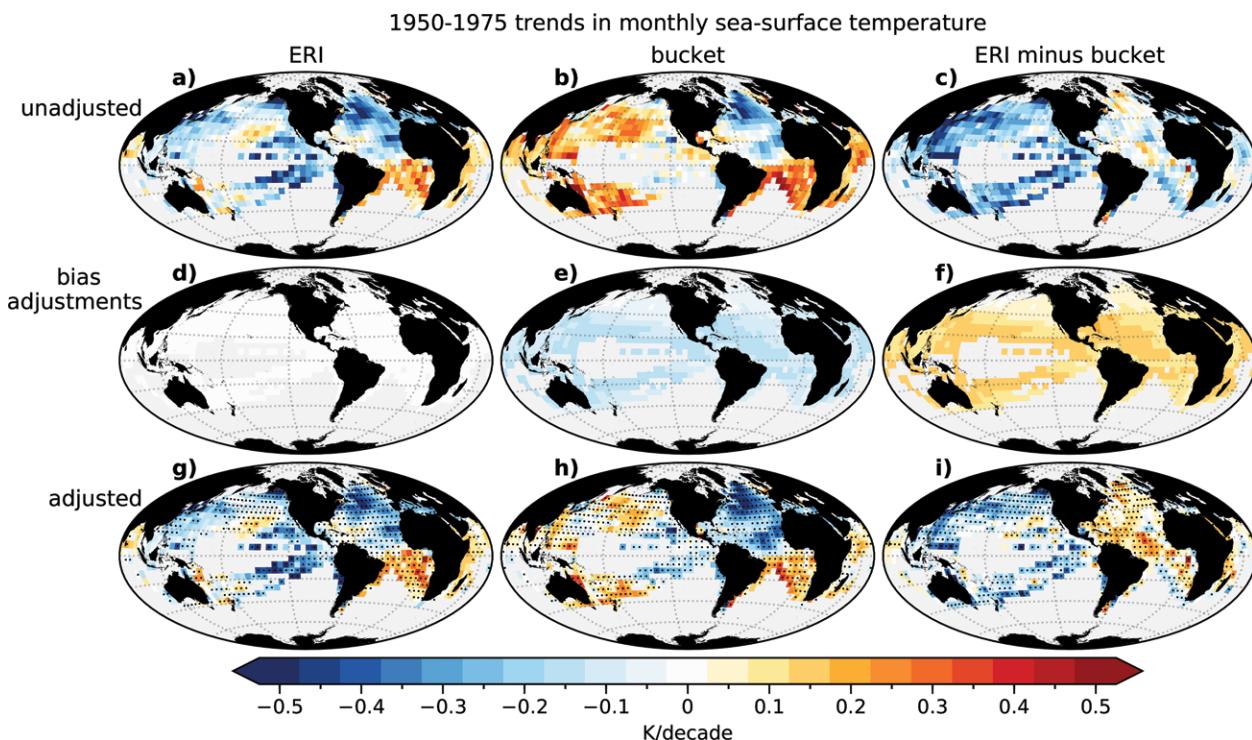


FIG. 3. LINEAR LEAST SQUARES TRENDS IN SEA SURFACE TEMPERATURE ANOMALIES FROM 1950 TO 1975 FOR THE INDICATED DATASETS. (from left to right) Trends in the ERI dataset, bucket dataset, and their differences. (from top to bottom) (a)–(c) Trends in the unadjusted data, (d)–(f) trends in the bias adjustments (i.e., the mean of the trends calculated for all 100 realizations of the HadSST3 bias adjustments), and (g)–(i) trends in the adjusted data. Trends were not computed where fewer than 50% of the time steps from 1950 to 1975 were available. Grid boxes are stippled where the (g),(h) trends and (i) trend differences have the same sign across all 100 bias adjustment realizations.

After 1975, the agreement between the ERI and bucket data improves as the magnitude of the adjustments decreases and the overall quality and consistency of the observations increases. Nevertheless there remain notable differences in the adjusted ERI and bucket SST trends over the 1976–2006 period, especially over the south-central Pacific sector (Figs. 4g–i) and during austral winter (Figs. ES4g–i).

Importantly, the amplitudes of the trends in the adjusted ERI and bucket data over the 1950–75 period are comparable to the differences between them (cf. the red and blue bars with the pink bars in Figs. 2k–o and Figs. 3g,h with Fig. 3i), which points to the scale of the uncertainty in the adjusted data. As indicated by the whiskers on the pink bars in Figs. 2k–o and the stippling in Fig. 3i, the 100 bias adjustment realizations cannot account for these differences and thus do not entirely characterize the bias uncertainties in the trends. When averaged over large spatial domains, the amplitudes of the trends are also comparable to the trends in the bias adjustments themselves (cf. the middle and bottom rows of Fig. 2). This is key, as the bias adjustment schemes are subject to considerable uncertainty, particularly prior to 1980 (Kennedy et al. 2011b; Kent et al. 2017).

In the case of HadSST3, the bias adjustment schemes are derived from metadata contained in ICOADS2.5 and historical documentation (Kennedy et al. 2011b). However, the metadata are frequently

incomplete, and thus various sources of bias are not known with confidence, including bucket type, the speed of the ship, the depth from which water for the engine room is drawn, and whether a datum is derived from a bucket or ERI measurement in the first place (Kent and Taylor 2006; Kent et al. 2017). For example, the HadSST3 bias adjustments assume 40%–80% of the SST data from 1960 to 1980 are derived from bucket measurements, whereas a recent reassessment of measurement type suggests the fraction of bucket measurements over this time is consistently closer to 40% (Carella et al. 2018).

The uncertainties in the bias adjustment scheme applied to HadSST3 data can be inferred from the time series in Fig. 1 as follows [see also Kent et al. (2017) and Carella et al. (2018)]. The unadjusted ERI and bucket time series can be decomposed as follows:

$$ERI_{\text{unadjusted}} = ERI_{\text{true}} + ERI_{\text{true bias}} \text{ and} \quad (1)$$

$$B_{\text{unadjusted}} = B_{\text{true}} + B_{\text{true bias}}, \quad (2)$$

where ERI_{true} and B_{true} are the “true” SST data in the absence of measurement bias and $ERI_{\text{true bias}}$ and $B_{\text{true bias}}$ are the “ideal” bias adjustments. Since the bucket and ERI data used here are collocated in space, it follows that $ERI_{\text{true}} = B_{\text{true}}$ over area averages large enough to suppress sampling and measurement uncertainties (Kennedy et al. 2011a; Carella et al. 2018), and therefore

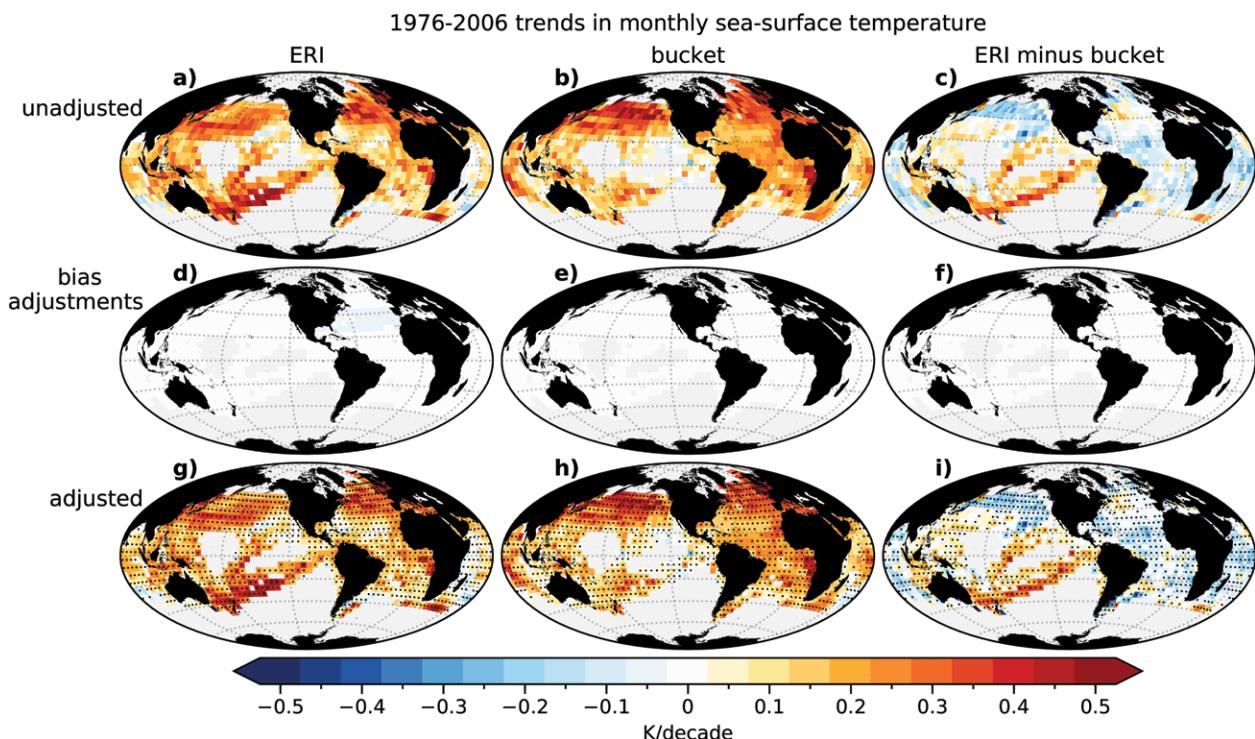


FIG. 4. As in Fig. 3, but for the 1976–2006 period.

$$\text{ERI}_{\text{unadjusted}} - B_{\text{unadjusted}} = \text{ERI}_{\text{true bias}} - B_{\text{true bias}}. \quad (3)$$

The uncertainty in the bias adjustments applied to the bucket and ERI data (and hence to HadSST3) can thus be estimated by comparing 1) the differences between the unadjusted ERI and bucket data with 2) the differences between the ERI and bucket bias estimates (i.e., the negative of the ERI and bucket bias adjustments). If the bias adjustments are ideal, then the time series given by items 1 and 2 should be identical. Note that the time series can also be identical if there is a common bias in the ERI and bucket measurements; the series being identical is a necessary but not sufficient criterion for ideal adjustments.

Figure 5 shows the results of the above calculation for the domains considered in Figs. 1 and 2. The orange lines indicate the differences between the ERI and bucket bias estimates averaged over all 100 pairs of adjustments (the range given by the 100 realizations is indicated by orange shading); the black lines indicate the differences between the unadjusted ERI and bucket data. Again, 1) if the bias adjustments are ideal, then the black and orange lines should overlap each other; and 2) if the 100 realizations of the bias adjustments characterize the uncertainty in the adjustments, then the black lines should lie within the regions of orange shading. Overall, the adjustments required to bring ERI and bucket data into agreement (black lines) are clearly much larger and much more variable than the mean of the actual bias adjustments applied to the HadSST3 data. The inferred uncertainties in the bias adjustments are comparable to the amplitude of the observed decadal variability in the SST field.

The uncertainties in decadal-scale variability indicated in Figs. 1–5 also affect the two other major historical SST datasets based on the ICOADS2.5 archive: the Centennial In Situ Observation-Based Estimates of SST (COBE-SST2; Hirahara et al. 2014) developed by the Japan Meteorological Agency, and version 4 of the Extended Reconstructed SST dataset (ERSST.v4; Huang et al. 2015; Liu et al. 2015) released by the U.S. National Climatic Data Center. The ERSST.v4 and COBE-SST2 datasets are included to provide a point of comparison with SST data that employ very different bias correction schemes: The bias adjustments applied in HadSST3 and COBE-SST2 are both based on information about measurement type as inferred from the metadata; the adjustments applied in ERSST.v4 are based only on comparisons with nighttime marine air temperature (NMAT) data, which require their own bias adjustments (Rayner et al. 2003; Kent et al. 2013; Kennedy 2014). In general, over the 1950–75 period, the trends in the

bias-adjusted COBE-SST2 data are similar to those in the HadSST3 data, whereas the trends in the bias-adjusted NMAT and ERSST.v4 data are somewhat weaker than those in the HadSST3 data, particularly over the Atlantic Ocean and in the global mean (Figs. 2k–o; see appendix for details of the analysis).

SO WHAT? The results shown here reveal a level of regional uncertainty in observed SSTs that is not widely acknowledged in the climate dynamics literature. In our view, it should be. Confidence in observed decadal variability derives from confidence in the bias adjustments applied to the SST data. And as shown here, the uncertainty in the bias adjustment schemes is frequently comparable to the amplitude of the observed decadal variability itself. The uncertainty has important implications for our understanding of the role of aerosols in twentieth-century climate change (Tett et al. 2002; Lamarque et al. 2010; Booth et al. 2012; Myhre et al. 2013; Folland et al. 2018), since aerosols are believed to have contributed to the absence of global warming during the mid-twentieth century (Kobayashi et al. 2015; Laloyaux et al. 2017; Taylor et al. 2012; Flato et al. 2013). It also has important implications for quantifying the amplitudes of patterns of decadal-scale variability, particularly over the problematic North Pacific sector (Figs. 1, 2) and in association with Pacific and Atlantic decadal variability (Mantua et al. 1997; Mantua and Hare 2002; Newman et al. 2016; Delworth and Mann 2000; Baines and Folland 2007; Knight et al. 2006; Semenov et al. 2010).

The findings indicate notable shortcomings in our ability to accurately classify SST measurement methods and to quantify the associated biases. Complicating matters is that measurement biases vary not only from one measurement method to the next but also within the individual methods: ERI biases can vary between individual ships and recruiting countries (Kent et al. 1993); bucket biases depend on the bucket type, and the transition from canvas to rubber buckets for a given recruiting country is highly uncertain (Kennedy et al. 2011b). Additionally, the metadata necessary to identify ships are often missing from ICOADS (Carella et al. 2018), and the proportions of recruiting countries can change substantially over time, especially before ~1970 (Fig. ES5; Thompson et al. 2008). For example, the large differences between trends in the bucket and ERI data over the Pacific sector relative to those over the Atlantic sector (Figs. 2l–o, 3i) are potentially due to differences in the types of bucket and ERI measurements used in each region, as implied by the differences in recruiting countries between the two sectors

(Fig. ES6). Not surprisingly, despite recent advances (Freeman et al. 2017; Carella et al. 2018; Hausfather et al. 2017; Cowtan et al. 2017; Hirahara et al. 2014), it may take years to resolve the discrepancies between the ERI and bucket time series highlighted here.

What is the best way forward? The recent review of SST biases by Kent et al. (2017) concludes with a series of recommendations for improving the reliability of historical SST bias estimations, especially after World War II. These include improving the metadata and volume of observations in the ICOADS archive, improving the classification of measurement methods from documentation and analyzing data characteristics, improving the physical and statistical models used to estimate SST bias, and entraining more scientists into the field of SST bias adjustment. Novel analyses could include clustering of observations by individual ship or recruiting country, which may help isolate the bias variations within each measurement method. Our results make clear the critical importance of the recommendations in Kent et al. (2017) for improving our understanding of twentieth-century climate variability.

ACKNOWLEDGMENTS. We thank the three anonymous reviewers for their insightful comments, which greatly improved the paper. L.L.B.D. and D.W.J.T. were supported by the NSF Climate and Large-Scale Dynamics Program (AGS-1547003). J.J.K. was supported by the Joint U.K. BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). E.C.K. was supported by the NERC (NE/J020788/1).

APPENDIX: METHODS. The HadSST3 data (Kennedy et al. 2011b) and nighttime marine air temperature data (Kent et al. 2013) were obtained from the Met Office Hadley Centre (<https://metoffice.gov.uk/hadobs>). Subsequent to this study, the ERI-only and bucket-only data are also published on the Hadley Centre website. The unadjusted ICOADS sea surface temperature observations (Woodruff et al. 2011) were

obtained from the National Center for Atmospheric Research (<https://rda.ucar.edu/>). The Japan Meteorological Agency (COBE-SST2; Hirahara et al. 2014) and National Climatic Data Center (ERSST.v4; Huang et al. 2015; Liu et al. 2015) sea surface temperature data were both obtained from the Earth System Research Laboratory Physical Sciences Division (<https://esrl.noaa.gov/psd>). To accommodate comparisons with the Hadley Centre data, the COBE-SST2 and ERSST.v4 data were 1) regressed onto the $5^\circ \times 5^\circ$ resolution HadSST3 grid and 2) had their respective monthly 1961–90 climatologies subtracted (to match the HadSST3 climatology period). The gridded NMAT,

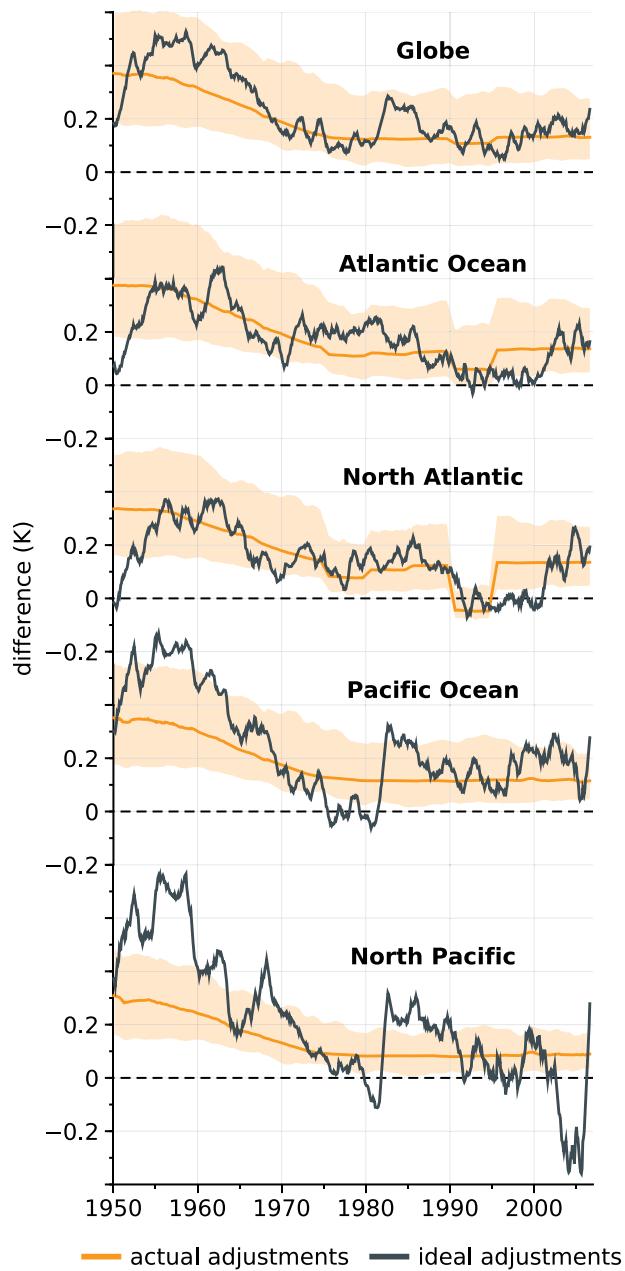


FIG. 5. ESTIMATING THE UNCERTAINTY IN THE BIAS ADJUSTMENTS FOR THE INDICATED REGIONS. The orange lines represent the differences between the bucket and ERI bias estimates averaged over all 100 realizations of the bias adjustments, and shading indicates the range of the differences. The black lines represent the differences between the unadjusted ERI and bucket data. The latter correspond to the ideal bias adjustments required to bring the bucket and ERI data into agreement. See Eq. (3) and accompanying discussion in the text. The time series are smoothed with a 12-month running mean to remove the seasonal cycle from the bias adjustments (see Folland and Parker 1995).

HadSST3, COBE-SST2, and ERSST.v4 data were matched to the spatial coverage of the collocated ERI-only and bucket-only data. The coordinate boundaries used for each spatial average are shown in Table A1 (grid boxes were weighted by the cosine of the central latitude and the ocean fraction within each box). Note that all data used in this study are in anomaly form with respect to the 1961–90 base period.

REFERENCES

Baines, P. G., and C. K. Folland, 2007: Evidence for a rapid global climate shift across the late 1960s. *J. Climate*, **20**, 2721–2744, <https://doi.org/10.1175/JCLI4177.1>.

Booth, B. B. B., N. J. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232, <https://doi.org/10.1038/nature10946>.

Carella, G., J. J. Kennedy, D. I. Berry, S. Hirahara, C. J. Merchant, S. Morak-Bozzo, and E. C. Kent, 2018: Estimating sea surface temperature measurement methods using characteristic differences in the diurnal cycle. *Geophys. Res. Lett.*, **45**, 363–371, <https://doi.org/10.1002/2017GL076475>.

Cowtan, K., R. Rohde, and Z. Hausfather, 2017: Evaluating biases in sea surface temperature records using coastal weather stations. *Quart. J. Roy. Meteor. Soc.*, **144**, 670–681, <https://doi.org/10.1002/qj.3235>.

Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dyn.*, **16**, 661–676, <https://doi.org/10.1007/s003820000075>.

Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

Folland, C. K., and D. E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, **121**, 319–367, <https://doi.org/10.1002/qj.49712152206>.

—, O. Boucher, A. Colman, and D. E. Parker, 2018: Causes of irregularities in trends of global mean surface temperature since the late 19th century. *Sci. Adv.*, **4**, eaao5297, <https://doi.org/10.1126/sciadv.aao5297>.

Freeman, E., and Coauthors, 2017: ICOADS Release 3.0: A major update to the historical marine climate record. *Int. J. Climatol.*, **37**, 2211–2232, <https://doi.org/10.1002/joc.4775>.

Hartmann, D., and Coauthors, 2013: Observations: Atmosphere and surface. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 159–254.

Hausfather, Z., K. Cowtan, D. C. Clarke, P. Jacobs, M. Richardson, and R. Rohde, 2017: Assessing recent warming using instrumentally homogeneous sea surface temperature records. *Sci. Adv.*, **3**, e1601207, <https://doi.org/10.1126/sciadv.1601207>.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, <https://doi.org/10.1175/JCLI-D-12-00837.1>.

Huang, B., and Coauthors, 2015: Extended Reconstructed Sea Surface Temperature version 4 (ERSST.v4). Part I: Upgrades and intercomparisons. *J. Climate*, **28**, 911–930, <https://doi.org/10.1175/JCLI-D-14-00006.1>.

Karl, T. R., and Coauthors, 2015: Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, **348**, 1469–1472, <https://doi.org/10.1126/science.aaa5632>.

Kennedy, J. J., 2014: A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Rev. Geophys.*, **52**, 1–32, <https://doi.org/10.1002/2013RG000434>.

—, N. A. Rayner, R. O. Smith, D. E. Parker, and M. Saunby, 2011a: 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, <https://doi.org/10.1029/2010JD015218>.

—, —, —, —, and —, 2011b: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. *J. Geophys. Res.*, **116**, D14104, <https://doi.org/10.1029/2010JD015218>.

TABLE A1. Boundaries used for spatial averages of gridded SST data.

Region	Boundaries
Globe	60°S–60°N, 180°W–180°E
Atlantic Ocean	10°–60°N, 80°W–0° 60°S–10°N, 70°W–20°E
Pacific Ocean	10°–60°N, 100°E–100°W 10°S–10°N, 100°E–70°W 60°–10°S, 150°E–70°W
North Atlantic	20°–60°N, 80°W–0°
North Pacific	20°–60°N, 120°E–100°W

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

—, and A. Kaplan, 2006: Toward estimating climatic trends in SST. Part III: Systematic biases. *J. Atmos. Oceanic Technol.*, **23**, 487–500, <https://doi.org/10.1175/JTECH1845.1>.

—, —, B. S. Truscott, and J. S. Hopkins, 1993: The accuracy of voluntary observing ships' meteorological observations: Results of the VSOP-NA. *J. Atmos. Oceanic Technol.*, **10**, 591–608, [https://doi.org/10.1175/1520-0426\(1993\)010<0591:TAOVOS>2.0.CO;2](https://doi.org/10.1175/1520-0426(1993)010<0591:TAOVOS>2.0.CO;2).

—, S. D. Woodruff, and D. I. Berry, 2007: Metadata from WMO Publication No. 47 and an assessment of voluntary observing ship observation heights in ICOADS. *J. Atmos. Oceanic Technol.*, **24**, 214–234, <https://doi.org/10.1175/JTECH1949.1>.

—, J. J. Kennedy, D. I. Berry, and R. O. Smith, 2010: Effects of instrumentation changes on sea surface temperature measured in situ. *Wiley Interdiscip. Rev.: Climate Change*, **1**, 718–728, <https://doi.org/10.1002/wcc.55>.

—, N. A. Rayner, D. I. Berry, M. Saunby, B. I. Moat, J. J. Kennedy, and D. E. Parker, 2013: Global analysis of night marine air temperature and its uncertainty since 1880: The HadNMAT2 data set. *J. Geophys. Res. Atmos.*, **118**, 1281–1298, <https://doi.org/10.1002/jgrd.50152>.

—, and Coauthors, 2017: A call for new approaches to quantifying biases in observations of sea surface temperature. *Bull. Amer. Meteor. Soc.*, **98**, 1601–1616, <https://doi.org/10.1175/BAMS-D-15-00251.1>.

Knight, J. R., C. K. Folland, and A. A. Scaife, 2006: Climate impacts of the Atlantic multidecadal oscillation. *Geophys. Res. Lett.*, **33**, L17706, <https://doi.org/10.1029/2006GL026242>.

Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, <https://doi.org/10.2151/jmsj.2015-001>.

Laloyaux, L., E. De Boisseson, and P. Dahlgren, 2017: CERA-20C: An Earth system approach to climate reanalysis. *ECMWF Newsletter*, No. 150, ECMWF, Reading, United Kingdom, 25–30.

Lamarque, J.-F., and Coauthors, 2010: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application. *Atmos. Chem. Phys.*, **10**, 7017–7039, <https://doi.org/10.5194/acp-10-7017-2010>.

Liu, W., and Coauthors, 2015: Extended Reconstructed Sea Surface Temperature version 4 (ERSST.v4): Part II. Parametric and structural uncertainty estimations. *J. Climate*, **28**, 931–951, <https://doi.org/10.1175/JCLI-D-14-00007.1>.

Mantua, N. J., and S. R. Hare, 2002: The Pacific decadal oscillation. *J. Oceanogr.*, **58**, 35–44, <https://doi.org/10.1023/A:1015820616384>.

—, —, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1079, [https://doi.org/10.1175/1520-0477\(1997\)078<1069:APICOW>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2).

Myhre, G., and Coauthors, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 659–740.

Newman, M., and Coauthors, 2016: The Pacific decadal oscillation, revisited. *J. Climate*, **29**, 4399–4427, <https://doi.org/10.1175/JCLI-D-15-0508.1>.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, <https://doi.org/10.1029/2002JD002670>.

—, P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. J. Ansell, and S. F. B. Tett, 2006: 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, <https://doi.org/10.1175/JCLI3637.1>.

Semenov, V. A., M. Latif, D. Dommengen, N. S. Keenlyside, A. Strehz, T. Martin, and W. Park, 2010: The impact of North Atlantic–Arctic multidecadal variability on Northern Hemisphere surface air temperature. *J. Climate*, **23**, 5668–5677, <https://doi.org/10.1175/2010JCLI3347.1>.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.

Tett, S. F. B., and Coauthors, 2002: Estimation of natural and anthropogenic contributions to twentieth century temperature change. *J. Geophys. Res.*, **107**, 4306, <https://doi.org/10.1029/2000JD000028>.

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

Woodruff, S. D., and Coauthors, 2011: ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967, <https://doi.org/10.1002/joc.2103>.

AMS Members

Give a
great gift
at a
great price

Looking for the perfect
present for the weather
enthusiast in your life?
Want to make a valuable
contribution to your local
library or community
college?



Send a subscription to
Weatherwise magazine for
just \$24.95*—That's nearly
50% off the list price!

Contact Member Services by
e-mail at
amsmem@ametsoc.org or
by phone at **617-227-2425**
to place all of your Weatherwise
orders today!

**Written for a general audience,
Weatherwise offers a colorful and
nontechnical look at recent discoveries
in meteorology and climatology.**

Check out the latest table of contents at
www.weatherwise.org.

Want your own?
Then order a personal subscription
at the same great price.