Random Error in Space-Time Bin Averages of Sea Surface Temperature Observations from Ships

Alexey Kaplan¹

¹Lamont-Doherty Earth Observatory, Columbia University: Palisades, NY, US

November 24, 2022

Abstract

Sea surface temperature (SST) observations made at ships are distributed irregularly in space and time and are affected by systematic biases and random errors. Such observations are often "binned": split into samples, contained within "bins" - grid boxes of a space-time grid (1oX10 monthly bins are used here), and their statistics are computed. Bin averages often serve as gridded representations of such data, thus requiring reliable uncertainty estimates, which for ship observations are particularly important because of their domination in the early observational records. Here ship SST observations for 1992–2010 are compared with an independent high-resolution satellite-based SST data set. To remove systematic biases, seasonal means were subtracted from the difference between bin-averaged data sets. In more than 66%(50%) of locations with binned temporal coverage exceeding 50%(66%), the magnitude of remaining anomalies agreed within 20%(10%) with random error model estimates. Separate estimates for sampling and measurement error components were obtained.

Random Error in Space-Time Bin Averages of Sea Surface Temperature Observations from Ships

Alexey Kaplan¹

 $^{1}\mathrm{Lamont-Doherty}$ Earth Observatory of Columbia University

Key Points:

1

2

3

4

5

6	•	Relative biases between SST data sets from ships and satellites, averaged to one
7		degree monthly bins, are estimated as climatological means
8	•	Magnitudes of difference anomalies between one degree monthly averages of SST
9		from ships and satellites agree with the random error model
10	•	Separate estimates are obtained for sampling and measurement error components
11		of the total error in bin averages of ship SST observations

Corresponding author: Alexey Kaplan, LDEO of Columbia University, P.O. Box 1000, 61 Route 9W, Palisades, NY 10964, USA; alexeyk@ldeo.columbia.edu

12 Abstract

Sea surface temperature (SST) observations made at ships are distributed irregularly in 13 space and time and are affected by systematic biases and random errors. Such observa-14 tions are often "binned": split into samples, contained within "bins" – grid boxes of a 15 space-time grid $(1^{\circ} \times 1^{\circ} \text{ monthly bins are used here})$, and their statistics are computed. 16 Bin averages often serve as gridded representations of such data, thus requiring reliable 17 uncertainty estimates, which for ship observations are particularly important because 18 of their domination in the early observational records. Here ship SST observations for 19 1992–2010 are compared with an independent high-resolution satellite-based SST data 20 set. To remove systematic biases, seasonal means were subtracted from the difference 21 between bin-averaged data sets. In more than 66%(50%) of locations with binned tem-22 poral coverage exceeding 50%(66%), the magnitude of remaining anomalies agreed within 23 20%(10%) with random error model estimates. Separate estimates for sampling and mea-24 surement error components were obtained. 25

²⁶ Plain Language Summary

Sea surface temperature (SST) is an important climate variable. SST observations 27 made at ships are distributed irregularly in space and time and are affected both by sys-28 tematic biases and randomly-varying measurement errors. To make them easier to use, 29 such data sets are often "binned", i.e., split into samples contained within "bins", which 30 usually are grid boxes of some space-time grid (monthly 1° longitude by 1° latitude bins 31 are used here), and the statistics of these binned samples are computed. Bin averages 32 often serve as gridded representations for data sets of ship observations; hence their un-33 certainty estimates have to be reliable. This is especially important since ship observa-34 tions dominate early on in the historical observational record. Ship SST observations for 35 1992–2010 are compared here with an independent high-resolution satellite-based SST 36 data set. To remove systematic biases, seasonal means were subtracted from the differ-37 ence between bin-averaged versions of these data sets, and the remainder was interpreted 38 as a sum of random errors. In more than 66%(50%) of locations where binned tempo-39 ral coverage exceeded 50%(66%), the remainder's magnitude agreed within 20%(10%)40 with the random error model estimates. Error components due to incomplete sampling 41 and due to measurement error were estimated. 42

43 1 Introduction

Sea surface temperature (SST) is one of the "essential" climate variables (Bojinski 44 et al., 2014), particularly well-suited for monitoring changes in the Earth's mean surface 45 temperature and very visible in the climate change debate (Hartmann et al., 2013). More 46 than two centuries of SST observations together with other in situ data for surface ocean 47 are assembled in the International Comprehensive Ocean-Atmosphere Data Set (ICOADS, 48 Woodruff et al., 1987; Freeman et al., 2017). These observations are irregularly distributed 49 in space and time. A typical preparatory step for their use in climate studies is "binning," 50 i.e., splitting them into subsamples, contained in non-overlapping spatiotemporal "bins", 51 usually grid boxes of a regular space-time grid (hereinafter, "bins" are monthly $1^{\circ} \times 1^{\circ}$ 52 grid boxes), and reporting statistical summaries of each bin's sample, e.g., number of ob-53 servations \mathcal{N}_o in the bin, their sample mean \mathcal{M}_o , standard deviation (SD) \mathcal{S}_o , etc. By 54 construction, each of these statistics forms a gridded field, albeit usually incomplete. In 55 lieu of averages over the complete bin's volume, which are generally unavailable, bin means 56 \mathcal{M}_o (a.k.a. "super-observations": Smith & Reynolds, 2005; Kennedy, 2014) are often used 57 as input data for objective analyses or data assimilation; hence having reliable uncer-58 tainty estimates for binned data averages is important. For ship data this importance 59 is especially high because of ships' domination, as an observational platform, in the early 60 part of historical data record. 61

⁶² If binned observations were independent and identically distributed (i.i.d.) with ⁶³ variance

63 64

$$\sigma_{\mathcal{B}o}^2 \stackrel{\text{def}}{=} \mathbb{E}\mathcal{S}_o^2,\tag{1}$$

the error SD estimate (ESDE) for \mathcal{M}_o would be

 $e_{\mathcal{M}o} \stackrel{\text{\tiny def}}{=} \sigma_{\mathcal{B}o} / \sqrt{\mathcal{N}_o} \,. \tag{2}$

Hereinafter label "def" above the "=" sign introduces an expression on the left of this 67 sign as denoting the expression on its right, and $\mathbb E$ denotes mathematical expectation. 68 Error estimates computed by (2), but with $\sigma_{\mathcal{B}o}$ calculated as the root-mean-square (RMS) 69 over the best data coverage period of all bins' SDs in that location, was introduced by 70 Kaplan et al. (1997) and used for objective analyses of historical SST observations by 71 Kaplan et al. (1998), Ilin and Kaplan (2009), and, with further modifications to σ_{Bo} es-72 timate, by Karspeck et al. (2012). Their analyzed fields and uncertainty estimates pro-73 vide some indirect validation for such uses of formula (2). 74

However, a direct comparison of the RMS difference between ICOADS bin means 75 and satellite SST data with error estimates based on (2), while showing general large-76 scale agreement between global patterns of error magnitude, had many regional and smaller-77 scale differences (Rayner et al., 2010, cf. their Figures 1e vs. 1f). There were many pos-78 sible reasons for this lack of detailed agreement: likely failure of the i.i.d. assumption, 79 since the binned samples included observations from different platform types (ships, moor-80 ings, drifting buoys), obtained by different observational methods and affected by biases 81 associated with particular methods of observations, platform types, or individual plat-82 forms (e.g., persistent thermometer biases on individual ships). Furthermore, the inter-83 pretation of that comparison was complicated by the dependence of the satellite-based 84 SST data set on the in situ observations themselves. 85

A high-resolution interpolated SST analysis product (available on a daily $0.05^{\circ} \times 0.05^{\circ}$ 86 grid), based on the satellite data, independent of the concurrent in situ SST observations, 87 and accompanied by verified uncertainty estimates, had become available several years 88 ago (Merchant et al., 2014). The actual RMS differences for 1992-2010 between bin-averaged 89 SST from ship observations and from this independent satellite-based analysis are com-90 pared here with their estimates based on the random error model and the analysis un-91 certainty estimates. While the bias structure of ship SST observations is complicated and 92 remains a subject of active research (Kent et al., 2017), it is hypothesized here that a 93 large part of biases in bin-averaged ship SST data can be approximated by its seasonally-94 dependent component. Once the climatological average is removed from the difference 95 between ship and satellite bin means, the residual anomaly is treated as a combination 96 of random errors. Separate estimates are obtained for sampling and measurement error 97 components in bin-averaged ship SST data. 98

Section 2 describes data sets used and their pre-processing for this study. Section 3
 presents error models, their estimates, and the technique of their comparison with the
 RMS of difference anomaly between bin-averaged versions of ICOADS ship SST and the
 satellite analysis product. Section 4 presents the results, which are discussed in section 5.
 Conclusions are given in section 6.

104 2 Data

105

2.1 High-resolution satellite SST analysis product

High-resolution globally-complete satellite SST data set, independent of in situ data
(Merchant et al., 2014), produced within the Climate Change Initiative (CCI) of the European Space Agency (ESA), is used. It is based on the consistent re-processing of major global streams of the infrared satellite SST data, namely, the data from (Advanced)

Along-Track Scanning Radiometer and from the Advanced Very High Resolution Ra-110 diometer missions, with the deliberate avoidance of product dependency on the concur-111 rent in situ SST observations (coefficients in SST retrievals were computed by optimal 112 estimation, based on the atmospheric radiative transfer simulations, rather than by best-113 fitting in situ SST observations). In addition to the more traditional "skin" SST, the time-114 adjusted temperature at 20 cm depth was also produced, by modeling the near-surface 115 thermally-stratified ocean layer. These temperature values with their uncertainty esti-116 mates were fed into the optimal interpolation system for the U.K. Met Office Ocean Sea 117 Surface Temperature and Sea Ice Analysis (OSTIA, Donlon et al., 2012; Roberts-Jones 118 et al., 2012, 2016), producing globally-complete ocean temperature fields at 20 cm depth, 119 $0.05^{\circ} \approx 6$ km spatial resolution, interpretable as local-time daily averages, with uncertainty 120 estimates, for 09/1991–12/2010. This product, known as ESA SST CCI Analysis, ver-121 sion 1.0, hereinafter will be referred to as "CCI Analysis" or simply "CCI." The period 122 of complete 19 years (1992–2010) and 75°S–75°N global ocean domain will be used. 123

124

142

144

2.2 Ship Observations of SST

Ship observations of SST in ICOADS (Release 3.0; Freeman et al., 2017) were iden-125 tified by the "Platform Type" indicator value (PT=5), corresponding to the "ship" ob-126 servational platform type, and put through the ICOADS own quality control (QC) sys-127 tem with settings for the "enhanced Monthly Summary Trimmed Group". For each ship 128 SST observation o that passed QC, its local time and date were computed and included 129 into its record (only Coordinated Universal Time and date are in the ICOADS own data 130 format). Then "match-up" SST value a^o and its uncertainty SD e^{ao} from the CCI Anal-131 ysis to this ship observation o were extracted and also added to the record for o. Match-132 up values from the CCI Analysis to the given ship observation are the analyzed SST value 133 and its uncertainty SD for the daily $0.05^{\circ} \times 0.05^{\circ}$ grid box within whose time-space lim-134 its this ship observation was made. (A small number of ICOADS ship SST observations 135 that passed QC but did not have CCI Analysis match-up values were excluded from this 136 study.) 137

138 2.3 Data Preparation

¹³⁹ Consider bin \mathcal{B} , representing a grid box of a regular monthly 1°×1° grid, and a sam-¹⁴⁰ ple \mathcal{B}_o of \mathcal{N}_o SST observations from ships that were taken within its space and time lim-¹⁴¹ its and successfully passed ICOADS QC:

$$\mathcal{B}_{o} \stackrel{\text{\tiny def}}{=} \{o_{1}, o_{2}, \cdots, o_{\mathcal{N}_{o}}\}$$

This "binned" sample is characterized by its mean \mathcal{M}_o and SD \mathcal{S}_o , as follows:

$$\mathcal{M}_{o} \stackrel{\text{def}}{=} \frac{1}{\mathcal{N}_{o}} \sum_{i=1}^{\mathcal{N}_{o}} o_{i}, \qquad \mathcal{S}_{o}^{2} \stackrel{\text{def}}{=} \frac{1}{\mathcal{N}_{o} - 1} \sum_{i=1}^{\mathcal{N}_{o}} \left(o_{i} - \mathcal{M}_{o} \right)^{2}.$$
(3)

¹⁴⁵ Note that S_o can only be computed if $\mathcal{N}_o > 1$. Therefore bins with only one ship obser-¹⁴⁶ vation ($\mathcal{N}_o=1$) form a special class of data samples: their means, but not variability can ¹⁴⁷ be estimated directly from their data. Dealing with this more complicated subset is left ¹⁴⁸ for further investigaton, and only bins with $\mathcal{N}_o \geq 2$ are considered in this study.

¹⁴⁹ Consider also a set \mathcal{B}_a of \mathcal{N}_a SST values from the CCI Analysis for all daily $0.05^{\circ} \times 0.05^{\circ}$ ¹⁵⁰ grid boxes contained within that same bin \mathcal{B} as above:

 $\mathcal{B}_a \stackrel{\text{def}}{=} \{a_1, a_2, \cdots, a_{\mathcal{N}_a}\}.$

Its statistics \mathcal{M}_a and \mathcal{S}_a are computed similarly to (3):

$$\mathcal{M}_a \stackrel{\text{def}}{=} \frac{1}{\mathcal{N}_a} \sum_{j=1}^{\mathcal{N}_a} a_j, \qquad \mathcal{S}_a^2 \stackrel{\text{def}}{=} \frac{1}{\mathcal{N}_a - 1} \sum_{j=1}^{\mathcal{N}_a} \left(a_j - \mathcal{M}_a \right)^2. \tag{4}$$

¹⁵⁴ Unless the land or ice cover are present within the bin \mathcal{B} , the number of data points in ¹⁵⁵ \mathcal{B}_a is quite large: typically, $\mathcal{N}_a \sim 20 \times 20 \times 30 \gg \mathcal{N}_o$ for ocean locations. In fact, con-¹⁵⁶ straint $\mathcal{N}_a \geq 2$ is satisfied automatically for all non-empty binned samples \mathcal{B}_a from the

157 CCI Analysis.

161

164

166

171

195

197

Recall that for each $o_i \in \mathcal{B}_o$, its CCI SST match-up a_i^o has been identified and stored in the record for o_i (Section 2.2). Therefore it is easy to assemble a sample of CCI Analysis match-ups to ship observations in \mathcal{B}_o :

$$\mathcal{B}_{ao} \stackrel{\scriptscriptstyle{\mathsf{der}}}{=} \{a_1^o, a_2^o, \cdots, a_{\mathcal{N}_o}^o\}$$

and to compute its statistics \mathcal{M}_{ao} and \mathcal{S}_{ao} analogously to (3). Additionally, differences between ship observations and their CCI Analysis SST match-ups

$$d_i \stackrel{\text{\tiny der}}{=} o_i - a_i^o, \quad i = 1, \cdots, \mathcal{N}_o \tag{5}$$

¹⁶⁵ are binned as well, resulting in the sample

$$\mathcal{B}_{d} \stackrel{\scriptscriptstyle{\mathsf{der}}}{=} \{d_{1}, d_{2}, \cdots, d_{\mathcal{N}_{o}}\}$$

and its bin statistics \mathcal{M}_d and \mathcal{S}_d .

168 It will prove useful to have bin statistics for CCI Analysis uncertainties pre-computed 169 as well. These are calculated in exactly the same way as was done above for correspond-170 ing SST values. Specifically, let

$$\mathcal{B}_{ea} \stackrel{\text{\tiny def}}{=} \{e_1^a, e_2^a, \cdots, e_{\mathcal{N}_a}^a\},\$$

where each e_j^a is the uncertainty SD for the CCI Analysis SST value $a_j \in \mathcal{B}_a$ and compute \mathcal{M}_{ea} , \mathcal{S}_{ea} analogously to (4). For the sample of the CCI Analysis uncertainty value match-ups to the ship observations in \mathcal{B}

$$\mathcal{B}_{eao} \stackrel{\text{def}}{=} \{ e_1^{ao}, e_2^{ao}, \cdots, e_{\mathcal{N}_o}^{ao} \}$$

where each e_i^{ao} is the uncertainty SD for the CCI Analysis SST value $a_i^o \in \mathcal{B}_{ao}$ and compute \mathcal{M}_{eao} , \mathcal{S}_{eao} analogously to (3).

Described above calculations of \mathcal{M}_x and \mathcal{S}_x are done for all monthly $1^{\circ} \times 1^{\circ}$ bins 178 with sample sizes larger than one, i.e., for bins with $\mathcal{N}_o \geq 2$, when x = o, ao, d, or eao, 179 and for all bins with non-empty samples when x = a or ea. Temporal attribution of 180 bin statistics $\mathcal{M}_x(y,m)$, $\mathcal{S}_x(y,m)$ is done using climatological (calendar) month m =181 $1, \dots, 12$ (January-December) and year $y = 1, \dots, 19$ (corresponding to 1992–2010). 182 Depending on location, statistics $\mathcal{M}_x(y,m)$, $\mathcal{S}_x(y,m)$ for x = 0, ao, d, and eao might 183 not be available for all (y, m) combinations. To simplify the formal treatment of their tem-184 poral averaging, for each location if there are $Y_m > 0$ years for month m with $\mathcal{N}_o \geq 2$, 185 other years are skipped, while available years are renumbered as $y = 1, \dots, Y_m$ for that 186 month. If only M < 12 climatological months are available, these are renumbered as well 187 from 1 to M. This change in temporal arguments is applied to statistics \mathcal{N}_o and \mathcal{M}_x , \mathcal{S}_x 188 for x = o, ao, d, and eao, as well as for \mathcal{M}_a and \mathcal{M}_{ea} , even though the latter two are 189 available for all (y,m) for bins in ocean locations; statistics $\mathcal{N}_a, \mathcal{S}_a, \mathcal{S}_{ea}$ remain attributed 190 to the full set of months and years in 1992-2010 period. 191

¹⁹² With these definitions, available bin averages of SST from ship observations \mathcal{M}_o ¹⁹³ and corresponding bin averages from CCI Analysis \mathcal{M}_a have the same temporal argu-¹⁹⁴ ments, with the timeseries length

$$N \stackrel{\text{\tiny def}}{=} \sum_{m=1}^{M} Y_m \,. \tag{6}$$

¹⁹⁶ Hence for their differences

$$d_{\mathcal{M}} \stackrel{\text{\tiny def}}{=} \mathcal{M}_o - \mathcal{M}_a \,, \tag{7}$$

¹⁹⁸ the full period RMS is computed as

201

$$\mathcal{D} \stackrel{\text{\tiny def}}{=} \left[\frac{1}{N} \sum_{m=1}^{M} \sum_{y=1}^{Y_m} d_{\mathcal{M}}(y,m)^2 \right]^{1/2} \,. \tag{8}$$

200 3 Methods

3.1 Models and assumptions

²⁰² CCI Analysis values a_j are estimates of water temperature at 20 cm depth, aver-²⁰³ aged over daily $0.05^{\circ} \times 0.05^{\circ}$ grid boxes. Corresponding "true" values t_j^a are averages of ²⁰⁴ true water temperature t at 20 cm depth over such grid boxes, so for values within bin ²⁰⁵ \mathcal{B}

218

221

226

228

229

232

$$t_j^a = a_j + \varepsilon_j^a, \qquad \qquad j = 1, \cdots, \mathcal{N}_a; \qquad (9)$$

$$\mathbb{E}\varepsilon_j^a = 0, \quad \mathbb{E}\left(\varepsilon_j^a\right)^2 = \left(e_j^a\right)^2, \quad j = 1, \cdots, \mathcal{N}_a, \tag{10}$$

where ε_j^a are the CCI analysis errors. These are assumed uncorrelated with the analyzed 208 values a_j , since the CCI Analysis is a form of optimal interpolation (Lorenc, 1986). Anal-209 ysis errors for different grid boxes, however, are not mutually independent, especially when 210 these are not greatly separated in time and space. CCI Analysis uses the increased range 211 (20–350 km) of spatial decorrelation scales of background error that resulted in improved 212 feature resolution (Roberts-Jones et al., 2016), hence the analysis error is likely domi-213 nated by spatial scales larger than $1^{\circ} \times 1^{\circ}$. Since the OSTIA background solution uses 214 day-to-day persistence and relaxes to reference climatology with the 30 day decorrela-215 tion time scale (Donlon et al., 2012), a near-perfect correlation of the analysis error within 216 monthly $1^{\circ} \times 1^{\circ}$ bins is assumed here: 217

$$\mathbb{E}\left(\varepsilon_{j}^{a}\varepsilon_{k}^{a}\right)\approx e_{j}^{a}e_{k}^{a}, \qquad j,k=1,\cdots,\mathcal{N}_{a}.$$
(11)

For conceptual simplicity, the same "truth" definition, as for the CCI Analysis (9), is used for ship observations as well:

$$o_i = t_i^{ao} + b + \varepsilon_i^o, \quad i = 1, \cdots, \mathcal{N}_o, \tag{12}$$

where t_i^{ao} is true 20 cm depth temperature averaged over the daily $0.05^{\circ} \times 0.05^{\circ}$ grid box containing ship observation o_i , bias b is assumed constant within each $1^{\circ} \times 1^{\circ}$ monthly bin, thus it does not depend on i in (12). Measurement errors ε_i^o are assumed independent of true temperature variations t_i^{ao} and i.i.d. within each bin, with

$$\mathbb{E}\varepsilon_i^o = 0, \quad \mathbb{E}\left(\varepsilon_i^o\right)^2 = \sigma_o^2, \quad i = 1, \cdots, \mathcal{N}_o, \tag{13}$$

where σ_o^2 is an (unknown) measurement error variance.

Introduce bin average and intra-bin variance of true SST

$$\theta \stackrel{\text{\tiny def}}{=} \mathcal{M}_{ta} , \qquad \upsilon^2 \stackrel{\text{\tiny def}}{=} \frac{\mathcal{N}_a - 1}{\mathcal{N}_a} \mathcal{S}_{ta}^2 , \qquad (14)$$

where \mathcal{M}_{ta} and \mathcal{S}_{ta} are the mean and variance of the set of true SST values in the bin

231 B:

$$\mathcal{B}_{ta} \stackrel{\scriptscriptstyle{\mathrm{der}}}{=} \{t_1^a, t_2^a, \cdots, t_{\mathcal{N}_a}^a\}$$

Note that v definition above changes the denominator of S_{ta}^2 from \mathcal{N}_a-1 to \mathcal{N}_a , since (14)

uses discrete analogues of the true value integrals over the bin, rather than statisticalestimates.

Another important assumption is that times and locations of ship observations are 236 random and uniformly distributed over the bin's volume. Hence the true SST match-237 ups to them form a set of \mathcal{N}_o equiprobable draws 238

$$\mathcal{B}_{tao} \stackrel{\text{\tiny def}}{=} \{t_1^{ao}, t_2^{ao}, \cdots, t_{\mathcal{N}o}^{ao}\}$$

 $\varepsilon_s \stackrel{\text{\tiny def}}{=} \mathcal{M}_{tao} - \theta$

from the full set \mathcal{B}_{ta} of the true SST values in the bin. Based on statistical theorems that 240 lay the foundation of the classical Monte Carlo method for evaluating definite integrals 241 (e.g., Section 3.2 of Robert & Casella, 2004), sample mean \mathcal{M}_{tao} and variance \mathcal{S}_{tao}^2 of 242 this set of random draws \mathcal{B}_{tao} are unbiased estimates of the true mean and variance of 243 the bin 244 245

$$\mathbb{E}\mathcal{M}_{tao} = \theta \,, \quad \mathbb{E}\mathcal{S}^2_{tao} = v^2 \,, \tag{15}$$

and the error of sample mean, a.k.a. sampling error, 246

247

249

252

254

256

257

260

263

265

269

271

239

has variance 248

 $\mathbb{E}\varepsilon_s^2 = v^2/\mathcal{N}_o$ (for detailed derivation see Section 2.10 of Cochran, 1997). 250

3.2 Single bin statistics 251

3.2.1 CCI Analysis samples

Averaging equations (9) over j and using (14), obtain 253

$$=\mathcal{M}_a + \mathcal{M}_{\varepsilon a},\tag{17}$$

(16)

where $\mathcal{M}_{\varepsilon a}$ is the CCI Analysis error, averaged over the bin. Based on (10) and (11), 255

θ

$$\mathbb{E}\mathcal{M}_{\varepsilon a} = 0,$$

$$e^{2} \quad \stackrel{\text{def}}{=} \mathbb{E}\mathcal{M}^{2} \quad = \quad 1 \quad \sum_{\alpha}^{\mathcal{N}_{a}} \mathbb{E}(e^{a}e^{a}) \approx \quad 1 \quad \sum_{\alpha}^{\mathcal{N}_{a}} e^{a}e^{a} = \quad 1 \quad \left(\sum_{\alpha}^{\mathcal{N}_{a}} e^{a}e^{a}\right)$$

$$e_{\mathcal{M}a}^{2} \stackrel{\text{def}}{=} \mathbb{E}\mathcal{M}_{\varepsilon a}^{2} = \frac{1}{\mathcal{N}_{a}^{2}} \sum_{j,k=1}^{\mathcal{N}_{a}} \mathbb{E}\left(\varepsilon_{j}^{a}\varepsilon_{k}^{a}\right) \approx \frac{1}{\mathcal{N}_{a}^{2}} \sum_{j,k=1}^{\mathcal{N}_{a}} e_{j}^{a} e_{k}^{a} = \frac{1}{\mathcal{N}_{a}^{2}} \left(\sum_{j=1}^{\mathcal{N}_{a}} e_{j}^{a}\right)^{2} = \mathcal{M}_{ea}^{2}.$$
(18)

Subtracting (17) from (9), averaging squares of both sides over j, find, using (14), 258 for mathematical expectation of both sides 259

$$\frac{\mathcal{N}_a}{\mathcal{N}_a - 1} v^2 = \mathbb{E}\mathcal{S}_a^2 + \mathbb{E}\mathcal{S}_{\varepsilon a}^2 \,, \tag{19}$$

where $S_{\varepsilon a}^2$ is the sample variance of the CCI analysis error in the bin. Using (10), (11), 261 and (18), derive 262

$$\mathbb{E}\mathcal{S}_{\varepsilon a}^{2} = \frac{1}{\mathcal{N}_{a} - 1} \sum_{i=1}^{\mathcal{N}_{o}} \mathbb{E}(\varepsilon_{i}^{a})^{2} - \frac{\mathcal{N}_{a}}{\mathcal{N}_{a} - 1} \mathbb{E}\mathcal{M}_{\varepsilon a}^{2} = \mathcal{S}_{ea}^{2} \,.$$
(20)

Substituting (20) into (19), find that 264

$$\hat{v}^2 = \frac{\mathcal{N}_a - 1}{\mathcal{N}_a} \left(\mathcal{S}_a^2 + \mathcal{S}_{ea}^2 \right) \,, \tag{21}$$

is an unbiased estimator of v^2 . 266

Subselecting from equations (9) those for the CCI analysis match-ups a^{o} to ship 267 observations from \mathcal{B}_{o} , 268 t_i^a

$$a^{ao}_i = a^o_i + \varepsilon^{ao}_i, \quad i = 1, \cdots, \mathcal{N}_o,$$
(22)

analogously to the derivation of (21), find that 270

$$\hat{v}_o^2 = \mathcal{S}_{ao}^2 + \mathcal{S}_{eao}^2 \,, \tag{23}$$

is another unbiased estimator of v^2 , but based on a small subset of CCI Analysis grid 272

points (\mathcal{N}_o match-ups to ship observations) than the full set of \mathcal{N}_a CCI Analysis points 273 in \mathcal{B} , used in (21). 274

3.2.2 Ship observations sample

Averaging both sides of (12) over i, obtain

$$\mathcal{M}_o = \mathcal{M}_{tao} + b + \mathcal{M}_{\varepsilon o},\tag{24}$$

(25)

where $\mathcal{M}_{\varepsilon o}$ is the bin mean of measurement errors with, based on (13),

 $\mathbb{E}\mathcal{M}_{\varepsilon o}^2 = \sigma_o^2/\mathcal{N}_o\,.$

Subtracting (24) from (12), averaging squares of both sides over i, find for mathematical expectation of both sides, using (13), (15), and recalling (1),

$$\sigma_{\mathcal{B}o}^2 = \mathbb{E}\mathcal{S}_{tao}^2 + \mathbb{E}\mathcal{S}_{\varepsilon o}^2 = \upsilon^2 + \sigma_o^2 \,. \tag{26}$$

Inserting t_i^{ao} from (22) into (12), obtain, recalling (5),

$$d_i = b + \varepsilon_i^o + \varepsilon_i^{ao}, \qquad i = 1, \cdots, \mathcal{N}_o.$$
(27)

By taking sample variances of both sides of (27) and considering their expectations, find an unbiased estimate of ship SST measurement error σ_o^2 :

$$\hat{\sigma}_o^2 = \mathcal{S}_d^2 - \mathcal{S}_{eao}^2 \,. \tag{28}$$

3.2.3 Bin mean differences

For differences between bin-averaged ship observations and CCI Analysis, defined by (5):

$$d_{\mathcal{M}} = b + \varepsilon_{d\mathcal{M}} \,, \tag{29}$$

292 where

275

277

282

284

287

288

291

293

295

296

297

301

$$\varepsilon_{d\mathcal{M}} \stackrel{\text{\tiny def}}{=} \varepsilon_s + \mathcal{M}_{\varepsilon o} + \mathcal{M}_{\varepsilon a}$$
 ,

and based on (16), (18), (25), and (26),

$$\mathbb{E}\varepsilon_{d\mathcal{M}} = 0, \qquad e_{d\mathcal{M}}^2 \stackrel{\text{def}}{=} \mathbb{E}\varepsilon_{d\mathcal{M}}^2 = \sigma_{\mathcal{B}o}^2/\mathcal{N}_o + \mathcal{M}_{ea}^2.$$
(30)

3.3 Statistics for a temporal sample of bins

3.3.1 Actual RMS differences

²⁹⁸ Consider a temporal sample of bin statistics for a certain location of the bin. Due ²⁹⁹ to (29), straight RMS \mathcal{D} of differences $d_{\mathcal{M}}(y,m)$, calculated by (8), is affected by bias ³⁰⁰ b. Bias estimate $\hat{b}_c(m)$ is obtained by climatological averaging of $d_{\mathcal{M}}(y,m)$:

$$\hat{b}_c(m) = \frac{1}{Y_m} \sum_{y=1}^{Y_m} d_{\mathcal{M}}(y,m), \quad m = 1, \cdots, M.$$
 (31)

The RMS of the differences $d_{\mathcal{M}}$ with the estimated bias removed, taking into account the reduction in the number of degrees of freedom (DOF) from (6) to

304
$$\sum_{m=1}^{M} (Y_m - 1) = N - M$$

305 becomes

$$\mathcal{D}' = \left[\frac{1}{N-M} \sum_{m=1}^{M} \sum_{y=1}^{Y_m} \left(d_{\mathcal{M}}(y,m) - \hat{b}_c(m)\right)^2\right]^{1/2} .$$
(32)

306



Figure 1. Comparison of monthly $1^{\circ} \times 1^{\circ}$ bin-averaged ($\mathcal{N}_{o} \geq 2$) ICOADS ship SST obserations with the ESA CCI Analysis for 1992-2010: (a) Percentage of ICOADS ship SST bins with $\mathcal{N}_o \geq 2$ among all bins with data ($\mathcal{N}_o \geq 1$); (b) RMS \mathcal{D} , °C, of difference $d_{\mathcal{M}}$ between bin-averaged ship and satellite data; (c) Zonal averages of differences $d_{\mathcal{M}}$, °C; (d) DOF in anomalies of binaveraged ship data (zero DOF grids are shown as missing data, in white); (e) RMS \mathcal{D}' , °C, of ship-satellite difference anomalies $d_{\mathcal{M}}$, \hat{b}_c ; (f) Estimate \mathcal{E} , °C, of ship-satellite difference anomaly RMS \mathcal{D}' ; (g) Relative difference ρ , %, between \mathcal{D}' and \mathcal{E} . Numbers at the end of panel labels are: for (a),(d) – global averages of displayed fields; for (b),(e),(f),(g) – global RMS of displayed fields.

3.3.2 Estimated RMS differences and errors

Based on (29) and (31), 309

$$\mathbb{E}\mathcal{D}^{\prime 2} = \frac{1}{N-M} \mathbb{E}\left[\sum_{m=1}^{M} \sum_{y=1}^{Y_m} \left(\varepsilon_{d\mathcal{M}}(y,m) - \frac{1}{Y_m} \sum_{q=1}^{Y_m} \varepsilon_{d\mathcal{M}}(y,m)\right)^2\right] = \frac{1}{N-M} \left[\sum_{m=1}^{M} \sum_{q=1}^{N} \left(\sum_{m=1}^{Y_m} \sum_{q=1}^{N} \sum_{m=1}^{N} \sum_{m=1}$$

$$= \frac{1}{N-M} \sum_{m=1}^{M} \left[\mathbb{E} \sum_{y=1}^{Y_m} \varepsilon_{d\mathcal{M}}(y,m)^2 - \frac{1}{Y_m} \mathbb{E} \left(\sum_{q=1}^{Y} \varepsilon_{d\mathcal{M}}(y,m) \right) \right]$$
$$= \sum_{y=1}^{M} \frac{\mu_m}{Y_m} \sum_{q=1}^{Y_m} e_{d\mathcal{M}}(y,m)^2,$$

$$= \sum_{m=1}^{M} \frac{\mu_m}{Y_m} \sum_{y=1}^{T_m} e_{d\mathcal{M}}(y,m)^2$$

where 313

314

319

326

308

310

$$\mu_m \stackrel{\text{def}}{=} (Y_m - 1)/(N - M), \quad m = 1, \cdots, M.$$
 (33)

=

is the portion of the total DOF due to each climatological month m (note that $\sum_{m=1}^{M} \mu_m =$ 315 1). Based on (30), 316

$$e_{d\mathcal{M}}(y,m)^2 = \sigma_{\mathcal{B}o}(m)^2 / \mathcal{N}_o(y,m) + \mathcal{M}_{ea}(y,m)^2$$

and 318

$$\mathbb{E}\mathcal{D}'^{2} = \sum_{m=1}^{M} \mu_{m} \sigma_{\mathcal{B}o}(m)^{2} / \mathcal{N}_{o}^{h}(m) + \sum_{m=1}^{M} \mu_{m} \mathcal{M}_{ea}^{q}(m)^{2} , \qquad (34)$$

where 320

321
$$\mathcal{N}_{o}^{h}(m) \stackrel{\text{def}}{=} \left[\frac{1}{Y_{m}} \sum_{y=1}^{Y_{m}} \mathcal{N}_{o}(y,m)^{-1} \right]^{-1}, \qquad \mathcal{M}_{ea}^{q}(m) \stackrel{\text{def}}{=} \left[\frac{1}{Y_{m}} \sum_{y=1}^{Y_{m}} \mathcal{M}_{ea}(y,m)^{2} \right]^{1/2}.$$

are harmonic $\mathcal{N}_o(y,m)$ and quadratic $\mathcal{M}_{ea}(m)$ means of \mathcal{N}_o and \mathcal{M}_{ea} , respectively, over 322 available years $y = 1, \cdots, Y_m$. 323

An estimate of $\sigma_{\mathcal{B}o}(m)^2$ is computed as pooled variance (Section 9.2.16 in Von Storch 324 & Zwiers, 2001) of binned samples over all available years $y = 1, \dots, Y_m$: 325

$$\hat{\sigma}_{\mathcal{B}o}^2(m) \stackrel{\text{\tiny def}}{=} \sum_{y=1}^{Y_m} \gamma(y,m) \mathcal{S}_o(y,m)^2, \quad m = 1, \cdots, M,$$
(35)

where weighting coefficients are 327

$$\gamma(y,m) \stackrel{\text{def}}{=} \left[\mathcal{N}_o(y,m) - 1 \right] / \Gamma(m), \quad y = 1, \cdots, Y_m, \quad m = 1, \cdots, M, \tag{36}$$

$$\Gamma(m) \stackrel{\text{def}}{=} \sum_{y=1}^{m} \left[\mathcal{N}_o(y,m) - 1 \right], \qquad m = 1, \cdots, M.$$
(37)



Empirical probability of $|\rho|$ as a function of DOF in anomalies of bin-averaged ship SST observations

Figure 2. Empirical probability (color) of $|\rho|$ calculated for 10%-wide segments of 0–100% interval (vertical axis) for each of 12-wide sub-ranges of the complete 1-216 range of the possible DOF in the climatological anomaly sample for 1992-2010 (horizontal axis). White lines are contours of cumulative empirical probability of $|\rho|$, conditional on the given DOF range, corresponding to the values of 0.5, 0.67, 0.95, and 0.99, as labels indicate.

Substituting estimate $\hat{\sigma}_{\mathcal{B}o}(m)^2$ from (35) for the value of $\sigma_{\mathcal{B}o}(m)^2$ in (34), obtain 330 an unbiased estimate for \mathcal{D}'^2 : 331 332

$$\mathcal{E}^2 \stackrel{\text{\tiny def}}{=} \mathcal{E}^2_{\mathcal{M}o} + \mathcal{E}^2_{\mathcal{M}a} \,, \tag{38}$$

where the terms in the right-hand side are estimates of error variances in bin averages 333 of ship observations 334

$$\mathcal{E}_{\mathcal{M}o}^2 \stackrel{\text{def}}{=} \sum_{m=1}^M \mu_m \hat{\sigma}_{\mathcal{B}o}(m)^2 / \mathcal{N}_o^h(m) \tag{39}$$

and of the CCI Analysis 336

$$\mathcal{E}_{\mathcal{M}a}^2 \stackrel{\text{def}}{=} \sum_{m=1}^M \mu_m \mathcal{M}_{ea}^q(m)^2 \,. \tag{40}$$

4 Results 338

335

337

Excluded from this study monthly $1^{\circ} \times 1^{\circ}$ bins with a single ship SST observation 339 constitute a surprisingly large percentage (31.8%) of all ICOADS 1992-2010 monthly $1^{\circ}\!\!\times\!1^{\circ}$ 340 bins with ship SST observations (with any $\mathcal{N}_o > 0$) Figure 1a shows local percentages 341 of bins included in this study ($\mathcal{N}_o \geq 2$) among all bins with ship SST data ($\mathcal{N}_o > 0$), iden-342 tifying better-sampled areas in North Atlantic and North Pacific Oceans and along ship 343 tracks. Figure 1b shows RMS \mathcal{D} of differences $d_{\mathcal{M}}$ between bin averages of ship SST ob-344 servations and CCI Analysis for 1992-2010 (see equations (7) and (8)). 345



(a) ERMSE $\mathcal{E}_{\mathcal{M}o}$ of bin-averaged ship SST, 0.67°C (b) ERMSE $\mathcal{E}_{\mathcal{M}a}$ of bin-averaged CCI, 0.31°C Figure 3. Components of estimated RMS difference anomaly between bin-averaged ship SST observations and CCI Analysis: (a) ERMSE $\mathcal{E}_{\mathcal{M}o}$ of bin-averaged ship SST, °C; (b) ERMSE $\mathcal{E}_{\mathcal{M}a}$ of bin-averaged CCI Analysis SST, °C; (c) Intra-bin SD $\hat{\sigma}^*_{\mathcal{B}o}$ of ship observations, °C; (d) Average error reduction factor $1/\sqrt{\mathcal{N}^*}$; (e) Sampling ESDE \hat{v}^*_o from the CCI Analysis match-ups to ship observations, °C; (f) Measurement ESDE $\hat{\sigma}^*_o$ of ship SST, °C; (g) Sampling ESDE \hat{v}^* from the full CCI Analysis, °C; (h) ship SST random ESDE from Kent and Challenor (2006, their Figure 2), °C; (i) same as (f), but in 30°×30° averages, °C; (j) Relative difference ρ between (h) and (i), %. Numbers at the end of panel labels indicate displayed fields' global RMS.

These differences have substantial mean and seasonal components, as seen in the time-latitude plot of zonally-averaged $d_{\mathcal{M}}$ (Figure 1c). Subtracting from $d_{\mathcal{M}}$ their climatological mean reduces the DOF by one for each climatological month, represented in the data (Figure 1d), but even accounting for the reduced DOF, the RMS \mathcal{D}' of $d_{\mathcal{M}}$ anomaly, calculated by (32) and shown in Figure 1e, is appreciably smaller than \mathcal{D} (8% global RMS reduction).

The difference $d_{\mathcal{M}}$ anomaly is interpreted here as the sum of random errors in bin averages of ship observations and of CCI analysis; the estimate \mathcal{E} of its RMS \mathcal{D}' , based on this model, is computed by equation (38) and shown in Figure 1f. It matches \mathcal{D}' pattern (Figure 1e) in many details. To aid their visual comparison, their difference

 $|\mathcal{E}|$

$$ho = (\mathcal{D}' - \mathcal{E})$$

is expressed as the percentage of the estimate \mathcal{E} and is shown in Figure 1g, where large areas of the actual and estimated RMS agreeing within 10% or so are clearly seen.

The areas of poor agreement in Figure 1g appear to colocate with areas of smaller DOF in Figure 1d. To quantify this relationship, the empirical probability of $|\rho|$ in 10% intervals is shown in Figure 2 for different 12-wide DOF ranges of $d_{\mathcal{M}}$ anomalies (1-12, 13-24,..., 205-216). As DOF increases, $|\rho|$ concentrates more in its interval of smallest values. For more than 67% of points where DOF exceeds 50% of its maximum value (108= $0.5 \times 12 \times (19-1))$, $|\rho| < 20\%$; for more than half of the points, where for DOF exceeds 144 (2/3 of its maximum), $|\rho| < 10\%$.

The variance of difference anomaly between bin-averaged ship SST and CCI Analysis is modeled by (38), as a sum of squares of two components: estimated RMS error (ERMSE) $\mathcal{E}_{\mathcal{M}o}$ of bin-averaged ship observations, calculated by (39) and shown in Figure 3a, and ERMSE of bin-averaged CCI Analysis $\mathcal{E}_{\mathcal{M}a}$, calculated using (40) and shown in Figure 3b. The former clearly dominates: the CCI analysis error represents only 17.6% of the global variance in the total ERMSE \mathcal{E} (Figure 1f).

As seen from (39), ERMSE for bin-averaged ship observations averages over the climatological month m products of intra-bin variance estimates $\hat{\sigma}_{\mathcal{B}o}(m)^2$ with inverse harmonic means $1/\mathcal{N}_o^h(m)$ of observational counts. Figures 3c,d show square roots of these quantities averaged over available climatological months:

356

$$\hat{\sigma}_{\mathcal{B}o}^* \stackrel{\text{\tiny def}}{=} \left[\sum_{m=1}^M \mu_m \hat{\sigma}_{\mathcal{B}o}^2(m)\right]^{1/2}, \qquad 1/\sqrt{\mathcal{N}^*} \stackrel{\text{\tiny def}}{=} \left[\sum_{m=1}^M \mu_m / \mathcal{N}_o^h(m)\right]^{1/2}, \tag{41}$$

where μ_m are defined by (33).

Figure 3c shows, in effect, the ESDE for the bin-averaged ship SST, if all monthly bins in the given location only had single observations in them, while 3d shows the ESDE reduction factor due to the multiple observations. Because of $N_o \ge 2$ constraint, all values shown in 3d do not exceed $\sqrt{1/2} \approx 0.71$; their global RMS is 0.57, and the reductions to much smaller factors are relatively rare: the interquartile range is 0.51–0.64, and only 3.1% of shown grid boxes have a reduction factor below 0.3. As seen from (26), the intra-bin variance $\sigma_{\mathcal{B}o}^2$ of ship observations consists of sampling and measurement error variance components. Using (23), (28), and pooled estimates like (35), these components can be estimated separately; with averaging analogous to (41), obtain

$$\hat{v}_{o}^{*2} = \sum_{m=1}^{M} \mu_{m} \sum_{y=1}^{Y_{m}} \gamma(y,m) \left[\mathcal{S}_{ao}(y,m)^{2} + \mathcal{S}_{eao}(y,m)^{2} \right],$$
(42)

$$\hat{\sigma}_{o}^{*2} = \sum_{m=1}^{M} \mu_{m} \sum_{y=1}^{Y_{m}} \gamma(y,m) \left[\mathcal{S}_{d}(y,m)^{2} - \mathcal{S}_{eao}(y,m)^{2} \right], \qquad (43)$$

where γ is defined by (36),(37). The intra-bin sampling \hat{v}_o^* and measurement $\hat{\sigma}_o^*$ ESDE for ship observations, computed by (42) and (43) are shown in Figures 3e,f. As with $\hat{\sigma}_{\mathcal{B}o}^*$, these are essentially ESDE components for a single observation, which are reduced by the factor $1/\sqrt{\mathcal{N}^*}$ (Figure 3d), when more observations are available.

³⁹⁴ 5 Discussion

388

389

395

399

404

5.1 Sampling error

Estimate \hat{v}_o^* , given by (42) is based on the match-ups of the CCI Analysis SST and its uncertainty to the ship SST observations, a relatively small data sample. An estimate, based on the equation (21) that uses full CCI Analysis and its uncertainty

$$\hat{v}^{*2} = \frac{1}{228} \sum_{m=1}^{12} \sum_{y=1}^{19} \frac{\mathcal{N}_a(y,m) - 1}{\mathcal{N}_a(y,m)} \left[\mathcal{S}_a(y,m)^2 + \mathcal{S}_{ea}(y,m)^2 \right]$$

is shown in Figure 3g. Expectedly, this estimate is larger (by about 10% in areas of high DOF numbers) and smoother than the one based on the incomplete data (Figure 3e).
It has the uncanny similarity in pattern, but generally is larger than the estimate presented by Kennedy et al. (2011, their Figure 1d).

5.2 Measurement error

Kent and Challenor (2006) used the semivariogram method to estimate SST mea-405 surement error in 1970–1997 ICOADS data from ships. They identified pairs of ship SST 406 observations made at the same hour and within 300 km of each other; squared differences 407 between paired observations were binned by distance to construct the semivariogram; 408 a linear fit to its points was extended towards zero distance separation to obtain the mea-409 surement error variance as the semivariogram's nugget. Ship measurement ESDE in $30^{\circ} \times 30^{\circ}$ 410 averages from Kent and Challenor (2006, their Figure 2) is compared here with the mea-411 surement error estimates $\hat{\sigma}_o^*$, averaged to the same $30^{\circ} \times 30^{\circ}$ grid (Figure 3h,i). Two es-412 timates have a great deal of similarity (their pattern correlation is 0.75), despite the dif-413 ferences in the study period and estimation method. Relative difference ρ , shown in Fig-414 ure 3j has global RMS of 18.0%, with $|\rho| \leq 10\%$ in most of grid boxes. Grid boxes with 415 $|\rho| > 10\%$ are generally in the areas of poor data coverage (cf. Figure 1d). 416

Kent and Berry (2008) introduced the measurement error model for marine obser-417 vations that combines random error with a "platform-dependent" bias or "micro-bias", 418 with the randomly distributed value over the platforms (ships). For this kind of error 419 structure, if a bin contains many observations from a relatively small number of plat-420 forms, the error variance of its mean decreases inversely-proportionally to the number 421 of platforms, rather than to the total number of observations. However, since moving 422 ships, even at 14 knots (a relatively slow speed for modern ships), would cross the equa-423 torial $1^{\circ} \times 1^{\circ}$ bin in less than six hours (a typical time interval between ship observations), 424 multiple observations from the same ship would not typically appear in the same bin, 425 thus making equation (2) usable in this study. 426

Kennedy (2014, Table 1) listed published in 1965-2011 ship SST measurement ESDE 427 that did not separate micro-biases from the purely random error parts. There are 19 es-428 timates there, ranging from 0.11° C to 3.5° C, with the median of 1.2° C, and $1-1.3^{\circ}$ C in-429 terquartile range. Depending on the way of averaging measurement error estimates and 430 especially on the averaging domain, global estimates can change appreciably. (Kent and 431 Challenor (2006) report their global ESDE for ship SST random error as 1.2° C, if weighted 432 by ocean area, and 1.3°C, if weighted by number of observations.) Estimates $\hat{\sigma}_{o}^{*}$ here 433 can average to the global RMS of 1.14°C (Figure 3f), 1.13°C (Figure 3i), or 1.21°C, if 434 the latter is constrained to the exact domain, where estimates in Figure 3h (global RMS) 435 of 1.26°C) are available. 436

437 6 Conclusions

Differences for 1992-2010 between monthly $1^{\circ} \times 1^{\circ}$ bin averages (for bins that con-438 tain more than one observation) of ICOADS ship SST and of the ESA SST CCI Anal-439 ysis are presented here as the sum of their climatological bias component and remain-440 ing residuals (anomalies) with magnitudes that agree with the random error model in 441 the areas of sufficient data coverage. The model assumes for ship observations the i.i.d 442 measurement and sampling errors within bins and high intra-bin correlation for the CCI 443 Analysis uncertainty. Location-dependent estimates of ship SST measurement and sam-444 pling error were obtained. Estimates of sampling error are similar in pattern, but larger 445 than those previously published. Ship SST measurement error is consistent with previ-446 ous estimates in spatial pattern and global RMS (1.13-1.21°C, depending on the aver-447 aging domain and procedure). 448

449 Acknowledgments

To the memory of M. Benno Blumenthal (1959-2018), creator of Data Library and Ingrid software, with which all figures in this Letter were calculated and plotted. Discussions with P.C.Cornilon, E.C.Kent, C.J.Merchant, J.J.Kennedy, and N.A.Rayner are gratefully acknowledged. ICOADS, Release 3.0, is obtained at

- 454 https://rda.ucar.edu/datasets/ds548.0/
- 455 ESA SST CCI Analysis for 9/1991–12/2010, version 1.0, is obtained at
- https://catalogue.ceda.ac.uk/uuid/916986a220e6bad55411d9407ade347c
- Supported by grants OCE-1853717 from NSF and NA17OAR4310156 from NOAA. LDEO
 contribution number xxxx.

459 References

467

468

- Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., & Zemp,
- M. (2014). The concept of essential climate variables in support of climate
 research, applications, and policy. Bulletin of the American Meteorological
 Society, 95(9), 1431–1443.
- Cochran, W. G. (1997). Sampling Techniques (3rd ed.). New York: John Wiley &
 Sons.
- ⁴⁶⁶ Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W.
 - (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158.
- ⁴⁶⁹ Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel,
- W. E., et al. (2017). ICOADS Release 3.0: a major update to the historical marine climate record. International Journal of Climatology, 37(5),
 2211–2232.

473	Hartmann, D. L., Tank, A. M. K., Rusticucci, M., Alexander, L. V., Brönnimann,
474	S., Charabi, Y. A. R., et al. (2013). Observations: Atmosphere and surface. In
475	Climate change 2013 the physical science basis: Working group i contribution
476	to the fifth assessment report of the intergovernmental panel on climate change
477	(pp. 159–254). Cambridge University Press.
478	Ilin, A., & Kaplan, A. (2009). Bayesian PCA for Reconstruction of Historical Sea
479	Surface Temperatures. In IJCNN: 2009 INTERNATIONAL JOINT CONFER-
480	ENCE ON NEURAL NETWORKS, VOLS 1-6 (p. 1138+). Int Neural Net-
481	work Soc; IEEE Computat Intelligence Soc. (International Joint Conference on
482	Neural Networks, Atlanta, GA, JUN 14-19, 2009)
483	Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., & Ra-
484	jagopalan, B. (1998). Analyses of global sea surface temperature 1856-1991.
485	Journal of Geophysical Research, 103(C9), 18567-18589.
486	Kaplan, A., Kushnir, Y., Cane, M. A., & Blumenthal, M. B. (1997). Reduced
487	space optimal analysis for historical data sets: 136 years of Atlantic sea surface
488	temperatures. Journal of Geophysical Research, 102(C13), 27835-27860.
489	Karspeck, A. R., Kaplan, A., & Sain, S. R. (2012). Bayesian modelling and ensem-
490	ble reconstruction of mid-scale spatial variability in North Atlantic sea-surface
491	temperatures for 1850-2008. <i>Quarterly Journal of the Royal Meteorological</i>
491	Society 138(662 A) 234-248
402	Kennedy, J. J. (2014) A review of uncertainty in in situ measurements and data
495	sets of sea surface temperature <i>Reviews of Geophysics</i> $52(1)$ 1–32
494	Kennedy I I Bayner N A Smith B O Parker D E & Saunby M (2011)
495	Reassessing biases and other uncertainties in sea surface temperature observa-
490	tions measured in situ since 1850: 1 Measurement and sampling uncertainties
497	<i>Journal of Coonducted Research: Atmospheres</i> 116(D14) D14103
498	Kont F C k Borry D I (2008) According to the marine observing system (as
499	mos)
500	
E 0.4	Kont B (V (hollonor P (C (2006) Loward octimating climatic trands in SST
501	Rent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST.
501 502	Rent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 22(3) 476 486
501 502 503	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486.
501 502 503 504	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO OLIANTIEVING BIASES
501 502 503 504 505	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SUBFACE TEMPERATURE Buildin of the
501 502 503 504 505 506	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Mateomological Society, 08(8), 1601–1616.
501 502 503 504 505 506 507	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Loreng, A. C. (1986). Applying methods for pumerical monther prediction. Quarterly
501 502 503 504 505 506 507 508	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Lowrongl of the Bawal Metaomlogical Society 110(74), 1177, 1104
501 502 503 504 505 506 507 508 509	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194.
501 502 503 504 505 506 507 508 509 510	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, C. K. et al. (2014). See surface temperature detects for alignetic applications.
501 502 503 504 505 506 507 508 509 510 511	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space A grave Climate Change Initiation (SCT)
501 502 503 504 505 506 507 508 509 510 511 512	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Consistence Data Lowmal. 1(2), 170–101.
501 502 503 504 505 506 507 508 509 510 511 512 513	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Bermer, N. A. Karlen, A. Kent, E. C. Burgelda, P. W. Brehen, B. C. K. S.
501 502 503 504 505 506 507 508 509 510 511 512 513 514	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate applications from Phase 1 of the surface state states for the state state states for the state states for the sta
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and biotoxical SST abcomptions. In J. Unit. D. F. Harrison, & D. Starway, (Data)
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Brownedinge of Quart Oke00. Systemed Oxean Observations. Operating of the formation.
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177–1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Dirivice WEDB 206. (Open Color 2002).
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177–1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, Surtaryken 21 25, 2000)
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SST. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009)
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Robert, C., & Casella, G. (2004). Monte Carlo Statistical Methods (2nd ed.). New
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Robert, C., & Casella, G. (2004). Monte Carlo Statistical Methods (2nd ed.). New York: Springer-Verlag.
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating chimatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Roberts-Jones, J., Bovis, K., Martin, M. J., & McLaren, A. (2016). Estimating back-
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Robert, C., & Casella, G. (2004). Monte Carlo Statistical Methods (2nd ed.). New York: Springer-Verlag. Roberts-Jones, J., Bovis, K., Martin, M. J., & McLaren, A. (2016). Estimating background error covariance parameters and assessing their impact in the OSTIA
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Robert, C., & Casella, G. (2004). Monte Carlo Statistical Methods (2nd ed.). New York: Springer-Verlag. Roberts-Jones, J., Bovis, K., Martin, M. J., & McLaren, A. (2016). Estimating background error covariance parameters and assessing their impact in the OSTIA system. Remote Sensing of Environment, 176, 117-138.
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Robert, C., & Casella, G. (2004). Monte Carlo Statistical Methods (2nd ed.). New York: Springer-Verlag. Roberts-Jones, J., Bovis, K., Martin, M. J., & McLaren, A. (2016). Estimating background error covariance parameters and assessing their impact in the OSTIA system. Remote Sensing of Environment, 176, 117-138. Roberts-Jones, J., Fiedler, E. K., & Martin, M. J. (2012). Daily, global, high-
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526	 Kent, E. C., & Challenor, P. G. (2006). Toward estimating climatic trends in SS1. Part II: Random errors. Journal of Atmospheric and Oceanic Technology, 23(3), 476-486. Kent, E. C., Kennedy, J. J., Smith, T. M., Hirahara, S., Huang, B., Kaplan, A., et al. (2017). A CALL FOR NEW APPROACHES TO QUANTIFYING BIASES IN OBSERVATIONS OF SEA SURFACE TEMPERATURE. Bulletin of the American Meteorological Society, 98(8), 1601-1616. Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 112(474), 1177-1194. Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). Geoscience Data Journal, 1(2), 179-191. Rayner, N. A., Kaplan, A., Kent, E. C., Reynolds, R. W., Brohan, P., Casey, K. S., et al. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs99: Sustained Ocean Observations and Information for Society (Vol. 2). ESA Publication WPP-306. (OceanObs' 09, Venice, Italy, September 21-25, 2009) Roberts-Jones, J., Bovis, K., Martin, M. J., & McLaren, A. (2016). Estimating back- ground error covariance parameters and assessing their impact in the OSTIA system. Remote Sensing of Environment, 176, 117-138. Roberts-Jones, J., Fiedler, E. K., & Martin, M. J. (2012). Daily, global, high- resolution SST and sea ice reanalysis for 1985-2007 using the OSTIA system.

- Smith, T. M., & Reynolds, R. W. (2005). A global merged land-air-sea surface temperature reconstruction based on historical observations (1880-1997). Journal of climate, 18(12), 2021-2036.
- Von Storch, H., & Zwiers, F. W. (2001). Statistical analysis in climate research.
 Cambridge University Press.
- Woodruff, S. D., Slutz, R. J., Jenne, R. L., & Steurer, P. M. (1987). A comprehen sive ocean-atmosphere data set. Bulletin of the American meteorological soci ety, 68(10), 1239–1250.