

2 **Sea surface temperature from a geostationary satellite by optimal**  
3 **estimation**

4  
5 C. J. Merchant<sup>a,\*</sup>, P. Le Borgne<sup>b</sup>, H. Roquet<sup>b</sup> and A. Marsouin<sup>b</sup>

6  
7  
8 <sup>a</sup> School of GeoSciences, The University of Edinburgh, UK

9 <sup>b</sup> Météo-France / Centre de Météorologie Spatiale, Lannion, France

10  
11  
12 \*Corresponding author : Tel +44-131-650-5097 ; fax : +44-131-662-0478.

13 E-mail address : [c.merchant@ed.ac.uk](mailto:c.merchant@ed.ac.uk) (C. J. Merchant).

14  
15 **Abstract**

16  
17 Optimal estimation (OE) is applied as a technique for retrieving sea surface temperature (SST) from  
18 thermal imagery obtained by the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) on Meteosat  
19 9. OE requires simulation of observations as part of the retrieval process, and this is done here using  
20 numerical weather prediction fields and a fast radiative transfer model. Bias correction of the simulated  
21 brightness temperatures (BTs) is found to be a necessary step before retrieval, and is achieved by filtered  
22 averaging of simulations minus observations over a time period of 20 days and spatial scale of 2.5° in  
23 latitude and longitude. Throughout this study, BT observations are clear-sky averages over cells of size

24 0.5° in latitude and longitude. Results for the OE SST are compared to results using a traditional non-  
25 linear retrieval algorithm (“NLSST”), both validated against a set of 30108 night-time matches with  
26 drifting buoy observations. For the OE SST the mean difference with respect to drifter SSTs is -0.01 K  
27 and the standard deviation is 0.47 K, compared to -0.38 K and 0.70 K respectively for the NLSST  
28 algorithm. Perhaps more importantly, systematic biases in NLSST with respect to geographical location,  
29 atmospheric water vapour and satellite zenith angle are greatly reduced for the OE SST. However, the OE  
30 SST is calculated to have a lower sensitivity of retrieved SST to true SST variations than the NLSST.  
31 This feature would be a disadvantage for observing SST fronts and diurnal variability, and raises  
32 questions as to how best to exploit OE techniques at SEVIRI’s full spatial resolution.

33

## 34 **1 Introduction**

35

36 The temperature of the ocean surface is routinely retrieved from broad-band infra-red brightness  
37 temperatures (BTs) observed by sensors on both geostationary and polar-orbiting satellites. Sea surface  
38 temperature (SST, also represented in this paper by variable  $x$ ) is one of the most precisely derived  
39 geophysical quantities from satellite observations. Nonetheless, users’ requirements for the accuracy,  
40 resolution and timeliness of SST become more demanding, and increasing attention is paid to understanding  
41 and correcting the differing bias characteristics of various SST products [e.g., Donlon et al, 2007]. In this  
42 study, we explore the nature of biases in SSTs obtained from the Spinning Enhanced Visible and Infra-  
43 Red Imager (SEVIRI) on board the meteorological satellite Meteosat-9. Meteosat-9 is a geostationary  
44 platform located in the equatorial plane at 0° longitude. SEVIRI observes a given location with a constant  
45 viewing geometry every 15 minutes, with ground resolution of ~5 km near nadir and increasing with  
46 increasing zenith angle towards the limb view. The present operational SST product is a “split-window”  
47 retrieval (defined below) based on radiative transfer modelling and empirical offset adjustment [Merchant  
48 and Le Borgne, 2004]. It is generated and distributed on a 3-hourly cycle by Météo-France in the context  
49 of the Ocean and Sea Ice Satellite Application Facility.

50

51 The advantages of the repeated observations of SST available from geostationary orbit are two-fold:  
52 greater daily fractional coverage (within the observed disk) than with a polar-orbiting sensor because of  
53 repeated opportunities to view the surface between the moving field of clouds; and the unique clarity with  
54 which large-scale diurnal variations in SST can be observed using hourly or higher temporal resolution  
55 [e.g., Merchant et al., 2008a].

56

57 The split-window retrieval technique is the “traditional” SST estimator, following the proposal of Anding  
58 and Kauth [1970]. The name refers to the fact that it employs two channels (nominally centred on 11 and  
59 12  $\mu\text{m}$ ) within the same band of relatively high atmospheric transmittance (the window between about 10  
60  $\mu\text{m}$  and 13  $\mu\text{m}$ ). For night-time retrievals, these are often augmented with a third channel around 3.8  $\mu\text{m}$ .  
61 This latter channel can greatly improve the accuracy and precision of SSTs because of the extremely non-  
62 linear variation of emitted radiance with temperature at this wavelength for a surface with at temperature  
63 in the terrestrial range; however, there is significant solar radiation at this wavelength which renders the  
64 channel very difficult to use for SST retrieval for day time scenes.

65

66 The usual form of the SST estimator (whether two-channel or three-channel) is a linear (or nearly linear)  
67 combination of brightness temperatures (BTs) [Anding and Kauth, 1970]:

68

69 (1) 
$$\hat{x} = a_0 + \mathbf{a}^T \mathbf{y}_o$$

70

71 where  $\hat{x}$  is the estimated SST,  $a_0$  is an offset coefficient,  $\mathbf{a}$  is a column vector of weighting coefficients  
72 and  $\mathbf{y}_o$  contains the observed BTs. The coefficients in the retrieval equation may be derived by regression  
73 of observed BTs to in situ measurements or by regression using BTs simulated using radiative transfer  
74 modelling. For two-channel retrievals, one widely adopted form [Walton et al., 1998] is the non-linear

75 SST (NLSST) algorithm, in which coefficients weighting the 11 and 12  $\mu\text{m}$  are a (weak) function of a  
76 prior SST, and this is the form used at Météo-France for SEVIRI SSTs.

77

78 Biases in SST may arise from: error in specifying retrieval coefficients from either forward modelling or  
79 instrumental biases [e.g., Merchant and Le Borgne, 2004]; undetected cloud; stratospheric aerosol; near-  
80 surface stratification (if the target is estimation of the bulk SST); and tropospheric aerosols. Apart from  
81 these biases related to calibration and quality control, linear and near-linear retrievals of SST are  
82 fundamentally subject to two other systematic contributions to error [Merchant et al., 2006a]: prior error  
83 (familiar from the theory of atmospheric sounding) and error arising from non-linearity of the physics of  
84 radiative transfer at infrared wavelengths.

85

86 The prior and non-linearity errors have complex spatial and temporal characteristics and are comparable  
87 in magnitude. Prior error arises as an intrinsic consequence of the form of the retrieval. Clear-sky BTs  
88 over the ocean are influenced by various geophysical quantities: principally by the skin SST and the total  
89 column water vapour in the atmosphere through which the SST is viewed by the radiometer; and more  
90 subtly by the leading modes of the vertical distribution of water vapour, the tropospheric lapse rate, the  
91 air-sea temperature difference, and (via their effects on emissivity and reflectivity) wind speed and  
92 salinity. Not all of the spatio-temporal variations in these geophysical quantities can be fitted  
93 simultaneously by variations in BT at only two or three surface-sensitive wavelengths; the result is that  
94 these variations manifest as geographically complex biases in retrieved SST. These biases are typically  
95 small – less than 1 K – but are not negligible relative to the demands of contemporary users [Donlon et  
96 al., 2007].

97

98 Merchant et al. [2008b] showed, for the case of the Advanced Very High Resolution Radiometer  
99 (AVHRR) on Metop-A, that optimal estimation (OE) with radiance bias correction could significantly  
100 address these biases present in traditional SST retrievals, as well as reducing single-pixel noise in SST.

101 Here, we undertake a comparable study in the context of SEVIRI on Meteosat-9. In comparison with the  
102 previous study, this paper will have a greater focus on the relative properties of coefficient-based and OE-  
103 based SSTs, to better put the properties of the latter approach in context. The issue of bias correction of  
104 simulated BTs used in the OE process was important in the case of Metop-A and is important in this  
105 study too; however, bias correction for a geostationary imager is different because each location is viewed  
106 always at the same zenith angle; thus the bias correction technique developed here is quite different to  
107 that of the previous paper. We also discuss the question of the most appropriate prior SST to use for OE,  
108 which was not fully addressed in the previous paper.

109  
110 Thus, this paper proceeds as follows. In the next section, we describe the data on which this study is  
111 based, including the available prior SSTs. Next, the operational NLSST algorithm is described and  
112 analyzed in terms of its biases compared to in situ observations, its degree of noise amplification, and the  
113 sensitivity of its retrieved SST to real changes in SST and atmospheric water vapour content. Then, a  
114 straightforward implementation of OE for SST is described and is likewise analyzed. Next, we  
115 demonstrate an approach to bias correction, of both BTs and prior SST, that is shown to be effective in  
116 reducing OE SST biases further. The paper concludes with some final discussion of the results and their  
117 implications.

## 119 **2 *Satellite, NWP and in situ data***

120  
121 The study exploits three months of data, at three-hourly intervals from 0000 UTC 1 February to 30 April,  
122 2008, extracted from the operational chain at the Centre de Météorologie Spatiale (CMS), Lannion,  
123 France. As part of these operations, full resolution SEVIRI imagery is screened for cloud. To render the  
124 data set for this study, we use observations averaged over clear-sky pixels to 0.5° resolution in latitude  
125 and longitude, on a 241 by 241 grid from 60° south to 60° north, and 60° west to 60° east. (These co-

126 ordinates are those of the centres of the outermost cells.) The number of clear-sky pixels contributing to  
127 each  $0.5^\circ$  cell is also retained, varying in the data set from 1 to 186; this affects the propagation of  
128 radiometric noise into cell averages and is informative about the prevalence of cloud cover at the  
129 observation time. Four thermal channels were extracted, namely those centred near 3.8, 8.7, 11 and 12  
130  $\mu\text{m}$ , all the BTs being consistent with the EUMETSAT definition of calibration that has been operational  
131 since 17<sup>th</sup> March 2008.

132  
133 Having all four surface-sensitive thermal channels allows calculation for night-time cells of an infra-red  
134 index for Saharan Dust [Merchant et al., 2006a]. The Meteosat-9 version of a simplified Saharan Dust  
135 Index (SDI) is:

$$(2) \quad \mathbf{SDI} = 1.39 + 0.53973(y_{3.8} - y_{8.7}) - 0.820135(y_{11} - y_{12})$$

136  
137  
138  
139 where  $y_\lambda$  is the BT for the channel centred near wavelength  $\lambda$ . Saharan dust is a significant feature of the  
140 disk viewed by Meteosat-9, and is associated with SST biases if not corrected for [Merchant et al.,  
141 2006b]. In this study, however, we retain only BTs where the SDI is less than 0.25; dust-related bias is  
142 not our focus here, and we thereby eliminate the most-affected data. (The SDI is an index that is by  
143 design comparable in magnitude for dust outbreaks to the  $0.55 \mu\text{m}$  aerosol optical depth (AOD), although  
144 it is only loosely correlated with AOD. SDI in excess of 2 is seen for a strong dust event, and clear-sky  
145 pixels are dust contaminated with high probability for  $\text{SDI} > 0.25$ ). Other than for this important  
146 screening step, the 3.8 and 8.7  $\mu\text{m}$  channels are not otherwise used in this study: the operational retrieval  
147 of SST and the OE SST proposed here both rely on the split-window channels only. Not switching from  
148 two-channel to three-channel retrieval for night-time cells has the advantage of consistency of retrieval  
149 throughout the diurnal cycle. Lastly, in order to minimize the complications introduced by the ocean  
150 surface's diurnal cycle of temperature, we use only observations for which the solar zenith angle exceeds  
151  $90^\circ$  (i.e., with the Sun below the horizon).

152

153 Optimal estimation involves forward modelling – simulation of the expected BTs in this case – before  
 154 solving the inverse problem of estimating the SST. The fast radiative transfer simulation is performed  
 155 here by RTTOV9 [Saunders et al., 2002]. The required inputs to the forward model are atmospheric  
 156 profiles of temperature and humidity, and the underlying surface temperature. Numerical weather  
 157 prediction (NWP) fields supply these profiles, and we obtain forecasts at three-hourly intervals and 0.5°  
 158 resolution from Météo-France’s ARPEGE forecasting system. The conclusions of this study are not  
 159 expected to depend on the source of NWP fields, although the geographical distribution of bias  
 160 corrections would undoubtedly be changed were other sources of NWP fields to be used. For each cell for  
 161 each 3-hourly slot, RTTOV9 is run on the ARPEGE profiles to provide a simulated BT for each channel  
 162 and (following Merchant et al. [2008b]) the partial derivatives of the BTs with respect to SST and total  
 163 column water vapour (TCWV, also represented by variable  $w$ ). We assume throughout that the NWP  
 164 fields used are sufficiently close to reality to give an effective point for local linearization of the  
 165 relationship between  $\mathbf{y}$ ,  $x$  and  $w$ . OE will be undertaken using a reduced state vector,  $\mathbf{z}(\mathbf{x}) = \begin{bmatrix} x \\ w \end{bmatrix}$ . Let  $\mathbf{F}$

166 represent RTTOV9. We can then define, for use later in the paper, the tangent linear

167 matrix,  $\mathbf{K} = \left[ \frac{\partial \mathbf{F}(\mathbf{x}_0)}{\partial \mathbf{z}} \right] = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{11}}{\partial w} \\ \frac{\partial y_{12}}{\partial x} & \frac{\partial y_{12}}{\partial w} \end{bmatrix}$ , where  $\mathbf{x}_0$  is the state (a vector containing the NWP profile

168 and SST) for a given cell and slot, around which  $\mathbf{K}$  provides a linearized forward model.

169

170 The NWP SST,  $x_0$ , is a model field into which in situ observations of SST from drifting buoys (“drifters”)  
 171 have been assimilated. Thus,  $x_0$  is not independent of the drifter SSTs against which our derived satellite  
 172 SSTs are to be validated. Because of this, we do not use the NWP SST as our prior for optimal  
 173 estimation, we use it only for forward modelling. This is a refinement compared to Merchant et al.  
 174 [2008b]. Our prior SST,  $x_a$ , where  $a$  stands for *a priori*, will be a satellite-only SST analysis routinely

175 produced at CMS. This analysis SST uses night-time SST observations from Meteosat-9, Metop and  
176 NOAA AVHRR and is produced for a nominal analysis time of 0000 h UTC.

177  
178 We do use the NWP profile to give the prior estimate of TCWV, so  $w_a = w_0$ .

179  
180 The operational chain gathers in situ SST observations taken around 0000 h UTC, from drifting buoys  
181 (“drifters” hereafter). Within each  $0.5^\circ$  cell where one or more drifter observations are available, the  
182 single drifter observation nearest in time to the slot is retained – i.e., there is no averaging of the in situ  
183 observations. In all satellite-drifter comparisons, we use only night-time satellite observations with low  
184 SDI. Doing satellite-drifter comparisons using only night time data minimizes biases associated with  
185 variable diurnal heating of the near-surface ocean; this is appropriate even when developing retrievals that  
186 apply for day and night [e.g., Merchant and Le Borgne, 2004], as here. The matching of satellite and buoy  
187 observations is therefore relatively loose (up to 12 hours difference time and located within a cell of  
188 dimension  $\sim 50$  km). Nevertheless, these satellite-drifter matches are adequate to examine the relative  
189 performance of different SST estimators. The match-up database (MDB) was implemented from the 21  
190 February, with no matches associated with the imagery collected before that point.

191  
192 A final source of SST observations that we use here is the Advanced Along Track Scanning Radiometer  
193 (AATSR). Night-time dual-view three-channel (“D3”) SSTs from AATSR have low noise ( $\sim 0.3$  K at  
194 pixel resolution) and low bias ( $\sim 0.2$  K, plus some effect of Saharan dust when present), and will be used  
195 as an additional reference to explore biases in SEVIRI SSTs. The principal limitation of AATSR SSTs is  
196 coverage: the scanning geometry used to obtain the dual-view capability limits the swath to  $\sim 500$  km. The  
197 sensor is on a polar-orbiting platform (Envisat) with a three-day repeat cycle and  $\sim 10.30$  h equatorial  
198 crossing time, so, typically a given cell will be observed every third night, and an AATSR SST will be  
199 available if there is some clear sky within the cell.

### 3 *Comparisons of available SSTs*

We first compare the CMS analysis SST and the ARPEGE NPW SST. Figure 1(a) shows the CMS analysis SST averaged over the three-month study period. The difference between this field and the NWP SST is shown in Figure 1(b). Consider the differences with relatively fine spatial structure seen off the coasts of Newfoundland and Argentina and south of South Africa. These are areas of strong spatial gradients in SST associated with boundary currents. When cloud cover permits, the satellite SSTs that inform the CMS analysis should place the locations of strong SST gradients accurately, while the sparser information from drifters assimilated into the NWP system may not. These differences therefore are likely at least partly to reflect sampling effects in the two data sets. The other striking feature is the area of negative bias of CMS relative to NWP SST across the Gulf of Guinea and the northeast tropical Atlantic, mostly north of the equator. This bias arises because all the satellites contributing to the CMS analysis tend to be negatively biased across this region, to varying degrees. We interpret this, along with the other more subtle large-scale differences in Figure 1(b), as a result of prior error in the operational SST retrievals. Saharan dust no doubt contributes to negative biases in this area, but is not the principal cause in this case: corrections [Merchant et al., 2006b] for Saharan dust are implemented at CMS, and in any case the mean SDI over this period, Figure 1(c), is too small and of the wrong geographical distribution to account for the major part of the tropical Atlantic bias. Nor is it merely the column amount of water vapour in the atmosphere, Figure 1(d), that accounts for the bias, since other equally humid areas do not show the same bias. Atmospheric conditions in the east tropical Atlantic are dominated by deep convective activity that propagates westward within the inter-tropical convection zone, which moistens the middle and upper troposphere. The vertical distribution of water vapour relevant to the satellite SST biases is that of clear-sky areas in-between convective and other cloudy areas, which one expects to be less humid than the mean water vapour profile of the region. There is extensive cirrus in this area and unscreened thin cirrus may affect satellite retrievals. In short, there are a number of possible factors that distinguish this region, and role of each in the SST biases is at present unclear.

227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251

To explore the biases in the operational SST from SEVIRI, we use the drifter SSTs in our data set. For this and all subsequent comparisons with drifters, we use only night-time data whose SDI is less than 0.25. This minimizes the complications for interpretation that would arise from having diurnal variability and significant Saharan dust present in the validation data. There are 30108 matches in this category. The spatial distribution of the matches in the match-up data base is uneven, as presented in Figure 1(e), and some areas have no in situ observations available. Nonetheless, many of the features of the CMS-NWP comparison are also seen in Figure 1(f), in particular the northeast tropical Atlantic bias. There are also contrasts with the CMS-NWP difference map, as one would expect given that sensors other than SEVIRI contribute to the CMS analysis. The tropical Atlantic negative bias extends across to the coast of South America. It should be borne in mind that some biased drifter observations can get through quality controls and into a data set such as this MDB, and it seems likely that this is the explanation of the area of apparently warm bias north of Egypt. (Our reasons to think this are not given for the sake of brevity.) This may also be the case for the isolated cell with mean difference exceeding +1.5 K in the Gulf of Guinea.

Figure 2 (a) and (b) verify using the MDB that the CMS analysis SST contains biases on large-scales (averages over  $10^\circ$  zones) at the level of a few tenths of kelvin, whereas for the NWP SST the large-scale biases are apparently an order of magnitude smaller. (Since NWP SST is obtained by assimilating drifter SSTs, this is not an independent test and should give small biases by design.) The standard deviation (SD) of the twelve zonal-mean differences in each plot is 0.17 K for the CMS SST and 0.02 K for the NWP SST. The plot for the operational SST shows that the SEVIRI SSTs entering the analysis have yet stronger zonal structure, with the SD of the zonal-mean differences being 0.37 K. The CMS analysis procedure moderates these biases.

252 To quantify the overall relationships between the SSTs, we present statistics in Table 1. The conventional  
 253 statistics, mean and SD, are supplemented by the median and the robust standard deviation (RSD). For a  
 254 Gaussian distribution, the median equals the mean, and the RSD equals the SD. Where outliers or non-  
 255 Gaussian wings to a distribution are present, the mean and SD can be quite significantly modified,  
 256 whereas the median and RSD are less affected and give a picture of the distribution of the central majority  
 257 of data. The median and RSD can thus be more informative about the behaviour of SST estimators under  
 258 ideal conditions, where all BTs used really are from clear-sky areas that are free of significant aerosol,  
 259 while the mean and SD describe more the behaviour of the whole system, including the effects of  
 260 unscreened cloud or aerosol contamination in the BTs used.

261

262 The operational SST is an example of the “non-linear SST” formalism, i.e., is:

263

$$264 \quad (3) \quad \hat{x} = a_0 + a_1 S + \begin{bmatrix} a_2 + a_3 S + a_4 S + a_5 x_c \\ -a_4 S - a_5 x_c \end{bmatrix}^T \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix}$$

265

266 where  $S = \secant(\theta) - 1.0$ ,  $\theta$  is the satellite zenith angle,  $y_{11}$  and  $y_{12}$  are the 11 and 12  $\mu\text{m}$  BTs, and  $x_c$  is a  
 267 climatological SST estimate, taken from the five-day fine-scale climatology of Casey and Cornillon  
 268 [1999]. The coefficients in ( 3 ) are defined by radiative transfer modelling followed by least-squares  
 269 regression [Le Borgne et al., 2006]. The radiative transfer model used was an earlier version of RTTOV  
 270 [Saunders et al., 2002], and it was applied to a radiosounding data base [Francois et al., 2002]. The above  
 271 formulation has a weight  $(a_2 + a_3 S)$  that multiplies the 11  $\mu\text{m}$  channel BT alone, and a factor  $(a_4 S + a_5 x_c)$   
 272 that multiplies the difference between the 11 and 12  $\mu\text{m}$  channels. In fact, operationally, smoothing of the  
 273 atmospheric correction is applied [Harris and Saunders, 1996] to reduce noise, by averaging  $y_{11} - y_{12}$  (for  
 274 clear-sky pixels, according to the cloud detection scheme) over a  $\sim 50$  km kernel centred on each pixel,  
 275 before multiplication by the factor  $(a_4 S + a_5 x_c)$  in the retrieval algorithm.

276

277 We re-apply the operational coefficients to the BTs in the MDB to create what we will refer to as the  
278 NLSST. Although we use only cells with  $SDI < 0.25$ , the SDI correction is operationally defined for the  
279 range  $0 < SDI < 1$ , so we also apply the SDI correction where applicable. The results are not identical to  
280 the operational SSTs: compare Figure 2(a) with 2(d). The differences arise only because the atmospheric  
281 correction smoothing cannot be applied to  $0.5^\circ$  resolution BTs. The absence of the smoothing makes the  
282 NLSST noisier (SD against drifters is 0.70 K for NLSST rather than 0.55 for the operational SSTs; see  
283 Table 1). This is most marked for cells with relatively few contributing pixels, as shown by Figure 3. The  
284 natural interpretation of this is that it is pixels that are surrounded by other pixels flagged as cloudy that  
285 are more likely to have some residual cloud contamination affecting their BTs. Smoothing of the  
286 atmospheric correction dilutes the effect of the additional noise and slight negative bias caused by  
287 undetected clouds. We should emphasize, therefore, that although using the same coefficients, the NLSST  
288 results presented here for comparison purposes do not represent the performance of the SST product  
289 delivered operationally.

290

291 We will use the NLSST hereafter as the point of comparison for the optimal estimation technique, since  
292 OE will be applied on exactly the same BTs as the NLSST, and thus like-for-like comparisons can be  
293 made. The zonal biases of the NLSST are shown in Figure 2(d). Figure 4 shows the NLSST behaviour  
294 with TCWV, satellite zenith angle, difference in brightness temperature ( $11 \mu\text{m}$  minus  $12 \mu\text{m}$ ) and SST.  
295 Because the equatorial bias of the NLSST coincides with the highest NWP TCWV values, there is a  
296 dependence of  $-0.6$  K in the mean over the range of water vapour, and this also appears as a negative bias  
297 for small zenith angles, large BT difference and the warmest bin for SST.

298

299 Four maps of mean properties of the NLSST are shown in Figure 5 to be compared with OE results. In  
300 (a), the mean NLSST compared to drifter SST is shown. The tropical Atlantic bias in validation is  
301 somewhat more exaggerated in the absence of atmospheric correction smoothing, as also seen in Figure 3.  
302 Figure 5(b) shows the random error in SST attributable to radiometric noise in the observations. This is

303 estimated as follows. First, we need estimates of the single-pixel radiometric noise for SEVIRI.  
 304 EUMETSAT [2001] reported that for the SEVIRI sensor on Meteosat-8, the noise equivalent differential  
 305 temperature is 0.11 K and 0.15 K for the 11 and 12  $\mu\text{m}$  channels respectively. Here, we will assume,  
 306 conservatively, that the noise is  $\varepsilon = \varepsilon_{11} = \varepsilon_{12} = 0.15$  K in each channel of the SEVIRI sensor on Meteosat-  
 307 9. Assuming independence of the radiometric noise between pixels, the cell-averaged random error in BT  
 308 is then  $\varepsilon/\sqrt{n_c}$ , where  $n_c$  is the number of clear-sky pixels available in cell  $c$ . Assuming independence of  
 309 radiometric noise between channels, the expectation of the noise-induced error in retrieved SST is:

$$310 \quad (4) \quad \sqrt{(a_2 + a_3S + a_4S + a_5x_c)^2 (\varepsilon_{11}^2 / n_c) + (a_4S + a_5x_c)^2 (\varepsilon_{12}^2 / n_c)}$$

311  
 312 This equation predicts that the SST retrieval for a cell is more noisy when  $n_c$  is small or the channel  
 313 radiometric noise is large, as one expects. The noise also depends on satellite zenith angle, via the term  $S$ .  
 314 Since, in the NLSST as formulated in ( 3 ), all the coefficients  $a_i$  tend to be positive, the retrievals become  
 315 more noisy towards the limb view. For example, the single pixel noise using ( 4 ) at nadir is 0.40 K and at  
 316  $60^\circ$  of satellite zenith angle is 0.52 K. The mean result of this combination of coverage and zenith angle  
 317 for this data set is Figure 5(b).

318  
 319 Having run RTTOV9 simulations for each cell, we are also able to quantify the sensitivity of the NLSST  
 320 to true changes in SST and to variations in TCWV, which are evaluated using

$$321 \quad (5) \quad \begin{aligned} \frac{\partial \hat{x}}{\partial x} &= (a_2 + a_3S + a_4S + a_5x_c) \frac{\partial y_{11}}{\partial x} - (a_4S + a_5x_c) \frac{\partial y_{12}}{\partial x} \\ \frac{\partial \hat{x}}{\partial w} &= (a_2 + a_3S + a_4S + a_5x_c) \frac{\partial y_{11}}{\partial w} - (a_4S + a_5x_c) \frac{\partial y_{12}}{\partial w} \end{aligned}$$

322  
 323  
 324 Figure 5(c) shows that the NLSST sensitivity to SST is between 0.9 and 1.0 for much of the disk, but  
 325 drops to as low as 0.5 for humid atmospheres viewed at oblique angles. For unusually dry atmospheres  
 326 observed at oblique angles, it happens that sensitivity greater than 1.0 can occur (up to 1.2), but this  
 327 doesn't happen on average. The sensitivity is not 1.0 because the change in BT associated per unit change

328 in SST locally (the partial derivative) is in general not the same as  $(BT_{\text{polar}} - BT_{\text{tropical}})/(SST_{\text{polar}} -$   
329  $SST_{\text{tropical}})$ , which, using an obvious nomenclature, represents the “total derivative” that is fitted by a  
330 global regression to derive retrieval coefficients. The underlying reason is that SST, TCWV and  
331 atmospheric transmittance are all spatially correlated. Regional SST coefficients [e.g., Shenoi, 1999] can  
332 be beneficial for retrieval over smaller spatial scales for which these partial and “total” derivatives are  
333 closer in magnitude.

334

335 The NLSST obtained for a given location and true SST is also sensitive to fluctuations in the humidity of  
336 the intervening atmosphere, Figure 5(d). The reason this sensitivity is not, as we would like, zero is  
337 similar to that given for the SST sensitivity above. It is not that the coefficients are somehow “wrong”:  
338 this effect is intrinsic to the NLSST formalism. For most locations, the sensitivity is negative, such that  
339 increased column water vapour reduces the retrieved SST. This effect has been observed in SEVIRI SSTs  
340 compared to SSTs at moorings in the equatorial Atlantic. A 20% variation in TCWV (see Figure 5(d))  
341 would, according to our estimate, give a decrease in NLSST of 0.5 to 1 K in this region. The dependence  
342 on water vapour is weakly positive for tropical areas in the extreme limb and for areas affected by dry air  
343 outflows from deserts: off the west coasts of the Sahara and Namib deserts, and over the Red Sea.

344

345 Regional coefficients may reduce the magnitude of sensitivity to TCWV, but would not eliminate it. Or—  
346 were an independent, and sufficiently accurate, estimate of  $w$  available—a local correction for TCWV  
347 fluctuations could, in principle, be designed. But the complexity of such ad hoc solutions would probably  
348 be comparable to that of performing optimal estimation, which gives a local solution that is optimized  
349 with respect to the available prior information and forward model.

350

## 4 Initial optimal estimator

The optimal estimation method broadly follows that of Merchant et al. [2008b]. The reader is referred to that paper for discussion in greater depth than given here, and for the equations for the optimal estimate, the propagated instrumental noise in the retrieved estimate, and the cost function (all referred to below).

Briefly, the OE is a *maximum a posteriori* (MAP) estimate [Rodgers, 2000]. We can consider MAP as the adjustment or nudging of the prior SST in the light of the difference between the observed BTs and the BTs that are simulated assuming the prior SST to be correct. The nudging is “optimal” in that it takes appropriate account of errors in combining the prior SST and the information in the observations. Appropriate estimation of errors is therefore crucial to making OE work well. If the error in the prior SST,  $e_{xa}$ , is small and the observations are noisy compared to  $e_{xa} \frac{\partial y}{\partial x}$ , the result will be little different to the prior. OE also assumes that the prior and the observations (relative to the forward model) are unbiased. Here, we already know this not to be the case for the prior SST, the CMS analysis SST, so we need to accommodate the bias in the formulation of the model for prior error. (We also will find in the next section that the observations and forward model are biased relative to each other, and will address this with a scheme for bias correction.)

The assumptions for our implementation of OE are summarized in Table 2, and are mostly comparable to Merchant et al. [2008b]. The main point of difference is in the use of a prior SST from a satellite-only analysis, which, as mentioned above, is distinguished here from the prior from NWP used for forward modelling, in order to ensure independence of the OE SST from our validation data. With this arrangement, the forward model is

$$(6) \quad \mathbf{F}(\mathbf{x}_a) \cong \mathbf{F}(\mathbf{x}_0) + \mathbf{K}(\mathbf{z}(\mathbf{x}_a) - \mathbf{z}(\mathbf{x}_0)) = \mathbf{F}(\mathbf{x}_0) + \frac{\partial \mathbf{y}}{\partial \mathbf{x}} (\mathbf{x}_a - \mathbf{x}_0)$$

375 We also need to account for the fact that the CMS analysis SST is biased, indeed, shares with the NLSST  
376 aspects of bias that we would wish the OE SST to improve upon. An estimate of the global random error  
377 is available from Table 1 (the SD compared to drifters), and is  $\sim 0.4$  K. But were we to use this for our  
378 prior error, most of the tropical Atlantic bias  $\sim 1$  K would likely be preserved in our result. So, for this  
379 initial implementation, we illustrate the effect of enlarging the prior error to accommodate bias. We  
380 calculate the mean square difference of the CMS and NWP SSTs for each cell over the 3-month period,  
381 smooth this using a 9-cell boxcar filter (so that small-scale areas of difference do not have an inordinate  
382 local effect). We then take the square root of the sum of the resulting field and the global random variance  
383 relative to drifters of  $0.4^2 \text{ K}^2$ . The result is shown in Figure 6(a). Note that the elevated prior errors  
384 obtained in the locations of western boundary currents may be inappropriate, since the disagreement of  
385 the CMS and NWP fields may be exaggerated by spatial sampling effects in locations of the strong  
386 thermal gradients.

387

388 We refer to the SST from this initial optimal estimator as “OE1 SST” from this point. The mean  
389 difference against drifter SSTs is shown in Figure 6(b). The tropical Atlantic biases are clearly improved  
390 relative to Figure 5(a). In terms of zonal bias, this is quantified in Figure 7(a). All zonal biases are now  
391 within 0.4 K of zero. The -1 K negative bias of NLSST for the zone covering  $0^\circ\text{N}$  to  $10^\circ\text{N}$  and the  
392 positive biases in the southern high latitudes have been markedly improved. The zonal structure is  
393 reminiscent of (but has not improved upon) that of the prior SST, Figure 2(a). A significant contributory  
394 factor in the improvement compared to the NLSST comes from reducing the impact of noisy  
395 observations, as shown in Figure 7(b). The cells with few contributing BTs, some fraction of which are  
396 suspected to be cloud contaminated, have much reduced bias and scatter SD than was the case for the  
397 NLSST. This works because the observation error is larger for small  $n_c$ , so the OE tends give a stronger  
398 weight to the prior SST, since this gives a lower-error solution.

399

400 If errors are Gaussian and well-estimated, the “cost” of the OE solution attained is a  $\chi^2$  distributed  
401 variable with a mean value equal to the number of independent channels of observation (here, two)  
402 [Rodgers, 2000]. Large values of cost indicate that the solution is far from the prior SST, or that the  
403 solution found is inconsistent with the BTs in some way, or both. The MAP estimate is, in fact, the  
404 solution that minimizes the cost, and solutions with relatively large cost are statistically unexpected given  
405 the assumed errors, and may therefore be outliers for which the assumed errors are inappropriate. The  
406 cost is therefore a useful indicator of the quality of retrieved SST. We see, for example, in Figure 7(c) that  
407 the SD is a strong function of cost, with the two thirds of data with  $\chi^2 < 2$  having SD of 0.43 K. OE1 also  
408 shows less dependence on TCWV, zenith angle, BT difference and SST than was the case for NLSST  
409 (not shown). Enough has been shown at this point to demonstrate that OE1 has lessened the impact of the  
410 least reliable observations and allows the least reliable retrievals to be identified. Since there remain  
411 undesirable bias characteristics evident in OE1, although less so than in the NLSST, we consider next the  
412 question of bias in the observations relative to the forward simulations, and explore whether addressing  
413 these can improve on the bias characteristics of OE1.

414

## 415 **5 Brightness temperature bias correction**

416 Statistics of simulated minus observed brightness temperature for all cells with observations within the  
417 three month period are given in Table 3. Given that we are seeking to limit our biases in retrieved SST to  
418 a few tenths of Kelvin, the biases in the split-window channels are significant. The differences between  
419 mean and median and between SD and RSD tell us that the distributions of simulation minus observation  
420 (hereafter, “S-O”) have significant positive tails—interpreted as cloud contaminated (cold) observations.  
421 We do not wish our bias correction scheme to adjust for cloud contamination, so the first step is to  
422 exclude from consideration these outliers as best we can. Cells with S-O falling outside the range given  
423 by the median  $\pm 3 \times$  RSD in any of the four channels are not used in the bias correction scheme. This

424 excludes 6.5% of cells and brings the conventional and robust statistics into much closer alignment, with  
425 the mean biases in the 11  $\mu\text{m}$  and 12  $\mu\text{m}$  channels becoming +0.18 and +0.47 K respectively.

426

427 The ideal bias correction scheme for our purpose will adjust the simulations for the following possible  
428 origins of non-zero S-O: systematic error in calibration, systematic error from the approximations  
429 inherent in the fast forward model, and systematic biases in the atmospheric profiles obtained from NWP.  
430 It must not correct for bias in prior SST, otherwise we will destroy the signal that has the potential to  
431 improve on the prior SST. For this reason, using for our simulations an SST field that is distinct from the  
432 prior SST is helpful, although not in itself sufficient. The NWP SST appears to be negligibly biased with  
433 respect to drifters on large scales, as we have seen. There may be bias at its full  $0.5^\circ$  resolution, however,  
434 because of the sparseness of coverage by in situ measurements in some areas, relative to this resolution.  
435 Because of its low resolution, there will definitely be local biases in the NWP SST field compared to the  
436  $\sim 5$  km scales of SEVIRI's resolution. The bias correction scheme we apply therefore smooths out spatial  
437 structure to  $2.5^\circ$  using boxcar spatial filtering. This scale may also be more appropriate to systematic  
438 biases in atmospheric profiles which we wish to correct for; certainly  $0.5^\circ$ , the full NWP resolution,  
439 would seem too fine for this purpose.

440

441 The correction of the atmospheric NWP also requires a separation by time scale. The scheme must  
442 average over the synoptic time-scale. It must not, for example, "correct" to-day's simulations in the light  
443 of yesterday's bias arising from mis-location of a humidity gradient associated with a weather front. On  
444 the other hand, we do wish to adjust for persistent tendencies of the NWP system to deviate from true  
445 atmospheric structure. A time-scale between the synoptic and seasonal is therefore appropriate, and we  
446 select 20 days. Through the three months of data studied, the field of 20-day running average S-O  
447 evolves. Figure 8 shows the bias correction obtained for days 21 and 41 of the study period. The former is  
448 based on S-O observed on days 1 to 20 and the latter on days 21 to 40, the data processing being  
449 summarized in Figure 9. There is sufficient similarity between Figure 8(a) and (b) to show that there is

450 some persistent spatial variability that is usefully captured using a 20 day interval. The range of predicted  
451 S-O in Figure 8, from  $-1$  to  $+1$  K, is comparable to the range of OE1 SST biases that we wish the scheme  
452 to improve.

453

454 This approach to BT bias correction is different to that adopted for Metop-A in Merchant et al. [2008b].  
455 There, bias correction was done using empirical fitting of S-O to analytic functions of  $x$ ,  $w$ ,  $\theta$  and  
456 latitude. Our original intention to follow that precedent was undermined by failure to find empirical fits  
457 that convincingly delivered the same efficiency of correction in the case of SEVIRI. Considering the  
458 relatively complex spatial patterns present in Figure 8, it is difficult to imagine what analytic functions  
459 could be found to fit them well. The approach taken here using persistence of S-O is well suited to a  
460 geostationary sensor, where a given location is always viewed at the same zenith angle.

461

## 462 **6 Optimal estimation after BT bias correction**

463 The BT bias correction improves the results of optimal estimation. This second implementation gives  
464 results we refer to as “OE2” SST. In Figure 10(a) we see that the zonal mean biases are reduced  
465 compared to OE1 (Figure 7) and are smaller than those of the NLSST, the SEVIRI operational SST, and  
466 the CMS analysis SST used as prior. This is the most important result: that with BT bias correction, OE  
467 on SEVIRI can partially correct biases in the satellite-only prior SST provided by the CMS analysis. We  
468 discuss this further below. The OE2 SSTs show less dependence on the number of pixels per cell than the  
469 operational scheme with smoothed atmospheric correction: compare Figure 10(b) and Figure 3(b). After  
470 BT bias correction, the dependence of OE2 minus drifter statistics on the retrieval cost is more like that  
471 seen in the case of Metop-A in previous work. There is little systematic variation of the mean OE2 minus  
472 drifter SST with factors such as water vapour (Figure 11), except for cold SST under very dry skies.

473

474 Summary statistics of the OE2 SSTs against drifters are given in Table 4, for all matches and for the  
475 97.5% of “best” data according to the retrieval cost. Although the results are evidently an improvement  
476 on the NLSST applied to the same BTs, the SD and RSD are greater than for the CMS analysis SST  
477 compared to drifters. But shouldn’t an optimal estimate always improve on the prior error variance by  
478 adding an independent observation? This is only “guaranteed” if the situation matches the assumptions  
479 underlying optimal estimation theory. OE2 violates several of these assumptions: (i) The error  
480 distributions for the observations relative to the forward model (S-O) are not Gaussian: although the  
481 outliers, including suspected cloud contamination, were filtered in the BT bias correction, that filtering  
482 has not been applied to the MDB. (ii) The prior SST error is not zero-mean, it is known to be biased, and  
483 the error model for the prior has had to accommodate this. (iii) Although some care has been taken to  
484 provide appropriate error estimates for  $\mathbf{S}_a$  (NWP error covariance) and  $\mathbf{S}_e$  (error covariance of  
485 observations relative to forward model), these have not been optimized for the system. (iv) Despite the  
486 bias in the tropical Atlantic, the CMS analysis SST has low noise at  $0.5^\circ$  scales: the analysis system can  
487 take advantage of assumptions of temporal correlation of SST and incorporates three-channel (lower-  
488 noise) SSTs from AVHRRs on polar-orbiting platforms.

489  
490 So, the more interesting result of the OE2 experiment is the fact that the results show less geographical  
491 bias than the CMS analysis SST (except for cold waters viewed by SEVIRI at high satellite zenith  
492 angles). The details of this can be studied in Figure 12, (a) to (c). Figure 12(a) shows OE2 minus drifter  
493 SST, and can be compared to Figure 5(a) for the NLSST and Figure 1(f) for the SEVIRI operational SST.  
494 Reduced bias, particularly in the tropical Atlantic but also elsewhere, is evident. Figure 12(b), the OE2  
495 minus NWP SST over the three month period, can be compared to Figure 1(b) for the CMS analysis.  
496 Over most of the disk the OE2 is closer to the NWP on average than the CMS analysis. The remaining  
497 biases in OE2 appear even less marked in the comparison with AATSR SSTs in Figure 12(c). The  
498 AATSR SSTs consist of dual-view three-channel retrievals averaged to  $0.5^\circ$  cells consistent with the rest  
499 of the data.

500  
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

Propagation of radiometric noise into the OE2 SST is much less (Figure 12 (d)) than with NLSST (Figure 5(b)). The OE2 SSTs are also much less sensitive to fluctuations in atmospheric water vapour, according to the estimate in Figure 12(e).

So far, good news, but it is important to consider Figure 12(f) carefully. This shows that the benefits identified above have come at a cost: the better conditioning of the OE2 retrieval (in terms of reduced bias, reduced noise and less sensitivity to water vapour) has been achieved by a heavier reliance on prior SST information. The OE2 sensitivity to true changes in SST at a given place is lower than with the NLSST, which is not a desirable feature. Lack of sensitivity is probably also the main factor behind the low noise propagation for tropical atmospheres at large zenith angles in Figure 12(d).

This lack of sensitivity arises mainly in areas of low transmittance because OE2 (MAP) gives the local minimum error solution given two available estimates of SST: from the BTs and from the prior. In low transmittance atmospheres, any changes in BT from SST variations are weak, and therefore this signal needs to be amplified by the retrieval. However, this amplification makes the estimated based on BTs very noisy compared to the (fairly narrow) error specified for the prior; therefore in these circumstances, the prior can be the dominant contribution to the retrieval.

Put another way, one can't expect an OE system to deliver sensitivity to large diurnal variability (~3 K amplitude) or large frontal steps in SST (~2 K) if the prior is a smooth SST field with an assumed error of 0.6 K.

Lack of sensitivity to SST is not ideal for looking at fronts or diurnal variation, but the OE2 could still be useful on the larger scales of this study, particularly when the OE2 SSTs are used as part of a system. If the SEVIRI observations on day  $n$  modify the prior SST field used on day  $n+1$  with weight  $i$  ( $0 < i < 1$ ),

526 then we can expect OE2 to be able to track the full amplitude of SST anomalies that persist for more than  
527  $\left(i \frac{\partial \hat{x}_{OE2}}{\partial x}\right)^{-1}$  days. For the equatorial region, this time scale is 8 days if  $i$  is 50%. This expectation will be a  
528 subject of future study.

## 7 Discussion

Optimal estimation can be interpreted as dynamic calculation of SST coefficients tailored to the observational context of a particular time and location, as estimated by the prior fields available. The resulting SST estimates are optimal in the sense that they provide the minimum error solution – if the prior fields and the forward-model-relative-to-observations are both unbiased and have well-characterized errors. Both this paper (for SEVIRI) and Merchant et al. [2008b] (for METOP) have shown that OE can bring significant benefit to the retrieval of SST, albeit at the cost of investing significant effort into bias correction and error characterisation. The benefits are improved retrieval statistics compared to drifting buoys (random errors and biases with respect to geographical location and parameters such as total column water vapour), reduced confounding of variations in atmospheric water vapour as SST variations, and the availability of a powerful additional statistic (the cost) for quality control.

However, for the case of SEVIRI, the instrumental noise, the forward modelling errors (including bias correction errors), and (away from nadir) the viewing geometry all conspire to give relatively low sensitivity of OE SST to true SST variations, Figure 12(f). As explained above, this is because the optimal estimate in these circumstances is one that relies heavily on the prior SST. This is the penalty paid for the other benefits that the OE SSTs have over the NLSSTs. Thus, there is a trade-off, and the merits of one form of retrieval compared to the other need to be considered in the light of the intended application of the SSTs.

In summary, we have shown that combination of bias correction scheme and optimal estimation presented here would have a positive impact on the bias, noise and water vapour sensitivity of SEVIRI SSTs, although in the sensitivity of the OE2 SSTs to rapid (diurnal) or local (frontal) changes in SST would be lowered in most areas. This suggests that the OE2 scheme is most useful at the larger scales and in the

554 context of an SST analysis system. All validation measures would show significant improvement with the  
555 OE2 SSTs compared to NLSST.

556

557 The obvious question for further research is: how can optimal estimation be best applied to SEVIRI at full  
558 sensor resolution? Aspects to consider are: the nature of the prior SST, particularly its spatial resolution;  
559 how to use the differing lengths scales of variability in ocean and atmosphere to achieve noise reduction  
560 (equivalent to atmospheric smoothing); and how to design the OE to increase sensitivity to SST in order  
561 to accurately capture frontal gradients and diurnal variability.

562

## 563 **References**

564 Anding, D., & Kauth, R. (1970). Estimation of sea surface temperature from space, *Remote Sensing of*  
565 *Environment*, 1 (4), 217-220.

566 Casey, K. C., & Cornillon, P. (1999). A comparison of satellite and in-situ based sea surface temperature  
567 climatologies, *J. Climate*, 12 (6), 1848-1863.

568 Donlon, C. J., Robinson, I., Casey, K. S. and others (2007). The Global Ocean Data Assimilation  
569 Experiment High-resolution Sea Surface Temperature Pilot Project, *Bulletin of the American*  
570 *Meteorological Society*, 88 (8), 1197-1213. doi:10.1175/BAMS-88-8-1197.

571 EUMETSAT (2001). MSG level 1.5 Image Data description, EUM/MSG/ICD/A05, issue2, November  
572 2001.

573 Francois, C., Brisson, A., Le Borgne, P., & Marsouin, A. (2002). Definition of a radio-sounding database  
574 for sea surface temperature brightness temperature simulations – application to sea surface  
575 temperature retrieval algorithm determination, *Remote Sensing of Environment*, 81, 309-326.

576 Harris A R and Saunders M A, (1996). Global validation of the along-track scanning radiometer against  
577 drifting buoys. *J. Geophys. Res.*, 101 (C5), 12127-12140.

578 Le Borgne, P., G. Legendre & A. Marsouin (2006), Operational SST retrieval from MSG/SEVIRI data,  
579 Proceedings of the 2006 EUMETSAT conference, Helsinki, Finland.

580 Merchant C. J. and Le Borgne P. (2004). Retrieval of sea surface temperature from space based on  
581 modeling of infrared radiative transfer: Capabilities and limitations, *J Atmospheric and Oceanic*  
582 *Technology*, 22 (11) 1734-1746.

583 Merchant, C. J., Horrocks, L. A., Eyre, J., & O'Carroll, A. G. (2006a). Retrievals of sea surface  
584 temperature from infra-red imagery: origin and form of systematic errors, *Quarterly J. Royal*  
585 *Meteorological Society*, 132, 1205-1223.

586 Merchant, C. J., Embury, O., Le Borgne, P., & Bellec ,B. (2006b). Saharan dust in night-time thermal  
587 imagery: detection and reduction of related biases in retrieved sea surface temperature, *Remote*  
588 *Sensing of Environment*, 104, 15-30.

589 Merchant, C. J., M. J. Filipiak, P. Le Borgne, H. Roquet, E. Autret, J.-F. Piolle, and S. Lavender (2008a),  
590 Diurnal warm-layer events in the western Mediterranean and European shelf seas, *Geophysical*  
591 *Research Letters*, 35, L04601, doi:10.1029/2007GL033071.

592 Merchant C. J., P. Le Borgne, A. Marsouin and H. Roquet, (2008b) Optimal estimation of sea surface  
593 temperature from split-window observations, *Remote Sensing of Environment*, 112 (5), 2469-2484.

594 Rodgers, C. D. (2000), Inverse methods for atmospheric sounding: Theory and Practice. World Scientific  
595 Publishing.

596 Saunders, R. W., Brunel, P., Chevallier, F., DeBlonde, G., English, S. J., Matricardi, M., & P. J. Rayer  
597 (2002), RTTOV-7 – Science and Validation Report, NWP Technical Report No. 387, Met Office,  
598 UK.

599 Shenoi S. C. (1999). On the suitability of global algorithms for the retrieval of SST from the north Indian  
600 Ocean using NOAA/AVHRR data, *International J. Remote Sensing*, 20 (1), 11-29.

601 Walton, C. C., Pichel, W. G. , Sapper, J. F., & May, D. A. (1998). The development and operational  
602 application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA  
603 polar-orbiting environmental satellites, *J. Geophysical Research*, 103, 27,999–28,012.

604

**TABLES**

605

**Table 1. Statistics of differences between various SST estimates in the match-up data base for this study.**

Difference	Mean / K	SD / K	Median / K	RSD / K
CMS minus NWP	-0.05	0.35	-0.06	0.30
CMS minus Drifter	-0.04	0.40	-0.05	0.30
NWP minus Drifter	+0.01	0.38	+0.01	0.28
Operational minus Drifter	-0.20	0.55	-0.21	0.45
NLSST minus Drifter	-0.38	0.70	-0.31	0.55
NLSST minus Operational	-0.17	0.48	-0.06	0.30

606

N = 30108

607

**Table 2. Summary assumptions implemented for optimal estimator.**

Aspect of	Assumptions implemented
optimal estimator	
Error in $x$ , $e_{xa}$	Random error in CMS SST against drifters (Table 1) combined with geographical bias of CMS SST compared to NWP SST from ARPEGE – see Figure 6(a).
Prior TCWV, $w$	NWP fields (ARPEGE)
Error in $w$ , $e_{wa}$ / $\text{kg m}^{-2}$	Error is [Merchant et al., 2008b]: $w(0.1 + (75 - w)/150)$
Prior error covariance, $\mathbf{S}_a$	$\begin{bmatrix} e_{xa}^2 & 0 \\ 0 & e_{wa}^2 \end{bmatrix}$
Radiometric noise in cell $c$ , $\varepsilon_\lambda$ / K	For both 11 $\mu\text{m}$ and 12 $\mu\text{m}$ channel [Merchant et al., 2008b]: $\sqrt{\frac{1}{n_c}} 0.15 \text{ K}$

---

RTTOV9	For both 11 $\mu\text{m}$ and 12 $\mu\text{m}$ channel: $\sec(\theta)0.15$ K, where $\theta$ is the satellite zenith
forward model	angle.
random error	
Error	
covariance of observations & model, $\mathbf{S}_\epsilon$	Uncorrelated: $\begin{bmatrix} 0.15^2 \left( \sec^2(\theta) + \frac{1}{n_c} \right) & 0 \\ 0 & 0.15^2 \left( \sec^2(\theta) + \frac{1}{n_c} \right) \end{bmatrix}$
Iterations	None; assume that prior $\mathbf{x}_a$ is always close enough that solution lies within linear range (i.e., $\mathbf{K}$ varies negligibly between prior and solution).
Skin and diurnal effects	Constant $-0.2$ K skin effect assumed, method of Merchant et al. [2008b]. Retrievals are sub-skin estimates, and thus should be unbiased compared to night-time drifter SSTs. Currently no model for diurnal warming effects for day-time retrievals.

---

608

609 **Table 3. Simulation-minus-observation statistics.**

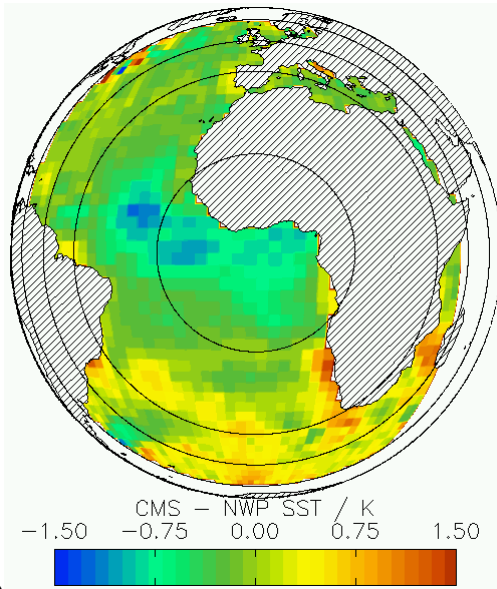
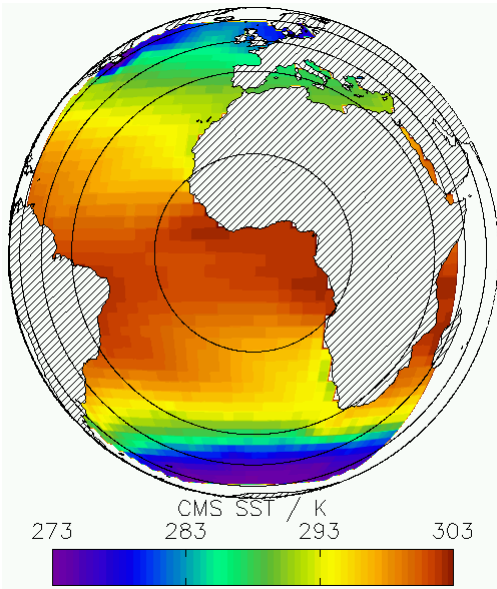
Channel	Mean / K	SD / K	Median / K	RSD / K
3.8 $\mu\text{m}$	-0.83	1.05	-1.03	0.64
8.7 $\mu\text{m}$	+0.31	1.40	+0.15	0.76
11 $\mu\text{m}$	+0.29	1.63	+0.12	0.86
12 $\mu\text{m}$	+0.58	1.55	+0.43	0.95

610  $N = 8.8 \times 10^6$  except for 3.8  $\mu\text{m}$  channel (night only) for which  $N = 4.2 \times 10^6$ 611 **Table 4. Statistics of the OE2 SST minus drifter SSTs.**

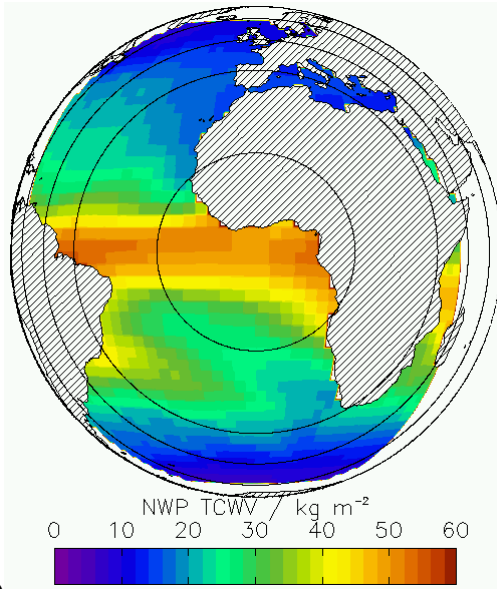
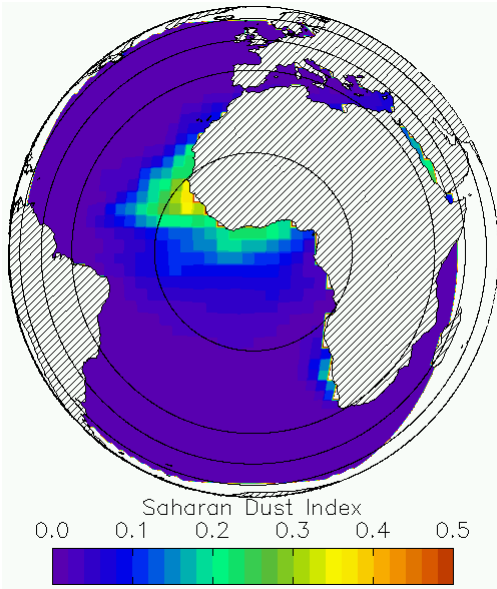
OE2 – drifter	Mean / K	SD / K	Median / K	RSD / K	N
Data subset					
All matches	-0.01	0.47	-0.01	0.34	30108 = 100%
$\chi^2 < 4$	+0.01	0.44	-0.00	0.33	29347 = 97.5%

---

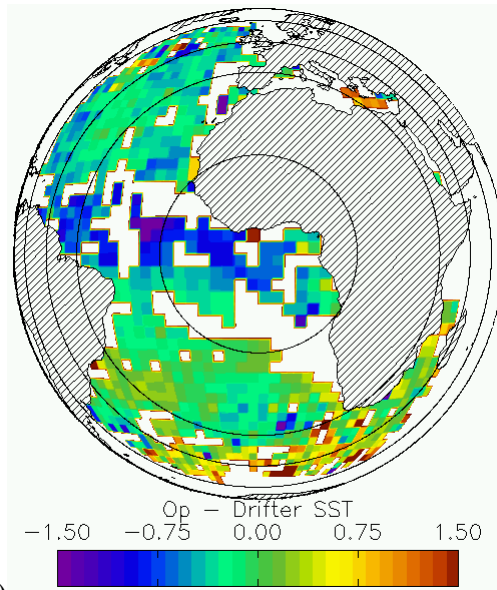
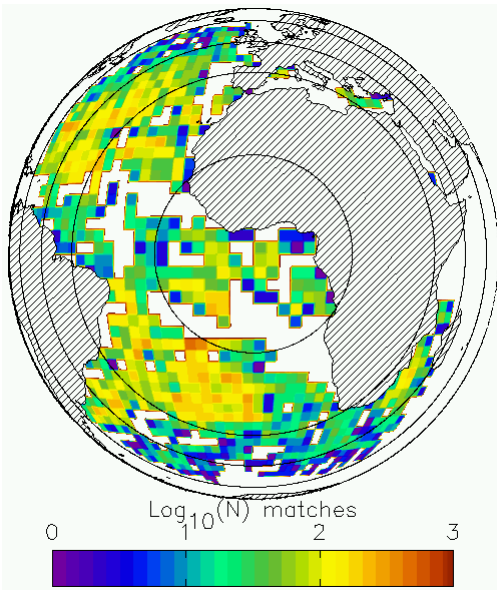




613 (a) (b)



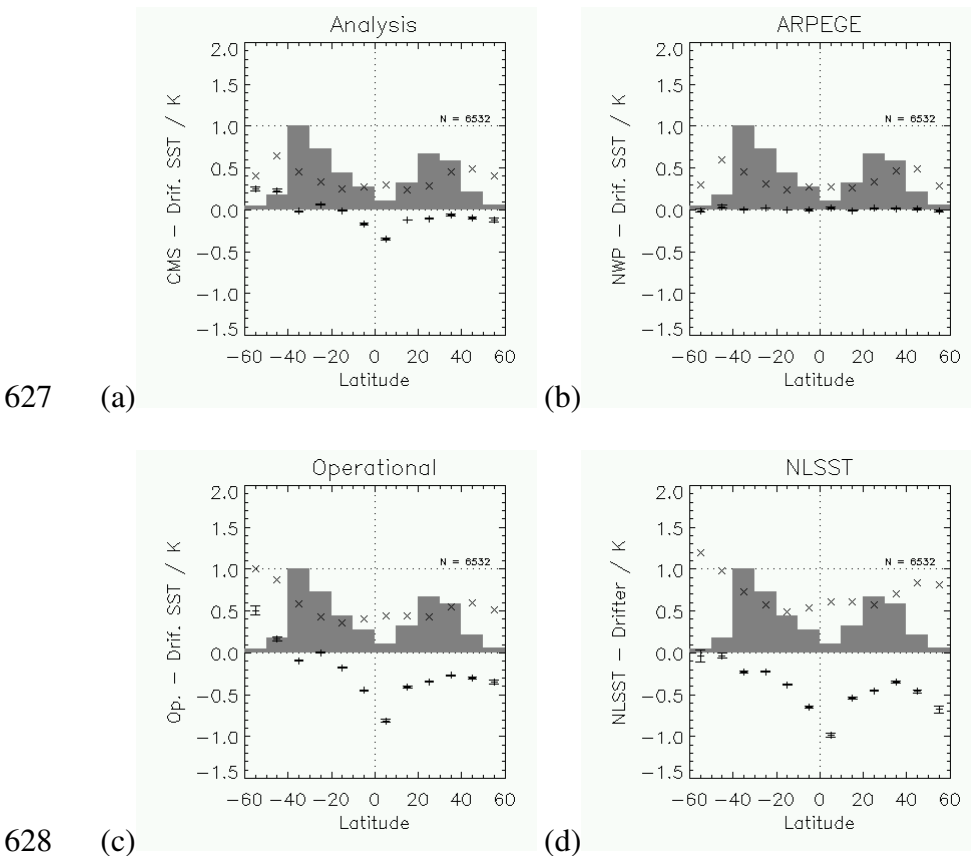
614 (c) (d)



615 (e) (f)

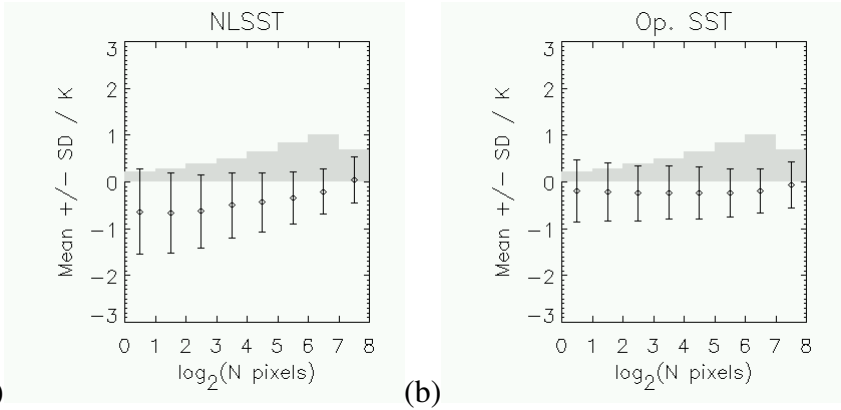
616 **Figure 1. [Previous page] (a) The daily 0000 h UTC SST analysis from CMS averaged over 1 February to 30 April,**  
 617 **2008 and presented at 2.5° resolution of latitude and longitude. The map projection is the Earth as seen from Meteosat-**  
 618 **9, and contours of satellite zenith angle at 24°, 48°, 60° and 72° are overlaid. (b) Averaged difference of the CMS**  
 619 **analysis SST from the NWP SST field obtained from the ARPEGE numerical forecasting system, into which in situ**  
 620 **SSTs have been assimilated. (c) Averaged Saharan dust index. (d) Averaged total column water vapour from ARPEGE.**  
 621 **(e) Base 10 logarithm of the number of matches per 2.5° cell: e.g., yellow (2.0 to 2.1 on the scale) indicates N between**  
 622  **$10^2 = 100$  and  $10^{2.1} = 125$  inclusive. No drifter observations are available in white areas. (f) Averaged difference of the**  
 623 **operational SEVIRI SST from drifter SST.**

624  
625  
626



629 **Figure 2. (a) CMS minus drifting buoy SST against latitude in 10° zones. Plus (+) symbols and associated error bars**  
 630 **show the mean and standard error of the bias for each zone. Crosses (x) show the zonal standard deviation. The filled**  
 631 **bars show the population of each zone relative to the most populated zone, which is normalized to 1, and which contains**  
 632 **N = 6532. (b) As (a) but for NWP SST. (c) As (a) but for operational SSTs from SEVIRI. (d) As (a), but NLSST with no**  
 633 **smoothing of atmospheric correction.**

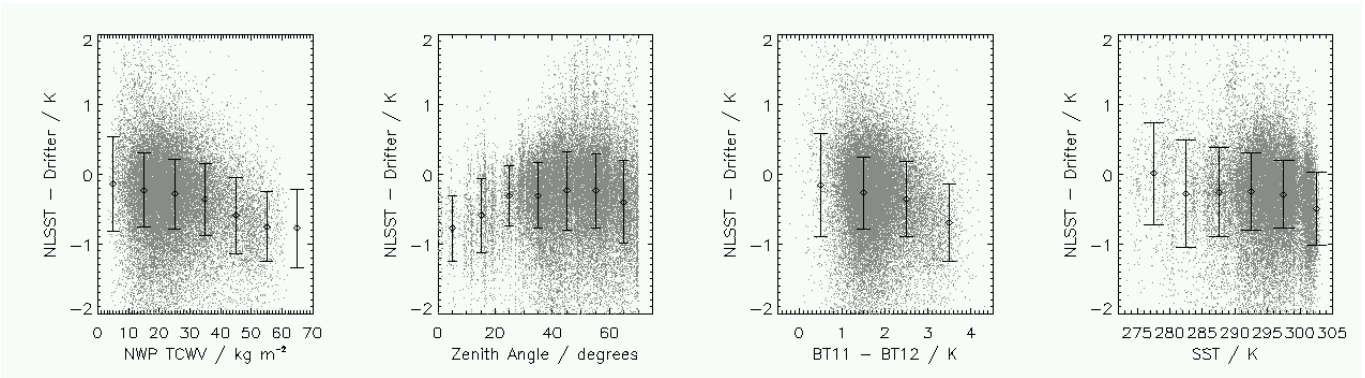
634  
635



636

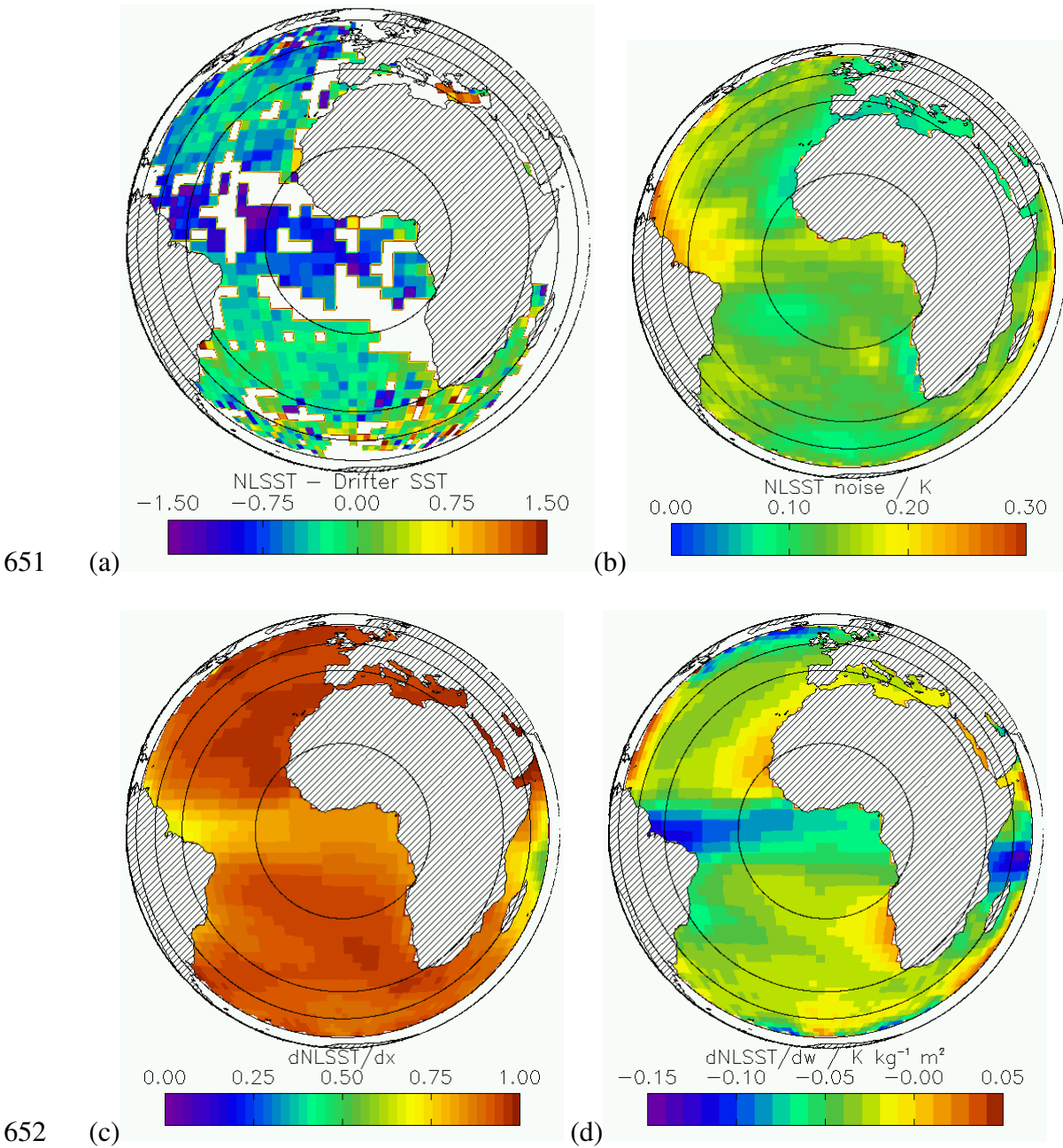
637 **Figure 3. Dependence of satellite minus drifter SST mean and SD on the number of pixels contributing to a cell, N.**  
638 **Diamonds: mean difference. Error bars: plus and minus 1 SD. Bar graph: number of cells in each range of N, relative**  
639 **to the most populous range which is normalized to 1. The horizontal axis is logarithmic to base 2, such that the third**  
640 **range (for example) includes cells where the number of pixels equalled or was greater than four ( $2^2$ ) and was fewer than**  
641 **eight ( $2^3$ ).** (a) NLSST minus drifter SST. (b) Operational SST minus drifter SST.

642  
643  
644



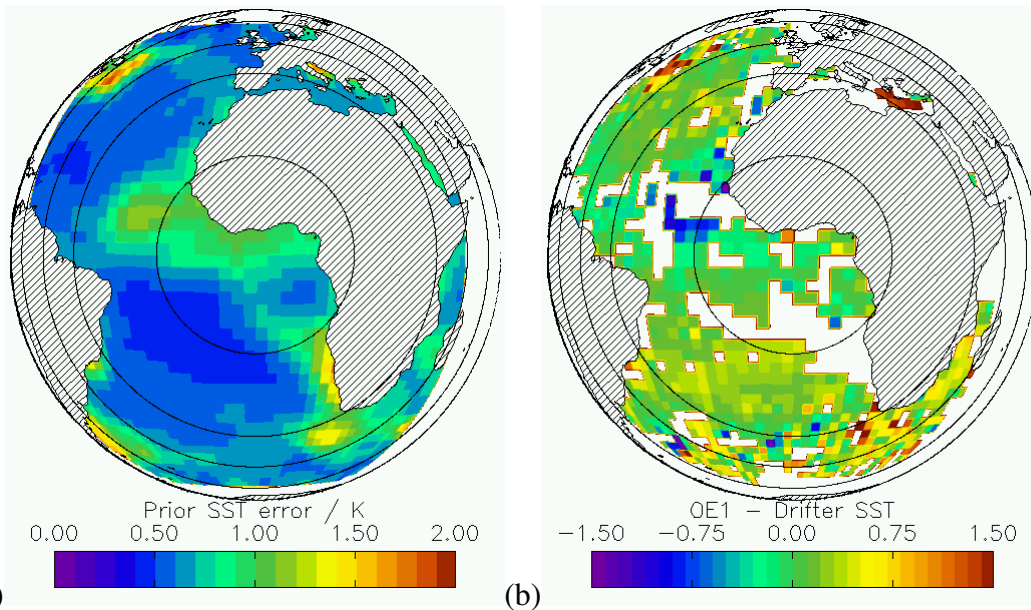
645

646 **Figure 4. Dependency of satellite minus drifter SST on (from left to right) total column water vapour, satellite zenith**  
647 **angle, brightness temperature difference between the 11 and 12  $\mu$ m channels, and SST. The satellite SST is the NLSST**  
648 **(see main text). Data points are in grey. The mean (diamonds) and  $\pm$ SD (bars) are plotted for bins of with 10 kg m<sup>2</sup>**  
649 **(TCWV), 10° (zenith angle), 1 K (BT difference) and 5 K (SST). (A small fraction of points in these and similar plots**  
650 **may be beyond the vertical range to allow sufficient discrimination of the majority of data.)**



653 **Figure 5. Mean properties of the NLSST over the three month study period. (a) Bias relative to drifters in the MDB. (b)**  
 654 **Random error in NLSST estimated by propagating radiometric noise in BTs through the  $0.5^\circ$ -cell-averaging and SST**  
 655 **retrieval process. (c) Sensitivity of NLSST to local variations in true SST, i.e., if a real change in SST occurs, this is the**  
 656 **fraction of that change registered in the retrieval. (d) Sensitivity of NLSST to local variations in TCWV, i.e., the false**  
 657 **temperature variation in the retrieval per unit change in column humidity.**

658

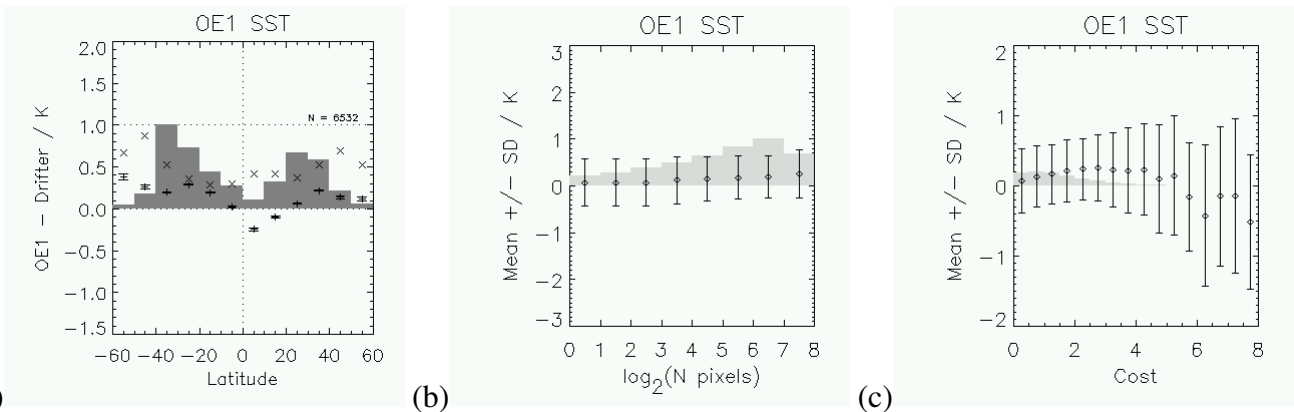


659 (a) (b)

660 **Figure 6. (a) Assumed prior error in SST. (b) OE1 SST compared to drifters.**

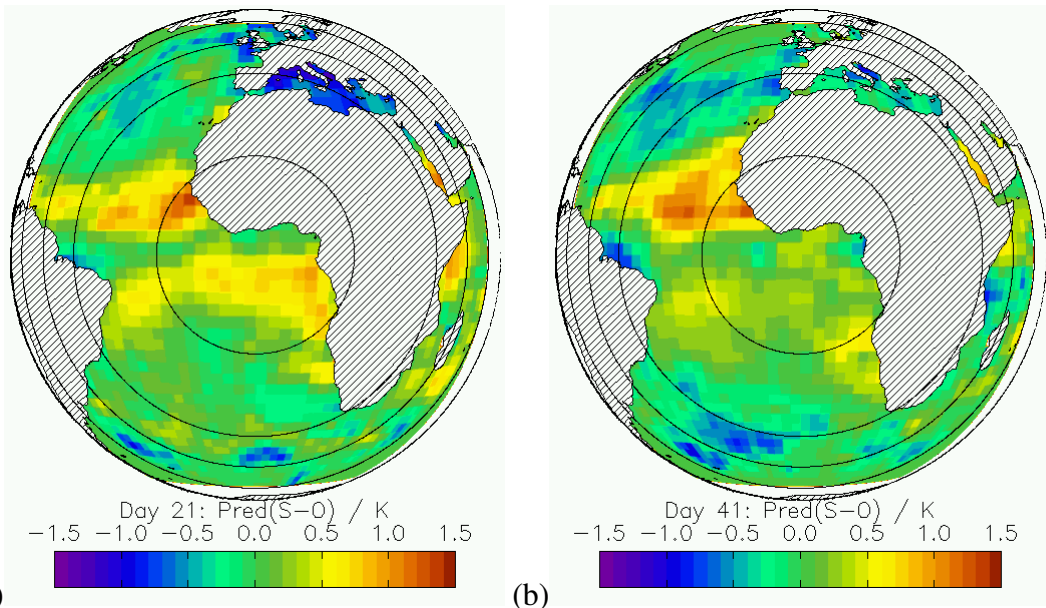
661

662



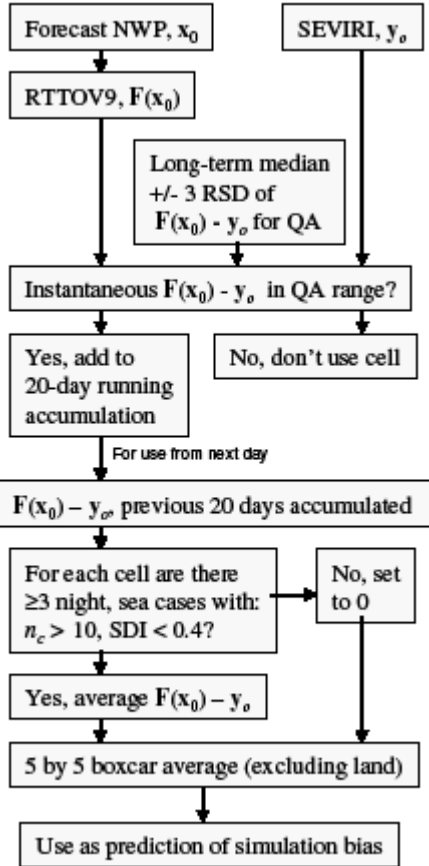
663 (a) (b) (c)

664 **Figure 7. (a) Zonal mean difference of OE1 SST from drifter SSTs; symbols as Figure 2. (b) As Figure 3(b), but for**  
 665 **OE1 SST. (c) Variation of OE1 minus drifter SST as a function of the cost of the MAP solution. Diamonds: mean in**  
 666 **each bin with respect to cost of width 0.5. Error bars: plus and minus one SD around mean. Bar chart in grey: relative**  
 667 **population per bin.**



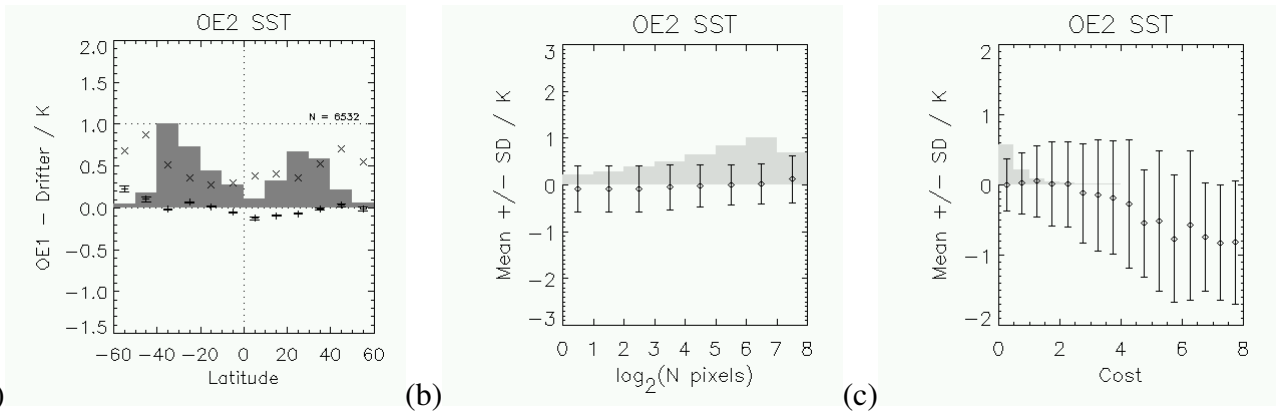
668 (a) (b)

669 Figure 8. Predicted S-O for days (b) 21 and (c) 41 of the study period, based on the previous 20 days' data in each case.



670

671 Figure 9. Schematic diagram of the procedure for estimating bias correction fields.



672

(a)

(b)

(c)

673

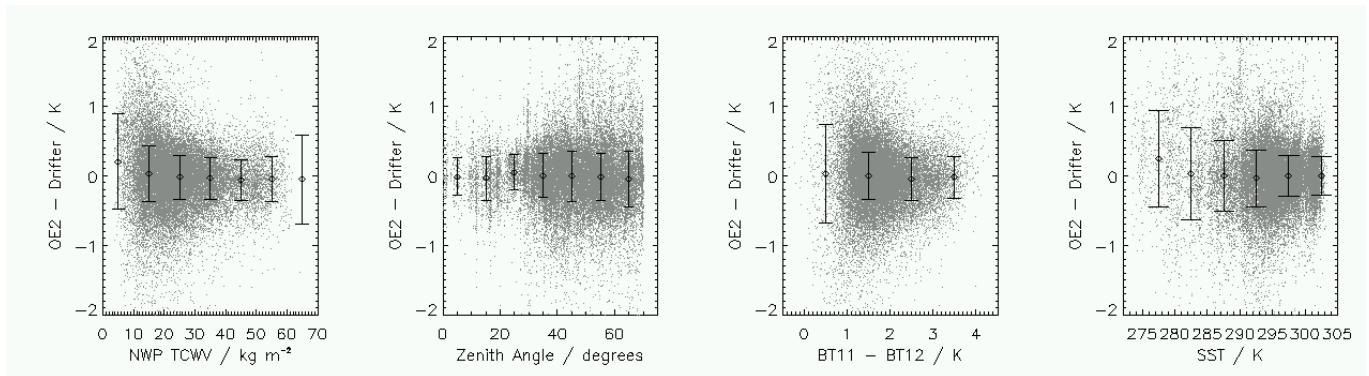
**Figure 10. (a) Zonal mean bias, standard errors, SD and relative population for OE2 SST minus drifter SST. Symbols**

674

**as in Figure 2. (b) Mean +/- SD of OE2 minus drifter as a function of number of contributing pixels in the  $0.5^\circ$  cell. (c)**

675

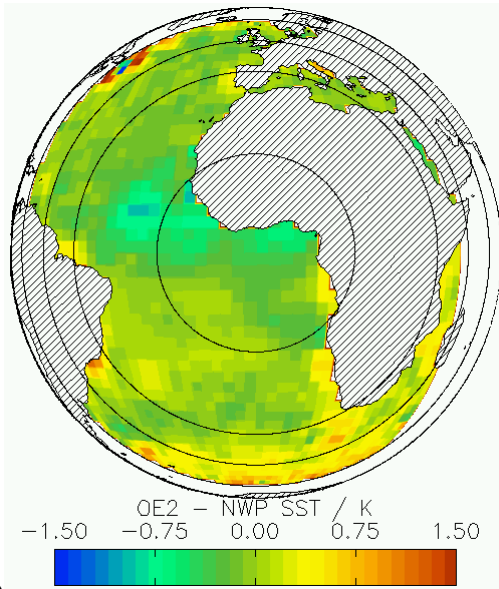
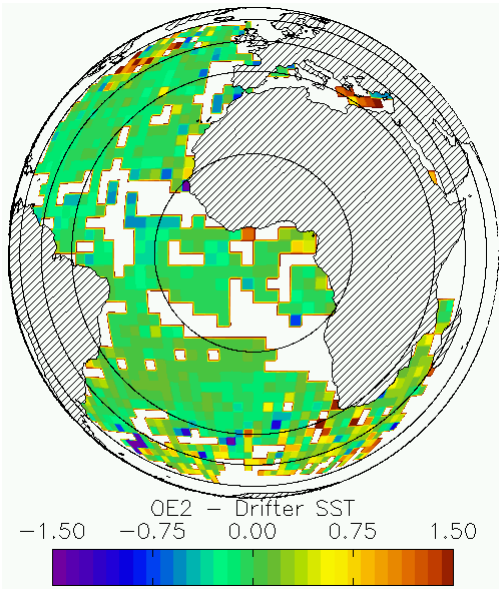
**Dependence of mean and SD on the cost of OE2 retrieval.**



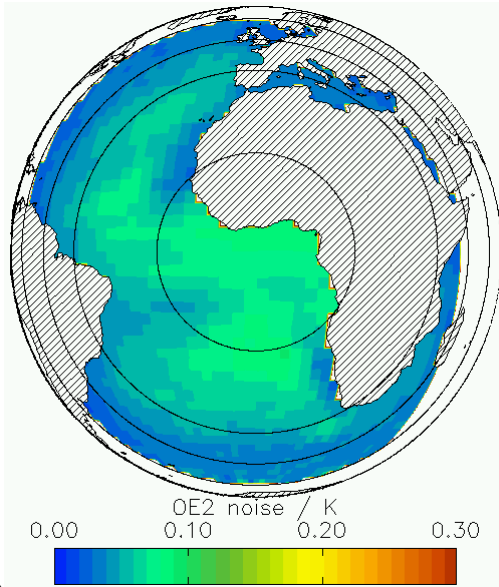
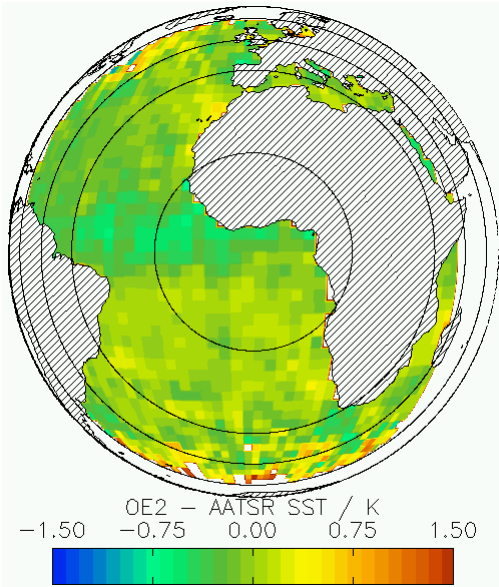
676

677

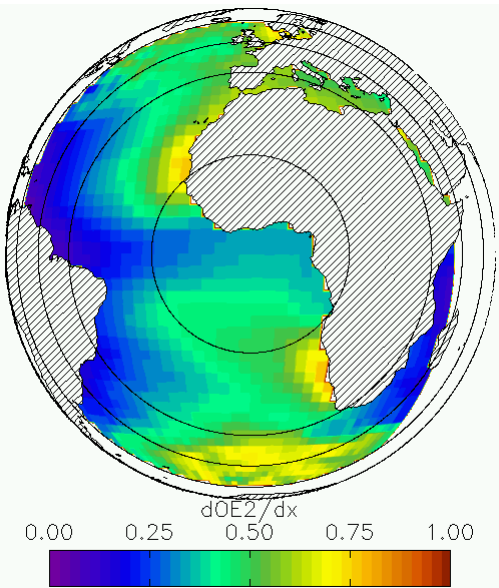
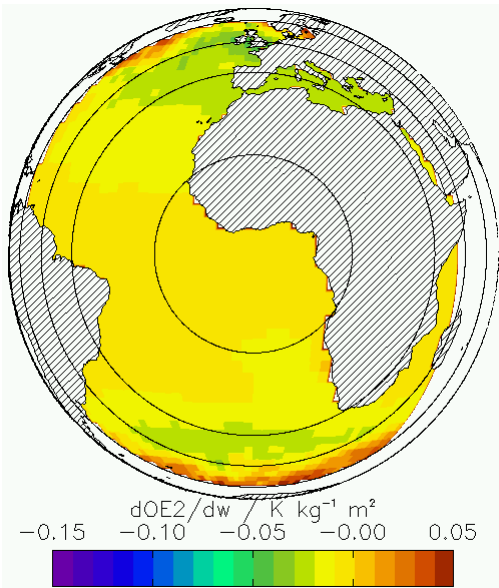
**Figure 11. As Figure 4, but for OE2 SST.**



678 (a) (b)



679 (c) (d)



680 (e) (f)

681 **Figure 12. [Previous page.] Aspects of OE2 SST. (a) Mean OE2 minus drifter SST. (b) Mean OE2 minus independent**  
682 **NPW SST. (c) Mean OE2 against independent satellite SST: dual-view three-channel night-time retrievals from the**  
683 **Advanced Along-Track Scanning Radiometer. (d) Radiometric noise propagation into OE2 SST estimates. (e)**  
684 **Sensitivity of OE2 SST to the total column water vapour. (f) Sensitivity of OE2 SST to changes in true SST.**

685