

Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data

Article

Accepted Version

Bulgin, C. E. ORCID: <https://orcid.org/0000-0003-4368-7386>, Embury, O. ORCID: <https://orcid.org/0000-0002-1661-7828> and Merchant, C. J. ORCID: <https://orcid.org/0000-0003-4687-9850> (2016) Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data. *Remote Sensing of Environment*, 117. pp. 287-294. ISSN 0034-4257 doi: 10.1016/j.rse.2016.02.021 Available at <https://centaur.reading.ac.uk/57700/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1016/j.rse.2016.02.021>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Sampling Uncertainty in Gridded Sea Surface Temperature Products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data

C. E. Bulgin^a, O. Embury^a, C. J. Merchant^a

^a*Department of Meteorology, University of Reading, Reading, UK*

Abstract

Sea surface temperature (SST) data are often provided as gridded products, typically at resolutions of order 0.05° from satellite observations to reduce data volume at the request of data users and facilitate comparison against other products or models. Sampling uncertainty is introduced in gridded products where the full surface area of the ocean within a grid cell cannot be fully observed because of cloud cover. In this paper we parameterise uncertainties in SST as a function of the percentage of clear-sky pixels available and the SST variability in that subsample. This parameterisation is developed from Advanced Along Track Scanning Radiometer (AATSR) data, but is applicable to all gridded L3U SST products at resolutions of $0.05\text{-}0.1^\circ$, irrespective of instrument and retrieval algorithm, provided that instrument noise propagated into the SST is accounted for. We also calculate the sampling uncertainty of ~ 0.04 K in Global Area Coverage (GAC) Advanced Very High Resolution Radiometer (AVHRR) products, using related methods.

Keywords:

1 **1. Introduction**

2 This paper addresses sampling uncertainty when deriving gridded sea sur-
3 face temperature products from satellite infrared imagery data. Remotely
4 sensed sea surface temperature data have uncertainties that should be quan-
5 tified for scientific applications. Typically, uncertainties in satellite retrieval
6 of sea surface temperature (SST) are quantified in a general sense via valida-
7 tion activities with reference to in-situ data (Donlon et al., 2007; GHRSSST
8 Science Team, 2010). In a companion paper, we present a method to estimate
9 context-specific uncertainties using physics-based models of uncertainty aris-
10 ing from different sources of error, evaluated for each SST retrieval. The un-
11 certainty estimates can be validated independently using in-situ data (Bulgin
12 et al., 2016). This is one component of the uncertainty budget in a grid cell
13 mean SST, which also includes random components from radiometric noise
14 (hereafter referred to as noise), locally systematic components that arise in
15 the SST retrieval step and uncertainty arising from unknown large scale sys-
16 tematic errors. A full discussion of these other components is provided in
17 Bulgin et al. (2016). This paper focuses on the derivation of an empirical
18 model of the uncertainty from spatially subsampling a grid cell for which an
19 area-average SST is to be estimated.

20 In this paper we will carefully distinguish the terms ‘error’ and ‘uncer-
21 tainty’, which are often used ambiguously. Error can be defined as the differ-
22 ence between an SST estimate (in this case from satellite data) and the true

23 SST (Kennedy, 2013; JCGM, 2008). In practice, the true SST is unknown
24 and therefore we cannot know the measurement error. We can instead cal-
25 culate the uncertainty, which is a measure of the dispersion of values that
26 could reasonably be attributed to that measurement error, we use a ‘stan-
27 dard uncertainty’ -ie, quoted uncertainties represent an estimate of the error
28 distribution standard deviation (JCGM, 2008). Although within this paper
29 the terms ‘error’ and ‘uncertainty’ are used according to these definitions,
30 usage differs in some cited references.

31 For many applications, SST data are not used or provided at the full
32 resolution of the sensor but are averaged over defined areas to produce a
33 gridded product. For large datasets with observations spanning many years,
34 this approach can be necessary to reduce the volume of data for some users.
35 Gridding in this way destroys more detailed information on the location of
36 measurements, and so a gridded SST value is taken as an estimate of the
37 average SST across the grid cell over some time period. Spatial sampling
38 uncertainty is present in gridded products, since the full grid cell may not
39 be observed (eg. because of partial cloud cover). If the gridded SST covers
40 a window of time (rather than being a measurement at a stated time) there
41 is also temporal sampling uncertainty, since the full time period may not be
42 observed (eg. one or two passes available during a day from which to make
43 a daily estimate). Temporal sampling issues are not discussed in this paper.

44 Sampling uncertainty has been widely considered in the construction of
45 global or regional SST records from in situ records for evaluating temperature
46 trends (Brohan et al., 2006; She et al., 2007; Rayner et al., 2006; Morrissey
47 and Greene, 2009; Jones et al., 1997; Folland et al., 2001; Karl et al., 1999).

48 In this context, sampling uncertainties arise from the number of observations
49 available in each grid cell and how well they represent the mean temperature
50 within the grid cell in both space and time (Jones et al., 1997). Sampling un-
51 certainty estimates consider the spatio-temporal correlation of measurements
52 at different locations within the grid cell (Morrissey and Greene, 2009), the
53 temporal variability in SST for each grid cell (Jones et al., 1997) and consis-
54 tency in observation depth (She et al., 2007).

55 Here we use data from the Advanced Along Track Scanning Radiometer
56 (AATSR) instrument to study sampling uncertainty in a gridded satellite
57 SST product. We calculate sampling uncertainties in data gridded at two
58 different spatial resolutions (0.05° and 0.1°) previously used in SST products
59 (eg. Embury and Merchant (2012); Merchant et al. (2014)). We separate
60 sampling uncertainty from other sources of uncertainty in SST so that it
61 can be estimated as a distinct contribution to the total uncertainty estimate.
62 We address only spatial sampling uncertainty because we aim to estimate
63 total uncertainty in SST in a grid cell at the stated time of the satellite
64 observations from a single overpass. We use the approach established in
65 this paper to consider sampling uncertainty in data provided at lower spatial
66 resolution than the native observations, for example in the case of Advanced
67 Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC)
68 products.

69 The remainder of the paper proceeds as follows. In Section 2 we discuss
70 the AATSR data and how they are used to synthesise sampling error distri-
71 butions. In Section 3 we derive steps for calculating sampling uncertainty.
72 In Section 4 we present our results using AATSR data and define a parame-

73 terisation for sampling uncertainty applicable over a range of spatial scales.
74 In Section 5 we consider uncertainties arising from GAC sampling from the
75 AVHRR instruments. We provide a discussion of the results in Section 6 and
76 conclude the paper in Section 7.

77 **2. Data and Methods**

78 Level 3 uncollated (L3U) satellite data products (the subject of this pa-
79 per) are defined as an average of the L2P data points of the highest quality
80 level that fall within the L3 grid cell (GHRSSST Science Team, 2010). The
81 gridded SST product as defined by the Group for High Resolution Sea Sur-
82 face Temperature (GHRSSST) specification is therefore a simple average of
83 the available observations as an estimate of the areal mean. Although other
84 methods could be considered for generating areal means, such as Kriging,
85 this is not the commonly accepted practice in this field.

86 When generating gridded SST products from infrared imagery, typically
87 only a subsample of the potential SST observations are available, predomi-
88 nantly due to cloud obscuring the surface, but occasionally due to a failed
89 retrieval or other problems with the observed data. If SST data points are
90 available covering the whole grid cell, an average SST can be calculated over
91 the grid cell. If a subset of points are available, the mean of these data may
92 differ from the true mean across the grid cell and therefore an element of
93 uncertainty is introduced into the mean of the available SSTs interpreted as
94 a grid cell mean. In this study, we mainly use data extracts from the Ad-
95 vanced Along Track Scanning Radiometer (AATSR) over clear-sky regions
96 in order to calculate the uncertainty introduced by estimating grid cell mean

97 SST from a subsample.

98 We extract 10 x 10 and 5 x 5 pixel samples globally which approximately
99 correspond to the size of 0.1 x 0.1° and 0.05 x 0.05° grid cells across the
100 tropics and mid-latitudes. AATSR has a pixel size of 1 km. At the equator,
101 5 km corresponds to 0.045 x 0.045° and at 60 degrees, this is 0.09 x 0.045°.
102 For the 10 km samples, these are 0.09 x 0.09° at the equator and 0.18 x
103 0.09° at a latitude of 60 degrees. Samples are selected from all latitudes
104 on the condition that all constituent pixels are classified as clear-sky using
105 the Bayesian cloud detection scheme applied to ATSR data in the Sea Sur-
106 face Temperature (SST) Climate Change Initiative (CCI) project (Merchant
107 et al., 2014). The 5 x 5 pixel cells are embedded in the 10 x 10 pixel cells
108 enabling a direct analysis of the impact of cell size on sampling uncertain-
109 ties. Subsamples of different numbers of clear-sky pixels (‘m’) are selected
110 from the full sample size (‘n’) using two methodologies to exclude pixels 1)
111 randomly and 2) using cloud-mask structures transposed from other cloudy
112 images. Random masks are compared with observed cloud-mask structures
113 to determine whether sampling uncertainties can be calculated accurately
114 using a more simple approximation. We calculate the sampling uncertainties
115 for all values of $m \Rightarrow 1$ and $m \leq n - 1$ for each cell size (5x5 or 10x10
116 pixel extracts).

117 The details of this approach are as follows. For each grid cell size and
118 value of ‘m’, we generated 500 random masks and extracted 500 realistic
119 cloud masks from other AATSR data screened using the SST CCI Bayesian
120 cloud detection scheme (Merchant et al., 2014). As noted above, all of the ex-
121 tracted samples are fully clear-sky, so neither mask corresponds to the cloud

122 conditions of any extract. However, the cloud masks obtained from other im-
123 ages have the spatial structures representative of cloud fields observed at the
124 scales of the imagery. Clear-sky samples were extracted from global AATSR
125 observations between 1st - 3rd January 2003 and sea surface temperatures
126 were calculated using an optimal estimation retrieval Merchant et al. (2014).
127 For each cell size we extracted 250,000 samples. We then applied each mask
128 (500 for each value of 'm') to each of the 250,000 extracts.

129 Figure 1 shows the global distribution of the 250,000 extracted samples.
130 These are classified according to the standard deviation of the SST over the 5
131 x 5 pixel cell to give an indication of the spatial distribution of sub-grid SST
132 variability. Clear-sky samples are extracted from orbit data globally with
133 the majority of extracts between 60° South and 60° North. These differences
134 between the masked and unmasked SSTs will be used to characterise sam-
135 pling uncertainty having accounted for the effect of SST noise on both the
136 full sample and subsample mean SST.

137 **3. Sampling Uncertainty Derivation**

138 This section presents the method of estimating sampling uncertainty from
139 these differences, accounting for the fact that the pixel SSTs are noisy. We
140 have to account for SST noise to develop a model for sampling uncertainty
141 that applies to sensors with different noise characteristics. Each mean SST
142 (of both a full extract and a subsample) will have an element of uncertainty
143 that ultimately derives from instrument noise in the observed brightness
144 temperatures from which the SSTs are estimated. To obtain a more accurate
145 sampling uncertainty we account for SST noise by the following method.

146 Considering first a single case, the mean SST across the full extract
 147 (\widehat{SST}_n) of ‘ n ’ pixels can be expressed as:

$$\widehat{SST}_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

148 where ‘ i ’ indexes the pixels for which ‘ x_i ’ is the (unknown) true SST. For
 149 a subsample of ‘ m ’ pixels, \widehat{SST}_m is:

$$\widehat{SST}_m = \frac{1}{m} \sum_{j=1}^m x_j \quad (2)$$

150 where the subscript ‘ j ’ represents the observations found in the subsample
 151 ‘ m ’. The subsampling error, ‘ E ’ is calculated by subtracting \widehat{SST}_n from
 152 \widehat{SST}_m .

$$E = \frac{1}{m} \sum_{j=1}^m x_j - \frac{1}{n} \sum_{i=1}^n x_i \quad (3)$$

153 Using the subscript ‘ h ’ to index only those observations that are not
 154 present in subsample ‘ m ’ (indexed using j) this equation can be rearranged
 155 to give:

$$E = \left(\frac{1}{m} - \frac{1}{n} \right) \sum_{j=1}^m x_j - \frac{1}{n} \sum_{h=1}^{n-m} x_h \quad (4)$$

156 This equation does not account for noise in the retrieved SST. In practice,
 157 we have only have an estimate (\hat{E}) of the true sampling error that is noisy
 158 because of SST noise in both \widehat{SST}_n and \widehat{SST}_m . Each retrieved SST, \hat{x}_i , is
 159 $\hat{x}_i = x_i + e_i$, where e_i is the error in the SST due to noise. We don’t know e_i
 160 explicitly, but we have an estimate of ϵ_i , which is the standard uncertainty

161 in a single pixel SST retrieval due to noise. The uncertainty due to noise in
 162 the extract mean is:

$$\epsilon_n = \frac{1}{n} \sqrt{\sum_{i=1}^n \epsilon_i^2} \quad (5)$$

163 with a similar expression for noise in a subsample mean, ϵ_m . The SST
 164 noise can be propagated through the form of equation (4) (Ku, 1966) to give
 165 the uncertainty in \hat{E} . Noise is negligibly correlated between pixels and the
 166 covariance term is therefore omitted.

$$\epsilon_E = \left[\left(\sum_{j=1}^m \left(\frac{\delta E}{\delta x_j} \right)^2 \epsilon_j^2 \right) + \left(\sum_{h=1}^{n-m} \left(\frac{\delta E}{\delta x_h} \right)^2 \epsilon_h^2 \right) \right]^{1/2} \quad (6)$$

$$= \left[\left(\frac{1}{m} - \frac{1}{n} \right)^2 \sum_{j=1}^m \epsilon_j^2 + \left(\frac{1}{n} \right)^2 \sum_{h=1}^{n-m} \epsilon_h^2 \right]^{1/2} \quad (7)$$

167 This uncertainty due to noise in \hat{E} is then subtracted from \hat{E} in variance
 168 space. Now, the sampling uncertainty (SU) we require is:

$$SU = [var(E)]^{\frac{1}{2}} \quad (8)$$

169 which we have to derive from realisations of \hat{E} for different extracts. E
 170 and \hat{E} are related by:

$$\hat{E} = E + e_E \quad (9)$$

171 where e_E is the error in the estimate of \hat{E} . Over multiple samples for a
 172 single given mask e_E is independent and uncorrelated with \hat{E} . Therefore, the
 173 variance in \hat{E} is equal to the sum of the variance in E and e_E .

$$var(\hat{E}) = var(E + e_E) \quad (10)$$

$$= var(E) + var(e_E) \quad (11)$$

$$= var(E) + \frac{1}{K} \sum_{k=1}^K \epsilon_{E_k}^2 \quad (12)$$

174 Where the k index represents different extracts. The variance of \hat{E} is
 175 estimated from the sample variance to give an unbiased estimate, as follows:

$$var(\hat{E}) = \frac{1}{K-1} \sum (\hat{E}_k - \langle \hat{E} \rangle)^2 \quad (13)$$

176 Here, K is the total number of extracts and $\langle \hat{E} \rangle$ the mean \hat{E} . Therefore,
 177 the sampling uncertainty can be estimated, accounting for noise,

$$SU = [var(\hat{E}) - var(e_E)]^{\frac{1}{2}} \quad (14)$$

178 Therefore substituting in equations (12) and (13):

$$SU = \left[\left(\frac{1}{K-1} \sum \left(\hat{E}_k - \frac{1}{K} \sum_{k=1}^K \hat{E}_k \right)^2 \right) - \left(\frac{1}{K} \sum_{k=1}^K \epsilon_{E_k}^2 \right) \right]^{1/2} \quad (15)$$

179 We apply this equation for calculating sampling uncertainty to the data
 180 as described in the following section.

181 **4. Results**

182 *4.1. Sampling Uncertainty over Different Grid Sizes*

183 We consider first sampling uncertainties over 5 x 5 pixel extracts corre-
184 sponding to gridded SST products at a resolution of 0.05°. For each value of
185 ‘ m ’ between $2 \leq m \leq 24$ (number of pixels available in the subsample) we
186 apply each of our 500 masks to the 250,000 extracts, treating random and re-
187 alistic cloud masks separately. For each of the masked samples we calculate
188 the difference between the full sample and the subsample mean SST. The
189 case where $m = 1$ is considered in a following section (4.3).

190 As demonstrated in Section 3, sampling uncertainty is dependent on the
191 number of pixels ‘ m ’ available in the subsample. It is also likely that the
192 magnitude of the sampling uncertainty will be dependent on the underlying
193 SST variability within the grid cell. There may be a significant gradient in
194 SST within a grid cell, for example in coastal regions, areas of upwelling or
195 near SST fronts. We would expect subsampling to introduce higher uncer-
196 tainties in the SST estimate in such locations than in grid cells where the
197 SST is more homogeneous. Our analysis is based on clear-sky data extracts
198 to which we have applied our various cloud masks, so we can calculate the
199 SST variability over the full grid cell. However, when considering subsampled
200 data in satellite imagery, the SSTs of the obscured pixels are unavailable. We
201 therefore examine the sampling uncertainty dependence on SST variability
202 by calculating the SST standard deviation across the ‘ m ’ pixels available in
203 the subsample.

204 The SST standard deviation across the subsample, minus the uncertainty
205 due to noise (subtracted in variance space) is calculated using Equation (15)

206 for each of the masked extracts and binned in 0.1 K bands between 0-0.6
207 K giving six groups of data. SST noise is propagated into the sample and
208 subsample SSTs from the pixel level uncorrelated uncertainties in the SST
209 product. In each bin we have 500 sampling uncertainty curves from the ap-
210 plication of 500 different masks, which are then combined to give a weighted
211 mean (Figure 2). With such a large dataset, for extracts where ‘m’ is small,
212 we find some cases for which the variance in the estimated SST noise is
213 greater than the variance in the subsample SST just because of statistical
214 fluctuations. To avoid negative variance, in these cases, we set the subsample
215 SST variance to zero, as the SST variability across the grid cell is extremely
216 low.

217 In Figure 2, panel (a) shows the results for the application of random
218 masks and panel (b) the application of realistic cloud masks. We see that
219 in both cases (random and realistic cloud masks) sampling uncertainty in-
220 creases as the percentage of clear-sky pixels (those available in subsample
221 ‘m’) decreases. The larger the SST standard deviation in subsample ‘m’,
222 the larger the associated sampling uncertainty for any given percentage of
223 clear-sky pixels. For the random masks, the sampling uncertainty increases
224 approximately linearly with a decreasing percentage of clear-sky pixels until
225 a value of 30-35 % where a more exponential increase is evident. Where
226 realistic cloud masks are applied, the relationship between the percentage of
227 clear-sky pixels and sampling uncertainty is more linear, with the gradient
228 of the line increasing with increasing subsample SST standard deviation.

229 We can also plot the same sampling uncertainty data as a function of the
230 SST standard deviation with the SST due to noise removed in the subsample

231 for selected values of ‘ m ’, as shown in the bottom two panels of Figure 2, again
232 for random (c) and realistic (d) cloud masks. These plots demonstrate that
233 there is little difference between the application of random and cloud masks
234 for ‘ $m = 24$ ’ where only a single pixel is masked, as would be expected. In
235 the application of realistic cloud masks the gradient in sampling uncertainty
236 as a function of increasing subsample SST standard deviation is steeper than
237 where random masks are applied, with larger overall sampling uncertainties
238 even in regions of low subsample SST variability (0.0-0.1 K).

239 The higher sampling uncertainties associated with the application of re-
240 alistic cloud masks in comparison with random masks are not unexpected.
241 Cloud fields tend to have coherent (non-random) spatial distributions, which
242 vary with the type of cloud. Realistic cloud masks are more likely to mask
243 adjacent pixels than random masks, which would increase the sampling error
244 where sample SST variability is high across the grid cell. SST is spatially
245 correlated between pixels at a resolution of 1 km. Therefore for a grid cell
246 with a high SST standard deviation there is likely to be a strong gradient
247 across the cell rather than a randomly distributed SST field. This coupled
248 with realistic coherent cloud spatial distributions increases sampling uncer-
249 tainty in comparison with using random masks. These results suggest that
250 spatial sampling uncertainties cannot be well represented by masking pixels
251 at random.

252 To assess the effect of the cell size on the sampling uncertainty we also
253 consider 10 x 10 pixel extracts that approximately correspond to SST gridded
254 products at 0.1° resolution. Figure 3 shows the equivalent plots to Figure 2
255 for the larger cell size. For the 10 x 10 pixel cell we see a wider sample of

256 clear-sky percentages due to the increased number of pixels in the sample.
257 The shape of the sampling uncertainty curve in relation to the percentage
258 of clear-sky pixels is similar for both random and realistic cloud masks to
259 the 5 x 5 pixel equivalent. There is a larger discrepancy in the absolute
260 sampling uncertainties here, with much higher values in the application of
261 realistic cloud masks. For random masks the the shift from a linear to more
262 exponential curve in SU as a function of the percentage of clear-sky pixels
263 occurs at $\sim 20\%$ for this cell size due to the greater number of pixels in each
264 extract.

265 For the larger sample size, we see higher maximum sampling uncertainties
266 when applying realistic cloud masks due to smaller clear-sky percentages
267 being represented by whole numbers of pixels. When pixels are masked
268 randomly, the sampling uncertainty for a given percentage of clear-sky pixels
269 and subsample SST deviation is lower when calculated over the 10 x 10 pixel
270 cell than then 5 x 5 pixel cell. For any given percentage of clear-sky pixels,
271 more pixels are available in the subsample ‘m’ over the 10 x 10 pixel cell
272 than the 5 x 5 pixel cell. This increases the likelihood that the observations
273 in the subsample will be distributed across the entire sampled cell for broken
274 cloud or where the length scale of the cloud structure is of the order of the
275 cell size.

276 *4.2. Modelling Sampling Uncertainties*

277 In practice, when generating gridded SST products we cannot calculate
278 sampling error by comparing the sample and subsample means as we do
279 not have SST available for pixels obscured by cloud. We need therefore
280 to model sampling uncertainty as a function of the variables we do have

281 available: the percentage of clear-sky pixels and the SST standard deviation
282 across subsample ‘ m ’, accounting for noise. We consider each SST standard
283 deviation band separately and plot sampling uncertainty with respect to the
284 percentage of clear-sky pixels, as a function of the number of pixels in the
285 full grid cell extract.

286 We can model the sampling uncertainty for the 5 x 5 and 10 x 10 pixel
287 cells by fitting a cubic in the form $SU = ax^3 + bx^2 + cx + d$ to the data
288 where x is the percentage of clear-sky pixels. Figure 4 shows the 5 x 5 and
289 10 x 10 pixel data and sampling uncertainty model. The coefficients for each
290 subsample standard deviation for the 5 x 5 pixel grid cells are given in Table
291 1 and for the 10 x 10 pixel grid cells in Table 2. Figure 4 indicates that the
292 cubic fit (shown in solid lines) is a close match to the data (dashed lines).
293 The data are slightly noisier than the model (as would be expected) and this
294 is more obvious in the 10x10 cell where more percentages of clear-sky pixels
295 are represented. In Figure 4, we see that the SST variability is the dominant
296 factor determining the shape of the modelled sampling uncertainty curve. As
297 this increases, the gradient of the sampling uncertainty curve increases giving
298 larger uncertainties particularly for lower percentages of clear-sky pixels. The
299 effect of varying ‘ n ’ is important in the context of generating products, regu-
300 larly gridded in latitude/longitude where the number of pixels falling within
301 each grid cell may vary with latitude or instrument coverage or viewing ge-
302 ometry. The sampling uncertainty curves for the two cell sizes show little
303 deviation from one another suggesting that the impact of small variations in
304 pixel number between grid cells at these scales is likely to be negligible.

305 The modelled sampling uncertainties are calculated from data where the

Table 1: Cubic coefficients as a function of subsample SST standard deviation for a 5x5 pixel cell using realistic cloud masks.

SST Std. Dev.	a	b	c	d
0.0-0.1 K	$-1.67e^{-7}$	$3.51e^{-5}$	$-2.82e^{-3}$	$9.65e^{-2}$
0.1-0.2 K	$-1.86e^{-7}$	$3.95e^{-5}$	$-3.63e^{-3}$	0.15
0.2-0.3 K	$-1.31e^{-7}$	$2.74e^{-5}$	$-3.37e^{-3}$	0.2
0.3-0.4 K	$-9.94e^{-8}$	$1.86e^{-5}$	$-3.12e^{-3}$	0.23
0.4-0.5 K	$-5.51e^{-8}$	$6.57e^{-6}$	$-2.53e^{-3}$	0.25
0.5-0.6 K	$-3.26e^{-8}$	$-1.59e^{-6}$	$-1.94e^{-3}$	0.25

Table 2: Cubic coefficients as a function of subsample SST standard deviation for a 10x10 pixel cell using realistic cloud masks.

SST Std. Dev.	a	b	c	d
0.0-0.1 K	$-2.80e^{-7}$	$5.44e^{-5}$	$-3.7e^{-3}$	0.1
0.1-0.2 K	$-3.36e^{-7}$	$6.56e^{-5}$	$-4.81e^{-3}$	0.16
0.2-0.3 K	$-2.80e^{-7}$	$5.48e^{-5}$	$-4.72e^{-3}$	0.2
0.3-0.4 K	$-2.49e^{-7}$	$4.72e^{-5}$	$-4.63e^{-3}$	0.24
0.4-0.5 K	$-2.47e^{-7}$	$4.36e^{-5}$	$-4.61e^{-3}$	0.27
0.5-0.6 K	$-2.22e^{-7}$	$3.63e^{-5}$	$-4.33e^{-3}$	0.29

306 effect of noise has been removed from the subsample SST. The model is there-
 307 fore applicable to SSTs generated from any instrument or retrieval on the
 308 same scale provided that the propagation of uncertainty due to noise within
 309 the SST calculation has been correctly accounted for. This has been verified
 310 using the nadir only ATSR Reprocessing for Climate (ARC) coefficient based
 311 SST estimate as a comparison (Embury and Merchant, 2012). The propaga-

tion of noise differs in an optimal estimation and coefficient-based retrieval. So although the data are from the same instrument (on different days) this is a good test of the robustness of the methodology. We find a maximum RMSE of 0.017 K and maximum mean percentage difference of 0.16% over both extract sizes between the sampling uncertainty model presented here and the equivalent model generated using ARC data. In cases where the variance in the noise exceeds the SST variance (for low numbers of clear sky pixels), the SST variance should be set to zero in order to use the model (using the same approach as that adopted in the generation of the model).

4.3. *Calculating Sampling Uncertainties for a Subsample of 1*

So far, the discussion on sampling uncertainty as a function of subsample size has excluded the case where the subsample size ‘ m ’ is equal to one. Under these conditions sampling uncertainty cannot be calculated as a function of the subsample SST standard deviation. We therefore calculate the sampling uncertainty for the case ‘ $m = 1$ ’ across all extracts (using a weighted mean from 500 masks applied to each of the 250,000 extracts and report the mean value in Table 3, treating random and realistic cloud masks separately. The sampling uncertainties calculated will be weighted towards lower full sample SST standard deviations as there are more extracts with lower SST variability. In regions of high SST gradients these values are therefore likely to be an underestimate, and in very homogeneous regions an over-estimate. The overall tendency towards full samples with lower SST variability is however typical of the global distribution of gridded SST values as samples are extracted across the globe (Figure 1). The sampling uncertainty where ‘ $m = 1$ ’ is larger over the 10 x 10 pixel grid cell (0.141 K) than the 5 x 5 pixel grid

337 cell (0.103 K). Where the full sample SST variability is high, a single pixel is
 338 unlikely to represent well the mean SST across the grid cell, and the larger
 339 the grid cell, the less likely this is to be representative.

Table 3: Sampling uncertainties for the case where $m = 1$ for cell sizes of 5 x 5 and 10 x 10 pixels, using realistic cloud masks (125×10^6 samples).

Cell Size	Sampling Uncertainty
5 x 5 pixels	0.103
10 x 10 pixels	0.141

340 5. AVHRR GAC Type Subsampling

341 So far we have considered the case where the number of pixels in the
 342 subsample ‘ m ’ is governed purely by data availability, ie. only observations
 343 obscured by cloud are eliminated from the available subsample. In the case
 344 of Global Area Coverage (GAC) data from the Advanced Very High Res-
 345 olution Radiometers (AVHRR) (Robel et al., 2014), there is a predefined
 346 sub-sampling in the transmitted data. Observations are made at 1.1 km res-
 347 olution at nadir, but due to limitations to data transmission from the early
 348 AVHRR instruments and latterly for consistency in data records, the GAC
 349 product is provided at a nominal resolution of 4 km. This is achieved by sub-
 350 sampling four pixels along the first scan line and then skipping a pixel before
 351 subsampling the next four pixels. The next two scan lines are skipped before
 352 resuming the sampling pattern described for the first line. Each four-pixel
 353 subsample is then considered to be representative of a 15 pixel cell (5 pixels
 354 across track by 3 pixels along track) (Robel et al., 2014). The signal received

355 for each GAC pixel is the average brightness temperature or reflectance over
356 the four pixels from the 15 pixel cell. Cloud screening is carried out on this
357 average, rather than the constituent pixels, before calculating SST. This in-
358 troduces a further source of sampling-related uncertainty, the calculation of
359 which is beyond the scope of this paper. Here we consider the uncertainty
360 introduced by regularly subsampling four in every fifteen pixels, and inter-
361 preting the four pixel average as an estimate for the full 3 x 5 pixel area.

362 We use Full Resolution Area Coverage (FRAC) Metop-A data to calcu-
363 late the sampling uncertainty in GAC products. We take data from 33 orbits
364 spanning the Metop-A data record and different times of year. From these
365 orbits we identify all of the 5 x 3 pixel clear-sky extracts using the opera-
366 tional EUMETSAT cloudmask, which gives good global coverage of scenes
367 (Ackermann et al., 2007). We apply the OSI-SAF coefficient based SST re-
368 trieval algorithms to these clear-sky extracts considering day and night sep-
369 arately, determined using solar zenith angle thresholds of $< 80^\circ$ and $> 100^\circ$
370 respectively (Le Borgne et al., 2007). We follow the methodology outlined
371 in Section 3 to calculate the sampling uncertainty by taking a subsample of
372 the first four pixels in every extract, having accounted for uncertainties due
373 to noise, subtracted in variance space. We discard extracts where the opera-
374 tional cloud detection has seemingly failed to identify cloudy pixels, giving
375 extreme SST variations across the fifteen pixel cell. We set an upper limit
376 on the SST variation across a given grid cell of 2 K with the threshold deter-
377 mined using model SST data (unaffected by clouds) at $\frac{1}{48}th^\circ$ resolution from
378 Estimating the Circulation and Climate of the Ocean (ECCO2) (Menemenlis
379 et al., 2008). For the closest match to GAC sampling we extract samples of

380 3 x 2 pixels ($\sim 6 \times 4$ km). Over a global sample of $> 11 \times 10^6$ extracts we
381 find a maximum SST gradient of 2.06 K across the full samples.

382 We specify AVHRR GAC sampling uncertainties under daytime and night-
383 time conditions, for three satellite viewing angle bands (1-1.1, 1.1-1.5 and
384 1.5-3 in secant theta space), corresponding to approximately 0-25°, 25-50°
385 and 50-70°. The results are shown in Table 4, in addition to the number
386 of extracts included in each calculation. For the OSISAF NL retrieval algo-
387 rithm we find that the sampling uncertainty is ~ 0.04 K and for the OSISAF
388 T37_1 algorithm ~ 0.03 K at night. One possible reason for the difference,
389 given that retrieval noise is accounted for, is the effect of cloud contamination
390 which is not explicit in the uncertainty budget. The OSISAF NL algorithm
391 shows slightly lower sampling uncertainties during the day than at night
392 which may be due to diurnal warming reducing SST variability (Katsaros
393 et al., 2005).

394 The OSISAF T37_1 nighttime algorithm uses the $3.7 \mu\text{m}$ channel in addi-
395 tion to the 11 and $12 \mu\text{m}$ channels, which is less sensitive to any cloud which
396 may be present eg. at cloud edges etc. This may explain the reduced vari-
397 ance in subsample minus full sample SSTs, and the slightly lower sampling
398 uncertainties when using this algorithm. For the GAC data, there is little de-
399 pendence on atmospheric path length with sampling uncertainties decreasing
400 by ~ 0.002 - 0.006 K at the swath edge. This is due to greater overlap of pixels
401 at this viewing geometry effectively reducing the unsampled area across the
402 15 pixel cell.

Table 4: AVHRR GAC sampling uncertainties as a function of viewing zenith angle. SSTs are calculated using OSISAF coefficient based retrievals, with the NL algorithm applied at solar zenith angles $< 80^\circ$ and the T37.1 and NL algorithms applied at solar zenith angles $> 100^\circ$.

Time	Algorithm	Viewing Angle	Sampling Uncertainty	Number of Obs
Day	OSISAF NL	0-25°	0.045	4278613
Day	OSISAF NL	25-50°	0.042	3396053
Day	OSISAF NL	50-70°	0.038	1841405
Night	OSISAF NL	0-25°	0.049	3639976
Night	OSISAF NL	25-50°	0.047	3015435
Night	OSISAF NL	50-70°	0.045	1668920
Night	OSISAF T37.1	0-25°	0.028	3639976
Night	OSISAF T37.1	25-50°	0.029	3015435
Night	OSISAF T37.1	50-70°	0.023	1668920

403 6. Discussion

404 Sampling uncertainties are yet to be routinely characterised in gridded
405 SST products and the model presented here provides a method for calculating
406 these uncertainties, applicable to all SST retrievals at the same scales as
407 those studied here, where uncertainties due to noise have been removed.
408 The impact of cell size is shown to be less important than the subsample
409 SST variability in determining the sampling uncertainty and therefore these
410 modelled uncertainties can be applied to grid cells at different latitudes and
411 varying viewing geometries where the number of pixels falling within each
412 grid cell can show local variation.

413 The results presented in Section 4.1 highlight significant differences in
414 the sampling uncertainties calculated when applying randomly generated
415 and observed cloud masks to the extracted samples. Sampling uncertain-
416 ties calculated on the basis of random masking are an underestimate of the
417 true uncertainty, a consequence of the spatial structure of both clouds and
418 the underlying SST field. Observed masks more often eliminate clumps of
419 pixels delineating a cloud feature rather than random pixels across a given
420 cell. As the percentage of clear-sky pixels is reduced this increases the like-
421 lihood of masking a large coherent section of the image. In all but the most
422 homogeneous cases of SST, the mean temperature of the remaining section
423 is less likely to be representative of the whole cell than a random distribution
424 of pixels scattered across the grid cell, due to the coherent structure of the
425 underlying SST.

426 Sampling uncertainties are inherent in all gridded products generated
427 from a subset of available data, eg. AVHRR GAC SSTs, Level 3 data.
428 These data can be used for a variety of applications and both data users
429 and providers should be aware of the uncertainties introduced by subsam-
430 pling the higher resolution data.

431 **7. Conclusions**

432 In this paper we present a methodology for calculating sampling uncer-
433 tainty in gridded SST products once the uncertainty due to noise in the ob-
434 servations has been removed. We model sampling uncertainty as a function
435 of the percentage of clear-sky pixels within a given grid cell and the SST vari-
436 ability within those available pixels, considering cell sizes of 0.05° and 0.1° .

437 We establish that the dominant factor in determining sampling uncertainty
438 is the subsample SST standard deviation and that latitudinal variations in
439 the number of pixels falling within a given grid cell have a negligible effect.
440 Our model is applicable to SST retrievals from any instrument on the same
441 spatial scales, using any retrieval scheme providing that the propagation of
442 instrument noise through the retrieval is correctly accounted for. We also
443 consider the impact of routine subsampling of higher resolution data in the
444 provision of GAC AVHRR products. We characterise sampling uncertainty
445 as a function of atmospheric path length corresponding to viewing zenith
446 angle, as information regarding the SST variability within the subsample is
447 not provided within the GAC product. We find that sampling uncertainty
448 is typically of the order of 0.04 K. We recommend the inclusion of sampling
449 uncertainties in the uncertainty estimates provided with SST products, and
450 demonstrate the validation of the ATSR uncertainty budget including this
451 component in the companion paper (Bulgin et al., 2016).

452 **8. Acknowledgements**

453 The work undertaken in this paper was funded by the European Space
454 Agency Sea Surface Temperature Climate Change Initiative project. We
455 thank the ECCO2 project for making their model output available.

456 **References**

457 Ackermann, J., & Klaes, D., & McKernan, E., & Montagner, F., & Heine-
458 mann, T., & Schl'ussel, P., & Schraidt, R., & Fiedler, L., & Lang, R., &

459 Bonsignori, R., & (2007). ATOVS/AVHRR Onboard Metop-A: Validation
460 and Use of Products. EUMETSAT, Am Kavalleriesand 31, D-64295,
461 Darmstadt, Germany.

462 Bulgin, C. E., & Embury, O., & Corlett, G., & Merchant, C. J. (2016). In-
463 dependent uncertainty budget estimates for coefficient based sea surface
464 temperature retrieval from the Along-Track Scanning Radiometer instru-
465 ments. *Remote Sensing of Environment*.

466 Brohan, P. & Kennedy, J. J., & Harris, I., & Tett, S. F. B., & Jones, P. D.
467 (2006). Uncertainty estimates in regional and global observed temperature
468 changes: A new data set from 1850. *Journal of Geophysical Research*, *111*,
469 D12, 21pp.

470 Donlon, C., & Robinson, I., & Casey, K. S., & Vazquez-Cuevo, J., & Arm-
471 strong, E., & Arino, O., & Gentemann, C., & May, D., & LeBorgne, P.,
472 & Piollè, J., & Barton, I., & Beggs, H., & Poulter, D. J. S., & Merchant,
473 C. J., & Bingham, A., & Heinz, S., & Harris, A., & Wick, G., & Emery,
474 B., & Minnett, P., & Evans, R., & Llewellyn-Jones, D., & Mutlow, C., &
475 Reynolds, R. W., & Kawamura, H., & Rayner, N. (2007). The Global
476 Ocean Data Assimilation Experiment High-resolution Sea Surface Tem-
477 perature Pilot Project. *Bulletin of the American Meteorological Society*,
478 *88*, 8, 1197-1213.

479 Embury, O., & Merchant, C. J. (2012). A reprocessing for climate of sea
480 surface temperature from the along-track scanning radiometers: A new
481 retrieval scheme. *Remote Sensing of Environment*, *116*, 47-61.

- 482 Folland, C. K., & Rayner, N. A., & Brown, S. J., & Smith, T. M., & Shen,
483 S. S. P., & Parker, D. E., & Macadam, I., & Jones, P. D., & Jones, R. N.,
484 & Nicholls, N., & Sexton, D. M. H. (2001). Global temperature change
485 and its uncertainties since 1861. *Geophysical Research Letters*, *28*, 13,
486 2621-2624.
- 487 GHRSSST Science Team (2010). The Recommended GHRSSST Data Speci-
488 fication (GDS) Revision 2.0 Technical Specifications. Available from the
489 GHRSSST International Project Office, <http://www.ghrsst.org>, *120pp*.
- 490 Joint Committee for Guides in Metrology (2008). Evaluation of measure-
491 ment data - Guide to the expression of uncertainty in measurement.
492 www.bipm.org, *134pp*.
- 493 Jones, P. D., & Osborn, T. J., & Briffa, K. R. (1997). Estimating Sampling
494 Errors in Large-Scale Temperature Averages. *Journal of Climate*, *10*,
495 2548–2568.
- 496 Karl, T. R., & Knight, R. W., & Christy, J. R. (1999). Global and Hemi-
497 spheric Temperature Trends: Uncertainties Related to Inadequate Spatial
498 Sampling. *Journal of Climate*, *7*, 1144–1163.
- 499 Katsaros, K. B, & Soloviev, A. V., & Weisberg, R. H., & Luther, M. E.
500 (2005). Reduced Horizontal Sea Surface Temperature Gradients Under
501 Conditions of Clear Skies and Weak Winds. *Boundary-Layer Meteorology*,
502 *116*, *2*, 175-185.
- 503 Kennedy, J. J. (2013). A review of uncertainty in in situ measurements and
504 data sets of sea surface temperature. *Reviews of Geophysics*, *51*, *1*, 1-32.

- 505 Ku, H. H. (1966). Notes on the Use of Propagation of Error Formulas *Journal of Research of the National Bureau of Standards - C Engineering and*
506 *Instrumentation*, 70C, 263-273.
- 508 Le Borgne, P., & Legendre, G., & Marsouin, A. (2007). Operational SST Retrieval from METOP/AVHRR. *Proceedings of EUMETSAT Conference*,
509 10pp.
- 511 Menemenlis, D., & Campin, J., & Heimbach, P., & Hill, C, & Lee, T., &
512 Nguyen, M, & Schodlok, M., & Zhang, H (2008). ECCO2: High resolution global ocean and sea ice data synthesis.. *Mercator Ocean Quarterly*
513 *Newsletter*, 31, 13-21.
- 515 Merchant, C. J., & Embury O. & Roberts-Jones, J. & Fiedler, E. & Bulgin,
516 C. E. & Corlett, G. K. & Good, S. & McLaren, A. & Rayner, N. & Donlon,
517 C. (2014). Sea surface temperature datasets for climate applications from
518 Phase 1 of the European Space Agency Climate Change Initiative (SST
519 CCI). *Journal of Atmospheric and Oceanic Technology*, 21, 11, 1734-1746.
- 520 Morrissey, M. L., & Greene, J. S. (2009). A theoretical framework for the
521 sampling error variance for three-dimensional climate averages of ICOADS
522 monthly ship data. *Theoretical Applied Climatology*, 96, 235-248.
- 523 Rayner, N. A., & Brohan, P., & Parker, D. E., & Folland, C. K., & Kennedy,
524 J. J., & Vanicek, M., & Ansell, T. J., & Tett, S. F. B. (2006). Improved
525 Analyses of Changes and Uncertainties in Sea Surface Temperature Measured
526 In Situ since the Mid-Nineteenth Century: The HadSST2 Dataset.
527 *Journal of Climate*, 19, 446-469.

528 Robel, J., & Graumann, A., & Kidwell, K., & Aleman, R., & Ruff, I.,
529 & Muckle, B., & Green, P., & Brown, S., & Throwe, J., & Planet,
530 W., & Knoll, J., & Fineberg, D., & Winston, W., & Robinson, G.,
531 & Miller, T., & Goodrum, G., & Askew, J., & Sapper, J., & Pat-
532 terson, S., & Harrod, E, & Brown, E., & Wark, D., & Ramsay, B.,
533 & Krimchansky, S., & Flynn, L., & Wrublewski, T., & Ross, D. &
534 Mo, T., & Graumann, A., & Bowman, D., & Sullivan, J. & Peterski,
535 N., & Hollinger, M., & Clouse, W., & Weinreb, M. & Rao, N., &
536 McMillin, L., & Goldberg, H., & Kleespies, T. (2014). NOAA KLM
537 User's Guide with NOAA-N, N Prime, and MetOp Supplements. U.S.
538 Department of Commerce, National Oceanic and Atmospheric Admin-
539 istration, National Environmental Satellite, Data and Information Service.
540 [http://www1.ncdc.noaa.gov/pub/data/satellite/publications/podguides/N-](http://www1.ncdc.noaa.gov/pub/data/satellite/publications/podguides/N-15%20thru%20N-19/pdf/0.0%20NOAA%20KLM%20Users%20Guide.pdf)
541 [15%20thru%20N-19/pdf/0.0%20NOAA%20KLM%20Users%20Guide.pdf](http://www1.ncdc.noaa.gov/pub/data/satellite/publications/podguides/N-15%20thru%20N-19/pdf/0.0%20NOAA%20KLM%20Users%20Guide.pdf).

542 She, J., & Hoyer, J. L.,& Larsen, J. (2007). Assessment of sea surface tem-
543 perature observational networks in the Baltic Sea and North Sea. *Journal*
544 *of Marine Systems*, 65, 314-335.

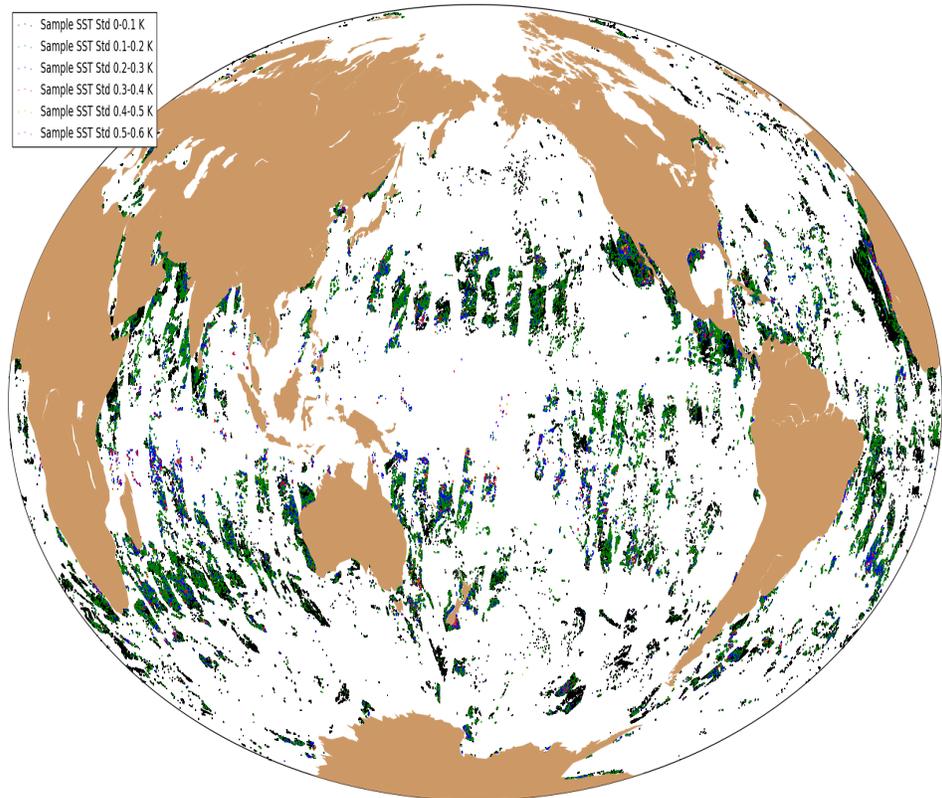


Figure 1: Global distribution of 10 x 10 pixel clear sky sea surface temperature samples extracted from AATSR data between 1st-3rd January 2003. Samples are colour coded according to the SST standard deviation across the sample cell.

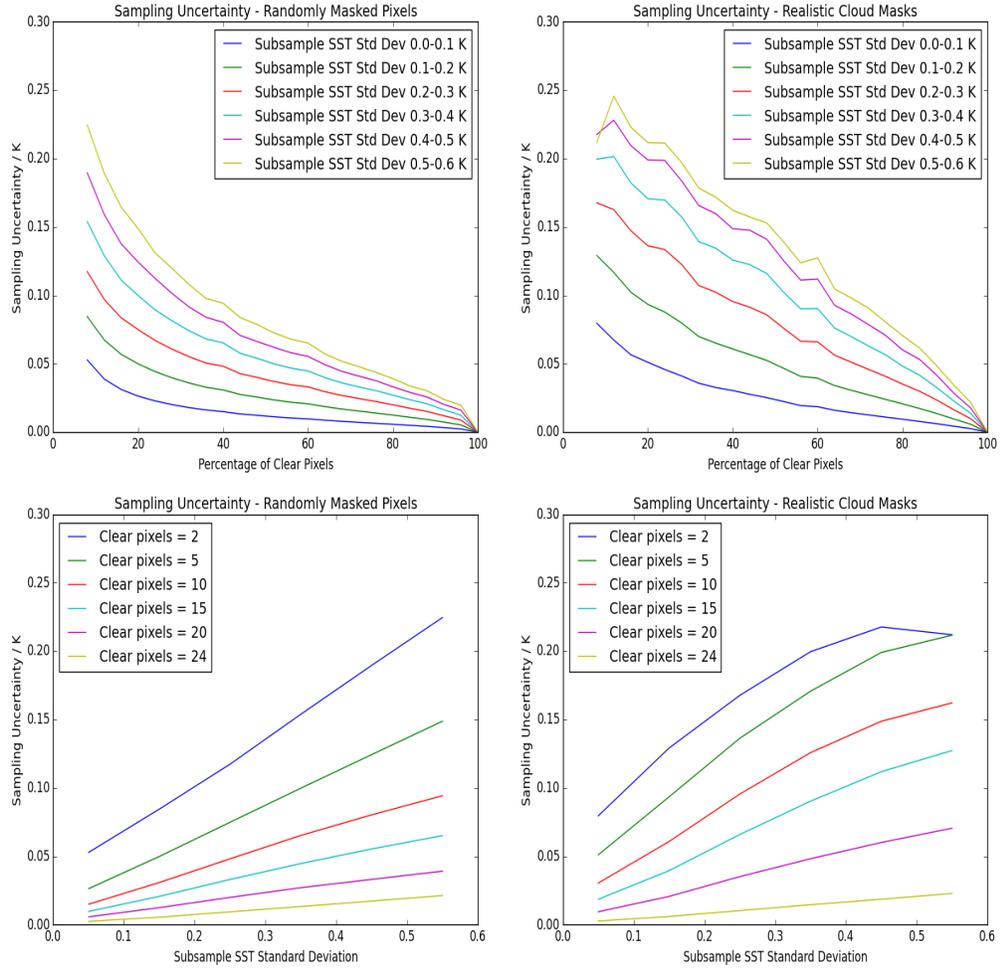


Figure 2: Top: Sampling uncertainties as a function of clear-sky pixel percentage over a 5 x 5 pixel cell with the application of randomly generated (left) and observed (right) cloud masks. Data are separated into six subsample SST standard deviation bands between 0-0.6 K. Bottom: Sampling uncertainty as a function of the subsample SST standard deviation with application of randomly generated (left) and observed (right) cloud masks. Results are presented for a number of clear-sky pixels.

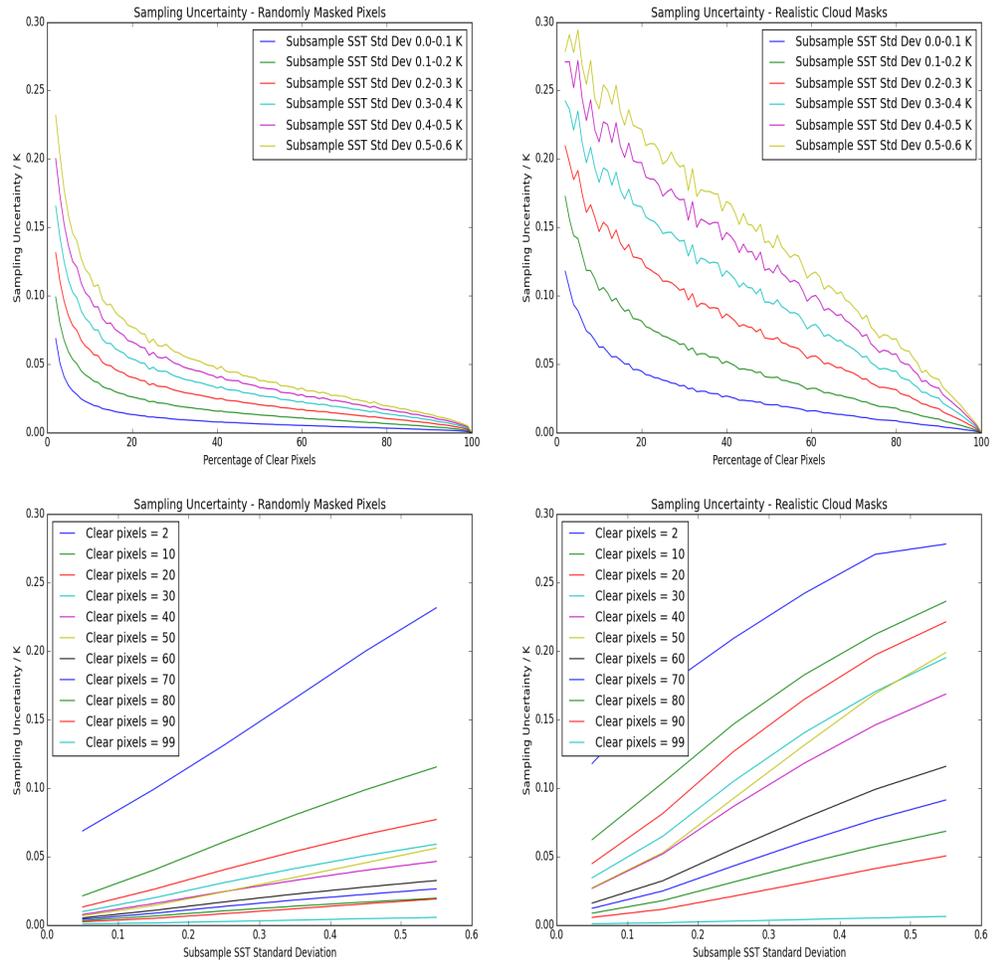


Figure 3: Top: Sampling uncertainties as a function of clear-sky pixel percentage over a 10 x 10 pixel cell with the application of randomly generated (left) and observed (right) cloud masks. Data are separated into six subsample SST standard deviation bands between 0-0.6 K. Bottom: Sampling uncertainty as a function of the subsample SST standard deviation with application of randomly generated (left) and observed (right) cloud masks. Results are presented for a number of clear-sky pixels.

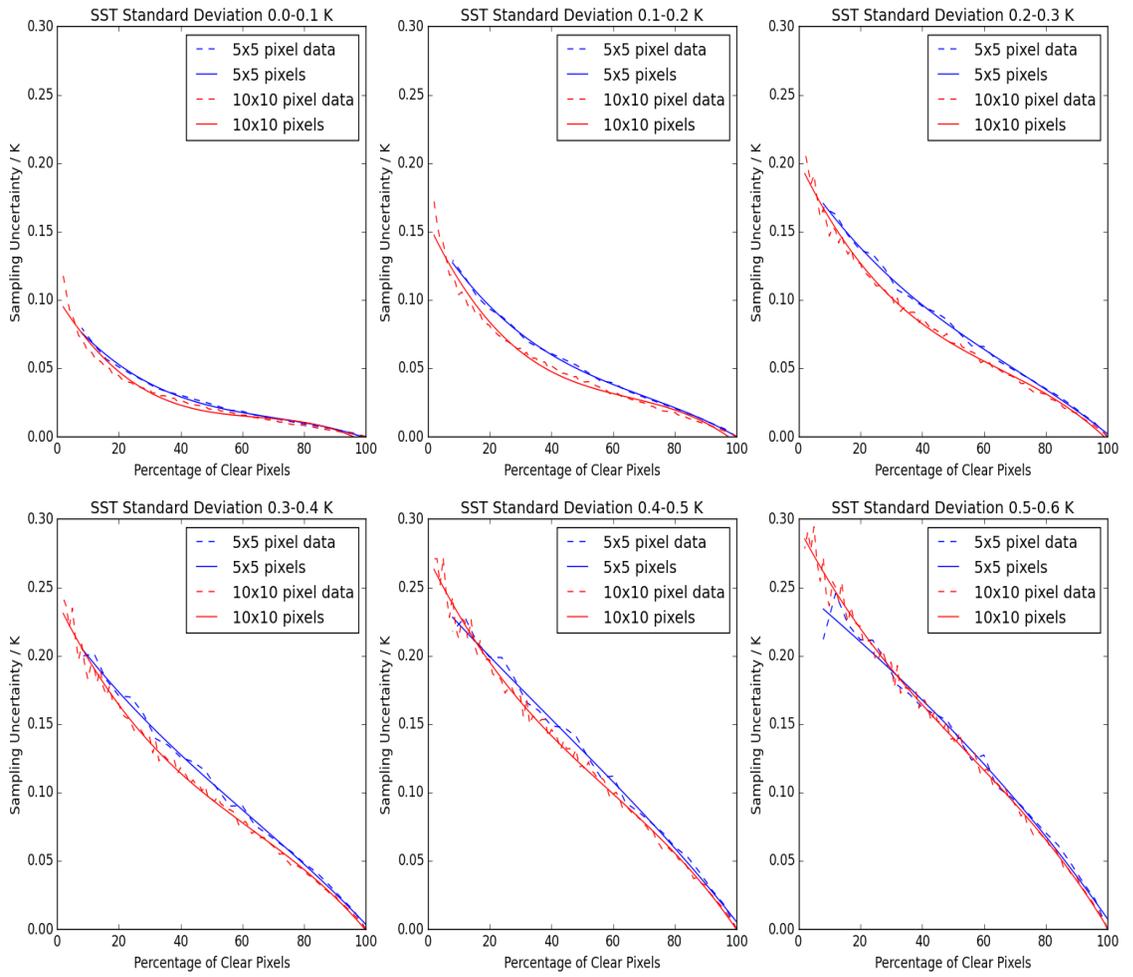


Figure 4: Modelled sampling uncertainties for 5 x 5 and 10 x 10 pixel cells over six subsample SST bands ranging between 0-0.6 K. Data for 25 and 100 pixel cells are over-plotted in each panel.