

volumes of data – the thicknesses of the arrows in Fig. 1 – than those in 2nd group. There are, of course, many exceptions to this categorization but it, together with the five examples presented below, helps to appreciate the motivation for a better understanding of the uncertainty associated with the value of one pixel relative to that of ‘nearby’ pixels - the precision of the retrievals.

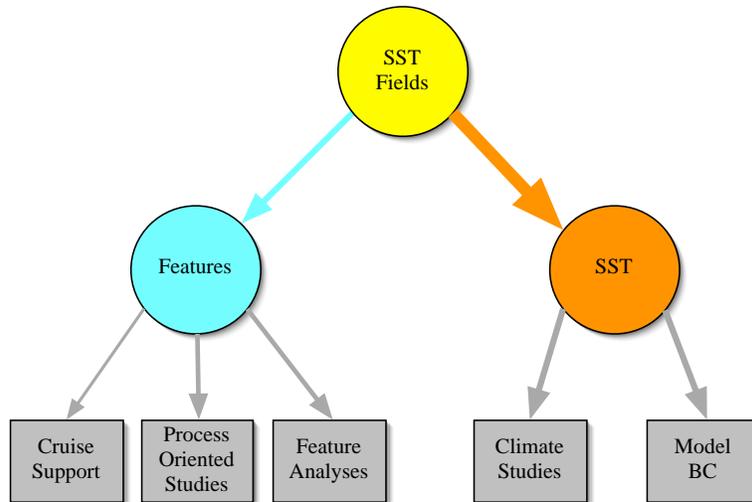


Figure 1. Uses made of satellite-derived SST fields fall into two broad categories.

Fig. 2 shows (left) the variance of the 4 km MODIS SST product (circa 2010) values calculated for each image in 2008 and then averaged for the year and (right) the same for the 4 km Pathfinder fields (circa 2010). Fig. 3 is a similar plot for the gradient magnitude calculated with the Sobel gradient kernel applied to the same fields used for Fig. 2. The ‘best’ quality flags were used for both MODIS and Pathfinder. The uncertainty of the SST values in these fields were quoted at very nearly the same value, on the order of 0.4K (I think, it’s been a while since I put these plots together). So, in principle, these plots should look quite similar. Well, you don’t need me to tell you that they don’t. The problem is that there was, and still is, no information suggesting that these fields should look different. Is it due to differences in the cloud masks, in the pixel-to-pixel noise or some other factor.

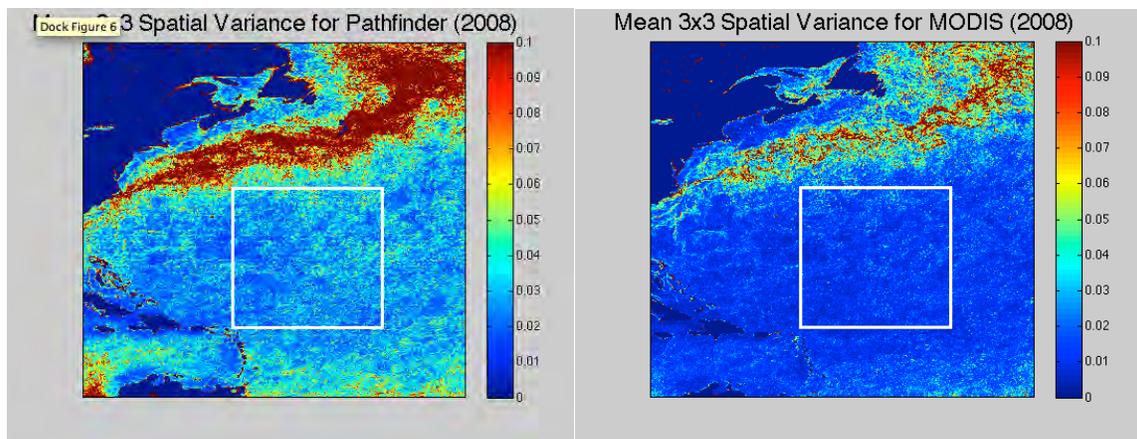


Figure 2. 3x3 variance averaged over 2008 for Pathfinder SST (left) and MODIS SST (right).

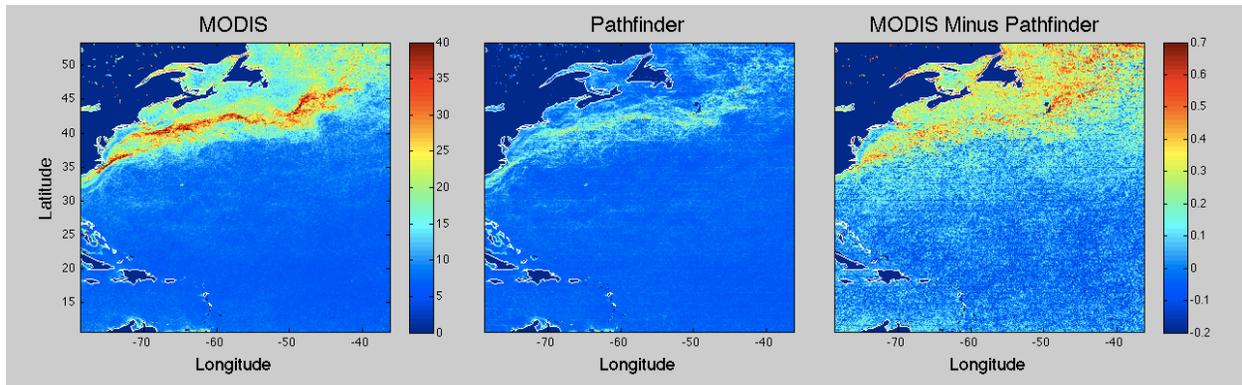


Figure 3. SST gradient magnitude averaged over 2008 for MODIS (left), Pathfinder (center) and two times MODIS minus Pathfinder over MODIS plus Pathfinder.

The second example is shown in Fig. 4. It shows SST for an approximately 400x300 km area with no cloud cover from AMSR-E, MODIS, SEVIRI and ATSR. The last three were averaged to the AMSR-E grid. (I'm sorry but I don't have the details of the time, location or exactly how the remapping was done but the image still shows the point quite clearly. We can reproduce this if there is agreement that such an image would be useful here.) The important point about this image is the significant difference in features evident from one image to the next for data from four different satellites purporting to be observing the same geophysical field.

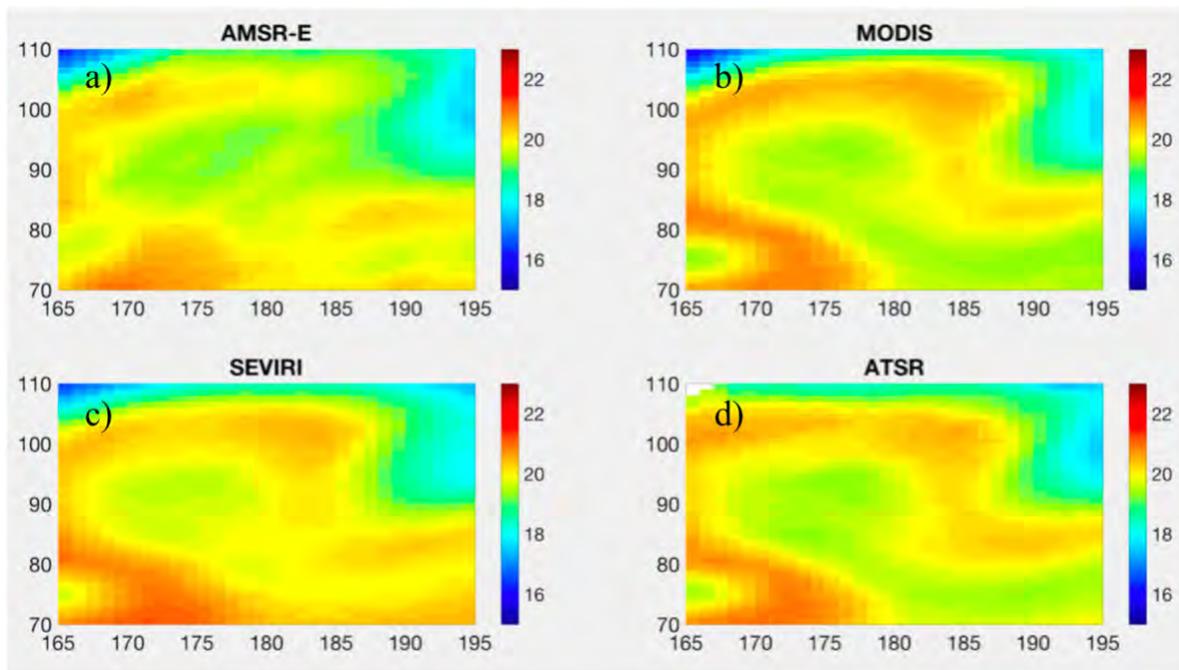


Figure 4. (a) SST obtained from AMSR-E for a 40x30 pixel region, (b) MODIS SST mapped to the same projection as AMSR-E, (c) the same for SEVIRI and (d) for ATSR.

The third example is shown in Fig. 5. It represents the probability of finding a front in regridded SEVIRI SST fields for cloud free conditions. Fronts were located with the Cayula-Cornillon edge detection algorithm. And, no, the regular sort of semi-circular bands west of 55°W and

south of 40°S, are not the result of some compression process associated with internal waves triggered from somewhere around 47°S, 50°W – I wish they were because would be pretty cool! They result from the resampling of the original data. What is interesting in this regard is that these patterns are not readily apparent in the SST fields, they only emerge when the frontal fields are averaged over a large number of images, approximately 3,000 in this case, if I remember properly. And, again, there is nothing in the statistical representation of the uncertainty in these fields that would lead the user to expect such behavior.

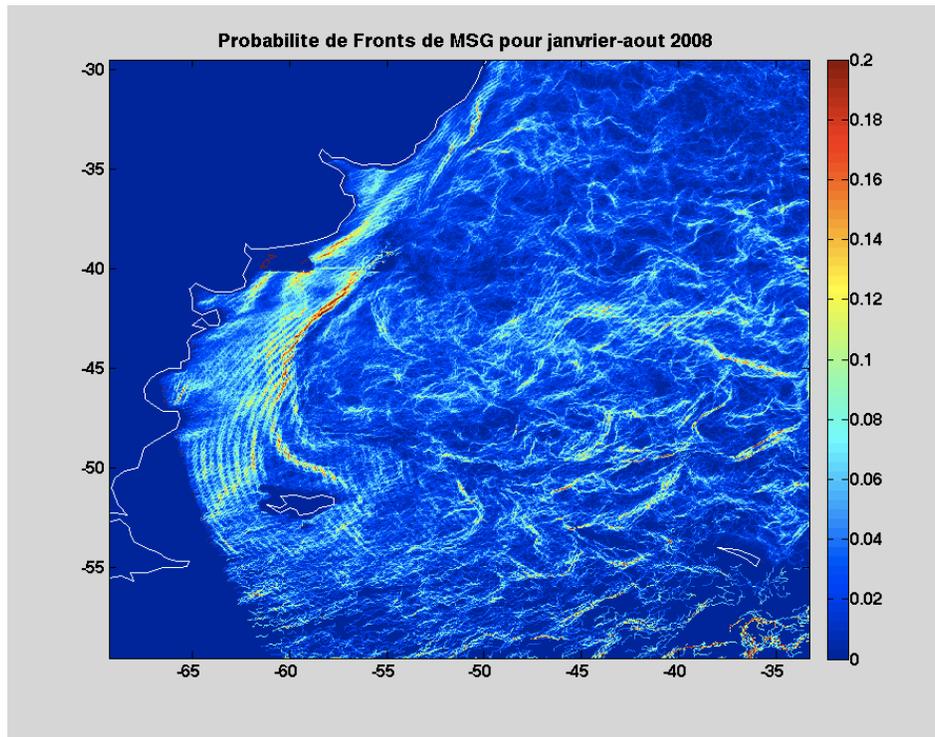


Figure 5. Frontal probability determined from SEVIRI SST fields for the period January through August 2008 using the Cayula-Cornillon edge detection algorithm.

The next two issues to be discussed are based on MODIS Aqua L2 SST fields produced by, and recently obtained from, the Ocean Biology Processing Group (OBPG) at NASA's Goddard Space Flight Center. We present the first of these in some detail because of the important issues it raises with regard to the proposed work of this task team. (For a more detailed description of the problem the reader is referred to GHRSSST XXII presentation S2-ID-038.) The statistics discussed in the following were obtained from 128x128 pixel regions (referred to as cutouts) extracted from the L2 global dataset of SST fields for 2003-2019. Only cutouts that were at least 95% clear were considered. Variograms were obtained for each cutout in the along-scan and along-track directions and these variograms were then combined for all cutouts falling in 200 km x 200 km x 5 day space-time, non-overlapping bins covering the globe. The standard deviation of the pixel noise – the precision – for each bin was obtained by extrapolating to zero the 4th order polynomial fit to the square root of the variogram for pixel separations of 1 to 20 km. The resulting along-scan σ_s are plotted geographically in Fig. 6.

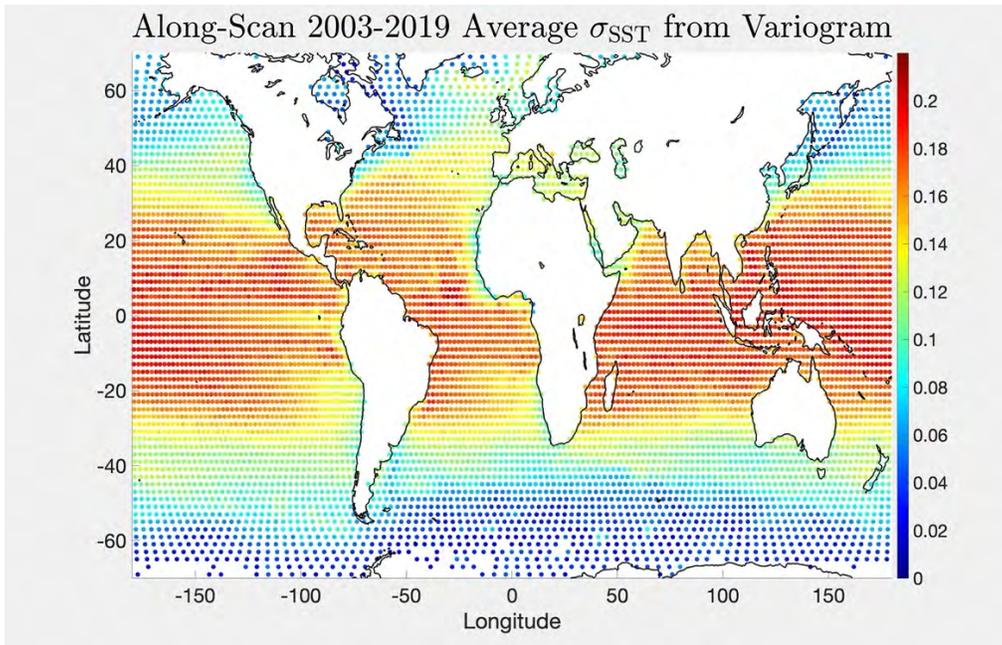


Figure 6. σ averaged over 2003-2019 in 200km x 200km bins. Each dot is at the center of the bin.

Fig. 7 shows the along-scan σ s plotted versus the mean SST value of the cutouts in the bin. (Results for the along track direction are similar except that σ is a bit larger because of scan-to-scan calibration noise.) The strong, nearly linear, dependence on mean SST came as a surprise. The best fit straight line to the data is of the form: $\sigma = 0.0048 \times \overline{SST} + 0.031$

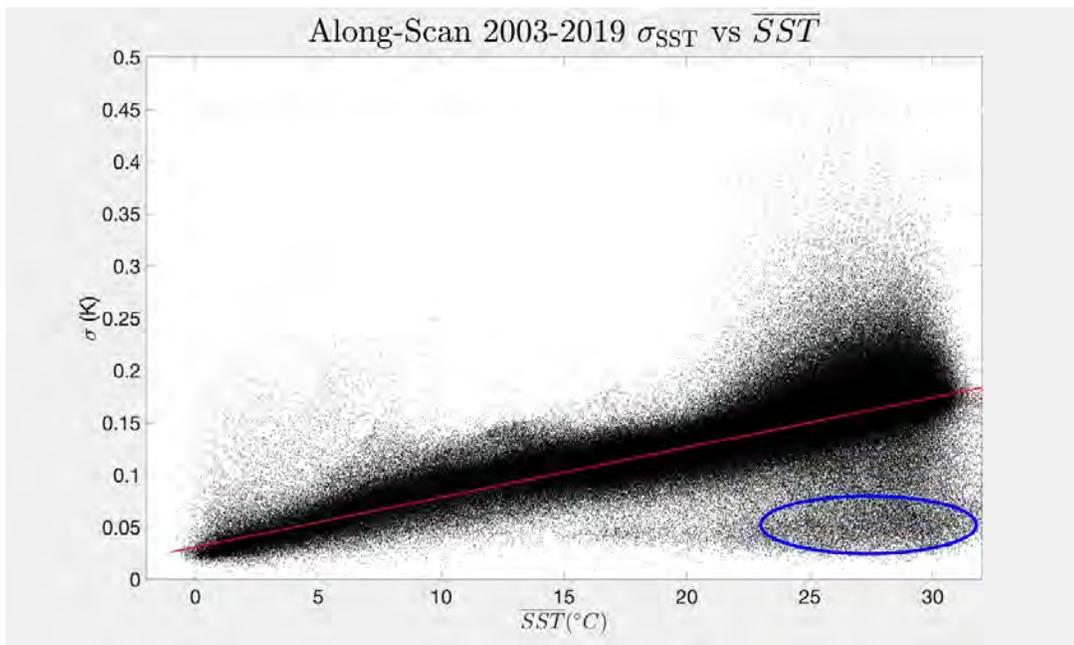


Figure 7. Standard deviation of 128x128 pixel cutouts from the along-scan variograms binned to 200km x 200km x 5day Red line fit to the data. Blue ellipse anomalous region off

A similar fit to the along-track data is given by: $\sigma = 0.0054 \times \overline{SST} + 0.038$

This behavior results in a substantial difference in the precision of retrievals at low temperatures (generally higher latitudes), values around 0.03 K for SST~0°, compared with that at higher temperatures (generally lower latitudes), values of around 0.19 K for SST~30°.

In addition to the temperature dependence of σ , there are a significant number of anomalously low values of σ at relatively high temperatures – the blue ellipse in Fig. 7. Plotted in Fig. 9 is a 2D histogram of the geographic distribution of the anomalous values for the rectangle shown in Fig. 8 ($\sigma < 0.085\text{K}$ & $SST > 21.5^\circ\text{C}$). The tight regional distribution of these values is puzzling; remember we are describing characteristics of the precision here and, not only does it depend on SST but there are a subset of the pixels with values that are atypical for the associated SSTs.

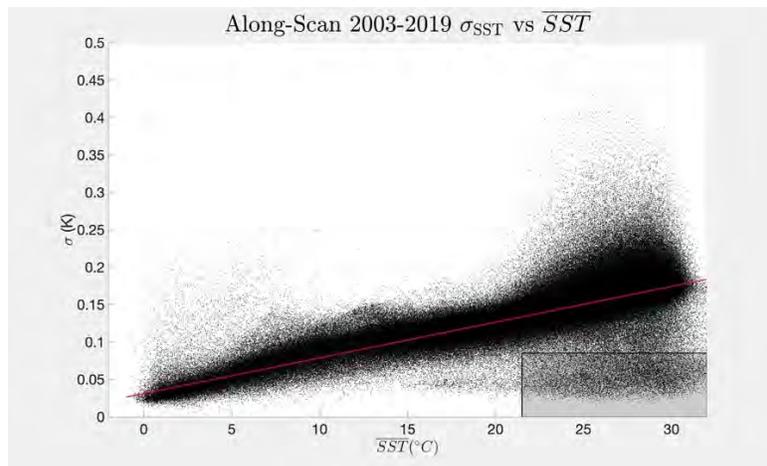


Figure 8. Same as Fig. 7 with the addition of the (σ, SST) region shown in gray for the values histogrammed in Fig. 9.

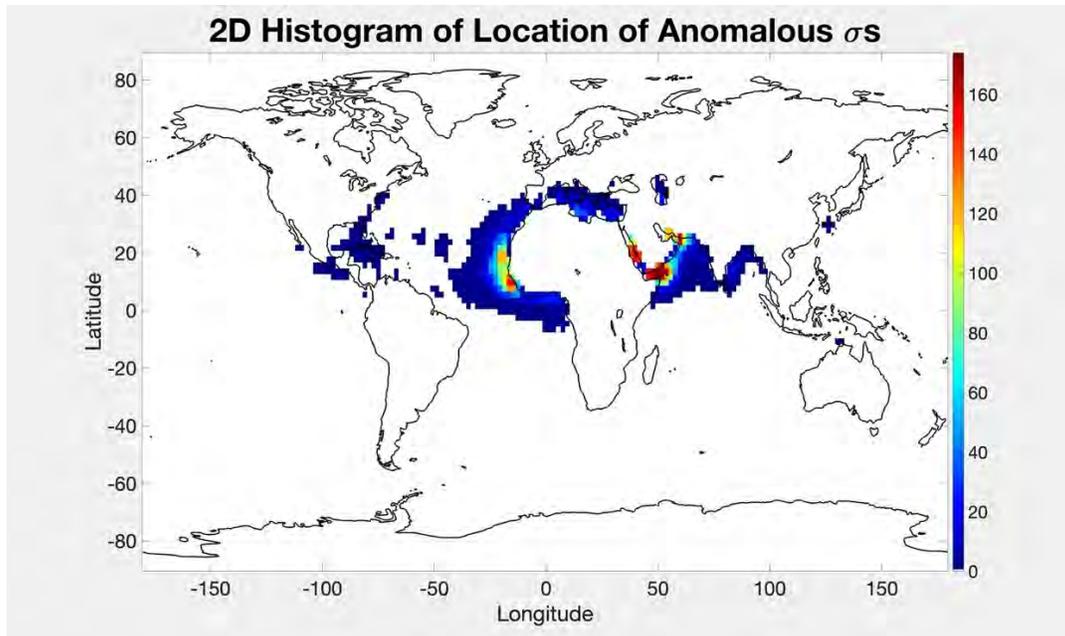


Figure 9. Geographic distribution of anomalous σ s.

The reason for emphasizing this case is two fold. First, noise at the levels similar to those found here can have a significant impact on the retrieval of typical ocean gradients. To demonstrate this Wu et al. (2017) simulated 10,000 3x3 pixel squares for a given gradient in x, added Gaussian white noise to each of the elements, applied the 3x3 Sobel gradient operator in x and y to these squares and then determined the mean gradient and the standard deviation of the gradient. The results of this analysis for gradients ranging from 0.001 K to 0.3 K, values typical in the ocean, and for levels of instrument noise ranging from 0.001 K to 0.3 K are shown in Fig. 10. It is clear that gradients over the entire range simulated will be impacted by the noise in the MODIS SST retrievals with the impact varying quite dramatically from one region in the ocean to another.

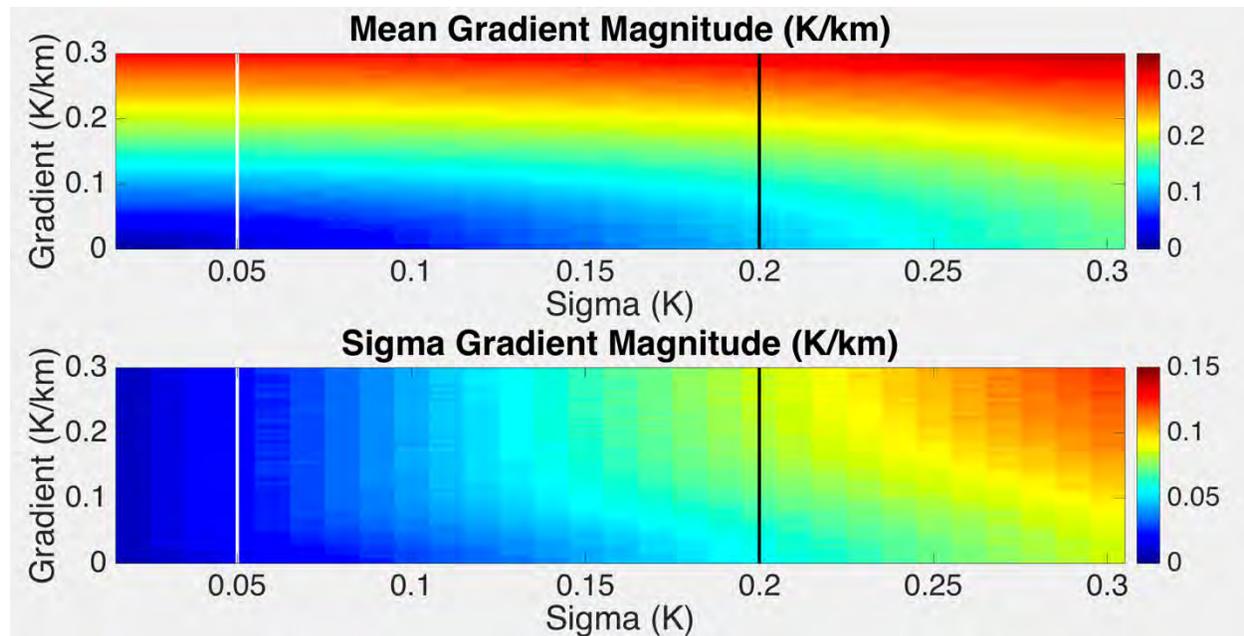


Figure 10. The gradient magnitude and its standard deviation obtained from simulated fields with the indicated gradient in the x-direction – no imposed gradient in the y-direction – with Gaussian white noise of magnitude sigma added to the fields. The vertical white and black lines indicate the approximate range of σ .

Second, both the temperature dependence of σ and more importantly, its spatial dependence suggests that estimates of instrument/retrieval noise from basics, the *error propagation* approach, outlined under Task 3, will benefit from estimates of this noise obtained directly from the data, the 2nd method outlined in Task 3. Specifically, the two approaches are complimentary in that they approach the problem from different perspectives hence results from one inform the other.

The final example we discuss relates to the pixel flags used to mask MODIS L2 SST fields obtained from OBPg. The quality mask employed by most users is derived from a number of flags several of which flag pixels as bad if they fail a variance test of the brightness temperatures. This results in a significant fraction of good pixels being flagged as bad in the quality fields. A post-processing algorithm has been developed to correct this problem – to unmask improperly masked pixels. Figs. 11-13 show an SST field for 19 June 2010 of the waters off the northeast coast of the United States. The first figure shows this field with the original cloud mask. The second figure is the same image following application of the ‘unmasking’ algorithm and the third figure highlights in black the corrected pixels.

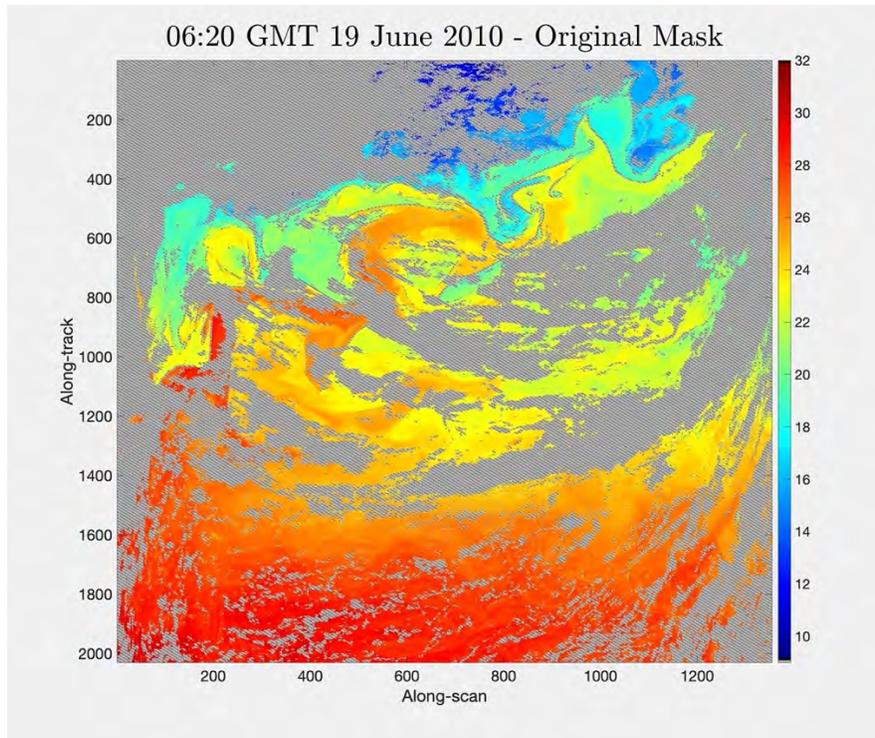


Figure 11. SST field retrieved from MODIS Aqua Brightness temperature by the Ocean Biology Processing Group of NASA's Goddard Space Flight Center. Original mask shown as gray pixels with slanted black lines.

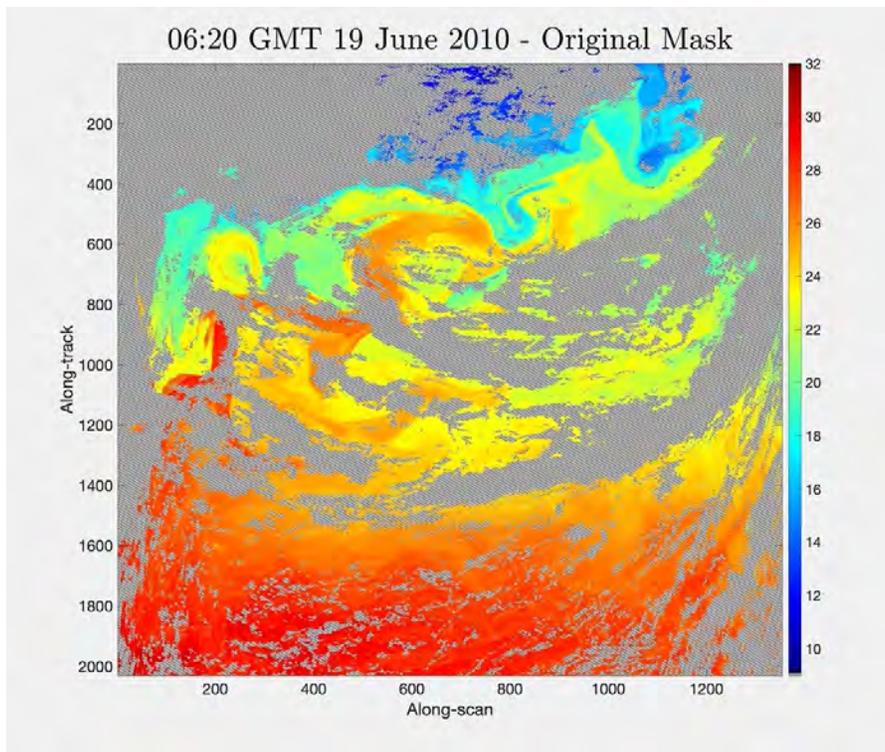


Figure 12. As in Fig. 11 except most of the improperly masked pixels have been 'unmasked'.

06:20 GMT 19 June 2010 - Showing Recovered Pixels

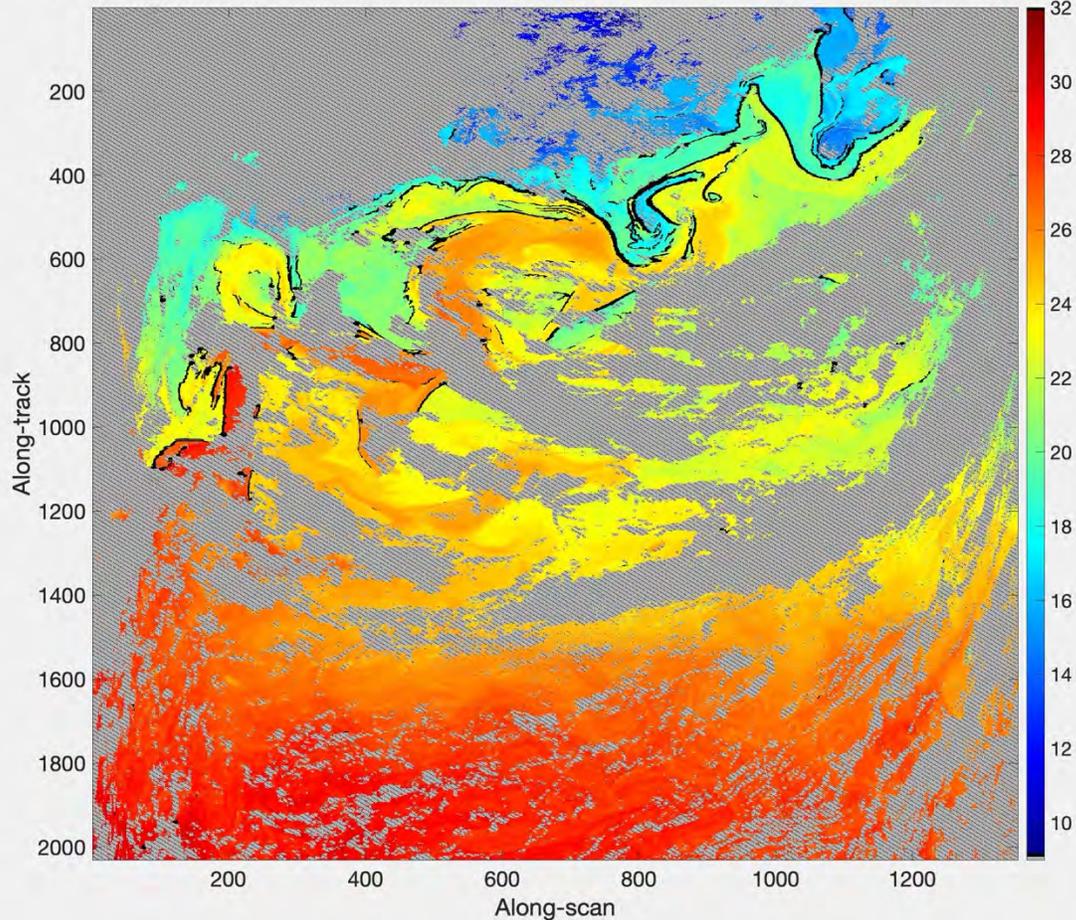


Figure 13. As in Fig. 11 except with corrected pixels shown in black.

Needless to say, this problem is devastating for those interested in studying strong fronts; cloud masking remains one of the most significant problems for those interested in feature analysis.

The examples presented above are not unique; SST products are ripe with the suppression or introduction of ‘features’ in the ocean. This presents serious obstacles for those studying processes in the upper ocean on scales of a few kilometers to scales of tens of kilometers. The objective of this task team is to identify the effects leading to these differences and then to quantify them.

In the following we propose using nomenclature introduced in section 3.1 of Mittaz et al. (2019, doi.org/10.1088/1681-7575/ab1705), specifically, the nomenclature related to the terms ‘error’, ‘uncertainty’ and ‘effects’ as well as the different classes of errors: ‘independent’, ‘structured’ and ‘common’.

Task 1: Determination of Feature Classes

Our first task is to identify the problems we encounter, which we group into three categories: those associated with uncertainties in the sensors and their calibrations, i.e., the uncertainty of pixel values in the L1 products used to obtain the L2 or higher fields, uncertainties introduced by the retrieval process – the algorithms processing the L1 fields into the L2 fields – and uncertainties introduced in resampling and averaging involved in processing from L2 to L3 fields. Diurnal warming can result in significantly different structure of SST fields obtained from satellites overflying a region within a few hours of each other. We do not address such differences here in that these differences are geophysical; they are not due to errors in the sensor or in the retrievals. Following is a classification of the sorts of issues, which contribute to the uncertainty of feature detection:

1. Uncertainties in L1 Fields – Errors in these fields result from a number of sources: detector noise, quantification, calibration,... These errors occur at several scales:
 - a. Pixel-to-pixel – Example shown in Figs. 2 & 3.
 - b. Across one or more scan lines – Also give rise to example shown in Figs. 2 & 3.
 - c. Cloud screening errors – This may also contribute to the 1st sample shown above.
 - d. By location in the orbit – Both this and the next issues are not likely to affect retrievals at the meso- or submesoscale so we will ignore them.
 - e. Over time
 - f. ...
2. Uncertainties Introduced by the Retrieval Process
 - a. Atmospheric properties – The example shown in Fig. 4 – the difference in the fields, which are very nearly synoptic – likely results from atmospheric contamination of the retrieved SST fields.
 - b. ...
3. Uncertainties Introduced by the resampling/averaging process.
 - a. Resampling errors – Close inspection of the center panel of Fig. 3, indicates horizontal stripes separated by approximately 18 pixels. This is an example of an error introduced as part of the resampling process. Fig. 5 shows another example of a resampling problem.
 - b. Averaging errors.
 - c. ...

Task 2: Effects Contributing to the Uncertainty of the Given Feature Classes

In this section we take a first crack at enumerating the ‘effects’ contributing to the classes of problems highlighted in the previous section.

1. Uncertainties in L1 Fields
 - a. Pixel-to-pixel
 - i. instrument errors – noise in the detectors.
 - ii. calibration errors – these problems depend on the sensor and how they are calibrated in-flight. For example, AVHRR calibrates by scan with some averaging. This results in a virtually no contribution to the pixel-to-pixel

uncertainty in the along-scan direction compared with the contribution in the along-track direction. Other sensors, VIIRS for example, make use of 16 detectors per scan, each with their own calibration. This gives rise to pixel-to-pixel errors.

- iii. ...
 - b. Across one or more scan lines
 - i. calibration errors – VIIRS for example, make use of 16 detectors per scan, each with their own calibration. This gives rise to a peak in uncertainty every 16 scan-lines.
 - ii. ...
 - c. Cloud screening errors
 - i. This will generally result in pixel-to-pixel errors but can also give rise to larger scale problems.
2. Uncertainties Introduced by the Retrieval Process
- a. Atmospheric properties –
 - b. ...
3. Uncertainties Introduced by the resampling/averaging process.
- a. Resampling errors –
 - b. Averaging errors –
 - c. ...

Task 3: Putting It All Together

There are two broad categories of approaches, which might be used, to determine the uncertainty in satellite-derived SST fields of importance in the study of mesoscale and smaller features; one based on the propagation of known errors through the processing system and, the other, based on the analysis of the retrieved SST fields. We address each of these below.

Error Propagation

As noted in the introduction, we are interested not in absolute SST but in spatial SST features, such as Sobel gradients capturing fronts, we want to know how different sources of satellite errors play into those derived features. For example, while we might expect that increased sensor noise increases uncertainty in SST gradient, it would be nice to know by how much so that the requirement to measure gradients can then inform requirements on noise in images. Also, as previously noted, not all error sources have the characteristic of error-independence we tend to associate with “noise”, the most obvious being that the contribution to error from calibration cycles will be correlated along scan lines, and less correlated (or even not correlated) along the track (between different scans). So, methods need to deal with variously correlated errors. But there is nothing mysterious – all we have to do is apply propagation of uncertainty systematically. The uncertainty maths is always the same: the work to do relates to defining the image features we are wanting to explore, and in understanding the error covariance of the SST image.

Uncertainty propagation to image feature quantities

Here is the uncertainty maths which always applies. Consider a “feature operator”, T . This is a function that operates on a subset, \mathbf{x} of SST image pixels (e.g., a $n \times n$ box centred on a target image pixel). \mathbf{x} simply lists the SSTs in the pixels operated on. The calculated feature for a particular subset is f . Let’s assume that the error covariance properties of \mathbf{x} are understood, as they can be from applying FIDUCEO methods (doi.org/10.1088/1681-7575/ab1705). The image error covariance properties are summarised by uncertainties and error correlations, which I choose to represent as follows:

- Uncertainty matrix: the diagonal values contain the total standard uncertainty (1 sigma spread of errors from all sources) in each image pixel:

$$\mathbf{U} = \begin{bmatrix} u_1 & 0 & \dots \\ 0 & u_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

- Correlation matrix: the Pearson correlation coefficient of errors between each pair of pixels:

$$\mathbf{R} = \begin{bmatrix} 1 & r_{1,2} & r_{1,3} & \dots \\ r_{1,2} & 1 & r_{2,3} & \dots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

From these relatively intuitive properties of the errors in the image, we can calculate the error covariance matrix

$$\mathbf{S} = \mathbf{URU}$$

\mathbf{S} is what we need for propagating uncertainty to an image feature, along with the sensitivity of the image feature to the value of each pixel. Assuming the feature is a single quantity, the sensitivity is a column vector \mathbf{c} :

$$\mathbf{c} = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \right]^T$$

(Note that there is no assumption here that f is analytically differentiable: if it is not, it can be calculated by perturbation. It is usually assumed that the feature is continuous, but even if discontinuous; e.g., there is a threshold-dependent aspect) the “typical” sensitivity could be found by perturbation over many instances (some of which cross the threshold and some of which don’t.) The uncertainty in the image feature (first order estimate) is always:

$$u_f = \sqrt{\mathbf{c}^T \mathbf{S} \mathbf{c}}$$

We may be interested in two or more related image features. An example would be the

gradient in the north and east directions. We could calculate the uncertainty in each image feature separately using the above, but it is useful to do something a little more sophisticated. Define a sensitivity matrix (dual-feature example):

$$\mathbf{C} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots \end{bmatrix}$$

This enables the calculation of the error covariance matrix of the two features:

$$\mathbf{S}_f = \mathbf{CSC}^T$$

\mathbf{S}_f is most easily understood by decomposing it: $\mathbf{S}_f = \mathbf{U}_f \mathbf{R}_f \mathbf{U}_f$ where \mathbf{U} is diagonal. The values of the on the diagonal of \mathbf{U} are the uncertainty for each image feature quantity. The off-diagonals of the correlation matrix give the degree to which errors between feature quantities are correlated (e.g., answering the question: if the error in the northward gradient is positive do we get a positive, zero or negative tendency for the error in the eastward gradient?)

Example: gradient operator

Consider calculating the gradients (in units of image-unit per pixel) in an image using a typical Sobel convolution on an image \mathbf{A} :

$$f_x = \frac{1}{8} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * \mathbf{A}$$

$$f_y = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

For a given centre pixel in the image, x_i , this can be written

$$\mathbf{f} = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \mathbf{T}\mathbf{x} = \frac{1}{8} \begin{bmatrix} 1 & 0 & -1 & 2 & 0 & -2 & 1 & 0 & -1 \\ 1 & 2 & 1 & 0 & 0 & 0 & -1 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_{i-1,j-1} \\ x_{i-1,j} \\ x_{i-1,j+1} \\ x_{i,j-1} \\ x_{i,j} \\ x_{i,j+1} \\ x_{i-1,j-1} \\ x_{i+1,j} \\ x_{i+1,j+1} \end{bmatrix}$$

(It is not necessary to include $x_{i,j}$ above, but it is written here to show the full 3 by 3 box.) For linear filters, when formulated like this, the transformation matrix \mathbf{T} is the sensitivity matrix. Thus, the error covariance of the gradients is (solution in general):

$$\mathbf{S}_f = \mathbf{T}\mathbf{S}\mathbf{T}^T$$

In the simplest case where pixels are subject only to independent constant noise of size u , this equals (unrealistic special case):

$$\mathbf{S}_f = u^2 \mathbf{T}\mathbf{T}^T = u^2 \frac{1}{64} \begin{bmatrix} 12 & 0 \\ 0 & 12 \end{bmatrix} = \begin{bmatrix} \left(\frac{\sqrt{3}}{4}u\right)^2 & 0 \\ 0 & \left(\frac{\sqrt{3}}{4}u\right)^2 \end{bmatrix}$$

and so we can see that the two components have errors that are independent (off-diagonals are zero—although only for independent pixel noise), and that the uncertainty in the component gradients is the image noise attenuated by a factor of $\sqrt{3}/4$.

A possible approach for the task team

Calculating the uncertainty for any chosen image feature always follows the above approach, and in that sense is not a difficult question. So, for the task team, progressing on image-feature uncertainty involves focussing on two tasks:

- Identifying for which image features it is useful to quantify uncertainty. By “identifying” I mean, ultimately, “writing down the equation for the image feature in terms of an operator to be applied to a list of image pixels \mathbf{x} ”
- Quantifying the image error covariance \mathbf{S} for a given SST image, so that the uncertainty/covariance in the image property can be calculated.

The definition of the useful image property is thus conceptually separated from L1 and L2-retrieval issues, since these are wholly described by determining \mathbf{S} . This separation should help to focus discussions into two streams: what do we want to know about the errors in SST image features? and how do we calculate what we want to know from our understanding of the instrument and retrieval?

The second question, determining \mathbf{S} , is by far the harder question, (i) because it is intrinsically a hard question, and (ii) because it changes every time we address a different SST product. But I think the FIDUCEO reports have set out a good framework for this, by addressing L1 error characterisation and the propagation to the L2 retrieval (see below).

Here are some obvious thoughts about which SST image features it may be relevant to calculate uncertainty for:

- Gradient components, using various gradient operators perhaps
- Total gradient magnitude
- Direction of SST gradient
- Pixel-to-pixel SST difference at different scales/orientation of pixel separation (is this “spatial precision”?)

Further thought: we should be clear when inventing names for image-feature uncertainties. “Uncertainty in SST difference between adjacent pixels in the across track direction” is unambiguous but rather long—if the task team wants to coin shorter names we should keep a shared table of “short name” and “long name descriptive of quantity” to avoid confusion.

What is available from FIDUCEO?

From FIDUCEO we have an uncertainty model for AVHRR L1 brightness temperatures that includes the main error sources and describes their across-track and along-track error covariances. And we know how to propagate those errors to an AVHRR L2 SST image (by whatever SST retrieval method) and how to inject the additional uncertainty arising from the retrieval process. So, for the case of AVHRR L2 imagery, we can determine S for specific cases in detail. More broadly, the FIDUCEO methods to analyse other instruments and their retrievals are generic, and are documented in FIDUCEO reports and papers.

Estimates of the Uncertainties Based on the SST Fields Themselves

An alternate approach to determining the uncertainty of importance to feature detection and analysis is that making use the variability in the SST fields themselves. This is exemplified by the work of Wu et al. (2017, 2019). In both of these approaches the uncertainty of the SST difference between adjacent pixels is determined from the statistical difference in SST values as a function of day/night, along-scan/across-scan and year for each of the AVHRRs carried on NOAA polar-orbiting satellites as well as the same for SST in the summer of 2015 obtained from VIIRS carried on Suomi-NPP. The approaches outlined in these studies were based on the Fourier Transform of the images, i.e., they were performed in wavenumber space, and on the variogram, performed in cartesian space. These are only presented as examples here; other approaches may be investigated as well.

Another approach based on the data would be the use of the SLSTR Tandem Phase (Mittaz et al. 2020) to determine the pixel-to-pixel uncertainty of the SLSTR interments and then use the tandem mission to test other approaches based on the data such as that of Wu et al. (2017).

References

- Mittaz, J., C.J. Merchant and E.R. Woolliams, 2019. *Applying principles of metrology to historical Earth observations from satellites*, **Metrologia**, 56, 032002
- Wu, F., P. Cornillon, B. Boussidi and L. Guan, 2017, *Determining the Pixel-to-Pixel Uncertainty in Satellite-Derived SST Fields*, **Remote Sens.**, 9(9), 877; <https://doi.org/10.3390/rs9090877>.

Wu, F., P. Cornillon, L. Guan and K. Kilpatrick, 2019, *Long-Term Variations in the Pixel-to-Pixel Variability of NOAA AVHRR SST Fields from 1982 to 2015*, **Remote Sens.**, 11(7), 844; <https://doi.org/10.3390/rs11070844>.