University of Reading

Thermal remote sensing of sea surface temperature

Book or Report Section

Accepted Version

Merchant, C. J. ORCID: https://orcid.org/0000-0003-4687-9850 (2013) Thermal remote sensing of sea surface temperature. In: Kuenzer, C. and Dech, S. (eds.) Thermal Infrared Remote Sensing: Sensors, Methods, Applications. Remote Sensing and Digital Image Processing, 17. Springer Netherlands, Dordrecht, pp. 287-313. ISBN 9789400766389 doi: 10.1007/978-94-007-6639-6 Available at https://centaur.reading.ac.uk/38526/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://link.springer.com/chapter/10.1007/978-94-007-6639-6_15
To link to this article DOI: http://dx.doi.org/10.1007/978-94-007-6639-6

Publisher: Springer Netherlands

Publisher statement: The original publication is available at www.springerlink.com.

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

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Thermal Remote Sensing of Sea Surface Temperature

Christopher J. Merchant

School of GeoSciences, The University of Edinburgh, Edinburgh, UK

Abstract Sea surface temperature has been an important application of remote sensing from space for three decades. This chapter first describes well-established methods that have delivered valuable routine observations of sea surface temperature for meteorology and oceanography. Increasingly demanding requirements, often related to climate science, have highlighted some limitations of these approaches. Practitioners have had to revisit techniques of estimation, of characterising uncertainty, and of validating observations—and even to reconsider the meaning(s) of "sea surface temperature". The current understanding of these issues is reviewed, drawing attention to ongoing questions. Lastly, the prospect for thermal remote sensing of sea surface temperature over coming years is discussed.

How does sea surface temperature vary?

Thermal remote sensing is a powerful technique to obtain global, frequent observations sea surface temperature (SST). Surface temperature across the oceans varies with time (e.g., Robinson, 2004), responding, for example, to the daily cycle in heating by the Sun (e.g., Fairall et al., 1996), to the passage of the seasons, and to changes in upwelling or vertical mixing driven by the wind blowing across the sea surface (e.g., Munk, 1950). Surface waters are constantly moving: in ocean currents and eddies; and, near coasts, with tides and river outflows. Surface water advection changes SST over time at a given location.

Scales of a kilometre and upwards are globally observed with radiometers from space (e.g., Donlon et al., 2010). Viewing the ther-

mal structure of the ocean surface at such scales, one may observe relatively sharp boundaries in SST (Figure 1). These "fronts" are signatures in SST of the convergence of surface water masses, convergence that can arise from a number of processes of oceanographic interest. More generally, thermal remote sensing can reveal any phenomenon that measurably alters the radiometric SST, provided that the surface temperature signature occurs on a length scale longer than the sensor's spatial resolution and persists for long enough relative to the time-sampling properties of the observing system. Note that the relevant sampling rate is not the rate of acquisition of images, but rather the rate at which a cloud-free observation at a given location is typically obtained. Examples of oceanographic phenomena and their SST magnitude and spatiotemporal scales are given in Table 1.

Table 1. Selected ocean phenomena and the magnitude and scales of their SST signature. (Extracted and adapted from Robinson (2004), which provides a more complete listing.)

Phenomenon	Magnitude / K	Length scale / km	Time scale
Climatological variation across oceans	35	10^4	
El Nino and interannual variability	0.5 to 5	500 – 5000	Months to years
Tropical instability waves	0.5 to 5	200 – 2000	Months to years
Meanders and eddies on major fronts and boundary currents	1 to 8	5 to 2000	Weeks to months
Diurnal warming cycle	0.1 to 5	5 to 1000	Hours
Coral bleaching events	0.3 to 3	20 to 200	Days
Coastal wind induced phenomena	0.2 to 2	1 to 100	Hours

Basis in physics of sea surface temperature remote sensing

All remote sensing depends on a remotely observable signal that reflects variations in the phenomenon of interest. To observe SST from space, the radiance at the top of atmosphere must change in response to changes in surface temperature (e.g., Deschamps and Phulpin, 1980). Figure 2 shows the spectral sensitivity of the top-of-atmosphere brightness temperature (BT, y) to SST (x) variations—

i.e., it is a plot of the variation with wavelength, λ, of $\partial y_\lambda/\partial x$. This has been simulated using the physics of thermal radiative transfer encapsulated in a radiative transfer model. Here, the radiative transfer (RT) model makes calculations of the emission, scattering, and absorption of thermal radiation at the surface and through the full vertical profile of the atmosphere, wavelength-by-wavelength. The spectral BT sensitivity has been calculated for an example of mid-latitude and of tropical conditions, in both cases for a cloud-free nadir view of the ocean. A change in temperature of 1 K of an ideal radiating surface (a black body) would change the spectral BT observed by a radiometer (under a vacuum) by 1 K—that is, the sensitivity would be $\partial y_\lambda/\partial x = 1$. A change in sea surface temperature leads to a change in BT observed at the top-of-atmosphere that is smaller, because of a number of factors listed and described in Table 2. The key point is that sensitivity is high at wavelengths where the atmosphere is relatively transparent to the passage of electromagnetic radiation.

Figure 2 presents two parts of the spectrum useful for SST remote sensing: the atmospheric windows in the near infra-red (around 4 μm), and the mid infra-red (between 8 and 13 μm, albeit interrupted by an ozone absorption feature around 9.7 μm). Although described as windows, there is variability in sensitivity with wavelength, which would be even more striking if plotted with finer spectral resolution. Within the windows, individual molecular absorption features reduce spectral sensitivity close to zero over narrow intervals of wavelength, while there are also some intervals a few cm^{-1} without such lines. An example of the latter is the micro-window at 2616 cm^{-1} (3.823 μm). In such micro-windows, the clear-sky attenuation of temperature by the atmosphere can be just a few tenths of degree kelvin (except when the presence of atmospheric aerosols decreases the transmittance). Most sensors from which SSTs are derived, however, have channels of width of order 100 cm^{-1}. This allows higher spatial resolution with reasonable noise characteristics, but requires that atmospheric effects must be accounted for when inferring surface temperature (e.g., McMillin et al., 1984).

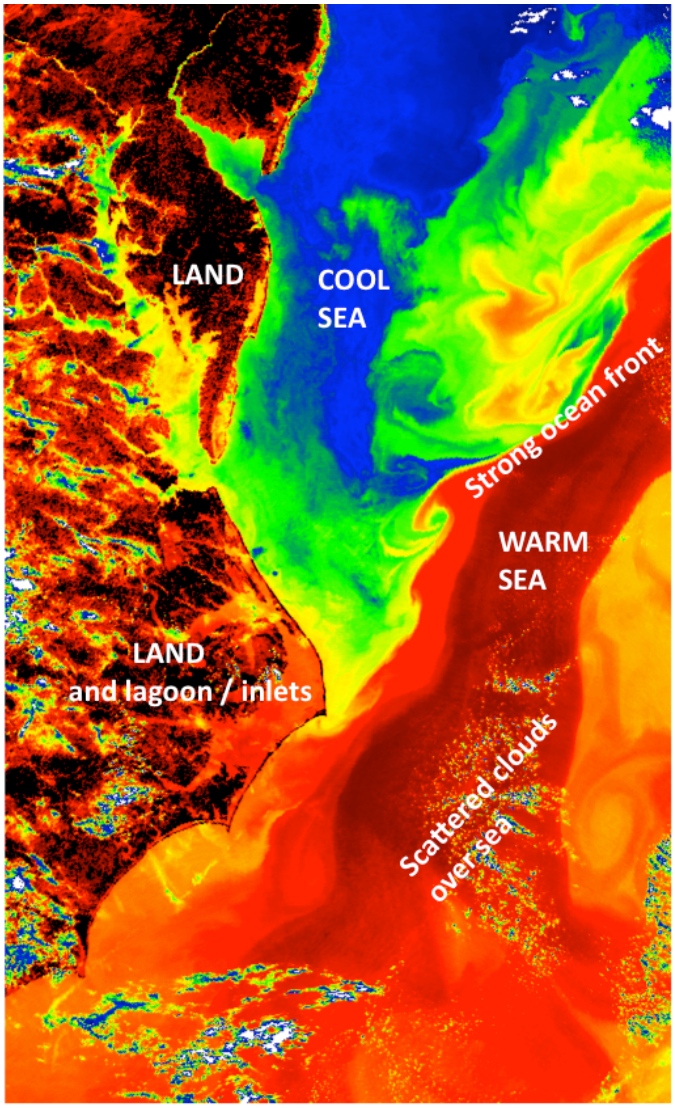

Figure 1. Image at thermal window wavelength of 11 μm of US Eastern seaboard including Cape Hatteras and Pimlico Sound, obtained by the second Along Track Scanning Radiometer. The width of the image is about 512 km and the pixel resolution is 1 km. Ocean features are reasonably well resolved, with smooth contiguous variations in surface temperature reflected in the image brightness temperature, except where there are scattered, cooler clouds that are often not fully resolved. Land is also more heterogeneous, and, this being a day time image, warmer than the sea in many areas. Image obtained from URL1 and adapted by the author.

Table 2. Factors that affect the top-of-atmosphere clear-sky spectral brightness temperature and its sensitivity to surface temperature variations. Summarized from Embury et al. (2012a).

Factor	Nature of influence
Sea surface emissivity	If emissivity is less than 1, emitted radiance is correspondingly less than the ideal Planck (black body) radiance. Sea water emissivity is generally high (0.96 to 0.99) for near-nadir observations at wavelengths relevant to thermal remote sensing. For flat seawater, spectral emissivity depends on temperature, salinity and the angle of view (emissivity reduces markedly at angles beyond about 55°). Under wind-roughened conditions the water is not flat, which modifies the effective emissivity as a function of wind speed.
Sea surface reflectivity	Downward atmospheric and (for near infra-red) solar radiance can be reflected at the surface. Reflectivity depends on the same factors as emissivity, and increases as emissivity decreases.
Radiatively active gases	The surface-leaving radiance is absorbed by gases in the atmosphere, to a greater or lesser degree according to wavelength. In decreasing order of impact on SST remote sensing, the most relevant gases are: water vapour, carbon dioxide, dinitrogen oxide, methane, CFC-12, nitrogen, CFC-11 and nitric acid. These gases also emit radiation by virtue of their temperature. For the most part, this does not wholly offset the absorption, since the atmosphere is mostly colder than the underlying surface. Nevertheless, the effect of the atmosphere is to introduce a source that is not directly dependent on the surface temperature, and therefore to reduce the BT sensitivity. At wavelengths where absorption is very efficient, the BT becomes independent of the surface temperature. Of the significant radiatively active gases, water vapour is by far the most variable.
Aerosols	Particles in the atmosphere absorb and emit radiance, as with gases, and may also significantly scatter radiance (in to or out of the view of a satellite). The radiative properties and concentrations of aerosols are much more variable and much less understood than for gases. Aerosol impacts of BTs can range from negligible (in clean air), to highly significant (e.g., dust storms).
Solar radiance	Solar radiance can be reflected by the surface and scattered by the atmosphere into the view of a satellite. This is usually very significant for day time observations at near infra-red wavelengths, and can be marginally significant at longer thermal wavelengths when there is strong specular (mirror-like) reflection (known as sun glint).

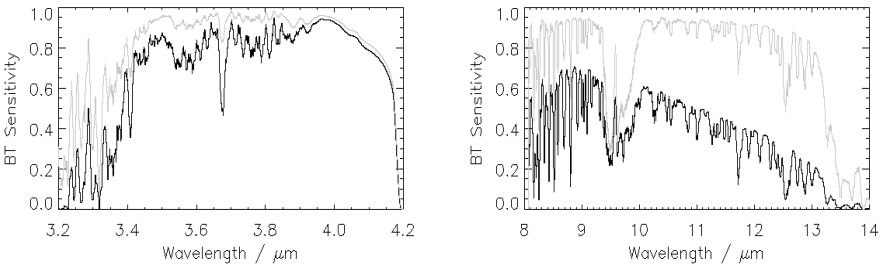

Figure 2. Spectral brightness temperature (BT) sensitivity (the response at the top of the atmosphere per unit change in surface temperature, in units of K K^{-1}). Grey lines: for a mid-latitude case with low total column water vapour (TCWV). Black lines: for an equatorial case, high TCWV. Left panel: a near-infrared window used for SST remote sensing (usually for night-time scenes only), presented at a spectral resolution of 10 cm^{-1}. Right panel: the mid-infrared window, spectral resolution of 3 cm^{-1}.

We can expect more accurate, less noisy estimates of SST when using observations with higher sensitivity, other factors being equal. This is because the SST signal is then greater in proportion to instrumental noise and signals associated with variations in the atmospheric state. Sensitivity to SST is greater in mid-latitude conditions than tropical conditions largely because there is less absorption related to water vapour in the atmosphere (Merchant et al., 2009). The near-infrared window maintains relatively high sensitivity even under tropical conditions, and is particularly useful for observing equatorial SSTs. However, this window is usually only used for night-time scenes, because of the complication of significant solar irradiance at these wavelengths during the day. The sensitivity across the mid-infrared window is highly responsive to the total amount of water vapour. Thus, when relying on the mid-infrared window for daytime SST, we can expect larger uncertainties for regions of high total column water vapour (TCWV).

The physics of atmospheric radiative transfer for these wavelengths is quantified with great precision in the spectroscopic databases exploited by line-by-line RT models (Rothman, 2010). Moreover, the sea surface is relatively simple and homogeneous. The thermal emission and reflection of the surface can also be well simulated. In the absence of significant aerosol, clear-sky RT modelling relevant to typical SST sensors gives BTs that seem to be physically realistic to of order 0.1 K. This is comparable to the calibration uncertainty and noise for “good” SST sensors. Therefore, RT-based

approaches to SST estimation are feasible and are currently used in practice, in addition to empirical approaches.

Sea Surface Temperature Retrieval

Simple empirical estimators

The process of estimating SST from a number of clear-sky brightness temperature observations is usually termed SST retrieval. Most retrieval methods have been based on defining coefficients for a weighted combination of BTs. This is a simple, computationally efficient approach. At least two BTs are required (Anding and Kauth, 1970), since it is necessary to infer both the SST (explicitly) and the impact of the atmosphere on BTs (implicitly). The difference SST minus BT is called the atmospheric correction, the idea being that this is the temperature that must be added to the BT to correct the net attenuating effect of the atmosphere. Despite the term, some of the difference is due to non-ideal emissivity of the sea surface.

The minimum of two BT observations required to retrieve SST must be at wavelengths that (i) have adequately high sensitivity to SST and (ii) are differentially absorbed by atmospheric water vapour. Of the relevant absorbing gases in the atmosphere, only water vapour is extremely variable, with total column water vapour (TCWV) varying from almost zero up to ~60 kg m^{-2}. The atmospheric correction generally increases with increasing TCWV, and does so more rapidly for wavelengths that are more effectively absorbed by water vapour. Consequently, the atmospheric correction for either of the BTs is approximately linearly related to the difference between the BTs:

$$x - y_1 \propto y_1 - y_2 \tag{1}$$

Figure 3 illustrates the degree to which proportionality holds. The SST can thus be estimated by an expression:

$$\hat{x} = a_0 + a_1 y_1 + a_2 (y_1 - y_2) \quad (2)$$

where a_0, a_1 and a_2 are retrieval coefficients.

Wavelengths between about 10 and 13 μm have high sensitivity that varies progressively across that window (Figure 2), reflecting differential water vapour absorption. This is the principal window used by SST sensors, and is usually split between two channels centred around 11 and 12 μm respectively. Eq. 2 in this case describes a split window retrieval. This equation and variants of it have been widely used to obtain SST from meteorological sensors (e.g., McClain et al., 1985; Walton et al., 1998). The coefficients need to be specified for each sensor, since the spectral responses of the nominal 11 and 12 μm channels are inevitably somewhat different between sensors. The split window equation is appealing because it is physically intuitive (Barton, 1995): the SST is the 11 μm BT scaled up a little to compensate for non-ideal emissivity (a_1 is usually a little more than 1.0), with an offset, a_0, added (interpreted as compensating for the absorption of non-varying trace gases), plus a term that accounts for the highly variable water vapour absorption (whose impact is proportional to the difference between the BTs).

The values of the coefficients have usually been defined empirically, by regression between BTs and matched *in situ* SST observations. This is discussed further in a later section. Here, we review the ways in which the simple split window equation has been elaborated.

The first elaboration is with respect to satellite view angle. The total column of water vapour encountered by radiance passing through the atmosphere on a slant path exceeds the TCWV by an approximate factor of approximately $\sec(\theta)$, where θ is the satellite zenith angle (angle of the slant path to the vertical at the surface). The degree to which sea surface emissivity is less than 1 increases markedly with angles beyond about 55°. A common approach has been to fit the combined BT impact of these effects by having some coefficients depend on $S = \sec(\theta) - 1$, such that $a_0 \rightarrow a_0 + b_0 S$ etc. (e.g., Walton et al., 1998).

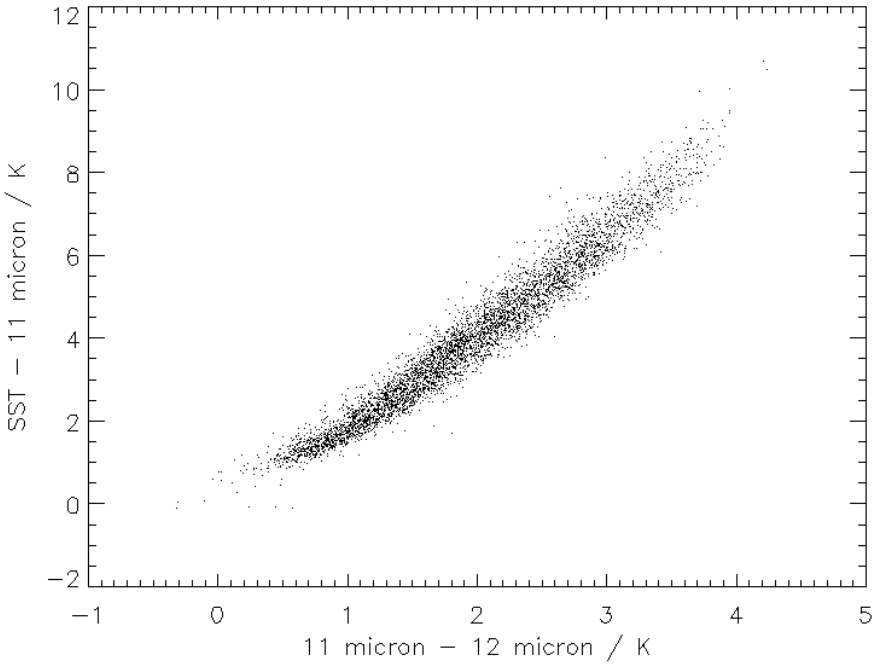

Figure 3. Nadir atmospheric correction for 11 μm brightness temperatures against difference in brightness temperature between 11 and 12 μm channels, for a typical split-window sensor. Each point represents an observation made at a particular location and time from a global sample. The distribution is close to linear for differences between the 11 and 12 μm brightness temperatures greater than about 1 K. The scatter for a given brightness temperature difference arises from factors such as variable vertical distribution of water vapour, troposphere-sea temperature difference, etc. These factors have systematic geographical variations, which tends to create coherent geographical biases in any SST retrieval based on fitting a function to such a distribution.

The second elaboration is adaptation to use additional channels. In addition to the split window channels of 11 and 12 μm, BTs observed in the near-infrared window between about 3.6 to 4 μm are useful for SST estimation. BTs in this region have mainly been used for night-time scenes, when solar-reflected radiance is absent. There is a high degree of sensitivity in this range (Figure 2) to surface temperature, and in addition the dependence of radiance on temperature is extremely steep for terrestrial temperatures $T \sim 285$ K—around T^{14}. This strong non-linearity reduces the impact of non-unity emissivity, of contaminants in the field of view such as sub-pixel undetected clouds, and of radiometric noise of a given magnitude. Therefore, on a well-designed sensor, a channel centred at 3.7 to 3.9 μm

can be particularly "clean" for SST, which is a colloquial way of expressing that the information content on SST is very high. Compared to the near-infrared window, the high-sensitivity region around 8.7 μm behaves more similarly to the usual 11 and 12 μm channels. Nonetheless, it can be useful if present.

Inclusion of additional BTs has been done using various equations, most of which can be re-expressed in the form:

$$\hat{x} = a_0 + b_0 S + \sum_i (a_i + b_i S)\, y_i \qquad (3)$$

Some investigators impose additional restrictions on the empirical fit of the coefficients in eq. 3 by choosing specific forms of equation (e.g., Li et al., 2001). For example, a three-channel algorithm with form:

$$\hat{x} = a_0 + b_0 S + (a_1 + b_1 S) y_{3.7\mu m} + a_2 \left(y_{11\mu m} - y_{12\mu m} \right) \qquad (4)$$

is eq. 3 with the additional constraints that $a_2 = -a_3$ and $b_2 = b_3 = 0$. Sometimes a physical argument is put forward to justify a specific form of equation. Imposing such additional constraints may seem at odds to adopting an empirical approach to determine coefficients, but if the empirical dataset relating BTs and SSTs is small, imposing such additional constraints may avoid over-fitting. Additional constraints can also modify the sensitivity of the estimator to factors such as atmospheric aerosols (discussed further below).

As well as observations at additional wavelengths, "channels" can be added by having sensors view the sea surface at more than one view angle, i.e., near-nadir and off-nadir (forwards or backwards along the track).

The third elaboration relates to use of non-linear terms. The residuals (retrieved minus *in situ* SSTs) of purely linear estimators such as eq. 3 usually display coherent systematic variations if plotted against latitude, longitude-within-a-latitude-zone, TCWV, BT differences, and so on. These reflect the non-linearity evident in Figure 3, and sensitivity to geographical variations in the broad vertical structure of water vapour and temperature in the atmosphere. A wide range of non-linear estimators have been proposed, for example:

- banding of coefficients by latitude or other regional optimisation (e.g., Minnett et al., 1990)
- banding of coefficients by BT difference (e.g., Kilpatrick et al., 2001)
- banding of coefficients by prior TCWV or retrieved TCWV (e.g., Barton, 2011)
- inclusion of a term that modifies coefficients via a prior SST (e.g., Pichel et al., 2001)
- quadratic dependence on BTs or prior TCWV (e.g., Emery et al., 1994)

These generally offer modest benefit to retrieval accuracy. Simple functional forms do not reflect the underlying origins of the systematic residuals, which are the non-linearity of the physics of RT and geographical variations in atmospheric structure (Merchant et al., 2006).

The final elaboration is to use alternatives regression methods. The usual means of defining retrieval estimators empirically has been ordinary least squares fitting (multiple linear regression). Neural nets in principle seem an attractive way to deal with the non-linear aspects of the retrieval problem, but are yet to demonstrate good success. 'Genetic' algorithm identification seems to converge on a form rather similar to the split window formulation.

Approaches involving radiative transfer modelling

Progress has been made in recent years using RT modelling to improve retrieval accuracy and precision.

The main use of RT has been to define coefficients that look rather like the simple empirical estimators discussed above. Instead of empirical matches between *in situ* SST measurements and satellite observations, RT-based coefficients are derived by regressing simulated BTs to the SSTs used as input to the simulations. The simulations are driven using atmospheric profiles obtained from radiosondes or numerical weather prediction (NWP) systems. The RT-based approach is compared to the empirical approach in Table 3.

Table 3. Contrasting advantages and disadvantages of deriving SST retrieval coefficients by empirical means compared with derivation by radiative transfer simulation. See also Merchant and Le Borgne, 2004.

Aspect	Empirical approach	RT-based approach
Instrument calibration and characterisation	Reduced need to understand instrument characteristics and calibration, since many calibration issues are empirically accounted for in the coefficients.	Need sensor to be well characterised (accurate spectral response functions available) and calibrated (ideally to ~0.1 K). In absence of this, significant effort is required to bias-correct simulated BTs to match observations.
Spatio-temporal sampling (representativity)	Can only match locations where *in situ* measurements are already present. This gives no formal basis for assessing the accuracy of retrieved SSTs in areas/periods with few or no *in situ* measurements. *In situ* coverage has greatly improved since the early 2000s, but high latitude seas remain underrepresented.	The spatio-temporal sampling is in the control of the investigator. Access to an NWP re-analysis provides a consistent atmospheric data set that can be sampled across all epochs and with no gap regions.
Nature of sea surface temperature	Satellite SSTs are regressed to the SST at the depth typical of the *in situ* measurements (tens of cm in the case of drifting buoys). However, BTs are sensitive to skin SSTs. Thus, empirical methods conflate different forms of SST that do not bear a simple relation to each other. An approach which addresses this issue is to restrict the regression cases to those where skin-depth differences are thought to be well understood (e.g., moderate wind stress at night).	The simulation can be done using skin SST. The coefficients are then unambiguously retrieval coefficients for skin-SST, the geophysical variable to which the BTs are sensitive.
Independence from *in situ* data sets	No independence (fully tied to *in situ*).	Independence possible for best characterized instruments. For reasonably well calibrated instruments, RT-based coefficients can be tuned by adjusting only the offset coefficient.
Difficulty of defining algorithm	Requires a statistically sound number of satellite-*in situ* matches (for every required stratum of latitude/TCWV/view angles/etc). Implies no retrieval scheme is available at launch.	Requires a RT simulation capability, sampled NWP profiles, and commensurate computing power. Retrieval scheme can be defined prior to launch.

One strength of approaching SST retrieval with RT simulations is the enhanced ability to diagnose and solve problems. An example is how to adapt SST retrieval to the presence of stratospheric aerosol (Merchant et al., 1999). Occasionally, major volcanic eruptions penetrate the stratosphere and create a haze of sulphuric acid droplets that persists at altitudes of order 20 km for a year or two. This stratospheric aerosol layer has climatic impacts, and also affects remote sensing at visible and infra-red wavelengths. The aerosol absorbs infra-red radiation, and causes BTs to be reduced. Let's assume that the impact per unit aerosol optical depth, τ, on the BTs of a typical three-channel sensor is $\frac{\partial \mathbf{y}}{\partial \tau} = \left(\frac{\partial y_{3.7\mu m}}{\partial \tau} \frac{\partial y_{11\mu m}}{\partial \tau} \frac{\partial y_{12\mu m}}{\partial \tau} \right)^T$. Here, and hereafter, **y** is used for a column vector containing the BTs to be used in a SST retrieval. For a given observation (at a particular view angle etc.), the retrieval equation can be conveniently written also using vector notation as:

$$\hat{x} = a + \mathbf{a}^{\mathrm{T}}\mathbf{y} \tag{5}$$

where, compared to eq. 3, $a = a_0 + b_0 S$; the first element of the coefficient column vector, **a**, is equal to $a_1 + b_1 S$; and so on. The summation in eq. 3 is achieved in eq. 5 by the matrix multiplication of row vector $\mathbf{a}^{\mathrm{T}}$ and the column vector **y**. Use of matrix algebra may at first seem unnecessary, but is a powerful tool for expressing and analysing retrieval algorithms.Written in this form, it is clear that the impact of the stratospheric aerosol on the retrieved SST will be:

$$\delta\hat{x} = \tau \mathbf{a}^T \frac{\partial \mathbf{y}}{\partial \tau} \tag{6}$$

at least over the range for which the BT depression is linear in the optical depth. Using radiative transfer, $\frac{\partial \mathbf{y}}{\partial \tau}$ can be calculated from knowledge of the properties of the sulphuric acid droplets (concentration and size distribution). Knowing this, the different sensitivities to stratospheric aerosol of different retrieval formulations can be understood using eq. 6. Moreover, we readily formulate the requirement for an SST retrieval algorithm to be robust (i.e., insensitive) to stratospheric aerosol. It is:

$$\mathbf{a}^{\mathrm{T}} \frac{\partial \mathbf{y}}{\partial \tau} = 0 \tag{7}$$

This property can be designed into retrieval coefficients by imposing eq. 7 as a linear constraint when deriving coefficients by regression. This has been done successfully for a dual view sensor in relation to the major eruption in 1991 of Mount Pinatubo in the Philippines (Merchant and Harris, 1999). The addition of an extra constraint like this means that a useful set of robust coefficients can only be found for three or more channels (at different wavelengths and/or view angles). The relationships between other perturbations to observed BTs and the resulting SST bias can be analysed with a similar approach. This illustrates that understanding a retrieval problem by simulation can lead to useful insights.

More recent RT-based approaches emphasize simulation of BTs for the particular context of an observation (rather than for a spatio-temporal sample, as when defining coefficients). To achieve this in near-real time, an operational centre requires routine access to NWP forecast fields and a fast simulation capability (Le Borgne et al., 2011). For retrospective processing, NWP re-analysis fields may be used. Either way, a prior estimate for the BT based on simulation is obtained for every satellite pixel. These simulated BTs can then be used in a variety of ways to give improved SST estimates.

If simulated BTs, $\mathbf{y}_b$ (where subscript b indicates BTs simulated using prior or 'background' information), are used with SST retrieval coefficients, a simulated SST estimate is obtained $\hat{x}_b = a + \mathbf{a}^{\mathrm{T}}\mathbf{y}_b$. But the simulation of $\mathbf{y}_b$ assumes a background SST as input to the RT model, x_b. The difference $\hat{x}_b - x_b$ is then an estimate of retrieval bias for the circumstances of the retrieval (i.e., for the circumstances embodied in the NWP information). If this is a good SST bias estimate, then an SST estimate, $\hat{x}'$, that improves upon the original estimate, $\hat{x}'$, can be obtained:

$$\hat{x}' = \hat{x} - (\hat{x}_b - x_b) = x_b + \mathbf{a}^{\mathrm{T}}(\mathbf{y} - \mathbf{y}_b) \tag{8}$$

Practical experience has shown that this is indeed a beneficial extension to coefficient based retrieval, reducing geographical biases and the standard deviation of discrepancies in validation (Le Borgne et al., 2011). It requires that simulated and observed BTs have, if

necessary, been tuned to have no relative bias on average. Eq. 8 shows that simulation-based bias correction (central expression) is equivalent to adjusting the background SST in the light of the discrepancy between observed and simulated BTs (rightmost expression).

In the above approach, the coefficients, **a**, are defined as retrieval coefficients that operate on BTs in the usual way. A further alternative is to keep essentially the same equation (adding only an offset coefficient), and then to redefine coefficients specifically for an 'incremental' retrieval (Petrenko et al., 2011). This involves regressing differences between background and *in situ* SSTs against differences between matched BT observations and simulated BTs. The relative calibration of sensor and forward model is empirically included within the incremental coefficients in this approach.

Despite their practical successes, none of the above methods (whether purely empirical or including RT) is formally optimal when viewed from the standpoint of inverse theory (Rodgers, 2000). Inverse theory gives a coherent framework for analysing SST retrieval as an inverse problem, in terms of a fundamental understanding of how much information is truly present about SST in a given set of BTs. Different 'optimal estimates' of SST can, in principle, be defined, that optimize clearly defined aspects of the SSTs obtained. Where the information content of the observations is high for SST, a useful estimator is the maximum likelihood (ML) estimate, which returns the most likely SST given the observations. In an ML retrieval, the background information is used, effectively, as a linearization point for an incremental retrieval with context-specific coefficients (derived dynamically using RT). However, for a traditional split window retrieval using 11 and 12 μm channels, the ML estimate is not always useful, particularly in tropical regions and/or at high satellite zenith angle when the 12 μm channel in particular becomes nearly insensitive to SST in comparison to observational noise. In these circumstances, the BTs fundamentally contain insufficient information on SST to obtain a good retrieval without relying on prior information. (This prior information may be embedded in empirical retrieval coefficients or explicitly represented by NWP profiles—either way, its presence is unavoidable.) Thus, for split window retrieval, an appropriate optimal estimate is the maximum a

posteriori (MAP) estimate. A simplified MAP formulation has been shown to be useful for split window retrieval (Merchant et al., 2008). A further benefit of optimal estimation techniques is that they naturally output a goodness-of-fit indicator that gives useful insight into retrieval quality.

For a truly optimal MAP retrieval, several relatively onerous conditions need to be met, including an unbiased RT simulation capability (for both BTs and their partial derivatives) and a thorough understanding of instrumental noise and background error characteristics. In some areas of thermal remote sensing—nadir sounding of trace gases in the atmosphere, perhaps—the inverse problem is sufficiently delicate that a formal optimal approach is virtually obligatory. In the case of SST, strong practical success has been obtained for three decades using more direct, intuitive methods. Nonetheless, renewed interest in driving down SST retrieval uncertainties and in understanding biases has prompted new activity in exploiting RT and in optimal estimation.

Evaluating SST retrievals

The quality of an SST retrieval scheme is typically evaluated in validation by considering 'error statistics', usually the mean and standard deviation of discrepancy between the satellite and matched *in situ* SSTs (e.g., Donlon et al., 2009). The mean discrepancy is often interpreted as 'bias', but this needs to be done with care. There are real geophysical differences to be expected between satellite and *in situ* measurements (and between different types of *in situ* measurements). No measurements are perfect, and the validation data (*in situ* measurements) can also contribute errors to the discrepancy between satellite and *in situ*. Here, 'bias' is avoided in preference to the more descriptive term 'mean discrepancy'.

Between the mid-1990s and the mid-2000s, drifting buoys that routinely report SST became progressively more numerous (Meldrum et al., 2009). For recent years, it is possible to calculate a statistically sound, geographically resolved mean discrepancy compared to drifting buoys for most of the global oceans (Figure 4). This is a great benefit to development of SST remote sensing. The uncer-

tainty of SST calibration across the ensemble of drifting buoys seems to be about 0.2 K (1 σ value; Castro et al. 2012), so in regions of Figure 4 where only a few different drifting buoys have contributed, the mean discrepancy could reflect buoy calibration errors rather than systematic error in satellite SST. In principle, the thermistor technology used in drifting buoys could be calibrated to better than 0.05 K uncertainty (on deployment). New requirements for estimating SST for climate (see later) arguably justify the expense associated with this improvement in the drifting buoy programme. Nonetheless, with the current drifting buoy network, geographical biases in satellite SST can now be inferred (with due care and interpretation) from maps of mean discrepancy covering a long enough time period. Figure 4 is based on 20 years of observations (1991 to 2011), with most matches being obtained in the last decade. The satellite SSTs are from a reprocessing for climate of Along Track Scanning Radiometer observations. The mean discrepancy maps have credible structure on length scales of 1000 km for a large part of the ocean. There is a noticeable variation in bias along the equator seen in Figure 4, particularly with negative biases in the tropical Atlantic. Similar or larger biases in tropical SSTs are a common feature of retrieval of SST by coefficients using the split-window channels in a single view (e.g., Merchant et al., 2009).

The precision of satellite SSTs refers here to the retrieval-error standard deviation. The information available to assess precision includes maps of standard deviation of discrepancy. The values in this measure are an upper limit on precision, since the *in situ* observations errors contribute to the spread of discrepancy. Assuming enough different drifting buoys contribute to a particular calculation of standard deviation of discrepancy, the satellite SST precision can be approximately inferred.

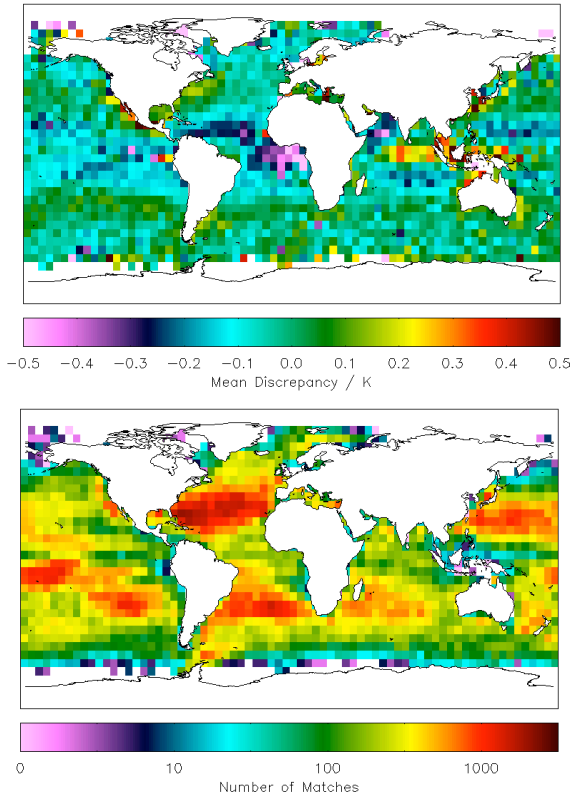

Figure 4. Upper panel: Mean discrepancy map for a single-view split-window SST estimator, adjusted for the SST skin effect and near-surface stratification, relative to drifting buoy SST. This is an estimate of the geographical variation in bias in the SST estimates. The SSTs are for the "nadir two-channel" retrieval for all matches found for a series of three Along Track Scanning Radiometers (ATSRs) between 1991 and 2011. (Note that the ATSRs are dual-view sensors; the geographical variations seen here in the tropics are greatly reduced when taking advantage of dual-view capability.) Lower panel: number of ATSR/drifting-buoy matches obtained over two decades. The marked variation in density of matches arises from the combination of drifting buoy deployment patterns and prevalence of cloud cover. Matches are particularly few in the highest latitudes and the tropical warm pool (round Indonesia), where both these factors are unfavourable.

In the context of climate applications of SST, an important quality is stability (e.g., GCOS, 2006). Stability is the constancy in time of the SST bias, or, equivalently, the additional uncertainty on any calculated climatic trends arising from (unknown) drift the calibration of the observing system. The current *in situ* observing system is not well equipped for assessment of stability of satellite SSTs since the SST calibration of drifting buoy and other *in situ* deployments has not been controlled with stability in mind. Long-term deployments of well-calibrated moorings in tropical seas (initially to monitor the El Niño region, and now worldwide; McPhaden et al., 2010) are useful for assessing stability, although geographically limited (Merchant et al., 2012). More recent near-surface (~5 m) measurements from Argo profiling floats (Freeland et al., 2010) may prove a useful global reference for stability as a longer time series accumulates.

The ideal for the long-term would be a network of SST reference sites of known, controlled stability, distributed at locations selected to allow assessment of global stability of SST. Research is needed to optimize such a network of reference sites to be effective and cost-effective. At reference site locations, both radiometric and sub-surface measurements of SST should be considered (Minnett and Corlett, 2012).

A further parameter to evaluate satellite SST is SST sensitivity (Merchant et al., 2009). SST sensitivity is the fractional response of the retrieved SST to variation in true SST, other factors (such as the atmospheric state) being equal. Ideally, the sensitivity should be 1 K K^{-1}, so that a true change in SST causes an identical change in retrieved SST. In general, this is not the case (Figure 5).

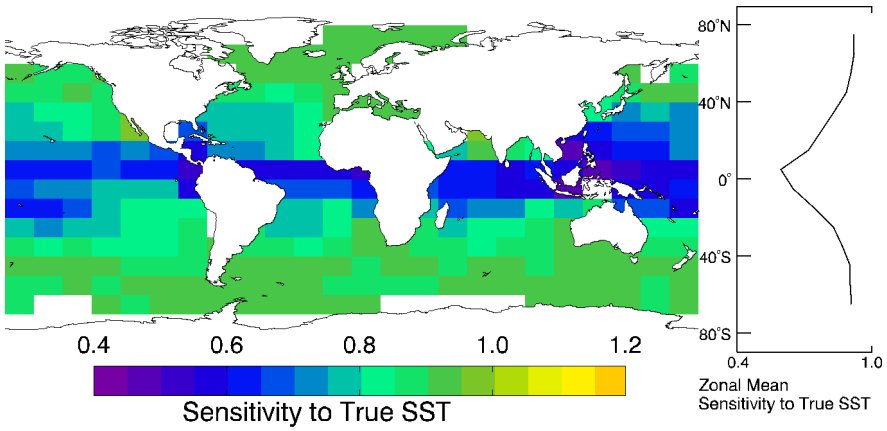

Figure 5. Change in retrieved SST per unit change in true SST, all other factors being held constant, for a split window SST estimate (non-linear SST retrieval applied to the Advanced Very High Resolution Radiometer on Metop-A). Reproduced from Merchant et al. (2009) with permission.

SST sensitivity is readily calculated for a coefficient-based retrieval as

$$\frac{\partial \hat{x}}{\partial x} = \mathbf{a}^{\mathrm{T}} \frac{\partial \mathbf{y}}{\partial x} \tag{9}$$

where the partial derivatives of BTs are calculated using RT simulation. Readers familiar with atmospheric sounding will recognize

that the averaging kernel is an indispensable tool for interpreting remotely sensed atmospheric profiles. If retrieving SST by optimal estimation, sensitivity is the diagonal term corresponding to SST in the averaging kernel matrix.

Where sensitivity of an SST estimate is much less than 1 K K^{-1}, it is expected that the strength of ocean thermal gradients is underestimated (an expectation that requires further validation at the time of writing). Likewise, diurnal variations in SST are attenuated by low-sensitivity estimators (Merchant et al., in press). Moreover, the SST sensitivity has a deep connection to the information content of the BTs (Rodgers, 2000). SST sensitivity is low when low BT sensitivity to SST reduces the signal-to-noise ratio, in which case the retrieval must depend more heavily on prior SST information. (The prior information may be explicit, as in optimal estimation, or may be implicitly embedded in SST retrieval coefficients.) Thus, the sensitivity can also be interpreted as the fraction of the information in a particular SST estimate that comes from the BTs. Figure 5 therefore illustrates the point made previously that typical split-window retrievals rely significantly on prior information in the tropics.

Meanings of 'sea surface temperature'

It is a general issue in remote sensing that the remotely sensed quantity is not identical to measurements made *in situ* of nominally the same quantity.

In the case of SST, a lot is understood about how different 'sea surface temperatures' relate. This understanding is the fruit of research cruises (e.g., Minnett et al., 2011) and profilers (e.g., Ward et al., 2004) that have undertaken intensive multiple observations, and of modelling the near-surface ocean and atmosphere using fundamental physics.

The thermal emission from the sea surface comes from a layer whose characteristic depth varies with wavelength (because the complex refractive index of water varies with wavelength; e.g., Hanafin and Minnett, 2005). The radiometric skin depth is ~10 μm at wavelengths around 12 μm, and ~100 μm at wavelengths around

4 μm. As well as this radiometric skin, the ocean surface also has a thermal skin. Heat can be transported through this thermal skin only by molecular heat diffusion, and not via the turbulent motions that are effective in moving heat within the bulk of the fluid. Heat flux is usually from ocean to atmosphere. A temperature gradient must be present within the thermal skin to transport heat via molecular diffusion. For a typical ocean-atmosphere heat flux, the drop in temperature across this thermal skin is of order 0.2 K. The radiometric temperature of the sea surface differs from the thermodynamic temperature of water below the diffusive layer (the 'sub-skin SST') both because the sea surface emissivity is less than 1, and because the temperature of the water within the radiometric skin depth is actually different from (usually cooler than) the sub-skin SST.

The sub-skin SST can be very close to the SST measured by drifting buoys (at a depth of order 20 cm), moored buoys (typically of order 1 m depth) or the top observation of conventional Argo profiles (around 5 m). Sub-skin SSTs and *in situ*-depth SSTs are equal when the near-surface is not thermally stratified, reflecting efficient mixing of the near-surface water by wind action. Although near-surface stratification is often small compared to SST uncertainties, it is sometimes considerable (e.g., Clayson and Weitlich, 2007). Stratification that causes sub-skin to *in situ*-depth SST differences may be caused by heavy rainfall creating a fresh 'lens' of water of a different temperature on the sea surface. However, near-surface stratification has been most clearly observed in satellite SSTs when caused by diurnal warming (e.g., Gentemann et al., 2008).

During the day, sunlight preferentially heats the upper centimetres to metres of the ocean, because most wavelengths of sunlight are absorbed by seawater over such distances. In the absence of wind, this will tend to cause a warm near surface layer—that is, it thermally stratifies the water (Fairall et al., 1996). Under persistently calm conditions (wind speed less than 1 or 2 m s^{-1}), this effect can warm the sub-skin SST by 6 or 7 K between sunrise and early afternoon (e.g., Gentemann et al., 2008). Wind action tends to act against stratification, by mixing the heat down, and an increase of wind can fairly rapidly erode diurnal stratification. Under wind speeds of about 6 m s^{-1} or more, the peak amplitude of the diurnal cycle in sub-skin SST is no more than a few tenths of kelvin.

In summary:

- The ocean thermal skin effect is generally present, so the skin SST is different (usually cooler) than the sub-skin SST. Measurements by thermal remote sensing are sensitive to this skin SST.
- The difference between sub-skin SST and SST at depths measured by drifting and moored buoys can range from negligible (e.g., night-time, windy conditions) to several degrees (high-insolation, sustained low-wind-speed conditions).

Satellite SSTs obtained using coefficients derived by regression to drifting buoys are sensitive to skin SSTs, but are tuned to remove the mean skin-depth difference present in the matched data set. This ignores the true variability of skin–depth differences, which then becomes part of the error budget for the satellite SST estimate.

Some investigators have restricted the *in situ* observations used for deriving empirical coefficients to situations likely to have negligible stratification. This can be done by specifying that the wind speed around the time of the satellite–*in situ* match must be above a threshold. The satellite SSTs can then be justifiably described as estimates of sub-skin SST on average, since the mean skin–sub-skin difference is tuned out.

Satellite SSTs retrieved using RT methods should return a true skin SST. This has the merit that the estimated quantity is the quantity to which the observations are sensitive. To compare such satellite SSTs to others then requires explicit account to be taken of skin–sub-skin and sub-skin–depth differences. This can be done using physical models (e.g., Kantha and Clayson, 1994) driven by heat flux and wind speed over the diurnal cycle (e.g. from NWP) to obtain an adjustment between skin (satellite) and depth (*in situ*) SST (e.g., Embury et al., 2012b). The complication with this approach is that any discrepancy after such an adjustment could arise from model errors as well as observational errors.

The discussion above focusses on geophysical differences in SST with respect to depth. Horizontal variability in SST is also important. *In situ* observations give the SST at a point in space. Satellite SSTs are estimates over an area of typically 1 to 30 km^2. Thus, there is point-to-pixel sampling variability in any satellite–*in situ*

comparison. Specialist research cruises and a few automated systems make radiometer measurements of SST. Radiometer measurements are attractive to compare with satellite SSTs because both respond to the skin SST, removing 'vertical' variability. Point-to-pixel issues remain in such comparisons, however (e.g., Wimmer et al., 2012). For both radiometer and well-calibrated subsurface SST, it seems to be difficult to reduce satellite–*in situ* discrepancies from geophysical variability to less than about 0.1 K (Castro et al., 2010).

The wider context for thermal remote sensing of SST

Operational SST production

To support weather forecasting (numerical weather prediction, NWP) and near-real time oceanography, SSTs are produced operationally (Donlon et al., 2010). This means regional-to-global, near-real time, high-availability, continuous generation and distribution of SST products. The requirements on timeliness and availability and the volume of data flow involved mean that operations tend to be undertaken by major agencies that can maintain 24 hr/day functions. The ultimate quality of SST products depends as much on the steps relating to calibration of observations and cloud-detection as on the SST retrieval itself.

Users of operational SST are varied. Some require visual interpretation of SST images, perhaps to locate fronts in real time. Such users may prefer 'level 2' imagery, where the SSTs are presented on the geographical locations at which they were observed (for the clear-sky areas available). Other users, requiring reduced data volumes and/or more convenient formats, prefer 'level 3 products', in which there is some regridding, and averaging or compositing in space and time. A level 3 product might comprise, for example, the daily average on a regular latitude-longitude grid of all the observations within each grid cell taken by a particular sensor.

For many applications, a spatially completed, gridded field is necessary – referred to as an 'analysis' or as a 'level 4 SST product'. To

improve the spatial sampling above that observed by a single sensor, blending data from multiple sources is usually performed for level 4 production (e.g., Reynolds et al., 2007; Donlon et al., 2012). This may include *in situ* observations and passive microwave (PMW) SSTs as well as SST from thermal remote sensing. Even so, spatio-temporal gaps will exist, requiring interpolation (gap-filling) – see Figure 6.

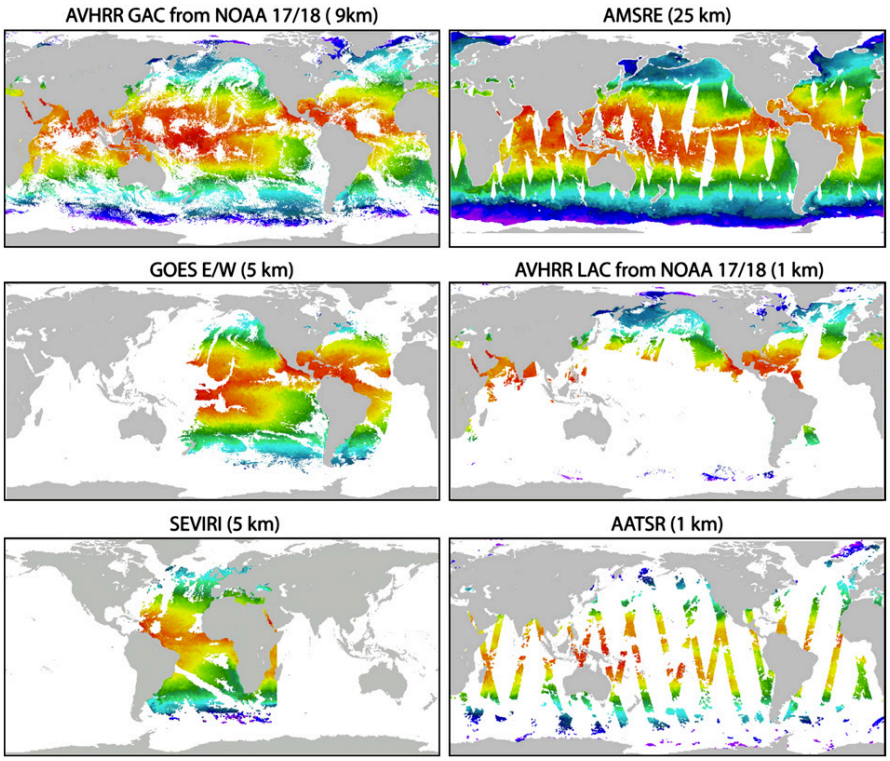

Figure 6. Typical example of the daily coverage of SST from six different SST data products, at different spatial resolutions, all from the same day. Reproduced from Robinson et al., 2012, with permission.

There are many approaches to SST analysis (combining data to cope with the different types of data, with different resolutions, gaps, uncertainties and biases). Irrespective of the method, the process of analysis will tend to create an SST field that, to a greater or lesser degree, has certain limitations. The analysis SST will tend to have poorer resolution of thermal features than the highest resolution data

used in the analysis (i.e., thermal gradients and contrasts will tend to be reduced by the analysis process). The results of the analysis are sensitive to the assumptions made about the relative uncertainties of different data sources, and the treatment of bias correction. The SST estimates given in the absence of observations (filling data gaps) are based on assumptions about the correlations of the unobserved SST anomalies to those observed nearby in time and space. Features in gap-filled areas may differ systematically from reality. In general, the SST error statistics across the analysis are far from uniform, although this variation is sometimes neglected in applications. These caveats have to be traded against the practical usefulness of having a spatially complete SST estimate.

Weather and ocean forecasting are major routine applications of near-real time SST products. The atmosphere and ocean interact, exchanging heat, mass (evaporation and rain) and momentum; in maritime climates, air temperature and humidity are partly determined by upwind SST. Surface winds interact with ocean thermal fronts, with the influence of SST changes propagating through the troposphere (Chelton et al, 2001). Meanwhile, cloudiness (via the strength of surface solar heating) and wind affect processes of mixing and stratification in the upper ocean. Numerical simulations for weather forecasting up to several days ahead are, at present, generally performed assuming the most recent level-4 SST analysis as a fixed boundary condition for the bottom of the atmosphere.

SST production for climate services

Table 4 lists some uses of SST products in the realm of climatology and climate services.

The demands on SST accuracy and stability for climate applications are onerous (GCOS, 2011). Consistent SST data sets for climate applications have often been provided via reprocessing projects (e.g., Kilpatrick et al., 2001). In such a project, an SST record for a particular time period is generated by re-deriving SST from the input satellite (and perhaps in situ) data streams in a manner that is consistent, and, hopefully, an improvement over previous products. However, some of the climate-related applications in Table 4 are de-

veloping into climate services that require an estimate of the thermodynamic state of the ocean that is constantly updated, while simultaneously being of 'climate quality' and consistent with a climate data record going back in time. An example is 'seamless' prediction of long-range weather, seasonal tendencies and future climate scenarios. This will require systems capable of delivering 'climate quality' SST with a relatively short delay from the time of acquisition (perhaps a few days).

Table 4. Established uses of SST in climatology

Use	Purpose(s)	Comments
Quantifying SST/climate variability and trends	Baseline knowledge of behaviour of atmosphere-ocean system, including geographical patterns of SST variability and corresponding weather anomalies. Assessment of long-term changes, including those associated with human forcing of climate.	Of order 100 years of SST required for variability at multi-annual and decadal scale, and for climate trend analyses. In satellite era, this can be addressed with much greater spatiotemporal detail than in pre-satellite era. Satellite SSTs are helpful in establishing modes (spatial patterns) of variability that can be exploited in filling gaps in historical data ("historical reconstruction").
Detection and attribution of climate changes	Assessment of climatic trends that are 'stand out' above climatic variability. Comparison of spatio-temporal progression of SST with expected evolution under different agents forcing change, to attribute which forcings have caused observed trends.	Usually requiring historical reconstructions of SST fields over of order 100 years.
Boundary condition of atmospheric reanalyses	Atmospheric reanalysis use a numerical weather prediction system retrospectively to infer the best estimate of past weather.	Reanalyses to date generally rely on prescribed, spatially complete SST fields. For recent decades these rely heavily on remotely sensed SST. Reanalyses are useful in generating consistent estimates of air-sea fluxes of heat and precipitation, amongst numerous other applications.
Boundary condition of climate model runs	Verification that climate models reproduce historical (e.g., 20^{th}C) climates when driven with best estimate SSTs.	SST field is prescribed, while atmosphere and land components evolve in the climate simulation in response. The ability to reproduce the land climate of the 20^{th}C given prescribed SST has been viewed as a basic test for validity of a given climate model.
Climate prediction	Forecasts from seasonal to decadal scales of the statistics of future weather.	Requires a coupled climate model (one in which ocean is interactive, not prescribed). Relatively slow response time of ocean mixed layer gives some level of predictability for seasonal forecasting. Seasonal forecasting requires that the simulation starts with as realistic as possible an estimate of state of the ocean (accurate 'initialisation'). On timescales of decades, predictability comes from dominance of the influence of forcings over internal variability.
Climate-ecological interactions	Research into the responses of plankton, fish and corals to SST variations.	Important in understanding present and future productivity and ecological health of the oceans.

International co-operation

There is a well-developed framework of international co-operation in operational remote sensing of SST involving many agencies worldwide. The Group for High Resolution SST (GHRSST) co-ordinates sharing of tasks, including routine inter-comparison of SST products, archiving and product distribution (Donlon et al., 2009). Common standards and data formats have been developed to increase the ease of use of SST from different sources. GHRSST also gives a forum in which the science of SST remote sensing is debated and advanced.

At time of writing, GHRSST has a project office supported by the European Space Agency and is formally linked to the Committee on Earth Observation Satellites (the international forum for co-ordination of civil remote sensing). At the GHRSST web site, www.ghrsst.org, links to operational and archive products (levels 2, 3 and 4) are available, with routinely updated visualizations of SST products and analyses (Martin et al., 2012; Dash et al., 2012), *etc*.

The SST sensor constellation and SST analysis

Table 5 illustrates the constellation of SST relevant sensors. Different classes of sensors / platforms have complementary technical capabilities and roles, with strengths in providing different aspects of the user requirements for remotely sensed SST. For example: geostationary sensors are well suited to resolve sub-daily variability in SST, complementing the higher-resolution less-frequent observations from lower-altitude polar orbiting instruments; dual-view sensors can give higher SST accuracy, but have poorer sampling because a dual-view swath width is unavoidably narrower than that for a traditional single-view imager. (The examples given in Table 5 do not include any SST-capable instruments whose products are not included within the GHRSST co-operative framework, although several such instruments are in flight.)

An important complement to the thermal sensing constellation is an SST capability at microwave frequencies (Wentz et al., 2000).

Passive microwave (PMW) sensors have some limitations relative to thermal remote sensing using infra-red (IR) wavelengths. First, PMW SSTs have lower spatial resolution. Because the size of footprint of a PMW sensor with a given antenna is inversely related to frequency, it is difficult to define a single resolution for the multi-frequency PMW retrievals; a reasonable indication is an area of order 2500 km^2, which is obviously a much greater than the 1 km^2 to 100 km^2 typical of IR imagers. Second, PMW SSTs are not available or have much increased uncertainty within about 50 km of land and sea-ice, because of sensitivity to the land or ice emission in the side-lobes of the antenna pattern. Third, PMW SSTs to date have to SST uncertainty (~0.5 K) comparable to the less capable IR imagers. A significant contribution in PMW SST uncertainty is the greater sea-state dependence in emissivity. Fourth, problems with radio frequency interference are degrading PMW SSTs across progressively more of the ocean in European seas and elsewhere.

Nonetheless, PMW SSTs are a very powerful addition to the sensor constellation for SST, because retrievals are available through non-raining clouds. PMW SSTs are particularly beneficial, therefore, in areas and periods of total cloud cover; they may provide the only satellite information about SST near a particular location for days or even weeks, sometimes with important consequences (Wentz et al., 2000).

The full constellation of sensors is beneficial to our ability to estimate the global distribution of SST at a given time – i.e., to the process of creating SST analyses. Synthesis of the different sources of SST information remains a challenge, especially in the face of increasing demands for high spatial resolution (e.g., 1 km globally) and sub-daily temporal resolution (e.g., 3 hourly, capturing the diurnal cycle). Given the sampling limitations of thermal remote sensing (because of clouds) and the resolution limitations of microwave remote sensing (~50 km, one or twice a day per sensor in the open ocean), the degree to which such demands can be met by the present constellation is a topic of ongoing research. ("Meaningfully" here means that the SST variations in the analyses are determined more by real information observed by the SST constellation, than by noise arising from observation uncertainty and the analysis system.) One clear direction for progress is to develop a more complete under-

standing of the uncertainties of different types of data and the degree of correlation of different components of error in time and space. SST analysis systems properly able to use such improved uncertainty characterization will (i) preserve as much as possible of the true information on SST held in the constellation of sensors, (ii) minimize the introduction of spurious features in the analysed SST, and (iii) deliver realistic estimates of the uncertainty in the analysed SST.

Table 5. Categories of infra-red SST sensors and examples used within the GHRSST framework.

Orbit	Channels	View	Role(s) in constellation	Example(s)
Geostationary	Split-window and near-IR (broad channels)	Fixed view of visible Earth disk	High temporal sampling: resolves diurnal cycle; maximizes spatial coverage by tracking gaps in cloud.	Spinning Enhanced Visible and Infra-Red Imager (SEVIRI)
Polar	Split-window and near-IR (broad channels)	Single view, across track	Near-global coverage on daily basis including high latitudes (before cloud screening). Main operational meteorological sensors, usually at least two in orbit (morning and afternoon).	Advanced Very High Resolution Radiometer (AVHRR)
Polar	As above plus additional SST relevant channels	Single view, across track	Extended capability relative to AVHRR-like channel set.	Moderate-resolution Imager Spectroradiometer (MODIS), near-IR bands centred on 3.95 and 4.05 μm. Visible/Infrared Imager Radiometer Suite (VIIRS)
Polar	Split-window and near-IR (broad channels)	Dual-view, across and along track	High-accuracy SST for climate and/or SST calibration reference. Poorer sampling from narrower swath. Greater robustness to aerosol contamination with dual view.	Advanced Along-Track Scanning Radiometer (AATSR) Sea and Land Surface Temperature Radiometer (SLSTR, from *c*. 2014)

Thermal remote sensing of SST is an integral part of the observation of the ocean, and indeed of the global environment. Space and

meteorological agencies are committed to maintaining the capability over the coming decades. This commitment is accompanied by ongoing improvement in our ability to sense the temperature of the ocean surface, bringing many benefits to society.

References

Anding, D., and Kauth, R. (1970), *Estimation of sea surface temperature from space*. Remote Sensing of Environment 1 (4), 217–220. DOI: 10.1016/S0034-4257(70)80002-5

Barton, I. J., (1995), *Satellite-derived sea surface temperatures: Current status*, Journal of Geophysical Research, 100, 8777-8790.

Barton, I. J., (2011), *Improving Satellite-derived Sea Surface Temperature Accuracies using Water Vapor Profile Data*, Journal of Atmospheric and Oceanic Technology, 28 (1), 85-93, doi:10.1175/2010JTECHA1502.1.

Castro, S. L., G. A. Wick and W. J. Emery (2012), *Evaluation of the relative performance of sea surface temperature measurements from different types of drifting and moored buoys using satellite derived reference products*. Journal of Geophysical Research, 117, C02029, DOI: 10.1029/2011JC007472.

Castro, S. L., G. A. Wick, P.J. Minnett, A.T. Jessup, and W. J. Emery, W. J., (2010), *The impact of measurement uncertainty and spatial variability on the accuracy of skin and subsurface regression-based sea surface temperature algorithms*. Remote Sensing of Environment 114 (11), 2666–2678. DOI: 10.1016/j.rse.2010.06.003.

Chelton D. B., S. K. Esbensen, G. Schlax and co-authors, (2001), *Observations of coupling between surface wind stress and sea surface temperature in the eastern tropical Pacific,* Journal of Climate, 14 (7), 1479-1498, 10.1175/1520-0442(2001)014<1479:OOCBSW>2.0.CO;2

Clayson, C. A. and D. Weitlich (2007), *Variability of tropical diurnal sea surface temperature,* Journal of Climate, 20 (2), 334-352, doi: 10.1175/JCLI3999.1.

Dash, P., A. Ignatov, M. Martin and co-authors, (2012), *Group for High Resolution Sea Surface Temperature (GHRSST) analysis fields inter-comparisons-Part 2: Near real time web-based level 4 SST Quality Monitor (L4-SQUAM)*, Deep Sea Research Part II – Topical Studies in Oceanography, 77-80 (Special Issue), 31-43, doi: 10.1016/j.dsr2.2012.04.002.

Deschamps, P. Y., and T. Phulpin (1980), *Atmospheric corrections of infrared measurements of sea surface temperature using 3.7 µm, 11 µm and 12 µm.* Boundary Layer Meteorology, 13, 131-143.

Donlon, C. J., K. S. Casey and co-authors (2009), *The GODAE high resolution sea surface temperature pilot project*, Oceanography, 22 (3), 34-45.

Donlon, C. & Co-Authors (2010). "Successes and Challenges for the Modern Sea Surface Temperature Observing System" in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2),* Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D.E. & Stammer, D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.24.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler and W. Wimmer (2012), *The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system*, Remote Sensing of Environment, 116 (Special Issue), 140-158, doi: 10.1016/j.rse.2010.10.017.

Embury, O.,C. J. Merchant, and M. J. Filipiak (2012a), *A Reprocessing for Climate of Sea Surface Temperature from the Along-Track Scanning Radiometers: Basis in Radiative Transfer*. Remote Sensing of Environment, pp32 - 46, DOI: 10.1016/j.rse.2010.10.016.

Embury, O., C. J. Merchant and G. K. Corlett (2012b), *A Reprocessing for Climate of Sea Surface Temperature from the Along-Track Scanning Radiometers: Initial validation, accounting*

for skin and diurnal variability, Remote Sensing of Environment, 62 - 78. doi: 10.1016/j.rse.2011.02.028.

Emery, W. J., Y. Y. Yu and G. A. Wick (1994), *Correcting Infrared Satellite Estimates of Sea Surface Temperature for Atmospheric Water Vapor Attenuation,* Journal of Geophysical Research, 99 (C3), 5219-5236, DOI**:** 10.1029/93JC03215.

Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson and G. S. Young (1996), *Cool-skin and warm-layer effects on sea surface temperature*. Journal of Geophysical Research, 101, 1295-1308.

Freeland, H. & Co-Authors (2010). "Argo - A Decade of Progress" in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2),* Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D.E. & Stammer, D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.32.

Gentemann, C. L., P. J. Minnett, P. Le Borgne, and C. J. Merchant (2008), *Multi-satellite measurements of large diurnal warming events*. Geophysical Research Letters 35 (22), L22602. DOI: 10.1029/2008GL035730.

Hanafin, J. A. and P. J. Minnett (2005), *Measurements of the infrared emissivity of a wind-roughened sea surface*, Applied Optics, 44 (3), 398-411, doi: 10.1364/AO.44.000398.

GCOS [Global Climate Observing System] (2006), *Systematic Observation Requirements for Satellite-based Products for Climate - Supplemental details to the satellite-based component of the GCOS Implementation Plan*, GCOS-107.

GCOS [Global Climate Observing System] (2011), *Systematic Observation Requirements for Satellite-based Products for Climate - Supplemental details to the satellite-based component of the GCOS Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC – 2011 Update*, GCOS-154.

Kantha, L. H. and C. A. Clayson (1994), *An improved mixed-layer model for geophysical applications,* Journal of Geophysical Research, 99 (C12), 25235-25266, doi: 10.1029/94JC02257.

Kilpatrick, K. A., G. P. Podesta and R. Evans (2001), *Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database,* Journal of Geophysical Research 106 (C5), 9179-9197.

Le Borgne, P., H. Roquet, and C.J. Merchant (2011), *Estimation of sea surface temperature from the Spinning Enhanced Visible and Infra Red Imager, improved using numerical weather prediction*. Remote Sensing of Environment 115, 55-66. DOI: 10.1016/j.rse.2010.08.004.

Martin, M., P. Dash, A. Ignatov and co-authors (2012), *Group for High Resolution Sea Surface temperature (GHRSST) analysis fields inter-comparisons. Part 1: A GHRSST multi-product ensemble (GMPE),)*, Deep Sea Research Part II – Topical Studies in Oceanography, 77-80 (Special Issue), 21-30, doi: 10.1016/j.dsr2.2012.04.013.

McClain, E. P., W. G. Pichel and C. C. Walton (1985), *Comparative performance of AVHRR-based multichannel sea surface temperatures*. Journal of Geophysical Research, 90, 11587-11601.

McMillin, L. M. and D. S. Crosby (1984), *Theory and validation of the multiple window sea surface temperature technique*. Journal of Geophysical Research, 89 (C3), 3655-3661.

McPhaden, M. & Co-Authors (2010). "The Global Tropical Moored Buoy Array in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2),* Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D.E. & Stammer, D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.61

Meldrum, D. & Co-Authors (2010). "Data Buoy Observations: The Status Quo and Anticipated Developments Over the Next Decade" in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2),* Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D.E. & Stammer, D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.62.

Merchant C J, Harris A R, Murray M J and Zavody A M (1999), *Toward the elimination of bias in satellite retrievals of skin sea surface temperature 1. Theory, modeling and inter-algorithm comparison*, J Geophys Res, 104, C10, 23565-23578.

Merchant, C. J., and A.R. Harris, (1999), *Toward the elimination of bias in satellite retrievals of sea surface temperature 2. comparison with in situ measurements*. Journal of Geophysical Research 104 (C10), 23579–23590. DOI: 10.1029/1999JC900106.

Merchant, C. J., A. R. Harris, H. Roquet, and P. Le Borgne (2009), *Retrieval characteristics of non-linear sea surface temperature from the Advanced Very High Resolution Radiometer*. Geophysical Research Letters 36, L17604. DOI: 10.1029/2009GL039843.

Merchant, C. J., L. A. Horrocks, J. R. Eyre, and A.G. O'carroll (2006), *Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors*. Quarterly Journal of the Royal Meteorological Society 132 (617), 1205–1223. DOI: 10.1256/qj.05.143.

Merchant, C. J., and P. Le Borgne (2004), *Retrieval of sea surface temperature from space, based on modelling of infrared radiative transfer: Capabilities and limitations*. Journal of Atmospheric and Oceanic Technology 21 (11), 1734–1746. DOI: 10.1175/JTECH1667.1.

Merchant C. J., P. Le Borgne, A. Marsouin , and H. Roquet (2008), *Optimal estimation of sea surface temperature from split-window observations*. Remote Sensing of Environment 112 (5), 2469-2484. DOI:10.1016/j.rse.2007.11.011.

Merchant, C. J., O. Embury, N. A. Rayner, D. I. Berry, G. Corlett, K. Lean, K. L. Veal, E. C. Kent, D. Llewellyn-Jones, J. J. Remedios, and R. Saunders (2012), *A twenty-year independent record of sea surface temperature for climate from Along Track Scanning Radiometers*, Journal of Geophysical Research, doi:10.1029/2012JC008400.

Merchant, C. J., P. LeBorgne, H. Roquet and G. Legendre (in press), *Extended optimal estimation techniques for sea surface temperature from the Spinning Enhanced Visible and Infra--Red Imager (SEVIRI)*, Remote Sensing of Environment.

Minnett, P. J., (1990), *The regional optimisation of infrared measurements of sea surface temperature from space*. Journal of Geophysical Research, 95, 13497-13510.

Minnett, P. J., M. Smith and B. Ward (2011), *Measurements of the oceanic thermal skin effect,* Deep Sea Research Part II – Topical Studies in Oceanography, 58 (6), 861-868, doi: 10.1016/j.dsr2.2010.10.024.

Minnett, P.J., and Corlett, G.K., (2012). *A pathway to generating Climate Data Records of sea-surface temperature from satellite measurements*, Deep-Sea Research II, 77–80, 44–51. http://dx.doi.org/10.1016/j.dsr2.2012.04.003.

Munk, W. H (1950), *On the Wind-driven Ocean Circulation*, Journal of Meteorology, 7 (2), 79-93.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton and co-authors, (2007), *Daily high-resolution-blended analyses for sea surface temperature,* Journal of Climate, 20 (22), 5473-5496, 10.1175/2007JCLI1824.1.

Robinson, I.S., J-F. Piollé, P. Le Borgne, D. Poulter, C. Donlon, and A. Olivier (2004), *Widening the application of AATSR SST data to operational tasks through the Medspiration Service*. Remote Sensing of Environment 116, 126-139. DOI: 10.1016/j.rse.20120.12.019

Robinson, I. S., (2004), *Measuring the Oceans from Space: The principles and methods of satellite oceanography*, Praxis Publishing Ltd, Chichester, UK, ISBN 3-540-42647-7.

Rodgers C. D., (2000) *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publishing Co Ltd, ISBN 081-02-2740-X.

Rothman, L. S. (2010), *The evolution and impact of the HITRAN molecular spectroscopic database*. Journal of Quantitative Spectroscopy and Radiative Transfer 111 (11), 1565–1567. DOI: 10.1016/j.jqsrt.2010.01.027

Petrenko, B., A. Ignatov, N. Shabanov, and Y. Kihai (2011), *Development and evaluation of SST algorithms for GOES-R ABI using MSG SEVIRI as a proxy*. Remote Sensing of Environment 115 (12), 3647–3658. DOI: 10.1016/j.rse.2011.09.003

Pichel, W., E. Maturi, P. Clemente-Colón, and J. Sapper (2001), *Deriving the operational nonlinear multichannel sea surface temperature algorithm coefficients for NOAA-15 AVHRR/3*. International Journal of Remote Sensing 22 (4), 699–704. DOI: 10.1080/01431160010013793.

URL1: http://www.atsr.rl.ac.uk/images/sample/atsr-2

Walton, C.C., W. G. Pichel, J. F. Sapper, and D. A. May (1998), *The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites*. Journal of Geophysical Research 103 (C12). DOI: 10.1029/98JC02370.

Ward, B., R. Wanninkhof, P. J. Minnett and M. J. Head (2004), *SkinDeEP: A profiling instrument for upper-decameter sea surface measurements*, Journal of Atmospheric and Oceanic Technology, 21 (2), 207-222, doi: 10.1175/1520-0426(2004)021<0207:SAPIFU>2.0.CO;2.

Wentz, F., C. Gentemann, D. Smith and co-authors, (2000), *Satellite measurements of sea surface temperature through clouds,* Science, 288 (5467), 847-850, doi: 10.1126/science.288.5467.847.

Wimmer, W., I. S. Robinson and C. J. Donlon (2012), *Long-term validation of AATSR SST data products using shipborne radiometry in the Bay of Biscay and English Channel*, Remote Sensing of Environment, 116 (Special Issue), 17-31, doi: 10.1016/j.rse.2011.03.022.