SEA SURFACE TEMPERATURE – A DYNAMIC FIELD: DEVELOPMENTS IN RESEARCH AND OPERATIONS

C J Merchant

Meteorology, The University of Edinburgh
JCMB 8204, King's Bldgs, Mayfield Rd, Edinburgh EH9 3JZ

ABSTRACT

This paper reviews the key themes in both research and operations of space-based sea surface temperature (SST). The paper reviews the suite of sensors for SST now available, including advances in geostationary coverage and microwave observations. The estimation process is discussed in terms of the screening and retrieval steps. The arguments for a fully Bayesian approach to screening are presented, and the limits to the accuracy of radiative transfer based retrievals briefly outlined. The question of how to combine observations in this data rich period is addressed, looking at merging techniques, and the fact that real geophysical variability in near-surface temperature structure must be properly treated.

1. CURRENT SENSORS FOR SEA SURFACE TEMPERATURE

The global oceans are less sparsely observed than ever before, and a contributing element of this data rich period is the suite of satellite-based sensors capable of sea surface temperature (SST). In particular, this present year has seen the advent of more complete coverage from infrared sensors on geostationary satellites, and in recent years nearly all-weather retrieval capability from microwave radiometry has been available.

The Advanced Very High Resolution Radiometer (AVHRR) [1] has been the "workhorse" of satellite based SST for two decades, and continuity of observation is planned in the context of both NOAA and Metop. AVHRR is the archetypal thermal sensor for meteorology, with its 1.1 km minimum resolution and channels in the thermal window regions used for SST (3.7, 11 and 12 micron). SST observations are made from AVHRR using statistically derived retrievals schemes, to an accuracy of 0.4 to 0.6 K.

A sensor that is somewhat similar to AVHRR, but with significantly improved calibration and noise specifications, is the Along Track Scanning Radiometer series, of which the current is the Advanced ATSR on Envisat. Although primarily a climate research sensor, AATSR is in the frame for operational use because the data are available in an operational mode – a Meteo product at 1/6 degree resolution. AATSR has a dual view of the ocean surface from the along track scanning arrangement, which has allowed the SSTs to be made highly robust (to ~0.1 K) against biasing by stratospheric volcanic aerosol – something that is not true of single view sensors, whose typical sensitivity is >1 K [2], although retrievals based on three channels may be rendered somewhat less sensitive [3]. Researchers at the University of Leicester [4] have shown that, as a by-product, these robust dual-view SSTs are also partly robust to the presence of Saharan dust (and probably other tropospheric aerosols too). See Fig 1 and 2.

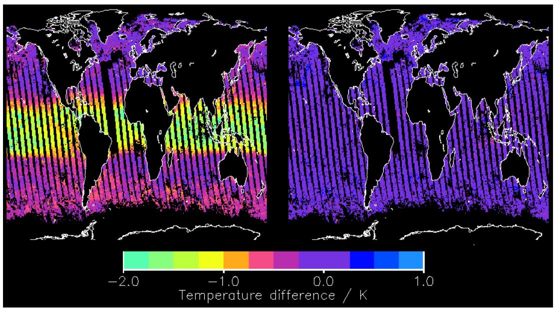

Figure 1. Showing lack of sensitivity of appropriately designed dual view SST coefficients to stratospheric volcanic aerosol, using ATSR data. The wide zone of bias in the original ATSR SSTs in the left panel mirrors the extent of the stratospheric aerosol layer 4 months following the Pinatubo eruption in 1991, and indicates the biases one might expect also from single view sensors such as AVHRR. The right panel shows the (relative lack of) bias for retrievals designed to be robust to stratospheric aerosol.

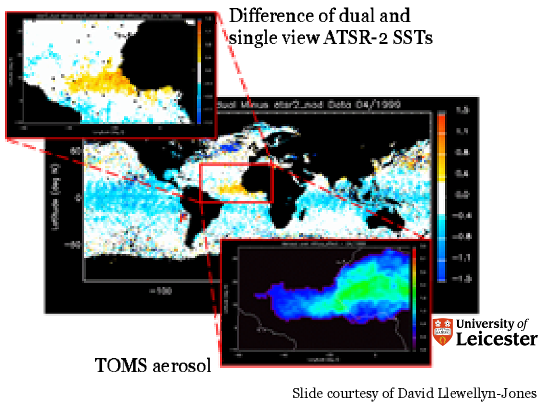

Figure 2. Dual view SSTs from ATSR-2 over the eastern Atlantic are warmer by 1 K or more than the single view SSTs of the same sensor over an area affected by Saharan dust. The cause is the sensitivity of the single view retrievals to this tropospheric aerosol, while dual view retrievals are relatively robust. AVHRR and other single view sensors will tend to suffer biases similar to that shown. Image provided by Prof David Llewellyn-Jones, and courtesy of him and John Remedios at the University of Leicester.

With the operations of Meteosat Second Generation and improved specification observations over Indonesia, geostationary SST coverage is available at 4 out of 5 stations required for near-global coverage, at a temporal rate of between a few hours and 15 minutes. All have traditional (AVHRR-like) SST sensors, and bring with them benefits from high temporal resolution, somewhat offset by lower spatial resolution (3 to 10 km, depending on the sensor the satellite view angle).

Microwave (MW) SST sensors, such as the TRMM Microwave Imager (TMI) and now the Advanced Microwave Scanning Radiometer-E (AMSR-E) on Aqua use the 4 – 10 GHz part of the spectrum, bringing the advantage of transparency of non-precipitating cloudy atmospheres (rainfall rate <3 mm hr^{-1}). Spatial resolution is relatively coarse (~50 km) with a retrieval precision of order 0.6 K, but TMI established the usefulness of the all-weather capability to meteorology, particularly in the context of hurricane forecasting [5], where the extent of cooling of the ocean surface by passing tropical cyclones can be detected through the extensive cloudiness associated with the system. MW SSTs need significant bias correction as a function of wind speed, for example using matches with in situ or infrared data – the variability of surface emissivity with wind speed being a source of potential error, together with the calibration system. The future Conical Microwave Imager Sounder (for NPOESS) will have two channels in the 4 – 10 GHz range and improved calibration which should yield better accuracy. MW SSTs cannot be obtained within ~100 km of the coast with current sensors because of contamination in the 5% sidelobes, and can be biased by islands of scale 1 km within the main field of view. But nonetheless, for non-coastal, lower resolution applications, MW SSTs are an excellent addition to the SST observing system.

2. DEVELOPMENTS IN SST ESTIMATION

From both IR and MW sensors, the process of SST estimation is in two steps: screening for pixels for which a valid SST can be obtained; the retrieval of the SSTs from the observations.

For MW sensors, the screening step should avoid land contamination, precipitating systems and contamination by telecommunications around 7 GHz. For IR sensors, the main difficulty is to screen out pixels for which a retrieval is not viable because of cloud, as discussed further below.

The archetypal approach to IR cloud screening is illustrated in Figure 3 taken from [6]. A sequence of comparisons between channels or combinations of channels is compared with thresholds that are predefined to encompass the observation “space” that is deemed to correspond to potentially valid clear-sky pixels.

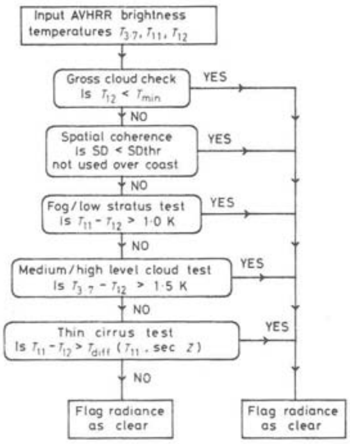

Figure 3. Example of threshold approach to cloud screening – here the early APOLLO system for AVHRR [6].

A pixel failing any test is screened as “cloudy” and a binary mask is produced. This approach is pragmatic and effective, but can be criticized on a number of levels. The thresholds and nature of the tests employed may be sensor specific and a matter of scientific judgement, and defining these tests and their thresholds can be a laborious research project. The thresholds embody in effect a climatology for expected clear sky radiances based either on inspection of imagery or simulations of radiance. But the thresholds are not necessarily optimal, and there is no formal way in which to quantify the severity of laxity of the screening obtained. In the operational context, the prior knowledge available in the form of forecast fields is much greater than is embodied in any fixed threshold scheme.

Refinements to the thresholding methodology include using forecast fields to make thresholds more dynamic. For example, forecast water vapour is used to vary the thresholds at the Nowcasting SAF; this sort of approach is current good practice. Similarly, the sharp binary distinction of clear and cloudy has been replaced by a more graded qualitative descriptor of quality of pixels for SST retrieval within the operations of the Ocean and Sea Ice SAF (Figure 4).

SAF-OSI: confidence flags

- Qualitative indicator, semi-probabilistic

1: cloudy (i.e., screened)
2: bad — Usable for visualization
3: acceptable
4: good — Used quantitatively
5: excellent

Figure 4: SST confidence flag at the Satellite Application Facility – Ocean and Sea Ice

At Edinburgh [7], we are prototyping an alternative approach to cloud screening that is particularly suited to the operational context; it will be used in SST operations at NOAA-NESDIS for GOES sensors. Instead of thresholding, we are developing a method that is fully probabilistic and physically based.

The approach is an application of Bayes' theorem, in which a prior expectation of cloudiness (climatological or taken from NWP cloud fraction) is modified in the light of the observation vector. The observation vector, $\mathbf{y^o}$**,** comprises radiances (for IR, expressed as BTs to linearize the problem somewhat) and perhaps texture measures, such as local standard deviation over 3x3 pixels. Prior knowledge of the background state (SST + atmospheric profiles), $\mathbf{x}^b$, is used together with a fast forward model for the sensor (such as RTTOV or OPTRAN) and knowledge of the instrumental error characteristics to define a probability density function for the prior (background) observation vector given clear sky (c). This pdf is gaussian. Very non-gaussian is the pdf given cloudy conditions, which must be predefined empirically – but only once for a generic type of sensors (e.g., for AVHRRs, ATSRs and GOESs). Bayes' theorem is applied to deduce an estimate of the clear sky probability of the pixel given the prior knowledge and the observation vector. See Figure 5.

Bayes' theorem

- Observation $\mathbf{y}^o \pm \varepsilon^o$ e.g. BTs, 3x3 SDs, ...
- Prior → FFM → $\mathbf{y}^b(\mathbf{x}^b) \pm \varepsilon^b$
- Calculate $P(\mathbf{y}^o \mid \mathbf{x}^b, c)$
- Empirically $P(\mathbf{y}^o \mid \mathbf{x}^b, \neg c)$

- Clear-sky probability is

$$P(c \mid \mathbf{y}^o, \mathbf{x}^b) = \left\{ 1 + \frac{P(\neg c) P(\mathbf{y}^o \mid \mathbf{x}^b, \neg c)}{P(c) P(\mathbf{y}^o \mid \mathbf{x}^b, c)} \right\}^{-1}$$

Figure 5. Bayes' theorem adapted to estimating clear-sky probability given observations and prior.

Calculating an estimate of probability of clear sky allows flexibility in trading coverage against clear sky confidence. For some users, such as those looking to assimilate satellite SSTs in ocean models, avoiding cloud-related cold biases is more important than maximising coverage, and only pixels with the highest clear sky probability could be used, e.g. >99.9%. A 99% threshold would probably be more typical for most users. Figure 6 compares standard (threshold-based) cloud screening and Bayesian cloud screening (P<99% is masked) for an ATSR-2 ocean scene. The Bayesian masking is using only climatological information as prior knowledge, and therefore this is a fair comparison with a thresholding approach, since the thresholds should be designed to embody climatological knowledge of clear-sky observations. RTTOV-7 is used here as the fast forward model.

There is no doubt that the ATSR-2 threshold-based cloud mask could be adjusted to be as efficient as the Bayesian mask in this, or in any other, example. The importance of the Bayesian approach is more its rigour, and, particularly important in an operational context, its flexibility

- to take advantage in a direct way any prior knowledge that is available
- to be updated to new sensors with minimal change (so long as the fast forward model is available)
- to continue operating optimally after unforeseen events, such as channel losses

The laborious art of designing a sequence of threshold tests that is both efficient and general is bypassed – this is effectively done dynamically by applying the fast forward model to the background state. Adapting the scheme to a different sensor principally involves changing the sensor-selection flag in the fast forward model (if the channels are very different, the empirical pdf of cloudy observations would also need to be updated; but this is still probably less effort than devising a new series of threshold tests).

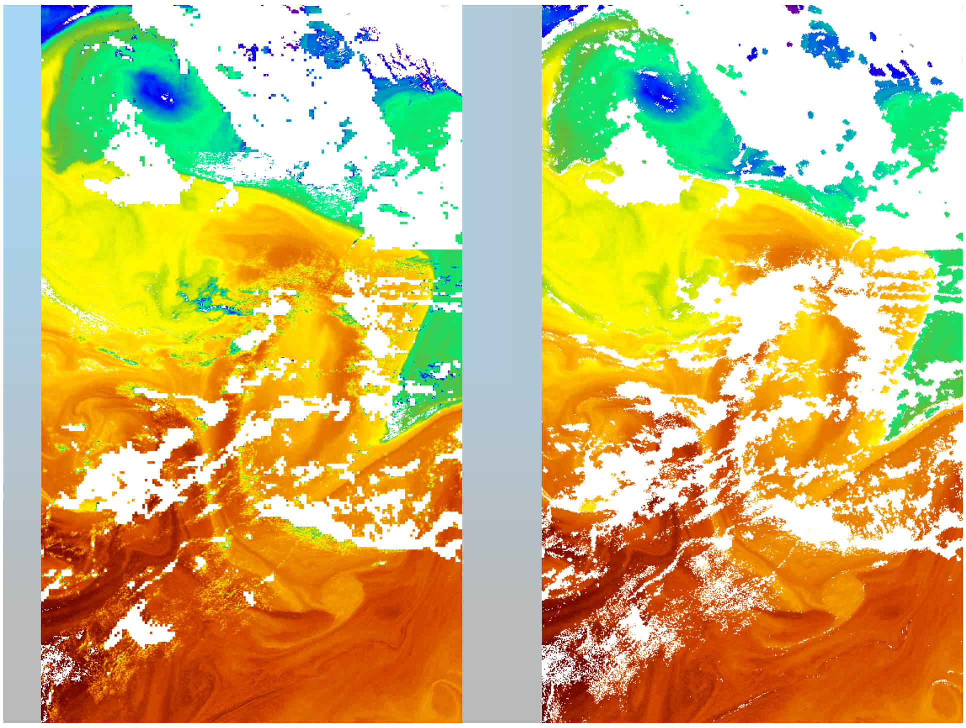

Figure 6. Standard (SADIST, threshold-based) cloud screening applied to an ATSR-2 scene (left) and the same scene masked using a Bayesian approach using climatological SST and TCWV as a prior. Scattered cloud regions are more successfully masked, as is an area of cloud over a cool patch of ocean. Strong thermal fronts tend to be mistaken for cloud because local standard deviation is included in the observation vector.

The other aspect of SST estimation is the retrieval once clear pixels have been identified. In the infrared case, this is increasing done using coefficients based on radiative transfer simulation for clear skies. This has a number of advantages over statistical regression based on matches:

- ‘day one’ retrieval capability
- globally representative
- retrieved parameter is well-defined (skin SST)
- no propagated buoy and match-up errors

The disadvantage of the physically-based approach is that it requires good instrumental characterization to be available and accurate simulation. For thermal window channels for SST, experience suggests that the systematic error in forward modelling a particular sensor channel from both these sources is of order +/-0.1 K. While this is pretty accurate, this systematic error is amplified by the factor $\sqrt{\sum_c a_c^2}$, where a_c is the SST retrieval coefficient for the cth channel. This means that RT-based SSTs can have biases of the order of +/-0.4 K for split window and +/-0.2 K for triple (i.e., using all three channels). Thus, to integrate the observations of different sensors into unified products requires an empirical bias adjustment of order a few tenths of kelvin [8].

3. INTEGRATING OBSERVATIONS

With the high level of coverage from geostationary sensors and the all-weather capability of microwave SST, in addition to the population of polar orbiters, there is the potential for a large number of observations of SST to be made for any particular place in the course of each day. This raises the question of how these diverse observations should be merged and used to create SST products that have enhanced accuracy and spatial and temporal coverage compared to any individual contributing data set.

As is clear from the previous section, the retrievals will need global bias correction on a common basis, which is straightforward in concept. Different sensors also display differing regional and seasonal bias patterns which are less easy to remove in the case of statistical schemes, although they can be successfully simulated in the radiative transfer framework.

After bias correction, merging of SST observation fields should proceed in a way the produces the 'best' quality final product. Two obvious strategies are (i) select the best (lowest error) estimate at each location (ii) use all the observations weighted appropriately to their error. The latter should be preferred, provided the bias correction and retrieval error estimation are done with sufficient care. The question, however, is complicated by the differing spatial scales and timing of the observations, as well as the data gaps that will inevitably exist. Optimal interpolation with a background from a prior analysis may be performed, or some other method of assimilation; practically, however, such approaches will be challenging given the number and diversity of inputs.

The Global High Resolution SST (GHRSST) project [9] is a co-ordinated international effort to address these questions as part of the Global Ocean Data Assimilation Experiment (GODAE). The project aims to use all routinely available data sources, including in situ and AMSR and AATSR.

There is another complication, namely, real geophysical variability of the radiometric temperature of the ocean. Until recently, this has effectively been treated as an extra source of noise on the satellite estimate of mixed layer temperature. The nature of this variability is described in the next section.

4. NEAR SURFACE TEMPERATURE VARIABILITY IN THE OCEAN

There are two effects that must be considered – cool skin effects and diurnal warm layer effects [10].

The cool skin effect is that the temperature of the air sea interface is a few tenths of kelvin cooler than the 'foundation' temperature a few millimetres below. Infrared observations are sensitive to temperature of the upmost 0.1 to 0.01 mm (at wavelengths around 4 and 11 microns respectively), while microwave observations sense the upmost ~1 mm. There are therefore systematic differences between the SSTs that IR and MW sensors 'ought' to retrieve. However, the variability is usually modest (skin effect in range 0.1 to 0.5 K) and the magnitude of the effect is similar to or smaller than the bias correction (which must be done with respect to foundation rather than radiometric SST).

More intermittent, but more significant when it occurs (up to ~3 K), is the diurnal warm layer effect. Unlike the skin effect, diurnal temperature excursions are not a quasi-instantaneous response to surface wind and solar irradiance, but depend on the integrated history over the day of the balance between stratifying effects (surface irradiances, freshwater influx) and mixing by wind stress (Figure 7). 1 K excursions are common in the tropics and summer hemisphere [11], and can be readily observed in SST differences between paired day and night overpasses (Figure 8).

Diurnal warming can be successfully modelled. Relevant observables are available from satellite retrievals routinely. In principle, then, an assimilation approach is feasible to constrain the trajectory of skin SST through the course of a day (and to infer the mixed layer temperature below the warm layer). This is probably an excessive effort for obtaining less 'noisy' SST, but may be interesting in itself in the context of better understanding air sea heat fluxes and other fluxes such as carbon dioxide.

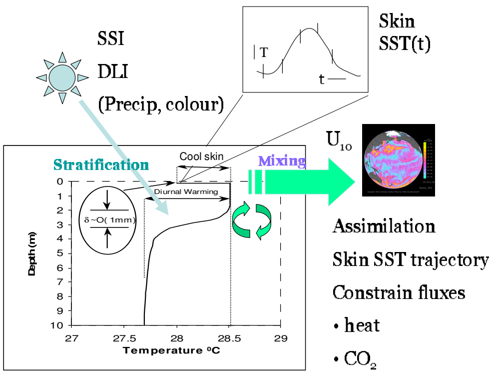

Figure 7. Diurnal and cool skin effects, and the observables relevant to these effects.

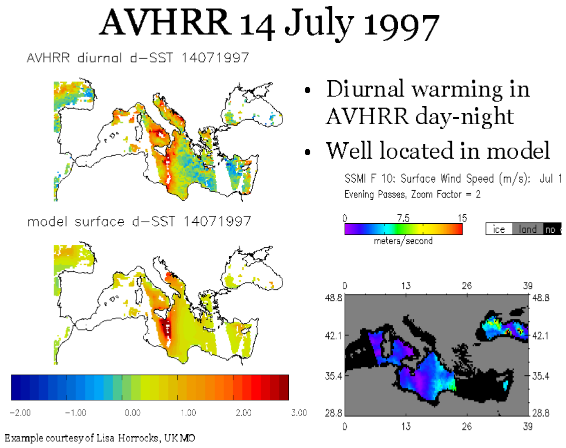

Figure 8. Upper left: difference in AVHRR 14 SST between day (2 pm) and night overpasses, showing a strip of diurnal warming north and south of Sicily. Lower left: model of diurnal warming at 1 m forced by NWP, showing the modelled warm layer is relatively well predicted (although the skin temperature excursion is exaggerated; not shown). Lower right: SSMI wind field showing that the diurnal warming correlates with a low wind speed region.

5. ACKNOWLEDGEMENTS

The presentation and this article were prepared with the contributions of many working in the field of satellite SST, whom I gratefully acknowledge: Craig Donlon (GHRSST), Chelle Gentemann (Remote Sensing Systems), Stephen Hallsworth (U. Edinburgh), Keith Haines (Earth Systems Science Centre), Lisa Horrocks (UK Met Office), Pierre LeBorgne (Meteo-France), David Llewellyn-Jones (U. Leicester), Peter Minnett (RSMAS), Alice Stuart-Menteth (Southampton Oceanog. Centre). Work at University of Edinburgh is funded by NOAA-NESDIS and the NERC, and the author has benefited from the EUMETSAT Visiting Scientist scheme.

6. REFERENCES

1. McCLAIN E P, W G Pichel and C C Walton (1985) Comparative performance of AVHRR-based multichannel sea surface temperatures. J Geophys Res 90, pp1587.
2. MERCHANT C J, A R Harris, M J Murray and A M Zavody (1999) Toward the elimination of bias in satellite retrievals of sea surface temperature 1. Theory, modelling and interalgorithm comparison. J Geophys Res 104, pp23565.
3. MERCHANT C J and P LeBorgne (2001) Designing OSI-SAF sae surface temperature retrievals to be robust to stratospheric aerosol. Proc EUMETSAT 2001 Met Sat Data Users' Conf, Antalya, Turkey October 2001, pp168.
4. LLEWELLYN-JONES D and Remedios J, personal communication, 2003.
5. WENTZ F, Gentemann C, Smith D and Chelton D (2000) Satellite measurements of sea surface temperature through clouds. Science 288 pp847.
6. SAUNDERS R S and Kriebel K T (1988) An improved method for detecting clear sky and cloudy radiances from AVHRR data. Int J Rem Sens, 9 pp123.
7. MERCHANT C J, MacCallum S, Harris A R and Maturi E (2003) Probabilistic physically-based cloud-screening for observing sea surface temperature from space. In preparation for Quat J Roy Met Soc.
8. MERCHANT C J and LeBorgne P (2003) Retrieval of sea surface temperature from space, based on modelling of infrared radiative transfer: Capabilities and limitations. Submitted to J Atmos Ocean Technol.
9. www.ghrsst-pp.org
10. FAIRALL C W, Bradley E F, Godfrey J S, Wick G A, Edson J B and Young G S (1996). Cool-skin and warm-layer effects on sea surface temperature J Geophys Res 101, pp1295.