Harmonization of space-borne infra-red sensors measuring sea surface temperature

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open access


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: https://www.mdpi.com/2072-4292/12/6/1048/pdf
To link to this article DOI: http://dx.doi.org/10.3390/rs12061048

Publisher: MDPI

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.
Harmonization of Space-Borne Infra-Red Sensors Measuring Sea Surface Temperature

Christopher J. Merchant 1, 2, *, Thomas Block 3, Gary K. Corlett 4, Owen Embury 1, 2, Jonathan P. D. Mittaz 1 and James D. P. Mollard 1

1 Department of Meteorology, University of Reading, Reading RG6 6BB, UK; o.embury@reading.ac.uk (O.E.); j.mittaz@reading.ac.uk (J.P.D.M.); j.mollard@reading.ac.uk (J.D.P.M.)
2 National Centre for Earth Observation, University of Reading, Reading RG6 6BB, UK
3 Brockmann Consult GmbH, 21029 Hamburg, Germany; tom.block@brockmann-consult.de
4 EUMETSAT, 1 Eumetsat Allee, 64295 Darmstadt, Germany; Gary.Corlett@eumetsat.int
* Correspondence: c.j.merchant@reading.ac.uk; Tel.: +44-118-378-7335

Received: 28 February 2020; Accepted: 21 March 2020; Published: 24 March 2020

Abstract: Sea surface temperature (SST) is observed by a constellation of sensors, and SST retrievals are commonly combined into gridded SST analyses and climate data records (CDRs). Differential biases between SSTs from different sensors cause errors in such products, including feature artefacts. We introduce a new method for reducing differential biases across the SST constellation, by reconciling the brightness temperature (BT) calibration and SST retrieval parameters between sensors. We use the Advanced Along-Track Scanning Radiometer (AATSR) and the Sea and Land Surface Temperature Radiometer (SLSTR) as reference sensors, and the Advanced Very High Resolution Radiometer (AVHRR) of the MetOp-A mission to bridge the gap between these references. Observations across a range of AVHRR zenith angles are matched with dual-view three-channel skin SST retrievals from the AATSR and SLSTR. These skin SSTs act as the harmonization reference for AVHRR retrievals by optimal estimation (OE). Parameters for the harmonized AVHRR OE are iteratively determined, including BT bias corrections and observation error covariance matrices as functions of water-vapor path. The OE SSTs obtained from AVHRR are shown to be closely consistent with the reference sensor SSTs. Independent validation against drifting buoy SSTs shows that the AVHRR OE retrieval is stable across the reference-sensor gap. We discuss that this method is suitable to improve consistency across the whole constellation of SST sensors. The approach will help stabilize and reduce errors in future SST CDRs, as well as having application to other domains of remote sensing.

Keywords: optimal estimation; satellite inter-calibration; harmonization; sea surface temperature; infra-red remote sensing; parameter estimation; retrieval theory; error covariance

1. Introduction

Sea surface temperature (SST) is the temperature of sea-water measured at or close to the air-sea interface. An analysis of the field of SST across the Pacific Ocean was first obtained by infrared (IR) remote sensing from space in 1970 [1] and in the same year was published the concept of better retrieving such SST estimates using two channels that are differently sensitive to atmospheric conditions [2]. During the intervening years, SST products informed by satellite observations have become widely used. Operational numerical weather prediction (NWP) and oceanography have become indispensable services to advanced economies, and these rely heavily on SSTs routinely obtained by a constellation of sensors on polar-orbiting and geostationary platforms [3]. Scientific applications in climatology and marine ecology demand SST information that is highly consistent and comparable over time and space. For example, understanding and monitoring the impact of thermal stress on coral reefs benefits from
long time series of stable SST observations, because corals are long-lived and adapted to the marine climate of past decades [4]. Increasing sophistication of retrieval methods [5–7] has delivered reducing SST uncertainty [8,9] and greater spatio-temporal resolution of SST analyses (analyses are gap-filled estimates of the SST field) [10,11].

SSTs retrieved from different sensors are combined when making either operational SST analyses or climate data records (CDRs). Bias adjustments between sensors are used in both processes, to avoid artefacts both in space (e.g., spurious fronts at the “joins” between differently biased data sources) and time (spurious temporal trends reflecting the different biases or sampling conditions of different sensors over time). Bias corrections imposed at higher levels of processing may introduce complex and poorly understood error characteristics. It is generally preferable to ensure that both radiances (or brightness temperatures) and swath (level-2) SSTs are as far as possible highly consistent. Brightness temperature harmonization methods for SST-relevant sensors have been proposed [12], aiming to improve usability of level-1 data (swath radiances). In this paper, we present a harmonization approach that specifically aims to minimize SST biases between sensors for the purpose of combining their observations in a CDR.

2. Context, Data and Methods

2.1. Gap in Dual-View Reference Sensors for SST

The context for this work is development of a multi-decadal CDR of SST within the European Space Agency (ESA) Climate Change Initiative (CCI) [13]. The SST CDR has been released in version 2 [14], and the work presented here is in preparation for version 3.

A feature of our objectives for the CDR is to maximize the degree of independence from in situ SST measurements [15], which is achieved by using selected sensors as on-orbit references. The reference sensors are the Along-Track Scanning Radiometers (ATSRs) [16] and the Sea and Land Surface Temperature Radiometers (SLSTRs) [17], chosen for their dual-view capability, low noise and two-point on-board thermal calibration system. The dual-view capability is extremely effective in delivering information that enables compensation for atmospheric influences on SST errors [18,19]. The ATSR-series observations covered the period August 1991 to April 2012, with three sensors, the last of which was advanced ATSR (AATSR). This series supported accurate, stable SST timeseries [20]. The SLSTR series started with the launch of the SLSTR-A unit on Sentinel 3A [21,22], with data available since November 2016.

There is therefore a gap between the sensors we wish to use as SST references, of 4 years 7 months. A necessary step towards harmonization of the SST constellation used in the CDR is therefore to bridge this gap with observations that are maximally consistent with the ATSR and SLSTR retrievals of SST. Achieving this is the specific focus of the work reported here.

There being no dual-view sensors available during the AATSR-SLSTR gap, we aim to harmonize a single-view sensor to the reference sensors. The requirements for this bridging sensor are that it must have at least a year of overlap with both AATSR and SLSTR and it must be in a stable orbit with a local equator crossing time close to the 1000 h time of AATSR and SLSTR and with a repeat pattern that ensures all-latitude coincidences between the single-view and reference sensors with good regularity. The possibilities fitting these criteria are the Advanced Very High Resolution Radiometer (AVHRR) on the MetOp-A platform and the Moderate-resolution Imaging Spectroradiometer (MODIS) on the Terra platform. This study addresses the MetOp-A AVHRR, which we hereafter refer to as MTA. Published information [23] suggests that the IR bands of MTA were (at the time of that assessment) stable and “in family”.

All of the target sensors (AATSR, SLSTR and MTA) for this study have channels with nominal wavelength centers of 3.7, 11 and 12 µm. Presently, the latter two channels are used for SST from daytime scenes, while all three channels are used for nighttime scenes. This avoids errors introduced by scattering of solar radiation in the 3.7 µm channel when viewing daytime scenes. The two channel
(so-called “split window”) retrieval configuration for SST has a long heritage [24,25], and split-window coefficient-based algorithms for this case are still proposed and discussed [26]. It is established that the two split-window channels have inadequate independent information content for retrieval of SST without sensitivity to atmospheric conditions [27] and the various more sophisticated retrieval methods all, implicitly or explicitly, attempt to compensate by introducing additional information from contextual (i.e., prior) knowledge, such as latitudinal variations in mean atmospheric conditions. In the SST CCI project, we have applied optimal estimation [5,28] for SST retrieval for AVHRRs. Optimal estimation (OE) explicitly exploits prior NWP fields of the atmosphere and fast radiative transfer modelling to obtain the weighting of observed brightness temperatures that gives the minimum error variance solution for the SST, assuming zero-mean biases in both the prior and the observed brightness temperatures. To bridge the dual-view sensor gap, we need to ascertain certain OE parameters (discussed further below) that bring SSTs obtained from MTA by OE into maximal consistency with SSTs from AATSR and SLSTR.

2.2. Data

To harmonize MTA with the dual-view sensors we use a dataset of clear-sky satellite-satellite matches [29]. Briefly, this multi-sensor match-up database (MMD) consists of matches between MTA and AATSR or SLSTR, where a match is defined as an acquisition from two sensors collocated to within a maximum allowed geodesic distance (here, 2 km) and a maximum allowed time difference (here 7200 s). We do not constrain the matches to near-nadir cases (so-called “simultaneous nadir overpasses”), since all satellite zenith angles need to be harmonized. Files output by the matching process contain all variables of the input satellite data extracted on a 7 × 7 pixel window around each match’s center-location. Each MMD match is augmented by ERA-Interim NWP data [30] interpolated to the AATSR/SLSTR spatiotemporal location using Climate Data Operators (CDO) [31].

To avoid large patches of data where the orbits of the sensors completely overlap, we use a three-dimensional (longitude, latitude and time) Sobol-Sequence based pseudo-random sampling strategy [32] to generate match-up locations. Furthermore, we remove all overlapping extraction windows from the dataset to ensure statistical independence of the matches. Matches are not kept where more than 10% of the extraction window pixels are over land.

The reference SST for each match is derived from the dual-view three-channel retrieval from the reference sensor, hereafter referred to as the D3 SST, extracted from the AATSR or SLSTR product. During the time-window of the MTA-reference match, the diurnal cycle in SST [33] will on average change the SST by up to ~0.1 K [34]. This is significant in this context. The retrieved SST of the reference sensor is therefore adjusted for the climatological diurnal cycle using a climatological parameterization informed by geographical location, time of day and wind speed [35]. Most adjustments are <0.05 K in magnitude. The diurnally adjusted D3 SST of the reference sensor is the reference SST for tuning the MTA retrieval.

Compared to other sources of SST that may be used as a reference, the choice of a reference SST based on satellite observations has strengths and weaknesses. The most obvious alternative would be SSTs measured in situ by drifting buoys. Although individual drifting buoys are not (for the most part) specifically calibrated or SI-traceable (with some exceptions [36]), it seems reasonable to assume that the multiple sensors across the drifting buoy network have a low average bias (because of compensating independent errors). However, drifting buoys measure the SST at a point location and at a depth of tens of centimeters (the exact depth varying with the drifting buoy design and sea state); they do not directly measure the skin SST over a pixel of ≥1 km², to which MTA is actually sensitive. The D3 SST has the advantage of representing a skin SST estimate on the appropriate spatial scale, and in three-way error analysis was found to have lower uncertainty that drifting buoys [37]. Disadvantages of using D3 SST include: using a single sensor means that systematic instrument and retrieval errors are not averaged down; because the MTA-reference matches are generally within 2 h of each other, any atmosphere-related errors in D3 SST may give rise to correlated errors in MTA SSTs.
Either drifting buoys (adjusted for skin effects) or a low-bias satellite SST, like D3 SST, could make sense, and here the choice of D3 SST is made because we want the bridging sensor to be independent from in situ measurements, like the D3 SST from the dual-view reference sensors.

Direct validation of the AATSR and SLSTR skin SSTs is possible using matches to in situ radiometer measurements. The point-to-pixel difference remains, but the in situ and satellite measurements are identically defined measurands. Radiometer matches are less numerous than matches to buoy measurements of SSTs, but the in situ uncertainty is lower, of order 0.1 K [38,39]. Evidence that the AATSR and SLSTR dual-view retrievals provide a good satellite reference for skin SST is obtained by plotting the median satellite-radiometer SST difference against satellite-radiometer time difference (Figure 1). The nighttime matches in Figure 1 are around 2200h UTC, at which time the ocean is generally gradually cooling, which is why there is a negative slope with respect to time difference. The intercepts at 0 h difference are for both AATSR and SLSTR within ±0.05 K of zero and the cross-over of the day and night curves occurs around 0 h time difference, as it should.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Comparison of (upper) advanced along-track scanning radiometer (AATSR) and (lower) sea and land surface temperature radiometer (SLSTR) dual-view skin sea surface temperature (SST) retrievals with matched in situ radiometer measurements. Median satellite-radiometer SST difference is plotted against satellite-radiometer time difference. The data are binned in time difference intervals of 7.5 min. The noisier appearance of the SLSTR-A curves is due to a lower accumulation of matches during SLSTR-A’s shorter lifetime to date.

Although drifting buoys are not used here as a reference, comparisons of the D3 SSTs to drifting buoys may be used as a further confirmation of the general validity of the D3 SSTs. The D3 SST from AATSR compares to SST from matched global drifting buoys (collected from the global telecommunication system and passing standard Met Office quality controls), adjusted for the skin effect, with median and robust standard deviation 0.01±0.18 K [40]. The D3 SST retrieval for SLSTR is derived identically to that of AATSR, except that SLSTR spectral response information is used. D3 SST from SLSTR, likewise adjusted for the skin effect, global drifting buoys with median and robust standard deviation 0.00±0.21 K [40].

These strands of evidence from radiometer and drifting buoy comparisons point to AATSR and SLSTR-A D3 SSTs having highly consistent properties, both being suitable to act as references for MTA at the start and end of its life respectively.

MTA-AATSR matches were obtained from 1 January 2010 to 8 April 2012 (end of AATSR overlap). The clear-sky subset of the matched data is determined using the results of probability-of-clear-sky calculations using a Bayesian method [41]. This method calculates the clear-sky probability given prior information \((x_a)\) and the satellite observations \((y)\) for every pixel, \(P_c = P(\text{clear}|x_a, y)\). A cloud mask may be defined by setting a single threshold on \(P_c\). To minimize the influence of clouds, we required...
min\( (P_c) > 0.9 \) for every pixel in a 3-by-3 box in both the reference sensor and the matched MTA. Since
the center pixels only are used for the harmonization calculations, this gives an additional buffer of a
pixel against the influence of residual cloud contamination. The central AATSR pixel in each match,
whose D3 SST acts as the SST reference, is required to have quality level 5 (best). The MTA satellite
zenith angle is restricted to <45\(^\circ\) for the SST harmonization (although SST retrievals at higher angles
remain useful). This is because of different bias effects that become apparent at greater angles, which
we hypothesize are a consequence of the absence in the forward model of enhanced reflection of
downwelling thermal irradiance \[42,43\]; the preferable solution is to develop the forward model rather
than to fold these effects into a bias correction. The result is a set of 51815 satellite-satellite matches to
be used for harmonization.

MTA-SLSTR matches were obtained for 2017 and equivalently filtered, generating 7371 matches
for SST harmonization.

For an independent test of the success of the harmonization process, a global sample of 94975 MTA
matches to drifting buoys are used, evenly spanning the period Nov 2006 to Dec 2018. The matches are
all for drifting buoy locations reported to be within the corresponding Metop-A pixel and within two
hours (mostly within one hour). To assess only cases with high probability of validity, all cases satisfy
\( P_c > 0.9 \). For comparison to the satellite skin SST retrieval, the drifting buoy SST is simply adjusted
by subtracting an average skin effect value of 0.17 K. Of the matches, 36% are night-time matches for
which 3-channel retrieval is possible, while for the remainder of matches the satellite SST is retrievable
from only the 11 and 12 \( \mu \)m channels.

2.3. SST Retrieval by Optimal Estimation

As in previous work \[5,44\] the formulation of OE for SST is in terms of a reduced state vector,
\( z^T = [x, w] \), where \( x \) is SST and \( w \) is total column water vapor (TCWV). The TCWV is varied by
scaling the NWP absolute humidity profile by a common fraction at all levels. The forward model for
brightness temperature given the prior information is RTTOV v11.3 \[45\], and this is run on the full (not
reduced) state vector to produce simulated values of brightness temperature, \( F \). RTTOV also generates
the partial derivative matrix, \( K \), which contains the rates of change of brightness temperature for each
channel with respect to the state vector variables. We reduce the dimensions of \( K \) to match the reduced
state vector by integrating the humidity partial derivatives to be with respect to fractional change in
TCWV. Given prior and observation-simulation error covariance matrices (discussed further below),
the OE retrieval, \( \hat{z} \) is calculated as an update to the prior state, \( z_a \), in the light of the difference
of the observations \( y \) from the simulations:

\[
\hat{z} = z_a + \left( K^T S_e^{-1} K + S_a^{-1} \right)^{-1} K^T S_e^{-1} (y - F).
\]

(1)

The necessary bias correction and error covariance parameters for Equation (1) are what we must
derive to harmonize MTA SSTs by OE with the dual-view reference sensors (as described in the next
section).

The D3 SSTs for AATSR and SLSTR are derived using retrieval coefficients rather than OE. This is
valid because of their additional, geometry-based information content for SST retrieval. The coefficients
are derived from simulations using a line-by-line layer-by-layer radiative transfer model (LBLRTM, \[46\]),
the advantage of this being smaller simulation errors than the fast model, RTTOV, that must be used for
OE. The basis in radiative transfer \[47\] and design of the coefficients \[48\] are as previously described
for AATSR and have been applied to the spectral response functions of SLSTR.

2.4. Parameter Estimation Method

The parameters to be derived for the MTA optimal estimator are:

- the brightness temperature bias in each of the 3.7, 11 and 12 \( \mu \)m channels, \( \beta \);
- the bias in the prior total column water vapor (TCWV) from NWP, \( \gamma \);
• the error covariance matrix of the difference between observed and simulated brightness temperatures, \( S_\epsilon \); and
• the error covariance matrix of the prior information, \( S_a \).

The method by which these parameters are estimated is fully described elsewhere [49], the difference here being the application to a satellite SST reference rather than an in situ reference. Briefly, the parameter estimation involves sequential iterative application of the following equations, which include the above terms, until convergence is obtained.

\[ \beta \] and \( \gamma \) are retrieved using Equation (1) adapted by extending the state vector to include these parameters as retrieved terms. Corresponding extensions to the tangent linear matrix and the error covariance matrix for the state are also made. In this mode, the state (SST and TCWV) and the parameters are retrieved for a match starting from the D3 SST as prior SST and the NWP water vapor. The D3 SST is assumed to be unbiased and individual values have relatively low uncertainty [8], which provides the information to constrain the bias correction values. Convergence is reached after ~100000 iterations of random draws (allowing repeat draws) from the set of matches.

Having reduced the prior and observation biases through estimation of bias correction values, new estimates for these are obtained using the following expressions:

\[
\hat{S}_\epsilon = \frac{1}{2} (d'_\epsilon d'_\epsilon^T + d'_\theta d'_\theta^T),
\]

\[
\hat{S}_d = \frac{1}{2} \langle (K^T K)^{-1} K^T (d'_s d'_s^T + d'_a d'_a^T) K (K^T K)^{-1} \rangle,
\]

where

\[
d'_\epsilon = y - F'(\hat{z}) - \langle y - F'(z) \rangle, \quad d'_\theta = y - F'(\hat{z}_a) - \langle y - F'(z_a) \rangle, \quad d'_s = F'(\hat{z}) - F'(z) - \langle F'(\hat{z}) - F'(z) \rangle.
\]

The difference terms, \( d'_\epsilon, d'_\theta \) and \( d'_s \), are combinations of observations, \( y \), and the forward model (bias corrected using \( \beta \)) \( F' \), run on the prior state \( (z_a) \) or the retrieved state \( (\hat{z}) \). These equations were introduced as diagnostic equations for data assimilation [50] and have been shown to be useful in iteratively finding error covariance matrix estimates [51].

The OE parameters are estimated as piecewise linear functions of variables on which they may plausibly depend. There is an element of expert judgement in selecting these variables. The brightness temperature bias correction parameter is determined as a piecewise linear function of \( \sec \theta \).

The rationale here is that errors in spectral response function can cause brightness temperature errors principally related to the path-integrated amount of water vapor viewed by an infra-red sensor in the window region, since water vapor is the main variable absorber in such channels. For compatibility, the observation error covariance is determined as a piecewise linear function of the same quantity. The prior TCWV correction and its error covariance are estimated as piecewise linear functions of TCWV.

3. Results

3.1. Harmonization to AATSR

Without determination of OE parameters, SST from MTA are biased relative to the reference SSTs from AATSR by an average of 0.23 K. The bias has a dependence on TCWV running from 0.14 K for retrieval through dry atmospheres to 0.31 K for the most humid clear-sky atmospheres. All these biases are outside the 0.05 K “1-sigma” target for measurement uncertainty in an SST CDR on 100 km scales [52]. SST harmonization is therefore required in order for MTA to function as a bridging reference sensor from AATSR to SLSTR.

After determining the OE parameters (Section 2.4), the global mean difference of MTA SSTs relative to the reference SSTs was 0.00 K. This is as expected, since this mean is assessed against the training data. (See the comparison to drifting buoy SSTs in Section 2.4 for an independent assessment.)
Note that the zero-bias is achieved via bias-correcting BTs and the prior TCWV, not by direct empirical regression to the target SSTs (as is often done). The mean differences are also calculated (not shown) for each decile of a number of other variables against which dependencies may typically arise. These variables are TCWV, satellite zenith angle, scan line element, latitude, wind, SST itself, day of year and time. The mean difference for every decile of every variable is within $\pm 0.045$ K of zero. The standard deviation of the differences is 0.22 K. The uncertainty for the D3 reference SSTs is expected to be in the range 0.12 to 0.16 K [8,37,48] and is estimated here (as the square root of an element in $S_f$) to be 0.13 K.

The OE parameters obtained from the harmonization to AATSR are now presented. Figure 2 (solid lines) shows the estimated uncertainty in and correction to the prior TCWV, as percentage for each decile of prior TCWV. Despite statistical fluctuations from decile to decile in the estimated uncertainty and correction, some plausible overall properties of the estimates are clear. The uncertainty in prior TCWV is largest in fractional terms, around 15%, for the driest atmospheres that mostly occur at high latitudes, while for the wettest atmospheres the fractional uncertainty is found to be less than 10%. Across the full range of TCWV, the estimated bias correction to prior TCWV is negative. Since the prior is a value for all-sky conditions (including the saturated areas of cloud) which is applied in the OE to cloud-cleared locations, it is expected that the prior TCWV should be too wet prior to correction. This expectation is consistent with these results, which suggest the clear-sky areas are 4% to 6% lower in TCWV than indicated by the prior.

Figure 2. Parameters for optimal estimation related to total column water vapor (TCWV), estimated per TCWV decile from harmonization of MTA to AATSR (solid lines) and SLSTR-A (dashed lines). Upper panel: prior TCWV uncertainty estimate, in percent. Lower panel: bias correction to be added to prior TCWV.

Figure 3 (solid lines) shows the estimated uncertainty and bias correction for the observation-simulation difference, addressing errors from both instrument and forward model. The estimates are made as a function of the integrated WV column observed along the slant path viewed by Metop-A. The results are fairly consistent between the 3.7 and 11 $\mu$m channels that have lower transmittances than the 12 $\mu$m channel, their estimated uncertainty being fairly constant around 0.15 K in comparison to an average of 0.29 K for the 12 $\mu$m channel. There is relatively high error...
correlation between the 11 and 12 µm channels, $r=0.93$, suggesting that forward model or calibration errors with a high degree of commonality are dominating over instrument noise (which is expected to be uncorrelated between channels).

![Observation-simulation uncertainty and correction](image)

**Figure 3.** Parameters for optimal estimation related to the difference in observed and simulated brightness temperature, for MTA channels at wavelengths around 3.7, 11 and 12 µm, estimated as a function of the slant column WV path along the view angle. Upper panel: uncertainty (arising from instrument and forward model errors). Lower panel: bias correction (to be added to forward model results). Solid lines: estimate with reference to AATSR. Dashed lines: estimate with reference to SLSTR-A.

3.2. Harmonization to SLSTR

Without determination of OE parameters, SST from MTA are biased relative to the reference SSTs derived from SLSTR by an average of 0.25 K. The bias has a dependence on TCWV running from 0.19 K for retrieval through dry atmospheres to 0.33 K for the most humid clear-sky atmospheres. These pre-tuning results are thus quantitatively similar to those from the matches to AATSR from ~6 years earlier.

After determining the OE parameters, the global mean difference of MTA SSTs relative to the reference SSTs was 0.00 K. As before, the mean differences are also calculated (not shown) for each decile of a number of other variables against which dependencies may typically arise. These variables are TCWV, satellite zenith angle, scan line element, latitude, wind, SST itself, day of year and time. The mean difference for every decile of every variable is within ±0.06 K of zero. The standard deviation of the differences is 0.22 K. The uncertainty for the SLSTR D3 reference SSTs is estimated here (as an element in $\tilde{S}_d$) to be 0.17 K, in line with expectations.

The dashed lines in Figure 2 show the results of the estimation of uncertainty in and correction to prior TCWV. The dashed lines in Figure 3 show the uncertainty in simulation-observation differences for the three thermal channels, and the brightness temperature corrections estimated with respect to SLSTR-A, as a function of the slant total column water vapor path. In all cases, the general magnitude, sign and variation of the parameters are broadly consistent with those obtained with reference to AATSR. The smaller set of satellite-satellite matches obtained between MTA and SLSTR-A limits the weight of interpretation that can be placed upon the differences in detail: the overall conclusion is that
MTA appears similar in comparison to both references, most obviously in the brightness temperature corrections obtained. The warm SST error apparent in untuned MTA SSTs arises from positive biases (relative to the references) in all thermal channels of MTA, being of order 0.2 K for the 3.7 and 11 μm channels, and of order 0.1 K for the 12 μm channel. The biases are more consistent between channels for low water-vapor scenes, and diverge when viewing more humid scenes.

3.3. Independent Assessment against Drifting Buoys over Gap Period

As an independent test of the OE parameters derived, we characterize the difference between MTA SSTs and matched, skin-adjusted drifting buoy SSTs. The MTA SSTs were derived by OE initialized from NWP data, whose prior uncertainty in SST relative to drifting buoys is 0.55 K.

Using the parameters derived from harmonization to AATSR D3 SSTs, the overall statistics satellite-minus-in-situ SST are shown in Table 1. Globally, they are satisfactory, the mean and median values being <0.1 K in magnitude. The standard deviation (SD) of 0.41 K is typical and reflects useful information gain relative to the prior. The value of the SD is influenced by the presence of outliers relative to a normal distribution, as shown by the significantly smaller robust SD (RSD) of 0.31 K. The outliers arise in both the satellite and in situ datasets. The estimated uncertainties for the OE retrievals are on average 0.21 K. Assuming uncertainty in the drifting buoy SST measurements of 0.2 K and geophysical variability of the skin effect of 0.05 K, the expected SD of difference between OE and drifting buoy SST is then \( \sqrt{0.21^2 + 0.2^2 + 0.05^2} = 0.29 \) K, which is similar in magnitude to the RSD.

Table 1. Statistics of difference between MTA SSTs derived by optimal estimation and matched drifting buoy SSTs (after skin adjustment).

<table>
<thead>
<tr>
<th>Harmonized Against</th>
<th>Mean/K</th>
<th>SD/K</th>
<th>Median/K</th>
<th>Robust SD/K</th>
<th>SST Sensitivity, %</th>
<th>Trend/K yr(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATSR</td>
<td>0.02</td>
<td>0.41</td>
<td>0.07</td>
<td>0.31</td>
<td>86</td>
<td>(-3.6 \times 10^{-3})</td>
</tr>
<tr>
<td>SLSTR</td>
<td>0.01</td>
<td>0.39</td>
<td>0.06</td>
<td>0.30</td>
<td>82</td>
<td>(-4.5 \times 10^{-3})</td>
</tr>
</tbody>
</table>

\(^1\) SD indicates standard deviation. The robust SD is calculated as the median absolute difference from median (“MAD”) scaled to match the value of SD for a Gaussian distribution. When the robust SD is less than the SD this indicates the presence of some outliers relative to a Gaussian distribution.

The OE retrieval result is influenced by the prior SST, which diminishes the sensitivity of the result to the true SST. It is desirable for the sensitivity to the true SST to be close to 100% \([6,27]\), and sensitivity is estimated within the OE framework \([5]\). Since the SST sensitivity is typically 86%, about 12% of any error in present in the prior SST is typically propagated to the retrieved value. Overall, the MTA SSTs are stable in time with respect to the in situ data, the least-squares trend fitted to the individual differences against time being \(-3.6 \) mK yr\(^{-1}\). The partitioning of this trend between the stability of the satellite and drifting buoy observing systems is not known, but the results suggest the satellite SST stability is close to the ambitious target confidence interval for stability of SST observation for climate applications stated by GCOS (±3 mK yr\(^{-1}\); \([52]\)).

The validation results using the parameters derived from harmonization of MTA SSTs to SLSTR D3 SSTs are highly consistent with the AATSR-based results (Table 1). This is consistent with a high degree of MTA SST stability combined with good agreement; it also is consistent with highly compatible relative calibration of AATSR and SLSTR-A, despite lack of overlap in time. At the expense of slightly lower SST sensitivity, the SLSTR-A-based results have marginally better validation statistics. The statistical (fitting) uncertainty in the least-squares trend is \(0.4 \times 10^{-3}\) K yr\(^{-1}\) and therefore the trend difference between the two harmonizations is not statistically significant.

Figure 4 shows greater detail of the performance of the satellite retrievals as a function of various independent variables of significance.
Figure 4. Dependence plots of satellite-minus-in-situ differences, calculated for deciles of independent quantities (~9500 matches in each decile): top left—satellite zenith angle; top right—latitude; middle left—total column water vapor; middle right—prior sea surface temperature; lower left—solar zenith angle at satellite observation time; lower right—time of observation in calendar years. Red line: mean difference for MTA SSTs using OE parameters harmonized against AATSR; red shading: corresponding standard deviation. Blue line: mean difference for MTA SSTs using optimal estimation (OE) parameters harmonized against SLSTR; blue shading: corresponding standard deviation.

Harmonized against either AATSR or SLSTR-A, the optimal estimates are highly stable against all independent variables: the dependencies are calculated for deciles of the independent variables, and the absolute mean differences for all deciles across all the variables are ≤ 0.1 K. Points to note from Figure 4 include: there is larger inter-decile variability against latitude and water vapor than for other independent variables, both arising from the same modest dependence of difference on TCWV; the SD is smaller for solar zenith angles corresponding to night-time (>90°), because of the extra information added by the 3.7 μm channel; there is a difference of about 0.1 K between day-time (two-channel) SSTs (mean ~ −0.05 K) and night-time (three-channel) SSTs (mean ~ 0.07 K).

4. Discussion

Being a single-view sensor with single-point on-board calibration, the AVHRR cannot match the performance of AATSR and SLSTR as an SST reference sensor. However, the results show that optimal estimation can be formulated with bias correction and well-estimated error covariance matrices to give results that are highly consistent with the dual view references, sharing the property of independence of the “calibration” of the retrieval from in situ data. This represents a significant step forward compared
to previous attempts to obtain SSTs from AVHRRs by both OE or regression-based coefficients, which have generally required empirical tuning to in situ data to achieve low (<0.1 K) biases.

MTA is shown to be a highly stable sensor for SST determination across the period 2007 to 2018, since OE formulations derived from either the AATSR overlap (using 2010 to 2012) or the SLSTR overlap (using 2017) both give stability of SST relative to drifting buoys around 0.004 K yr\(^{-1}\) across the whole period (although in principle both satellite and in situ observing systems could, by coincidence, have a much larger trend in common, a small relative trend between two highly independent observing systems is more likely explained by both observing systems being stable to a level comparable to their trend difference).

Part of the reason MTA delivers this SST stability is that it flies on a platform with an orbit that is controlled to maintain an ascending node time of 09.30 to within a few minutes. AVHRR instruments on platforms whose ascending node times drifted throughout their mission present additional difficulties for obtaining stable records [53,54].

The origins of the minor differences between “calibration” of the retrieval with reference to AATSR or SLSTR may include both statistical uncertainty and real instrument calibration effects. Given that the AATSR-based and SLSTR-based results are the best estimates available early and late in the MTA mission respectively, it is appropriate to interpolate the OE parameters linearly with respect to time across the reference-sensor gap.

Within the context of climate data record development, this work enables MTA SSTs to make an important contribution to overall stability and continuity of the SST climate data record. The combination of the ATSR-series, MTA (using the techniques herein) and the SLSTR-series gives a continuous record of satellite SST with very high levels of independence from in situ observations from 1991 to present. Dependence on in situ measurements of SST in this record is limited to the following: for the dual-view sensors, NWP fields are used to assist cloud-screening, and NWP fields have a dilute dependence on in situ SST; and in the case of MTA SSTs by OE, there is typical sensitivity to the NWP prior SST of 12% to 16%, which amounts to a more direct dependence on the in situ data used in constructing the NWP prior SST. Nonetheless, the dual-view SST retrievals are based on the physics of radiative transfer (not, as has often been done, by tuning to drifting buoys) and that basis in radiative transfer is propagated across to MTA through the parameter estimation.

MTA is additionally useful to bridge the AATSR-SLSTR gap because it is, like those references, in a mid-morning orbit. Another SST-capable mission in a controlled mid-morning orbit that overlaps AATSR and SLSTR is the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua platform (21.30 h). Additional thermal channels on MODIS are available that can easily be accommodated within the flexible framework of OE, and may improve the robustness of SST retrieval under (for example) desert dust conditions. We have not investigated MODIS as a candidate bridging sensor for purely practical reasons (computer resources), although we would be keen to do so in future.

The parameterization process gives estimates for two sources of bias: relative bias between sensor calibration and forward model in each of three thermal channels; and bias in the prior TCWV (when used for clear-sky areas). The former source is here characterized as a function of slant water vapor path and the latter as a function of vertical water vapor path (i.e. of TCWV itself). Since any bias in TCWV input to a forward model causes a bias in the simulated brightness temperature, there is potential for mis-attribution of bias between satellite/forward-model calibration versus prior TCWV. In particular, any component of calibration bias that lies along the same direction in multi-channel brightness temperature space as \( \frac{\partial y}{\partial w} \) cannot be distinguished from bias in the prior TCWV in a given match. To the degree that the iterative procedure has managed to separate these sources of bias, separation is possible because the two sources have different spatio-temporal patterns. The consistency of the results between the AATSR- and SLSTR-based harmonizations is encouraging, but is not a guarantee that the estimated prior TCWV biases (when used to represented clear-sky areas) are accurate. It remains possible that a component of the average calibration bias is co-linear with \( \frac{\partial y}{\partial w} \) and that an unknown portion of this bias is mis-attributed to prior TCWV bias.
A more ambitious development of the method presented here could be devised to use a larger constellation of contemporary sensors with pair-wise matches. The aim would be to parameterise individual sensor biases and derive a common estimate of prior state bias. Such a constellation approach would enable more confident discrimination between the calibration versus prior biases; we hope to pursue this possibility in future.

5. Conclusions

In developing CDRs from satellite observations it is often crucial to characterize and reduce systematic effects, since these typically become dominant sources of uncertainty for climate applications where multi-annual, regional-to-global scales are crucial. This work is a demonstration of the use of reference data to constrain biases, by deriving parameterized corrections for observation-minus-simulation biases and for biases in prior information. Moreover, in the context of optimal estimation, self-consistent prior and observation error covariance matrices are derived that make the OE retrieval closer to optimal.

Specifically, in the context of the CDR for sea surface temperature, we have demonstrated that a highly consistent SST record can be obtained from the dual-view reference sensors (ATSRs and SLSTRs) bridged across the AATSR-SLSTR gap using the MetOp-A AVHRR. This record is highly independent of in situ SST measurements, a feature that is unique to efforts within the ESA SST CCI and is useful in improving knowledge of the past evolution of this essential climate variable. We have harmonized the MTA SSTs with AATSR and SLSTR, and have assessed MTA SSTs against independent matched drifting buoy measurements. We find that the harmonized MTA SSTs have good validation statistics (biases <0.1, SD ≈ 0.4 K), good stability (relative trend ≈ −4 mK yr$^{-1}$), reasonable uncertainty estimates (not accounting for outliers), and high (~84%) SST sensitivity. These properties are delivered by brightness temperature and prior TCWV corrections that are physically plausible, in conjunction with new well-founded error covariance estimates.

The approach we have employed here will help stabilize and reduce errors in future climate data records of SST, as well as having application to other domains of remote sensing.


Funding: This research was funded by the European Space Agency Climate Change Initiative, grant number 4000109848/13/I-NB.

Conflicts of Interest: The authors declare no conflict of interest.

References


30. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 2011, 137, 553–597. [CrossRef]

31. Schulzweida, U.


