AATSR LST PRODUCT VALIDATION

Elizabeth Noyes, Simon Good, Gary Corlet, Xin Kong, John Remedios, and David Llewellyn-Jones

Department of Physics and Astronomy, University of Leicester, Leicester, LE2 1AJ, UK

ABSTRACT

Operational land surface temperature (LST) data from ESA’s Advanced Along-Track Scanning Radiometer (AATSR) are now available to users. The target accuracy of these data is 1.0 K at night, and 2.5 K during the day. We present the results of one experiment where the accuracy of the AATSR LSTs has been assessed over a site in Oklahoma, USA. The results indicate that the accuracy of the product is 1.2 K (standard deviation 1.2 K) at night and 0.5 K (3.3 K) during the day. The AATSR LSTs exhibit a seasonal bias over this site with the maxima of the cycle (AATSR minus ground) occurring during the summer. Similar seasonal biases are observed in comparisons between the AATSR LSTs and other data sets (e.g. MODerate Resolution Imaging Spectroradiometer). Investigations using the results of radiative transfer simulations suggest that the AATSR LST algorithm has a strong dependency on the concentration of atmospheric water vapour, which may be a contributing factor to the observed biases.

Key words: AATSR, LST, Validation, ARM.

1. INTRODUCTION

Demand for global observations of land surface temperature (LST) is increasing within the modern climate science community. As LST is sensitive to surface wetness, it can be a good indicator of the biophysical properties of the land, particularly if combined with vegetation indices [1, 2]. Repeated observations of LST can therefore be used to monitor long-term change in land use. LST is also a parameter that is potentially of great use within the modelling community, for example to constrain and validate geophysical models [3].

Several modern Earth-observing satellite sensors have the capability to provide regular, global observations of surface temperature. The instruments make measurements of the radiance at the top of the atmosphere in the infrared or microwave regions of the electromagnetic spectrum, from which LST can be ‘retrieved’. The focus of this paper is on data from the most accurate of the current infrared sensors, the Advanced Along-Track Scanning Radiometer (AATSR) on board the European Space Agency’s (ESA) Envisat platform. Since 10 March 2004, a global LST product has been produced operationally from data recorded by this instrument, which is now included in the reprocessed AATSR level 2 data (ATS_NR_2P) dating back to August 2002. These data are produced at 1 km resolution, with a target accuracy of 1.0 K at night and 2.5 K during the day. Full details of the algorithm utilised in the operational LST retrieval scheme can be found in [4].

The primary interest in satellite-derived LST data is focused around the geostationary sensors, which provide the information about the diurnal LST cycle that is required for many scientific applications. Arguably the greatest use for LSTs derived from an accurate polar-orbiting satellite instrument such as the AATSR, is as a validation/calibration source for LST derived from geostationary satellites. All satellite products must be rigorously validated by comparing them with equivalent ground based measurements to assess their accuracy and identify where improvements to the retrieval schemes are required. In practice, this is complex due to the up scaling of the point measurements to the scale of the satellite pixel (typically >1 km²). The larger pixel size of the geostationary (>3 km²) compared with polar-orbiting sensors (~1 km²) makes this even more challenging in the case of the former. A well validated polar-orbiting LST product could be a useful tool for assessing the accuracy of the equivalent geostationary products where in situ validation is not possible (e.g. at high latitudes where the geostationary footprint is at its largest). Polar-orbiting LST data also provide a consistent global reference data set with which to cross-calibrate the different geostationary LST data sets. For the AATSR, there is an additional benefit in the fact that this instrument is the third in the ATSR series, which spans almost 15 years in total. As the AATSR LST retrieval scheme can be applied to data from the ATSR-1 and -2, this provides a long-term consistent LST data set that could be used to ensure consistency between successive geostationary sensors.

In this paper, we present the results of an experiment where the accuracy of operational 1 km AATSR LST data has been assessed through comparison with in situ LSTs recorded over more than two years. We also present evidence to suggest that the AATSR LSTs are subject to seasonal variation in bias through these validation results, the comparison with other satellite and model data sets,
and the results of radiative transfer simulations. Finally, we present some issues relating to the auxiliary data currently used in the AATSR LST retrieval scheme that are compromising the accuracy and performance of the algorithm.

2. VALIDATION OF AATSR LSTS OVER THE ARM-OKLAHOMA SITE

Rigorous and continued validation of all satellite products is an absolutely essential part of any Earth observation mission. Ideally, the satellite data should be validated over long periods of time (> 1 year) to enable any seasonal effects to be identified. In this section, we present the results of a validation experiment over the Atmospheric Radiation Measurement (ARM) field site in Oklahoma, USA, where in situ temperature observations have been recorded continuously since December 2003. The site is located at 36.600°, -97.500°, and comprises cattle fields and wheat pasture. In the AATSR LST retrieval scheme, this site is classified as land type 12: broadleaf deciduous trees with winter wheat (see [4] for further information). Further details of the site can be found at http://www.arm.gov/sites/sgp.stm.

2.1. Validation Methodology

The in situ radiometric temperature data are recorded by a pyrgeometer located at the ARM site central facility and averaged in 1-minute bins. In this experiment, the in situ data used were matched to the exact time of the AATSR overpass and corrected for emissivity effects using data from the ASTER spectral library [5] and at-site observations of downwelling sky irradiance, also made using the pyrgeometer.

The ASTER spectral library contains reflectance spectra corresponding to visible and infrared wavelengths for almost 2000 natural and man-made materials. In this instance, average emissivities for soil (\(\epsilon_{soil}\)) and green vegetation (\(\epsilon_{veg}\)) were calculated using the spectral response of the pyrgeometer at the site. The in situ emissivity for each matchup was then estimated according to Equation 1:

\[
\epsilon = f\epsilon_{veg} + (1 - \epsilon_{soil})
\]

(1)

where \(f\) is the estimated fractional vegetation of the site, derived from AATSR-observed normalised differentiated vegetation index (NDVI - also an operational AATSR product) appropriate for the date of the matchup (see for example http://www.gis.usu.edu/mikew/awer6760/assignment9.html on how to derive \(f\) from NDVI).

Once the in situ LST has been calculated, a direct comparison is then made between this data point and the AATSR LST corresponding to the pixel nominally containing the in situ observation. Any matchups flagged cloudy in the AATSR data (see [6]) were rejected from this experiment.

2.2. Validation Results

The results are shown in Figure 1. Figure 1a shows the comparison for all data flagged cloud-free by the AATSR processor. The very cold outliers in these data suggest that some cloud is being missed by the operational cloud-screening algorithms; almost all of these cold outliers correspond to day time observations when fewer cloud test are implemented.

In light of this, additional cloud-screening was also carried out by eye and through further filtering of cold outliers lying below the mean AATSR-in situ LST difference minus 3\(\sigma\) (one \(\sigma\) = one standard deviation). The resulting ‘cloud-screened’ data set (50 day and 44 night) is shown in Figure 1b, where the bias (AATSR minus in situ) and standard deviation of these data are calculated to be 0.5 K and 3.3 K, respectively, for day time, and 1.2 K and 1.2 K, for night.

The bias in both the day and night time results is extremely encouraging. However, the high standard deviation is of some concern: for day time matchups, this reflects the seasonal cycle that is apparent in these data, where AATSR appears to be yielding LSTs that are cold biased in the winter and warm biased in the summer months. A seasonal cycle is also evident in the night time data, although the amplitude is not as large. This seasonal cycle is investigated further in the following sections of this paper.

3. SEASONAL BIAS IN THE AATSR LST DATA

3.1. Comparisons with other LST data sets

The results presented in Section 2.2 suggest a seasonal bias in the AATSR LST retrievals. A similar seasonal cycle is not observed in the equivalent experiment performed using LSTs derived from data recorded by the Moderate Resolution Imaging Spectroradiometer (MODIS) on board NASA’s Terra platform (Figure 2). Although the MODIS data do not exhibit any seasonal variation in bias, the average bias of these data when compared with the in situ LSTs is much greater than for the AATSR LST data. For MODIS, the bias and standard deviation are −4.0 K and 2.7 K, respectively for cloud-free day time matchups, and −2.0 K and 1.1 K for cloud-free night time matchups. This compares with biases for the AATSR of just 0.5 K and 1.2 K for day and night time matchups, respectively.

Biases of a seasonal nature are also seen in comparisons between MODIS and AATSR that have been performed
Figure 1. Validation results from the ARM field site in Oklahoma (biome 12) spanning the period December 2003 to December 2005 (inclusive). The y-axis shows the difference between the AATSR and collocated in situ LSTs. (a) shows the results where the data flagged cloudy by the operational cloud screening have been removed (b) shows the results after additional cloud screening. The red, dashed lines represent the day-time 2.5 K target accuracy of the AATSR LST product, and the blue, dashed lines the night time 1.0 K target accuracy.

Figure 2. Comparison between in situ LSTs obtained from the ARM Oklahoma validation site and co-located MODIS LSTs between July 2003 and June 2006. (a) shows all data points, and (b) shows the data after additional cloud-screening has been employed.
over a number of other sites located in Europe. Figure 3 shows the variation in the difference between AATSR and the equivalent MODIS LST data over a 0.25° box centred on 36.75°, 007.25°. For each matchup, the MODIS LSTs have been interpolated to the AATSR overpass times using a correction derived from the Spinning Enhanced Infrared and Visible Imager (SEVIRI) on board the geostationary satellite, Meteosat-8. Only data flagged cloud free by the operational processor for each sensor have been used in the averages; additional screening for cloud has been performed by fitting a seasonal cycle to the AATSR LST data and removing any extreme outliers (data lying outside of 1.5 standard deviations from the mean seasonal cycle).

The overall bias indicates that MODIS is colder than the AATSR by approximately 5.8 K (standard deviation 3.4 K) during the day and 3.6 K (2.2 K) at night. These figures are comparable with the differences in the bias of each LST product compared with the in situ data over the ARM validation site (MODIS is colder by 4.5 K during the day, and 3.2 K at night). There is a clear seasonal bias between the two sensors, where the peak of the cycle (AATSR minus MODIS) occurs during the summer months. Again, the cycle is apparent for both day and night time matchups, although the amplitude is smaller in the case of the latter. Similar results are obtained for nine other test sites located around Europe and North Africa.

The overall agreement between AATSR and SEVIRI is better than for MODIS, and we find that on average, SE-VIRI is colder than the AATSR by 1.3 K (standard deviation 2.2 K) during the day, and 0.1 K (1.3 K) at night. As for the AATSR ARM validation and AATSR-MODIS comparisons, a seasonal bias is apparent, where the maximum (AATSR minus SEVIRI) occurs during the summer months. Again, the amplitude of the cycle is smaller for the night time comparisons than during the day. Similar results are obtained for nine other test sites located around Europe and North Africa.

A third set of comparisons has been performed between AATSR LST data and night time 2.5°-gridded European Centre for Medium-Range Weather Forecasts (ECMWF) operational model skin temperatures. The comparison methodology utilised is similar to that described above, with the exception that the ECMWF data has not been averaged (as it is on a coarse-resolution model grid) and no temporal interpolation of the ECMWF data has been performed. Instead, the comparison site has been selected such that the time of the ECMWF observation is almost an exact temporal match to the time of the AATSR overpass (daily global ECMWF data fields are available at 00:00, 06:00, 12:00 and 18:00 GMT). In this instance, the 0.25° box utilised in the averaging of the AATSR data is centred on the ECMWF model grid point.

Figure 5 shows the results of the comparison between AATSR and ECMWF LSTs for a 0.25° box centred on 60.0°, 045.0°. The results demonstrate that the ECMWF data is on average 0.5 K warmer (standard deviation 2.4 K) than the equivalent AATSR LSTs. As for the AATSR-MODIS/SEVIRI comparisons, a seasonal cycle is apparent in these data, where the maximum occurs during the summer months (note the timescale on this plot runs from summer to summer, in contrast to the MODIS/SEVIRI plots which run from winter to winter).
Similar results are obtained for several other comparisons performed at other sites around the globe.

3.2. Investigation into Sources of Bias in the AATSR LST Retrievals

In this section, we report some preliminary investigations into sources of bias in the AATSR LST retrieval algorithm through the results of radiative transfer simulations. The methodology used in this experiment is that of [7].

In brief, AATSR top of atmosphere (TOA) brightness temperatures (BT) are simulated using a radiative transfer model for a given set of surface and atmospheric conditions at a particular location. In this case we have used the Oxford RFM (see http://www-atm.physics.ox.ac.uk/RFM/). The AATSR LST retrieval algorithm appropriate for the site conditions and time of year is then applied to these modelled BTs. For a perfect algorithm (under the test conditions), this ‘retrieved’ LST, hereafter referred to LST$_{retrieved}$, will be equal to the LST input into the radiative transfer model (LST$_{model}$). Any differences between these two quantities will therefore give an indication of the expected bias for the test conditions. The principal assumptions in this method are that we have an accurate radiative transfer model, and that the model parameters of surface temperature, channel-emissivity and atmosphere are realistic for that site for the chosen time of year. In the case of the latter, we have used ECMWF atmospheric and surface temperature data or the reference atmospheres provided by [8], with surface emissivities estimated from the ASTER spectral database and the AATSR instrument characteristics (channel response functions).

The results suggest that a seasonal bias is expected over this site as a result of atmospheric variation, where the AATSR LSTs are warm-biased during the summer months and cold-biased during the winter. Similar cycles are observed over a number of test sites examined in this experiment. The principal effects of the atmosphere on the AATSR channels used in the LST retrieval scheme are due to water vapour, atmospheric temperature and surface-air temperature difference. Figures 7, 8 and 9, respectively, show the sensitivity of the AATSR LST retrieval algorithm to changes in these parameters (in this instance for AATSR biome 12, but the results are similar for all biomes). In each case, the bias has been simulated for both the tropical and mid-latitude atmospheres provided by [8].

For Figures 7 and 8, the ideal response (gradient) would be zero. However, in both situations, we can see that this is not the case, and that the bias in the algorithm due to these parameters varies by 1-2 K. For the surface-air temperature difference, the ideal response would be unity, such that for every 1 K change in LST, the retrieved LST would also change by 1 K. From Figure 9 we see that this is not the case, and the bias is strongly dependent on the magnitude of this parameter. Considering the tropical atmosphere, for example, a bias of 1.1 K is predicted. However, under the test conditions, this bias would increase.
Figure 7. Simulated bias in the AATSR LST algorithm for biome 12 at tropical and mid-latitudes due to variation in total column water vapour. The values represented by the x-axis signify the % variation in total column water vapour (i.e. the entire water vapour profile has been perturbed by that amount).

Figure 8. Simulated bias in the AATSR LST algorithm for biome 12 at tropical and mid-latitudes due to variation in atmospheric temperature. The values represented by the x-axis signify the % variation in temperature at all heights (i.e. the entire temperature profile has been perturbed by that amount).

Figure 9. Simulated bias in the AATSR LST algorithm for biome 12 at tropical and mid-latitudes due to variation in surface-air temperature difference. The values represented by the x-axis signify the magnitude of the surface-air temperature difference (i.e. the atmospheric profile, including the near-surface air temperature, remains fixed and the underlying LST input into the model is increased or decreased by this amount).

The results of this study, although ongoing, suggest that the accuracy of the AATSR LST retrievals is very sensitive to variation in atmospheric conditions. This sensitivity is, at least in part, causing biases in the retrievals that are of a seasonal nature. Although not examined in this study, emissivity effects may also be a contributing factor to the seasonal bias in the AATSR LST retrievals observed in the real data. An investigation into these effects is currently underway.

4. AATSR LST AUXILIARY DATA ISSUES

Through the work carried out within the framework of this study, a number of issues relating to the auxiliary data used in the current AATSR LST retrieval scheme have become apparent. These issues are compromising the performance of the algorithm and it is strongly recommended that they be addressed as a matter of urgency to ensure that the AATSR LST data provided to users are of the highest possible quality.

The AATSR LST algorithm makes use of auxiliary data defining the land type (biome) and vegetation cover (fractional vegetation) for a given location. These data are used to appropriate the algorithm to the conditions of that location; further details can be found in [4]. At present, all the auxiliary data are tabulated at 0.5°. For land surfaces, this is an extremely coarse resolution and the prod-
uct would benefit from a higher resolution auxiliary data.

To illustrate some of the potential problems associated with using the current biome map, Figure 10 shows the current AATSR biome map over Tahoe and an equivalent 1-km biome map obtained from the University of Maryland (UMD - download from www.geog.umd.edu/landcover/1km-map/UMD-1km-latlon.img.Z), which has been derived from AVHRR data (together with Landsat data to characterise and ‘calibrate’ certain scenes). Although several of the biomes are common between both biome maps, the 0.5° resolution AATSR map does not even begin to capture the spatial variation of the land type. Moreover, Tahoe, which lies on the boundary between four AATSR biome cells is not classified correctly - if it were, all four of these biome cells would need to be classified as lake, despite the fact that Tahoe comprises only a tiny fraction of each of these cells. The effect of using the incorrect variant of the AATSR LST algorithm (determined by the biome classification) for a particular location can result in biases of several K, so obtaining the correct classification is imperative to the performance of the algorithm.

In addition to the resolution of the biome map, we have observed a problem with the land coverage of these data. For example, Figure 11 (left) shows the distribution of AATSR biomes over Europe. It can be seen that some of the land has not been assigned a biome classification in this image, which is resulting in absent LST data in the operational product (also shown in Figure 11). Closer inspection of the biome map shows in Figure 11 yields that many of the west coasts in Europe are not classified as land and that some of the ocean has been assigned a biome classification. Overall, it would appear that the biome map over Europe is offset approximately 1° to the East.

A solution to these problems would be the implementation of a higher-resolution biome map. For example, the UMD biome map is available at 1 and 8 km resolution, and is freely available from the web (see above). In addition, many of the biome classifications are the same as those utilised in the current scheme, so it would be possible to transfer the existing retrieval scheme to the new biome map.

Similar problems have been observed with the auxiliary fractional vegetation data. Figure 12 shows the current operational fractional vegetation map for August and an observed estimate of the actual fractional vegetation on 12 August 2004, derived from operational AATSR NDVI. Clearly the operational fractional vegetation map does not capture the true variation in fractional vegetation for this geographical region. For example, at approximately 39.241°, -900.297° the AATSR auxiliary data yields a value of f = 0.48, which is quite different from the value of f = 1.0 estimated from the observed data. This latter figure agrees well with in situ observations at this location on this day (César Coll, personal communication), indicating that the AATSR auxiliary data is inappropriate for this site on this date. Such inaccuracies have been observed to cause discrepancies in the AATSR LST product of the order of several K (César Coll, personal communication).

Although it may not be practical to utilise the AATSR NDVI data operationally to estimate the fractional vegetation parameter required for the LST algorithm, the product would benefit enormously from the use of a higher-resolution data set (e.g. <5 km).

5. CONCLUSIONS AND RECOMMENDATIONS

The AATSR is the most accurate infrared radiometer currently being flown in space. For the first time in the (A)ATSR mission, an operational LST product is available. The target accuracy of this product is 2.5 K during the day and 1.0 K at night.

In this study, we have validated more than two years of operational AATSR LST data with in situ data recorded over the ARM field site in Oklahoma. The results of the study indicate the accuracy (AATSR minus in situ) of the LSTs is 0.5 K (standard deviation 3.3 K) during the day and 1.2 K (standard deviation 1.2 K) at night under cloud-free conditions. These results suggest that the AATSR retrievals are performing better than the equivalent MODIS data over this site, where a bias of −4.0 K (standard deviation 2.7 K) and −2.0 K (standard deviation 1.1 K) is observed for day and night, respectively. However, there is a distinct seasonal bias in the AATSR LSTs that is not apparent in the MODIS data, where the AATSR are warmer than the in situ observations during the summer months, and colder during the winter. This cycle is observed in both the day and night time AATSR comparisons, although the amplitude is smaller in the case of the latter.

A seasonal bias of a similar nature is observed in the comparisons between the AATSR and equivalent data from the MODIS and SEVIRI sensors, and ECMWF model LST data over a number of sites around the globe. In each case, the seasonal bias (AATSR minus MODIS/SEVIRI/ECMWF) peaks during the summer months, which is consistent with the bias observed in the AATSR LST validation results for the ARM-Oklahoma site. As this seasonal bias is not observed in the MODIS validation results for this site, these results suggest that the AATSR may also be subject to seasonal biases over these additional sites. However, this cannot be confirmed without further in situ validation at these locations.

A seasonal bias in the AATSR LST data is also suggested through the results of radiative transfer simulations. These simulations indicate that the current retrieval scheme has very high sensitivity to atmospheric variability, namely water vapour, atmospheric temperature and the surface-air temperature difference. We have shown that the AATSR LSTs may be biased by several K from the true LST in each case, and with the natural annual
Figure 10. Comparison between the current AATSR LST biome map around Lake Tahoe (left) and the equivalent biome map (right) obtained from the University of Maryland (UMD). Lake Tahoe is centred on 39°, -120° - indicated by the cross on the AATSR biome map (classified as biome 0 in the UMD map). The AATSR biome for lake classification is biome 14. Note that the biome numbers used in each map are not equivalent.

Figure 11. Current AATSR biome map over Europe (left). Some of the land has not been assigned a biome (e.g. Wales, Sardinia, W. Sicily, W. Portugal). This is resulting in ‘missing’ LST data, as shown in this LST image of the UK (right). The cyan colours correspond to locations where no LST data is available.

Figure 12. Operational fractional vegetation (left - month of August) and estimated observed fractional vegetation, derived from the AATSR NDVI product over E. Spain on 12 August 2004.
variability of the atmosphere can result in a seasonal bias (AATSR minus true) with maxima occurring during summer months.

Although the overall bias observed in the product validation results is encouraging, the seasonal nature in these data should be addressed in order to improve the precision of the retrievals. The product would also benefit from improvements to the auxiliary data used in the current retrieval scheme. At 0.5°, the resolution of both the biome (land-type) and fractional vegetation data is too coarse. Improving the resolution (e.g., to at least 5 km) would improve the quality of the retrievals considerably, as using inappropriate values of these parameters has been observed to cause biases of several K in the LST retrievals. In addition, not all land is covered by the existing auxiliary data, resulting in gaps in the retrieved LSTs. Lastly, further validation, ideally with in situ LSTs recorded over long periods of time (> 1 year) over many more locations around the globe is required in order to gain a full picture of the performance of this product.

ACKNOWLEDGMENTS

This work was carried out under European Space Agency (ESA) contract 19054/05/NL/FF. The AATSR data were supplied by ESA and NEODC. The in situ radiometric data acquired over the Oklahoma ARM site were supplied by the Atmospheric Radiation Measurement (ARM) Program. The MODIS data used in this study were acquired as part of the NASA’s Earth Science Enterprise. The algorithms were developed by the MODIS Science Teams. The data were processed by the MODIS Adaptive Processing System (MODAPS) and Goddard Distributed Active Archive Center (DAAC), and are archived and distributed by the Goddard DAAC. The SEVIRI data were provided by the Land Surface Analysis Satellite Applications Facility (LSA-SAF). The 1-km UMD biome map used in this paper was obtained from the University of Maryland.

REFERENCES


