A MULTIPLATFORM LAND SURFACE TEMPERATURE DATASET USING AATSR, SEVIRI, MTSAT AND GOES-E FOR WACMOS-ET – ALGORITHMS AND VALIDATION RESULTS

JOÃO P. A. MARTINS[1,2], **ANA PIRES**[1,2], **PHILIPP SCHNEIDER**[3], **ISABEL F. TRIGO**[1,2], **CARLOS JIMENEZ**[4]

(1) Instituto Português do Mar e da Atmosfera (IPMA), Rua C do Aeroporto, 1749-077 Lisbon, Portugal
(2) Instituto Dom Luiz, University of Lisbon, IDL, Campo Grande, Ed C1, 1749-016 Lisbon, Portugal
(3) NILU - Norwegian Institute for Air Research, Kjeller, Norway
(4) Estellus, Paris, France

In the context of the ESA-funded European project WACMOS-ET, a global dataset of Land Surface Temperature (LST) was produced covering the 2005-2007 period. The LST dataset was generated using inputs from the Advanced Along-Track Scanning-Radiometer (AATSR), the Spinning Enhanced Visible and Infrared Imager (SEVIRI), the Multifunction Transport Satellite (MTSAT) and the Geostationary Operational Environmental Satellite - East (GOES-E) imagers. It is the aim of this study to minimize discrepancies among LST fields generated from different sensors through the use of a common approach for algorithm and auxiliary input data. The radiances from each sensor were processed using a common generalized split-window (GSW) algorithm, using data from the 10.8µm and 12.0µm channels (except for GOES-E), the same surface emissivity (monthly database from the University of Wisconsin-Maddison), and 3-hourly total column water vapour (TCWV) from ECMWF forecasts. Since the 12.0 µm channel is not available for GOES-E (GOES-12 onwards), a new mono-channel linear regression model was developed to incorporate explicitly emissivity and angle dependencies. An LST error estimate is provided for each instrument retrieval, based on the robustness of the model coefficients, sensor noise, and both emissivity and TCWV uncertainties. The datasets are provided on a sinusoidal grid with 1km spatial resolution for the polar orbiting sensor (AATSR) and with 5km spatial resolution for the geostationary sensors. For MTSAT, focus was given on Australia (the region of interest for the project), whereas for SEVIRI and GOES-E full disk products were provided.

The derived LST datasets have been validated against in situ stations and other independent LST satellite products (e.g. MODIS). Preliminary comparisons with ground stations showed that AATSR compares better with ground stations than MODIS. As expected, the agreement with in situ measurements is significantly better for night time than for day time, for all instruments. Nevertheless, all instruments show good performance, with error estimates and accuracy within the expected limits for LST products (biases around 0.2-0.3ºC and RMSEs between 1.4ºC and 3.2ºC)

This dataset was first designed primarily to serve as an input for evapotranspiration models but it will be freely available for download at the WACMOS-ET website. The use of this methodology may also be applied to future sensors, as it contributes to narrow the sources of uncertainty, favouring the optimal use of multi-sensor products.

1 INTRODUCTION

Land Surface Temperature (LST) is a key parameter of the surface radiative budget, as it measures the available energy at the surface-atmosphere interface. LST is a useful quantity for the scientific community, namely for those dealing with weather and climate numerical models. Accurate values of LST are also of special interest in a wide range of areas related to land surface processes, including meteorology, hydrology, agrometeorology, climatology and environmental studies. LST is very hard to quantify, even with in situ measurements, due to its high temporal and spatial variability as well as to strong directional effects.

Radiometers are often used over controlled sites with well-known surface properties and these measurements closely match those obtained by remote sensing techniques (Trigo *et al.*, 2008).

The WACMOS-ET project was motivated by the need to develop a predictive capability for terrestrial evapotranspiration to support both climate research and operational water management and agriculture. The goal was therefore to provide a Reference Input Dataset to derive and validate evapotranspiration (ET) estimates, at the global and regional scales, using as consistent as possible remotely sensed datasets as inputs (Michel *et al.,* 2015; Miralles *et al.*, 2015). LST is a vital parameter for some ET estimation methodologies, and the choice of a common algorithm to process different sensors is a problem of its own, as well as the assessment of its quality (Coll and Caselles, 1997; Yu *et al.*, 2008). In this paper, the choice of the algorithm and inputs are discussed, and the product accuracy is evaluated based on validation against independent data sets.

2   SENSOR CHARACTERISTICS AND CURRENTLY USED PRODUCTS

2.1   AATSR

One of WACMOS-ET main goals is to demonstrate the use of AATSR-derived LST fields for the estimation of evapotranspiration. AATSR was one of the 10 Earth-observing instruments on-board ESA's polar orbiter EnviSAT (ENVIronment SATellite), launched in March 2002. Its mission ended in 2012. Although the sensor provides a dual-view scan, only the nadir view is used here. The instrument has a zenith view angle (ZVA) ranging from 0º to 21.6º (Sòria and Sobrino, 2007), and provides a spatial resolution at nadir of 1 km by 1 km. The sensor has 2 bands in the atmospheric window region, centred around 10.9 and 12.1 µm.

The AATSR was initially designed to provide sea surface temperature (SST) maps, ensuring the production of a unique 10 year near-continuous data set at the levels of accuracy required (0.3 K or better). However, AATSR data is also being used to obtain LST on a global scale, particularly taking into account that SLSTR onboard Sentinel-3 (expected to be launched in 2015; Donlon *et al.*, 2012) will provide continuity of AATSR data and LST products. A Level 2 LST product was provided by ESA according to the AATSR Algorithm Theoretical Basis Document (Prata, 2002). The operational algorithm requires as inputs the following parameters: seasonally-dependent land cover classification, fractional vegetation and precipitable water. The first is based on the biomes provided by Dorman and Sellers (1989) in a of 1°×1° grid, while precipitable water data is based on the NVAP climatology at 0.5°×0.5° resolution and monthly intervals (Randel *et al.*, 1996). The spatial resolution of both fractional cover and precipitable water is currently one of the main problems in the retrieval of the in the official product.

2.2   SEVIRI / MSG

Meteosat Second Generation (MSG) is a joint project between ESA and the European Organisation for the Exploitation of Meteorological Satellites (Eumetsat). The first MSG satellite was launched in August 2002, entering into operational service with Eumetsat in early 2004 being then renamed Meteosat-8. The second MSG, Meteosat-9, was launched on 21 December 2005 and MSG-3 (Meteosat-10) was launched on 5 July 2012. Their geostationary orbit is centered over 3.5ºE, 9.5ºE, and 0ºE, respectively. The operational LSA-SAF LST based on data from the SEVIRI instrument (Spinning Enhanced Visible and Infrared Imager) onboard the MSG platform is available with a 15 minute temporal frequency and at the original satellite spatial resolution (3 km sampling distance at the sub-satellite point) and geostationary projection. The data set is available for the whole study period (2005-2007). The SEVIRI sensor encompasses 12 channels covering the visible (VIS) to the infrared (IR).

2.3   MTSAT

MTSAT-2 is the Japanese Meteorological Association (JMA) Multifunction Transport SATellite launched in February 2006, following the MTSAT-1R launched one year prior. Both MTSATs are geostationary satellites centered at 145ºE and 140ºE, respectively. Their Japanese Advanced Meteorological Imager (JAMI) onboard has 2 available split-window channels, centred around 10.8 µm and 12.0 µm, based on which the LST is computed, and which have a spatial resolution of 4 x 4 km and a temporal resolution of 3 hours. MTSAT data were processed for an area covering Australia and for a period covering June 2006-December 2007.

2.4 GOES-E

GOES-12 (Geostationary Operational Environmental Satellite) was launched in July 2001. In April 2003, GOES-12 became GOES-East at 75ºW, and was decommissioned in April 2010. Its mission ended August 2013. GOES-12 was the last of the third generation GOES satellites. In contrast to SEVIRI, AATSR and MTSAT, the imager onboard the GOES satellites (from GOES-12 onwards, http://www.oso.noaa.gov/goes/goes-calibration/change-channels.htm) does not include two thermal infrared channels in the thermal atmospheric window of the spectrum. Alternative methodologies to the split-window algorithms must then be applied in order to obtain LST, as detailed in section 3.1.

3 ALGORITHMS

3.1 Description of the candidate algorithms

LST is estimated here using linear regression methods (generalized split-window and mono-window algorithms) with surface emissivity as an explicit input. To derive formulations for the linear regressions, the radiative transfer equation is used as a starting point (e.g. Li *et al.*, 2013):

$$L_i = B(T_{b,i}) = \epsilon_i B_i(T_{sfc})\tau_i + L^\uparrow_{atm,i} + (1 - \epsilon_i)L^\downarrow_{atm,i}\tau_i \qquad (1)$$

Where L_i is the monochromatic radiance at the sensor level, $B(T_{b,i})$ and $B_i(T_{sfc})$ are the radiances emitted by blackbodies at temperatures T_b (brightness temperature) and T_{sfc} (surface temperature), ϵ_i is the monochromatic emissivity, τ_i is the monochromatic atmospheric transmissivity, and $L^\uparrow_{atm,i}$ and $L^\downarrow_{atm,i}$ are the upward and downward radiances emitted by the atmosphere. It is possible to derive formulations for linear regression methods expanding some of the terms of equation (1) using Taylor series. Split-window algorithms use these derived expressions for both channels in the atmospheric window. These methods have advantages with respect to other methods such as the so-called "physical" algorithms, which solve the radiative transfer equation directly, thus being much less computationally efficient. Since it is possible to use different approximations, there are a few different algorithms documented in the literature. Yu et al. (2008) gathered the most commonly used algorithms that retain explicit emissivity dependence and compared their performance (Table 1).

No	Formula	Reference(s)
1	$T_s = C + \left(A_1 + A_2\dfrac{1-\epsilon}{\epsilon} + A_3\dfrac{\Delta\epsilon}{\epsilon^2}\right)\dfrac{T_{11}+T_{12}}{2} + \left(B_1 + B_2\dfrac{1-\epsilon}{\epsilon} + B_3\dfrac{\Delta\epsilon}{\epsilon^2}\right)\dfrac{T_{11}-T_{12}}{2}$	(Freitas *et al.*, 2010; Wan and Dozier, 1996)
2	$T_S = C + A_1\dfrac{T_{11}}{\epsilon} + A_2\dfrac{T_{12}}{\epsilon} + A_3\dfrac{1-\epsilon}{\epsilon}$	(Caselles *et al.*, 1997; Prata and Platt, 1991)
3	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3(1 - \epsilon) + A_4\Delta\epsilon$	(Ulivieri *et al.*, 1994)
4	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3\dfrac{1-\epsilon}{\epsilon} + A_4\dfrac{\Delta\epsilon}{\epsilon^2}$	(Vidal, 1991)
5	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3(T_{11} - T_{12})(1 - \epsilon_{11}) + A_4 T_{12}\Delta\epsilon$	(Price, 1984)
6	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3\epsilon$	(Ulivieri and Cannizzaro, 1985)
7	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3\epsilon + A_4\dfrac{\Delta\epsilon}{\epsilon}$	(Sobrino *et al.*, 1994)
8	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3(1 - \epsilon_{11}) + A_4\Delta\epsilon$	(Coll *et al.*, 1997)
9	$T_S = C + A_1 T_{11} + A_2(T_{11} - T_{12}) + A_3(T_{11} - T_{12})(T_{11} - T_{12}) + A_4(1 - \epsilon_{11}) + A_5\Delta\epsilon$	(Sobrino *et al.*, 1993)

Table 1 - List of split-window algorithms tested in this work and their respective references. T_s denotes LST; C, A_1, A_2, A_3, B_1, B_2, B_3 are the regression coefficients, T_{11} and T_{12} are the brightness temperatures in the IR 108 and IR 120 channels, ϵ_{11} and ϵ_{12} their emissivities and ϵ the average of both emissivities.

In the same study, the inclusion of a path correction term with the form $D(T_{11} - T_{12})(\sec(ZVA) - 1)$, which avoids calibrating the model per classes of ZVA, was also tested. The algorithm choice in this study is based on the methodology by Yu et al. (2008), but it uses different calibration and validation databases (see section 3.2). For AATSR, SEVIRI, and MTSAT the single best split-window algorithm was chosen based on a comprehensive uncertainty analysis (see section 3.3). For GOES-E, a single mono-channel algorithm was used (Duguay-Tetzlaff *et al.*, 2015):

$$T_s = A \frac{T_{11}}{\epsilon_{11}} + B \frac{1}{\epsilon_{11}} + C\big(\sec(ZVA) - 1\big) + D \qquad (2)$$

3.2   Calibration and Verification Database

The split-window coefficients are trained and verified for a wide range of atmospheric and surface conditions and viewing geometries, using simulations by the MODerate spectral resolution atmospheric TRANSmittance model (MODTRAN; Berk *et al.*, 1999) for more than 15700 clear sky profiles from the database prepared by Borbas *et al.* (2005), referred to hereafter as SeeBor. The algorithm (MODTRAN4) provides a useful tool for quantifying the top-of-atmosphere spectral radiance, L_v, which is done for the bands corresponding to IR 108 and IR 120 channels, with a spectral resolution of 1 cm^{-1}. The integration of L_v weighted by the ith channel response function, $\phi_{i,v}$, within the band limits v_1 and v_2, provides its effective radiance:

$$L_i = \frac{\int_{v_{i,1}}^{v_{i,2}} \phi_{i,v} L_v dv}{\int_{v_{i,1}}^{v_{i,2}} \phi_{i,v} dv} \qquad (3)$$

The radiances are converted to equivalent blackbody brightness temperatures using the inverse of the Planck function:

$$T_{b,i} = \frac{hcv_i}{k \log(\frac{2hc^2v_i^3}{L_i}+1)} \qquad (4)$$

where h is the Planck constant, c is the speed of light an k is the Boltzmann constant. The SeeBor database includes the necessary information about atmospheric temperature and humidity, surface emissivity, skin temperature, pressure, total column water vapour and land cover and elevation. The simulations took into account the respective channel response function. For each profile, a random ZVA is assumed; for AATSR the angle interval is within the range of the nadir field-of-view (0º - 21.6º), while for the geostationary satellites the interval includes angles up to 75º for MTSAT and SEVIRI, and 60º for GOES, comprising most of the disk. A subset of this dataset was used to produce a calibration database, with 80 profiles geographically well distributed and with a uniform total column water vapour (TCWV) distribution in order to cover a wide range of atmospheric conditions. For each of these profiles, the skin temperature was perturbed by taking the 10 m temperature and adding perturbations in the range [-15 K 15 K] in 5 K intervals. Each of these cases was modelled using the following ZVA intervals: in the range 0-22º in 1º intervals for AATSR; in the range 0-75º in 5º intervals for MTSAT and SEVIRI, and 0-60º in 5º intervals for GOES. The emissivity of the IR 12.0 channel was set to vary between 0.935 and 0.995 in 0.015 intervals. For each of these values, the IR 10.8 emissivity was calculated by adding perturbations in the range [-0.03 0.024] in 0.006 intervals (excluding cases where it exceeds 1). In this way, the robustness of the database can be ensured with respect to atmospheric conditions, differences between skin and 10 m temperature, ZVA and also to the range of expected emissivities of the split-window channels. The remaining profiles in the database were used for validation of the calibrated coefficients, using random ZVAs, as described above.

3.3   Uncertainty analysis and model performance

In order to decide which split-window algorithm to use (see Table 1), a combination of algorithm performance and an uncertainty analysis was used. To evaluate model performance, the modelled values of the verification database were compared to the true value, also given by SeeBor (*Figure 1*). In that assessment, all the inputs were assumed as free of any uncertainty. We also compared the inclusion of the path correction term against simulations that used different coefficients for classes of ZVA, and concluded that its impact on algorithm performance is negligible (not shown). Therefore, a formulation using the path correction term is preferred as it is expected to provide smoother resulting LST fields. All algorithms show that larger errors occur when total column water vapor increases. There are a few models with similar performances, the notable exceptions being algorithms 2, 5 and 6 (see Table 1), with larger biases and RMSEs. The algorithm with the lowest bias and RMSE was algorithm 1.

The sensitivity to all the inputs was analysed in detail for all four sensors, but only a brief description focusing on AATSR will be given here, for the sake of brevity. Sensor noise errors were inferred by perturbing the

brightness temperatures obtained for each channel, and for each of the MODTRAN simulations of validation database, with values with a standard deviation corresponding to the noise-equivalent temperature reported for each channel. The perturbed values were then used to estimate LST with each of the algorithms and compared to the unperturbed estimates. The results show that errors tend to grow for cases with larger TCWV and ZVA, and all models show the same type of behavior.
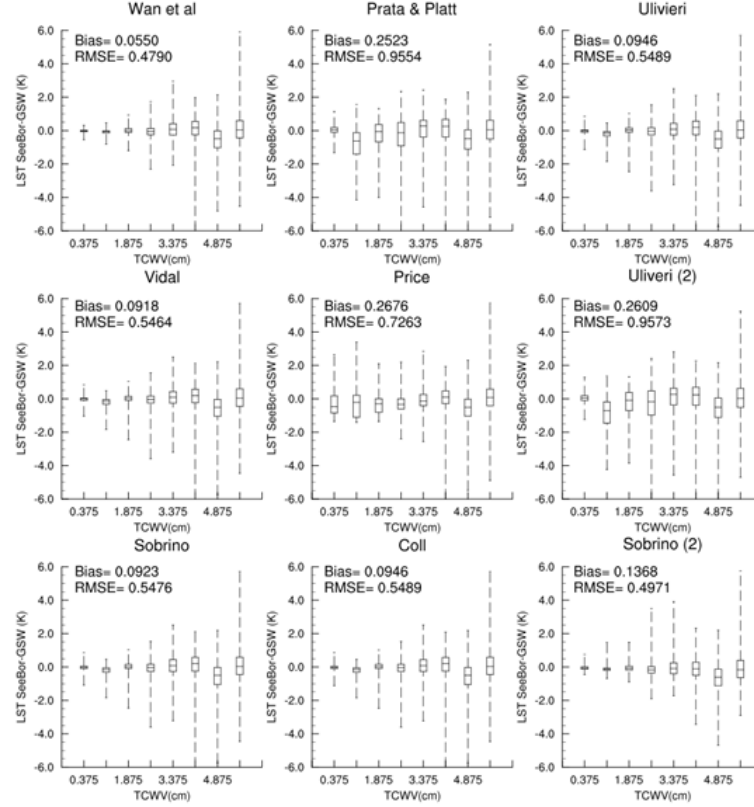
Figure 1 - **Validation results using a path correction term, for AATSR. Each bar is a boxplot of the differences between the LST in SeeBor and the corresponding GSW result, grouped by TCWV class. Also shown are the bias and RMSE for each model.**

To take this into account, perturbation values of the original emissivity values were varied as follows: for $0.8 < \epsilon \le 0.95$, gaussian perturbations in the range of ± 0.016; for $0.95 < \epsilon \le 0.98$, in the range of ± 0.012 and for $0.98 < \epsilon \le 1$, perturbations in the range of ± 0.007. In the case of emissivity the comparison of perturbed vs. unperturbed LSTs revealed that larger discrepancies occur for lower values of TCWV. This effect may be explained using equation (1): TCWV decreases the atmospheric transmissivity and increases atmospheric emission terms. So the lower the TCWV, the higher is the value of τ and the lower the terms $L_{atm,i}^{\uparrow}$ and $L_{atm,i}^{\downarrow}$. Therefore the value of the first term (surface emission) is amplified if TCWV is low, and the value of the last term (surface reflection of atmospheric emission) is reduced in the same conditions. Two of the tested algorithms are less sensitive to this effect (2 and 6).

In the case of the uncertainty due to TCWV, the error comes from the fact that an incorrect set of coefficients might be used. The error that is incurred by using the wrong set of coefficients is combined with the probability of that event, which is estimated by comparing ECMWF forecast fields to the corresponding analysis. In general, the ECMWF forecast fields do not fail by more than one class of TCWV, and only for very high values does the probability of using the neighboring class exceeds 30%. In general, errors due to TCWV do not exceed 0.20 K except for the last class (5.25 to 6 cm), which takes values of about 0.5 K.

The total uncertainty is obtained as follows:

$$\sigma_{Total} = \sqrt{\sigma_{Tb}^2 + \sigma_{\epsilon}^2 + \sigma_{TCWV}^2 + \sigma_{model}^2}, \tag{5}$$

5

where σ_{Tb}, σ_{ϵ}, σ_{TCWV} and σ_{model} are the uncertainties due to sensor noise, emissivity database, TCWV forecasts and regression model, respectively. All algorithms show a similar behaviour of increasing uncertainty with TCWV (except algorithms 2, 5 and 6 which had already shown poor regression performance). Although not shown here, it is thus very difficult to choose which of those is the best. The choice of the algorithm was then mainly based on the model performance (Figure 1): the model that showed the best performance is the one by Wan and Dozier (1996), with a path correction term:

$$T_s = C + \left(A_1 + A_2 \frac{1-\epsilon}{\epsilon} + A_3 \frac{\Delta\epsilon}{\epsilon^2}\right)\frac{T_{11} + T_{12}}{2} \\ + \left(B_1 + B_2 \frac{1-\epsilon}{\epsilon} + B_3 \frac{\Delta\epsilon}{\epsilon^2}\right)\frac{T_{11} - T_{12}}{2} + D\,(T_{11} - T_{12})(sec(ZVA) - 1) \tag{6}$$

The magnitude of errors associated to each of the sources described above, using the selected algorithm, and for AATSR, is shown in Figure 2. For small TCWV, the total error is mainly modulated by the emissivity contribution, whereas for larger values, the model error is the main source. The total uncertainty varies between 0.5 and 2.3 K. It is worth noting that for the geostationary platforms, the magnitude of the uncertainties is larger because the optical path has larger ranges, as the ZVA may reach as high as 75°. In those cases the error is roughly independent of ZVA, for ZVAs<50° but increases rapidly for higher angles. Cases for which the prescribed error surpasses 4K are not processed.
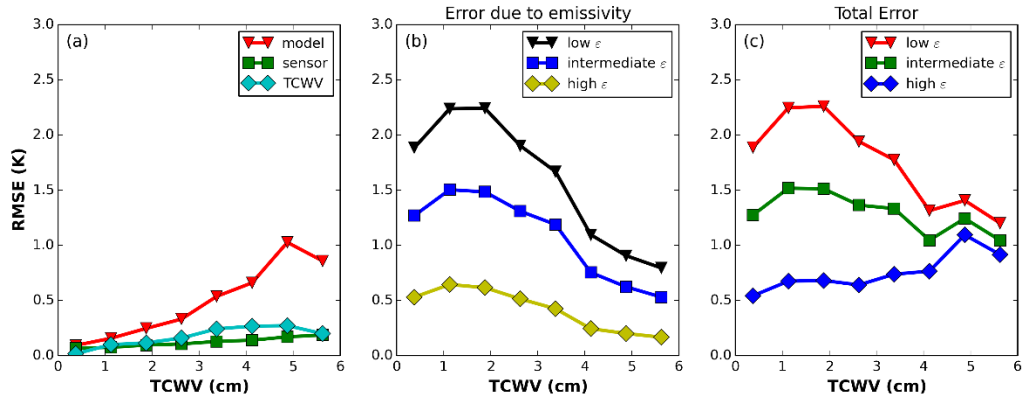

Figure 2 – LST uncertainty (K) as a function of TCWV (cm) for the chosen algorithm [Equation (6)] (a) contribution of model (in red), sensor noise (green) and TCWV forecasts (blue) (b) contribution of emissivity, separated by classes of emissivity (low, intermediate and high emissivity) (c) sum of all contributions [equation (5)]

4 PROCESSING ASPECTS

The previous section describes how to calibrate the multiple regression model that is used here to convert the radiances from each sensor to LST. It further shows how the value of the error that is assigned to each pixel was computed, depending on its emissivity, view angle and TCWV. The TCWV was retrieved from ECMWF 3-hourly global fields at 0.75° resolution: the 0h, 6h, 12h and 18UTC were analysis, the 3h and 9h are the steps 15 and 21 from the T-12 forecast and the 15h and 21h are the steps 15 and 21 from the T-0 forecast. For each pixel, the time of retrieval is read and the TCWV is interpolated in space and time for the pixel location. Spectral emissivity was retrieved from monthly files available at 0.05° resolution at the Global Infrared Land Surface Emissivity Database from the Univ. of Wisconsin (Seemann *et al.*, 2008). Solar angles are calculated using spherical geometry routines available from the LSA-SAF processing chain and adapted for each of the sensors (except AATSR which contained all the necessary angles in its L1 files).

Since all LST fields are calculated for cloud-free pixels, a cloud mask is required. In the case of AATSR, the L1 radiances include an operational cloud mask, so the user is referred to the AATSR documentation for a complete description of their cloud mask algorithms (Závody *et al.*, 2000). For the geostationary sensors, the algorithm used is the one adapted from the NWC-SAF set of algorithms. Basically the algorithm consists of a comparison of the IR radiances and VIS reflectances of each scene with background values. If the differences are higher than given thresholds, the pixel is considered as cloudy. Additional spatial texture tests are also applied to filter spurious results. Further details on the cloud mask can be found here: http://land.copernicus.eu/global/sites/default/files/products/GIOGL1_ATBD_Cloud_I1.00.pdf.

5 VALIDATION

The accuracy of the WACMOS-ET LST product was evaluated using 1) in-situ data and 2) intercomparison against other independent datasets. Again, for the sake of brevity and consistency with the previous section, focus will be given to the AATSR validation. An extensive analysis can be found in the project official validation report (available to download here http://aramis.obspm.fr/~jimenez/Docs/WACMOSET/WACMOSET_WP2130_approved.pdf):

5.1 In-situ data

Very few stations dedicated to LST validation exist. For the WACMOS-ET years only data from Gobabeb and Evora stations could be used. Gobabeb (22.33° S, 15.03° E) is located on large gravel plains (>900 km2) at an altitude of 408 m; these plains are sparsely covered by desiccated grass. Some non-continuous measurements were available for the WACMOS-ET period (in December 2007), since the station only started its nominal operations in January 2008. Evora (38.54° N, 8.00° W) is part of the global flux network of measurements (Fluxnet) as it is composed of sparse oak tree canopy (30-40 trees/ha) and a grassland soil. The station was not fully operational at the WACMOS-ET period, but some data was available for November and December 2007.

Therefore data from both the Surface Radiation (SURFRAD, see locations at http://www.esrl.noaa.gov/gmd/grad/surfrad/sitepage.html) and from the Atmospheric Radiation Measurement program (ARM) were used to estimate in-situ LST. It should be noted that these networks were not specifically designed to measure LST, which is why some stations do not exhibit homogeneous land cover at the scale of the satellite footprint, thus increasing the uncertainty of the emissivity values used in the estimation. In situ LST is estimated from the broadband longwave flux measurements through the equation:

$$LST = \left(\frac{LW_{out} - (1-\epsilon)LW_{in}}{\epsilon\sigma}\right)^{\frac{1}{4}}, \qquad (5)$$

where LW_{out} and LW_{in} are the upwelling and downwelling longwave radiation, respectively, ϵ is the broadband emissivity of the Earth's surface in the spectral range of measurements and σ is the Stefan Boltzmann constant ($5.6704 \times 10^{-8}\ W\ m^{-2}\ K^{-4}$). Broadband emissivity was estimated from the same global database used for the LST production (Seemann *et al.*, 2008), using the following equation (Ogawa *et al.*, 2003; Wang, 2005):

$$\epsilon = a\epsilon_{8.5} + b\epsilon_{11} + c\epsilon_{12}, \qquad (5)$$

where $\epsilon_{8.5}$, ϵ_{11} and ϵ_{12} are the narrowband emissivities at 8.5 μm, 11 μm, and 12 μm, respectively. The coefficients a, b, and c were found to be 0.2122, 0.3859, and 0.4029 respectively (Wang, 2005).

A comparison of AATSR LST with matchup data from the different in-situ stations during night time is shown in Figure 3. It can be observed that the majority of matchups follows the 1:1 line at all sites, but also that at most sites frequent negative outliers occur. These are the characteristics of failures of the used operational L1 cloud mask. Another indicator for this behavior is that a site located in the Nevada desert (Desert Rock), which sees relatively few clouds has a much smaller number of negative outliers than the typically more cloudy sites in the mid-latitudes (e.g. the Bondville site). During the day, those issues are not so serious because the cloud masking algorithm relies on extra information from the visible channels (not shown).

Because the cloud masking issue resulted in an unusually high number of outliers which in turn masked the true accuracy of the AATSR-based LST retrievals, a very simple filtering scheme was used to eliminate the most extreme outliers caused by an inefficient cloud masking procedure. This scheme evaluated the difference between in situ LST and AATSR LST and removed the matchups where this difference was at the same time negative and its value greater than two standard deviations from the mean. While this is a relatively crude method, it is effective here in the sense that it becomes possible to derive summary statistics, which are more representative of the actual accuracy of the AATSR LST product once the night time cloud masking issues are taken care of (Table 2). Although there is a small increase of the bias, both for the day time and in the night time, the RMSE becomes much lower, especially at night. Also shown are the respective values for MODIS, which show a pronounced negative bias and a much higher RMSE at night.
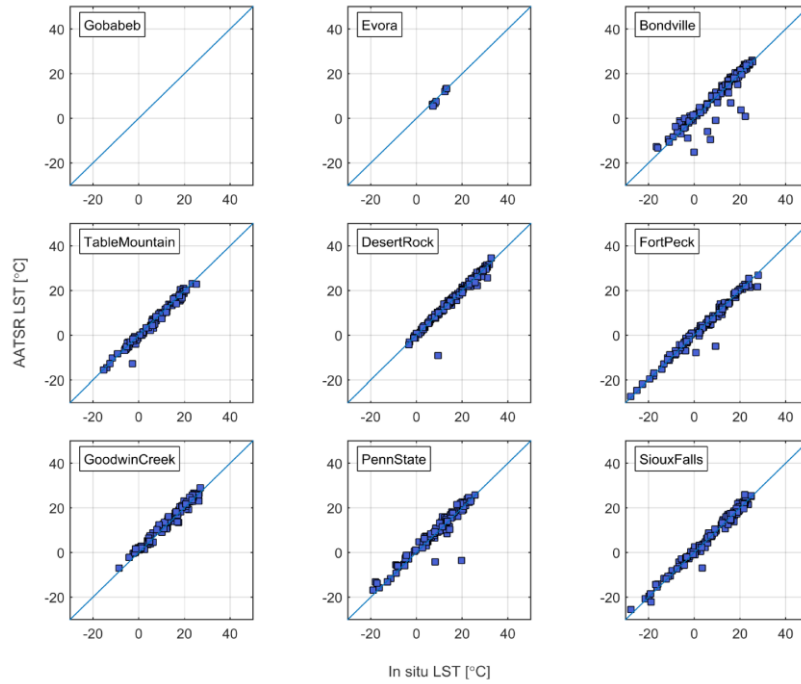
Figure 3 - Scatter plots of AATSR-derived nighttime LST against LST computed from observations at in situ stations

	Night			Day		
	WACMOS AATSR	WACMOS AATSR (without cloud outliers)	MODIS (for reference)	WACMOS AATSR	WACMOS AATSR (without cloud outliers)	MODIS (for reference)
Bias (K)	0.0	0.3	-3.0	1.0	1.4	-1.5
RMSE (K)	2.1	1.3	3.1	3.8	3.2	3.4

Table 2 – Summary statistics for night and daytime comparisons of the original AATSR, AATSR with cloud-filtered scheme and MODIS with in situ stations.

To further investigate the cloud masking issue, AATSR was compared to the corresponding product from the GlobTemperature project (http://www.globtemperature.info/), which uses the algorithm of Prata (2002) but is improved in terms of the spatial resolution of the auxiliary datasets such as fractional vegetation cover and global biome distribution, and uses an improved cloud mask. However, in order to make the summary statistics comparable, that cloud mask was replaced by the one used in WACMOS AATSR. The results in Table 3 show that the WACMOS dataset outperforms GlobTemperature especially during day time, with a RMSE reduction of more than 0.5K. During the night the improvements are more evident in terms of the bias, which is about 0.1K lower for WACMOS.

Product	Bias	StdDev	RMSE
WACMOS Night	0.06	2.26	2.25
GlobTemp Night	0.15	2.27	2.27
WACMOS Day	1.17	3.25	3.45
GlobTemp Day	1.20	3.81	3.99

Table 3 – Overall summary statistics of all nighttime and daytime matchups for 2007 at all in situ sites. Note that the same cloudmask was used for both products (same used in WACMOS-ET).

The validation of the geostationary sensors was challenging. MTSAT does not have any usable data from in situ stations on its disk, as the only station in Darwin (Australia) is located too close to the ocean. In addition, the site is in a very complex urban landscape and not well suited for LST validation. Comparisons with MODIS revealed an overestimation of 3.6ºC by MTSAT relative to MODIS (in terms of bias). LST from GOES-E was evaluated against in situ observations from SURFRAD stations. The results indicate a very good correspondence with an average bias of 0.21 °C. The RMSE, as an overall representation of error, was 2.56

°C. These intercomparisons are complicated by the fact that the observation times are sometimes significantly different as well as the fact that the MODIS LSTs were shown to be systematically cooler than in situ estimations by 1-2 °C. Regarding SEVIRI, no validation effort was not made so far. Inter-sensor comparisons will be made between the derived AATSR LST and the geostationary products.

6   CONCLUSIONS

The comparison of LST data from different remote sensors is generally a difficult task as each sensor uses its own algorithm and ancillary inputs, apart from the differences in the sensor itself, viewing geometries and observation times. The development of the WACMOS-ET LST dataset allowed the inter-comparison of LST data retrieved by different sensors but using the same inputs and similar algorithms (notable exception for GOES-E due to the unavailability of one of the split-window channels), thus reducing the sources of discrepancy between datasets.

The best retrieval algorithm was chosen from a list of 9 commonly used algorithms documented in the literature. The main criterion was to pick the best comparison to the "true" LST, from a validation database of over 15000 atmospheric profiles, viewing geometries and surface conditions. From the list of tested models, only three had poor uncertainty propagation performance and the others had similar performance, with uncertainties growing with TCWV. The major source of uncertainty is the surface emissivity. In fact, comparisons of emissivity databases reveal large discrepancies, especially over semi-arid and desert areas. These discrepancies naturally translate in differences in the retrieved LST, depending on the database that is used.

The comparisons of the AATSR LST with in-situ data showed very good agreement, especially when cloud contaminated pixels were removed with a simple statistical scheme. It was shown that the dataset frequently outperforms reference LST datasets such as MODIS and GlobTemperature (discussed the project validation report).

Furthermore, although not extensively discussed here, the validation of the geostationary datasets was more challenging due to 1) the unavailability of dedicated and non-dedicated LST ground stations (especially for the MTSAT disk) and 2) the sometimes large time discrepancy between sensors (especially for GOES-E).

7   REFERENCES

Berk A, Anderson GP, Bernstein LS, Acharya PK, Dothe H, Matthew MW, Adler-Golden SM, Chetwynd Jr. JH, Richtsmeier SC, Pukall B, Allred CL, Jeong LS, Hoke ML. 1999. MODTRAN4 radiative transfer modeling for atmospheric correction. *Proc. SPIE*, 348–353. DOI: 10.1117/12.366388.

Borbas EE, Seemann SW, Huang HL, Li J, Menzel WP. 2005. Global profile training database for satellite regression retrievals with estimates of skin temperature and emissivity. *International TOVS Study Conference-XIV Proceedings*.

Caselles V, Coll C, Valor E. 1997. Land surface emissivity and temperature determination in the whole HAPEX-Sahel area from AVHRR data. *International Journal of Remote Sensing* **18**(April 2013): 1009–1027. DOI: 10.1080/014311697218548.

Coll C, Caselles V. 1997. A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison. *Journal of Geophysical Research* **102**(D14): 16697. DOI: 10.1029/97JD00929.

Coll C, Valor E, Schugge T, Caselles V. 1997. A procedure for estimating the land surface emissivity difference in the AVHRR channels 4 and 5. *Remote Sensing Application to the Valencian Area*.

Donlon C, Berruti B, Buongiorno A, Ferreira M-H, Féménias P, Frerick J, Goryl P, Klein U, Laur H, Mavrocordatos C, Nieke J, Rebhan H, Seitz B, Stroede J, Sciarra R. 2012. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. *Remote Sensing of Environment* **120**: 37–57. DOI: 10.1016/j.rse.2011.07.024.

Duguay-Tetzlaff A, Bento V, Göttsche F, Stöckli R, Martins J, Trigo I, Olesen F, Bojanowski J, da Camara C, Kunz H. 2015. Meteosat Land Surface Temperature Climate Data Record: Achievable Accuracy and Potential Uncertainties. *Remote Sensing*. Multidisciplinary Digital Publishing Institute **7**(10): 13139–13156. DOI: 10.3390/rs71013139.

Freitas SC, Trigo IF, Bioucas-dias JM, Göttsche F. 2010. Quantifying the Uncertainty of Land Surface Temperature Retrievals From SEVIRI / Meteosat. **48**(1): 523–534.

Li Z-L, Tang B-H, Wu H, Ren H, Yan G, Wan Z, Trigo IF, Sobrino J a. 2013. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment* **131**: 14–37. DOI: 10.1016/j.rse.2012.12.008.

Michel D, Jiménez C, Miralles DG, Jung M, Hirschi M, Ershadi A, Martens B, McCabe MF, Fisher JB, Mu Q, Seneviratne SI, Wood EF, Fernández-Prieto D. 2015. The WACMOS-ET project – Part 1: Tower-scale evaluation of four remote sensing-based evapotranspiration algorithms. *Hydrology and Earth System Sciences Discussions*. Copernicus GmbH **12**(10): 10739–10787. DOI: 10.5194/hessd-12-10739-2015.

Miralles DG, Jiménez C, Jung M, Michel D, Ershadi A, McCabe MF, Hirschi M, Martens B, Dolman AJ, Fisher JB, Mu Q, Seneviratne SI, Wood EF, Fernaìndez-Prieto D. 2015. The WACMOS-ET project – Part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences Discussions*. Copernicus GmbH **12**(10): 10651–10700. DOI: 10.5194/hessd-12-10651-2015.

Ogawa K, Schmugge T, Jacob F, French A. 2003. Mapping land surface window (8-12 µm) emissivity from ASTER thermal data. *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, 3213–3215 vol.5. DOI: 10.1109/IGARSS.2003.1294733.

Prata AJ, Platt CMR. 1991. Land surface temperature measurements from the AVHRR. *Proceedings of the 5th AVHRR data users conference*.

Prata F. 2002. *Land Surface Temperature Measurement from Space : AATSR Algorithm Theoretical Basis Document*. Aspendale, Australia.

Price JC. 1984. Land surface temperature measurements from the split window channels of the NOAA 7 Advanced Very High Resolution Radiometer. *Journal of Geophysical Research*, 7231. DOI: 10.1029/JD089iD05p07231.

Randel DL, Greenwald TJ, Vonder Haar TH, Stephens GL, Ringerud MA, Combs CL. 1996. A New Global Water Vapor Dataset. *Bulletin of the American Meteorological Society* **77**(6): 1233–1246. DOI: 10.1175/1520-0477(1996)077<1233:ANGWVD>2.0.CO;2.

Seemann SW, Borbas EE, Knuteson RO, Stephenson GR, Huang HL. 2008. Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements. *Journal of Applied Meteorology and Climatology* **47**: 108–123. DOI: 10.1175/2007JAMC1590.1.

Sobrino JA, Li ZL, Stoll MP, Becker F. 1993. Determination of the surface temperature from ATSR data. *Proceedings of 25th International Symposium on Remote Sensing of Environment*.

Sobrino JA, Li Z-L, Stoll MP, Becker F. 1994. Improvements in the split-window technique for land surface temperature determination. *Geoscience and Remote Sensing, IEEE Transactions on* **32**(2): 243–253. DOI: 10.1109/36.295038.

Sòria G, Sobrino J a. 2007. ENVISAT/AATSR derived land surface temperature over a heterogeneous region. *Remote Sensing of Environment* **111**(4): 409–422. DOI: 10.1016/j.rse.2007.03.017.

Trigo IF, Peres LF, DaCamara CC, Freitas SC. 2008. Thermal Land Surface Emissivity Retrieved From SEVIRI/Meteosat. *IEEE Transactions on Geoscience and Remote Sensing* **46**(2): 307–315. DOI: 10.1109/TGRS.2007.905197.

Ulivieri C, Cannizzaro G. 1985. Land surface temperature retrievals from satellite measurements. *Acta Astronautica* **12**(12): 977–985.

Ulivieri C, Castronuovo MM, Francioni R, Cardillo A. 1994. A split window algorithm for estimating land surface temperature from satellites. *Advances in Space Research* **14**(3): 59–65. DOI: 10.1016/0273-1177(94)90193-7.

Vidal A. 1991. Atmospheric and emissivity correction of land surface temperature measured from satellite using ground measurements or satellite data. *International Journal of Remote Sensing* **12**(12): 2449–2460.

Wan Z, Dozier J. 1996. A Generalized Split- Window Algorithm for Retrieving Land-Surface Temperature from Space. *IEEE Transactions on Geoscience and Remote Sensing* **34**(4): 892–905.

Wang K. 2005. Estimation of surface long wave radiation and broadband emissivity using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature/emissivity products. *Journal of Geophysical Research* **110**(D11): D11109. DOI: 10.1029/2004JD005566.

Yu Y, Privette JL, Pinheiro AC. 2008. Evaluation of Split-Window Land Surface Temperature Algorithms for Generating Climate Data Records. *IEEE Transactions on Geoscience and Remote Sensing* **46**(1): 179–192. DOI: 10.1109/TGRS.2007.909097.

Závody AM, Mutlow CT, Llewellyn-Jones DT. 2000. Cloud Clearing over the Ocean in the Processing of Data from the Along-Track Scanning Radiometer (ATSR). *Journal of Atmospheric and Oceanic Technology*. American Meteorological Society **17**(5): 595–615. DOI: 10.1175/1520-0426(2000)017<0595:CCOTOI>2.0.CO;2.