Article

Assessment of Mono- and Split-Window Approaches for Time Series Processing of LST from AVHRR—A TIMELINE Round Robin

Corinne Myrtha Frey *, Claudia Kuenzer and Stefan Dech

German Remote Sensing Data Center (DFD), Earth Observation Center (EOC), German Aerospace Center (DLR), Oberpfaffenhofen, D-82234 Wessling, Germany; claudia.kuenzer@dlr.de (C.K.); stefan.dech@dlr.de (S.D.)
* Correspondence: corinne.frey@dlr.de; Tel.: +49-815-328-3473

Abstract: Processing of land surface temperature from long time series of AVHRR (Advanced Very High Resolution Radiometer) requires stable algorithms, which are well characterized in terms of accuracy, precision and sensitivity. This assessment presents a comparison of four mono-window (Price 1983, Qin et al., 2001, Jiménez-Muñoz and Sobrino 2003, linear approach) and six split-window algorithms (Price 1984, Becker and Li 1990, Ulivieri et al., 1994, Wan and Dozier 1996, Yu 2008, Jiménez-Muñoz and Sobrino 2008) to estimate LST from top of atmosphere brightness temperatures, emissivity and columnar water vapour. Where possible, new coefficients were estimated matching the spectral response curves of the different AVHRR sensors of the past and present. The consideration of unique spectral response curves is necessary to avoid artificial anomalies and wrong trends when processing time series data. Using simulated data on the base of a large atmospheric profile database covering many different states of the atmosphere, biomes and geographical regions, it was assessed (a) to what accuracy and precision LST can be estimated using before mentioned algorithms and (b) how sensitive the algorithms are to errors in their input variables. It was found, that the split-window algorithms performed almost equally well, differences were found mainly in their sensitivity to input bands, resulting in the Becker and Li 1990 and Price 1984 split-window algorithm to perform best. Amongst the mono-window algorithms, larger deviations occurred in terms of accuracy, precision and sensitivity. The Qin et al., 2001 algorithm was found to be the best performing mono-window algorithm. A short comparison of the application of the Becker and Li 1990 coefficients to AVHRR with the MODIS LST product confirmed the approach to be physically sound.

Keywords: LST; AVHRR; time series; accuracy; precision; sensitivity

1. Introduction

The estimation of land surface temperature (LST) from acquisitions of medium-resolution sensors has a long tradition. Data from the AVHRR (Advanced Very High Resolution Radiometer) sensors flown on NOAA-satellites is operationally available since the early 80s with 1 km spatial resolution at nadir. Given that the data has been archived, the generation of long time series is possible. This paper focuses on a comparison of different mono- and split-window algorithms to derive LST from AVHRR top of atmosphere (TOA) data under clear sky conditions. The comparison is performed against the background of future generation of long time series and their analysis. Included is not only a careful selection of algorithms based on their suitability for time series generation, but also a straightforward specification of related errors and sensitivities.

One way to estimate LST is to remove the attenuation effects of the atmosphere from TOA brightness temperatures. These effects are thereby determined quantitatively using radiative transfer

models (RTMs) and assumptions of the composition of the atmosphere. Given that the atmosphere is well characterized, this method can be very accurate. However, main constraint of such approaches is the high CPU usage of the RTMs, which makes them unsuitable for the processing of extensive time series with pixel-based approaches. More straightforward is the application of methods, which are based on pre-computed functions and coefficients. These allow an effective data processing. A selection of such methods is described in this article specifically for the case of AVHRR bands. A more complete overview of methods is given for example in [1]. Another fast option to retrieve LST are neural networks as described for example in [2].

The estimation of LST from AVHRR/2, AVHRR/3 (13 sensors on the NOAA satellites and now AVHRR/3 on Metop-A and Metop-B), as well as AVHRR-heritage sensors, can be accomplished taking advantage of the splitting of the thermal domain into two channels. The two channels are in wavelength ranges with different absorption features of mainly water vapour. As such, the difference of the two channels can be used to eliminate these atmospheric effects in the data. The two channels are usually located around 11 and 12 μm. Measured brightness temperatures of these two channels can be directly used with the respective surface emissivities to deduct LST. The split-window technique is simple and robust [3] and therefore favoured by many researchers and also operational providers of LST (e.g., Land Surface Analysis Satellite Applications Facility (LSA SAF) [4], Copernicus Land [5], NASA MODIS LST [6]). AVHRR/1 however carries only one band in the thermal domain, only mono-window algorithms are applicable here. These algorithms are missing any measured information about atmospheric absorption; correction procedures are fully dependent on external data. Due to reduction to one band the accuracy of mono-band algorithms is expected to be lower than split-window algorithms [7]. Nevertheless, [8] reported accuracies for mono-window algorithms which are comparable to what was found for split-window algorithms.

Early developments of mono-window algorithms were undertaken by [9], also [10] suggests in his review paper a single channel algorithm. Further work was presented by [11,12]. A more recent overview is given by [1] who address the mono-window algorithms as single-channel methods. The work of [8] complements the overview with an additional statistical mono-window algorithm. All these methods are based on precomputed coefficients and functions, which can vary for different states of the atmosphere.

In the literature, numerous split-window algorithms have been published, starting as early as 1970, when first studies were published about the estimation of Sea Surface Temperatures [13]. The studies were based on modelled and on measured data. A good overview is given in [1]. The authors divide the split-window algorithms into linear and non-linear split-window algorithms. Both these algorithms use the information of one of the two brightness temperatures as a general offset, as well as the difference between the two brightness temperatures to express the magnitude of absorption in the atmosphere. All algorithms come with a range of coefficients, which are determined empirically and reflect the influence of the emissivity, the difference of the emissivities of the two bands, the water vapour in the atmosphere and the sensor view angle. As such, expressions have been developed to empirically deduct the coefficients from these variables in a linear or also non-linear way (e.g., [9,14–18]). Another way of incorporating the influence of water vapour, the longer atmospheric path determined by the view angle, and also temperature itself to the expression, is to use different sets of coefficients for subranges of each variable. This approach was followed for example by [16,19], who simulated band brightness temperatures using a radiative transfer model for a variety of atmospheric conditions and a range of sensor view angles. The atmospheric conditions thereby covered given ranges of atmospheric water vapour, and initial guess LST or T_{air} broad enough, to reflect the most possible atmospheric conditions. The coefficients are then derived for subranges of atmospheric water vapour and initial guess LST or T_{air} from the modelled brightness temperatures and the input LST into the radiative transfer runs. As the band brightness temperatures usually are modelled using the sensors spectral response, the resulting coefficients are sensor-specific and cannot be transferred from one to another sensor without loss of accuracy. There have been studies comparing different

split-window algorithms ([3,20–22]), highlighting the advances and drawbacks of single algorithms and their implementation at that time. As pointed out by [8], the approach of creating sets of coefficients does also apply for mono-window algorithms.

Nowadays, the availability of long term time series of medium resolution data of sensors such as AVHRR enables researchers to investigate climate relevant trends in this multi-decadal Earth Observation data. This study therefore investigates the suitability of a variety of selected algorithms to long time series processing in the framework of the TIMELINE project at the Earth Observation Center (EOC) of the German Aerospace Center (DLR). The project aims in creating long and homogeneous time series from AVHRR/1, AVHRR/2 and AVHRR/3 starting in the early 80s. Most of the proposed split-window algorithms have been developed a few years back. This is why the accompanying coefficient sets are usually given only for one or more of the older AVHRR sensors. These coefficient sets might match perfectly to the one AVHRR sensor or to the set of sensors, but fail to give similar accuracies using the coefficients on all available AVHRR data, as each sensor features its unique spectral response curve. As such, the generation of a longer time series with coefficients from literature are prone to larger errors, which might introduce artificial anomalies like sudden steps in the time series. It is, therefore, of utmost importance to use updated coefficients. Nowadays, the focus of satellite product analysis is more and more shifted towards time series analysis. As such, LST products should be consistent and fit to the concept of climate data records (CDR) [3]. This work therefore includes the generation of new coefficient sets for each of the AVHRR sensor, to enable analyses that focus on time series applications. The generation of the coefficients follows the development of e.g., [16], by generating different coefficient sets for different ranges of columnar water vapour, temperature, mean band emissivity and—if applicable—the difference of the two band emissivities.

The focus of our analysis not only emphasises on the performance, as expressed by accuracy and precision measures, of the split-window algorithms themselves, but also on the sensitivity of the split-window algorithms to their input data (e.g., columnar water vapour). Performance assessments provide information on the magnitude of variation that can be expected from using one or the other split-window approach. The sensitivity analysis highlights the influence of the input variables into the split-window algorithms. This is important mainly in cases where the quality of an input variable is of varying quality. In such cases, reduced quality of input data caused by systematic and/or random effects is propagated to the LST generation—depending on its sensitivity. Both, the performance and the sensitivity measures can be further used for uncertainty and quality estimation of a final LST AVHRR time series product which relate product quality to state of the art user requirements (accuracy better than 1 K as given for example by GCOS ([23]). Current data providers of LST, which aim for long time series with high accuracy, are for example the LSA SAF, which will soon provide consistent Meteosat LST in a joint effort with CM SAF [24]. In addition, the Copernicus Land Programme provides LST from MeteoSat data ([4,5]). Furthermore, the GlobTemperature Project using (Advanced) Along-Track Scanning Radiometer (A)ATSR data ([25]), and NASA's MODIS LST product [26] provide LST time series. The latter product was used in the study for a comparison with one of the proposed algorithms to confirm the soundness and validity of the approach.

2. Data

This assessment was specifically designed for AVHRR. The AVHRR sensors (AVHRR/1, 2, and 3) are mounted on the NOAA-series of satellites, since 2006 AVHRR/3 is also onboard of the MetOp series. AVHRRs carries one band in the red (band 1), one in the near infrared (band 2), and one in the shortwave infrared (band 3) domain. Since AVHRR/3, band 3 can also contain acquired radiation in the mid infrared wavelength. Main bands for LST estimation are however the bands in the thermal infrared. AVHRR/2 and 3 contain two channels in the thermal infrared, while AVHRR/1 features only one band in this domain (Table 1).

Table 1. Spectral bands of the AVHRR sensors.

Channel	AVHRR/1 (NOAA-6,8,10) (μm)	AVHRR/2 (NOAA-7,9,11,12,14) (μm)	AVHRR/3 (NOAA-15,16,17,18,19 MetOp A and B) (μm)
1	0.58–0.68	0.58–0.68	0.58–0.68
2	0.725–1.1	0.725–1.1	0.725–1.0
3A	-	-	1.58–1.64
3B	3.550–3.93	3.550–3.93	3.55–3.93
4	10.5–11.5	10.3–11.3	10.3–11.3
5	10.5–11.5	11.5–12.5	11.5–12.5

Although the different AVHRR sensors measure in the same bands, their spectral responses are not identical, as each sensor is a unique instrument, which is subject to some form of degradation. The degradation of a sensor and therewith possible overall changes in the spectral responses are accounted for by calibration. However, no information about possible changes in the spectral form is available after lunch [27]. Figure 1 shows the spectral response curves of band 4 and band 5 of the different sensors. The response curves may differ substantially not only between the different models, but also within one model (especially AVHRR/3).
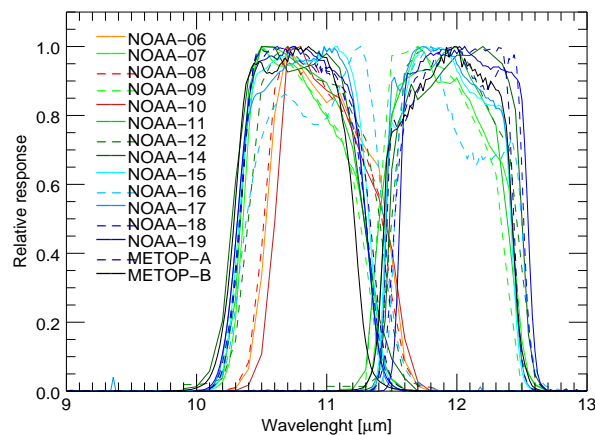

Figure 1. Spectral response curves of band 4 and 5 of the different AVHRR/1, 2, and 3 sensors onboard the NOAA and the MetOp satellite series. AVHRR/1 curves are given in red tones, AVHRR/2 curves in green tones, and AVHRR/3 curves in blue tones and black.

For the method development of this assessment, it was necessary to compile a dataset of atmospheric profiles. For this, a selection of a training database of global profiles (called SeeBor Version 5.0) was used [28]. This database consists of 15.704 global profiles of temperature, moisture and ozone at 101 pressure levels—all for clear sky conditions. The profiles were all quality checked and a surface temperature was assigned to each profile based on a physically based scheme. The profiles of the database are compiled from five different existing databases (NOAA-88, ECMWF 60L training set, TIGR-3, ozone sondes from 8 NOAA Climate Monitoring and Diagnostics Laboratory sites, and radiosondes from 2004 in the Sahara Desert) [28]. From this database a selection was made for this assessment, which ensured having profiles in all ranges of surface temperature and columnar water vapour (see Table 2), as well as from as many land cover types (referring to the International Geosphere-Biosphere Programme (IGBP)) as possible. Due to the uneven distribution of land cover classes in the profile database, the number of profiles for each land cover class differs. A total of 2662 profiles were selected, featuring the following characteristics: 811 profiles are taken from TIGR-3, 317 profiles from Radiosondes, 479 profiles from Ozonesondes, and 1055 profiles are taken from ECMWF. The profiles were taken in a range of different land cover classes as defined in Table 3.

Table 2. Minimum, maximum and mean atmospheric values from selected profiles.

	Minimum	Maximum	Mean
Surface temperature (K)	209.2	327.7	285.9
Wind speed (m·s^{-1})	0.0	20.1	6.6
Mean atmospheric temperature (K)	209.4	252.2	240.1
Mean columnar water vapour (kg/m^2)	0.1	87.4	31.1

Table 3. Assigned land cover classes to selected profiles.

Land Cover Classes	Number of Profiles
Evergreen Needleleaf Forest	414
Evergreen Broadleaf Forest	299
Deciduous Needleleaf and Broadleaf Forest	67
Mixed Forest	169
Open and closed Shrubland	504
Woody Savanna	146
Savannas	96
Grassland	203
Cropland	369
Cropland/Natural Vegetation Mosaic	276
Barren or Sparsely Vegetated	425
Urban	10
Snow and Ice	231
IGBP Water Bodies	772
Unknown	37
Total profiles	4031

Further, a land use classification [29] was used to retrieve emissivity values for the application of the algorithm on real AVHRR data. The emissivity estimation was done using the Vegetation Cover Method (VCM) from [30].

For the comparison with MODIS data, the MOD11A1 V6 Land Surface Temperature and Emissivity product (MOD11A1) was used. The data was retrieved from the USGS HTTP-Server [26]. The MODIS data, as well as the AVHRR data, were projected to a common 1 km lat/lon grid. The MODIS LST product is already cloud screened. FOR AVHRR we used the APOLLO (AVHRR Processing scheme over cLouds, Land and Ocean) tests. APOLLO is an algorithm that was designed in the late 1980s and continuously improved and extended afterwards [31,32]. Additionally, a buffer was laid around all clouds, as some border pixels were not detected by the algorithm.

3. Methods

Ten mono- and split-window approaches were tested to their suitability of long-term processing of LST from AVHRR data. Table 4 presents all the approaches tested and compared during this study. General to all algorithms is that new coefficients were generated for each of them. As such, only the approach was tested, but not the original published expression.

Table 4. Formulations of the tested mono- and split-window methods.

Mono-Window Algorithms			
Method	**Formula**		**Notes**
Price 1983 [9]	$LST = \frac{p_0 + \varepsilon_4}{p_1}(p_2 t_4 + p_3)$	(1)	Developed for AVHRR/1
Qin et al., 2001 [11]	$LST = \frac{1}{c}(p_2(1 - c - d) + t_3(p_1(1 - c - d) + c + d) - dT_{atm})$	(2)	Developed for LANDSAT TM
	$c = \varepsilon_4 \tau$	(3)	
	$d = (1 - \tau)(1 + \tau(1 - \varepsilon_4))$	(4)	
	$\tau = p_0 - p_1 cwv$	(5)	
Jiménez-Muñoz and Sobrino 2003 [12]	$LST = \gamma\left(\frac{1}{\varepsilon}(\psi_1 L_4 + \psi_2) + \psi_3\right)\Delta$	(6)	Developed for LANDSAT TM. Δ and γ are coefficients derived from an initial guess LST, $\psi_{1,2,3}$ are atmospheric functions, which are calculated on the basis of the wavelength and the columnar water vapour
Linear	$LST = \frac{p_0 t_4 + p_1}{\varepsilon_4{}^4}$	(7)	Linear expression, developed for this assessment
Split-Window Algorithms			
Method	**Formula**		**Notes**
Price 1984 (as given in Pozo Vázquez et al., 1997 [21])	$LST = (t_4 + p_0(t_4 - t_5))\left(\frac{p_1 - \varepsilon}{p_2} + p_3 t_5 d\varepsilon\right)$	(8)	Developed for AVHRR/2
Becker and Li 1990 [14]	$\alpha = 1 + p_1\frac{1 - \varepsilon}{\varepsilon} + p_2\frac{d\varepsilon}{\varepsilon^2}$	(9)	Developed for AVHRR/2
	$\beta = p_3 + p_4\frac{1 - \varepsilon}{\varepsilon} + p_5\frac{d\epsilon}{\varepsilon^2}$	(10)	
	$LST = p_0 + \alpha\left(\frac{t_4 + t_5}{2}\right) + \beta\left(\frac{t_4 - t_5}{2}\right)$	(11)	
Ulivieri et al., 1994 [17]	$LST = t_4 + p_0(t_4 - t_5) + p_1(1 - \varepsilon) - p_2 d\varepsilon$	(12)	Developed for AVHRR/2
Wan and Dozier 1996 [16]	$LST = \left(p_0 + p_1\frac{1 - \varepsilon}{\varepsilon} + p_2\frac{d\varepsilon}{\varepsilon^2}\right)\frac{t_4 + t_5}{2}$ $+ \left(p_3 + p_4\frac{1 - \varepsilon}{\varepsilon} + p_5\frac{d\epsilon}{\varepsilon^2}\frac{t_4 - t_5}{2}\right) + p_6$	(13)	Developed for MODIS
Yu et al., 2008 [3] (extension to [17])	$LST = t_4 + p_0(t_4 - t_5) + p_1(1 - \varepsilon) - p_2 d\varepsilon$ $+ p_3(t_4 - t_5)(\cos(VA) - 1)$	(14)	Developed for AVHRR/2 and AVHRR/3
Jiménez-Muñoz and Sobrino 2008 [33]	$LST = t_4 + p_1(t_4 - t_5) + p_2(t_4 - t_5)^2 + p_0$ $+ (p_3 + p_4 CWV)(1 - \varepsilon)$ $+ (p_5 + p_6 CWV)d\varepsilon$	(15)	Developed for ATSR2, AATSR, MODIS, AVHRR/3, IMAGER, SEVIRI

t_4 = Brightness temperature of AVHRR band 4 (K); t_5 = Brightness temperature of AVHRR band 5 (K); ε_4 = Emissivity of band 4; ε = Mean emissivity of band 4 and band 5; $d\varepsilon$ = Difference of band 4 and band 5 emissivity (band 4 − band 5); VA = View angle (°); CWV = Columnar water vapour (kg·m^{-2}); p_{1-6} = Ccoefficients, as described in Section 3.1. The coefficients are different for each equation.

3.1. Coefficients Generation

The coefficients for the mono- and the split-window algorithms were all retrieved in the same manner. The first step consists of the selection of profiles from the SeeBor V5 database, as described Section 2. MODTRAN 5.3 was then run using the selected profiles. Thereby, each profile was used multiple times with varying surface characteristics. First of all, the already existing LST values of each profile [28] were tripled as [LST − 5 K, LST, LST + 5 K] to increase the simulated LST variability. Then, the sensor view angle was set to the values [0, 20, 40, 60], the mean band emissivity to values of [0.89, 0.92, 0.96, 0.99] and the difference of the band emissivities to [−0.02, −0.01, 0.00, 0.015]. MODTRAN 5.3 was run for each combination of these values and for each band. Resulting radiances were saved in a database. In a subsequent step, the radiances were convolved over the spectral response of each of the AVHRR sensors and converted to brightness temperatures using Planck's radiation equation as described in the NOAA Polar Orbiter Data [34] User's Guide for AVHRR/1 and AVHRR/2 and NOAA KLM User's Guide for AVHRR/3 [35]. One half of the brightness temperatures were then used together with the associated input LST values for the coefficient estimation of the six split-window equations using least squares minimization. The other half was reserved for accuracy analysis.

The coefficient estimation was done for three different kinds of input cases, whenever possible:

- Case A: The coefficients were retrieved for each class of sensor view angle (VA).
- Case B: The coefficients were retrieved for each class of sensor view angle, as well as for 8 ranges of columnar water vapor (VA/CWV).
- Case C: The coefficients were retrieved for each class of sensor view angle, for 8 ranges of columnar water vapor and for 4 classes of surface temperature (VA/CWV/LST). Using the Jiménez-Muñoz and Sobrino 2008 algorithm and Qin et al., 2001 algorithm, the coefficients were retrieved for sensor view angle and surface temperature (VA/LST). The Yu et al., 2008 algorithm uses ranges of columnar water vapour and surface temperature only (CWV/LST). Case C has for each algorithm the maximal surface ranges necessary.

The hypotesis is that Case C will give the highest accuracies, while Case A shows the highest errors. Due to the large range of columnar water vapor and LST, not all possible combinations in Case C could be filled using the data from the selected profiles. In operational processing of LST therefore, also Case B and Case A might be of interest. Therefore Case A, B, and C were analysed.

According to [3,36], the temperature lapse rate influences the accuracy and precision of the split-window retrieval. They found, that the final accuracy decreases with increasing lapse rate and brightness temperature difference. As a consequence, the different set of coefficients were retrieved for daytime (LST − T_{air} > −2 K) and night-time (LST − T_{air} < 2 K) conditions, selecting only the appropriate profiles. As the conditions have an overlap of 4 K, the profiles also have a proper subset. It should be noted that the criteria for day and night discrimination are violated for example in case of advection of air masses. The discrimination between day and night conditions could have been expanded to the definition of different LST − T_{air} subranges. However, the present profile database does not provide enough samples to add this extra dimension. Additionally, accurate information about the skin temperature-T_{air} relation must be available before application of the algorithm to real data, which might be critial—especially for the past decades.

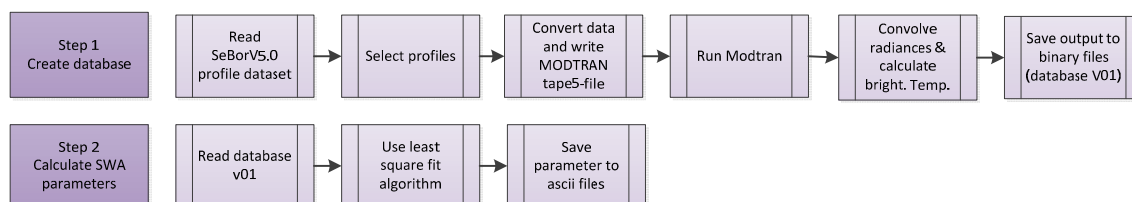Figure 2 shows a schematic overview over the coefficient retrieval.

Figure 2. Schematic overview of split-window coefficient retrieval.

3.2. Performance and Sensitivity Assessment

Performance, as expressed by accuracy and precision, was measured using the reserved second half of the modelled brightness temperature values from the coefficient retrieval. Thereby, for each profile and view angle, LST was calculated using the algorithms on the basis of the brightness temperature, and the atmospheric information from the profiles for parameter class selection (columnar water vapour and initial guess LST). Statistics was summarized for each class of view angle, initial guess LST, and columnar water vapour by comparing the initial LST and the output LST from the split-window algorithms. Such, the difference between the initial LST and the output LST from the split-window algorithms for each class was calculated. From these differences then statistical values were derived. The comparison was performed in form of mean absolute difference (MAD), root mean square (RMS), standard deviation (STDEV), and correlation coefficient (r^2).

The sensitivity analysis was undertaken by artificially altering the input parameters columnar water vapour, initial guess LST, mean band emissivity, and difference of band emissivity. In case of

the columnar water vapour and the initial guess LST, the wrong input values affect the choice of the parameter class and led therefore to a lower accuracy of LST. In case of wrong mean emissivity and wrong differences of band emissivity, the correct parameter classes are chosen, but the algorithm does not foresee such wrong values and therefore renders incorrect LST values. The magnitude of the over- or underestimation of LST is expressed in boxplot figures. Furthermore, total sensitivity is given for each algorithm.

3.3. Emissivity Estimation

To estimate LST using mono- or split-window algorithms, emissivity must be known a priori. For the comparison with MODIS data, the emissivity was estimated using the *Vegetation cover method* (VCM) from [30]. The VCM method was proposed for AATSR data. Due to the spectral similarity of the thermal bands of AVHRR and AATSR, the method could be employed for this study.

$$\varepsilon_k = \varepsilon_{kv} FVC + \varepsilon_{kg}(1 - FVC) + 4\langle d\varepsilon_k \rangle FVC\,(1 - FVC) \tag{16}$$

ε_k with k = 4, 5 is the emissivity of band 4 resp. 5, ε_{kv} is the emissivity of vegetation, ε_{kg} is the emissivity of the ground below the vegetation, $\langle d\varepsilon_k \rangle$ is the maximum cavity term, and FVC is the fractional vegetation cover. The coefficients ε_{kg}, ε_{kv} and $\langle d\varepsilon_k \rangle$ are dependent on land surface cover and the spectral band. Additionally, to the estimation of band emissivity, the method of [30] offers the possibility to estimate the related uncertainty of the emissivity. It should be noted that this expression does not account for the anisotropy of emissivity. As pointed out by [37] on the example of inorganic soils, this can lead to systematic errors in LST ranging between 0.4 K and 1.3 K for dry atmospheres.

4. Results

4.1. Performance

The different performance measures of Case A, B, and C are shown on the example of the revised Qin et al., 2001 and the Becker and Li 1990 algorithm only. The relation of the performances of the different cases is very similar for all algorithms; therefore the following detailed presentation of the results from the two revised algorithms is sufficient. Figure 7 will later give an overview over all algorithms. The left-hand side of Figures 3–6 show the r^2, the MAD, the RMS, and the STDEV separately for each view angle, columnar water vapour class and daytime coefficient sets. Thereby, case C shows the best agreement in all columnar water vapour classes. r^2 values are very high except for extremely humid conditions, and MAD values stay below 0.5 K (Becker and Li 1990) and around 1 K (Qin et al., 2001) except for very humid conditions and extreme view angles. Case B does show only slightly lower accuracies, except in the case of very low temperatures. There, this coefficient set is not sufficient. Case A shows the lowest accuracies in all cases, MAD values do hardly fall below 1 K. There is a general decrease with higher view angles and with higher columnar water vapour, a result that was expected from other studies [38]. The right-hand side of Figures 3–6 show the same data but this time split up by ranges of initial guess LST. The relations between cases A, B, and C as well as the influence of the view angle are the same. Furthermore, all algorithms work better in case of very low temperatures, which is due to the usually very low water content in the atmosphere in such conditions. Interestingly, the case B and case C give slightly better performance in high temperature conditions (310–320 K) than in moderate temperature conditions (270–290 K).
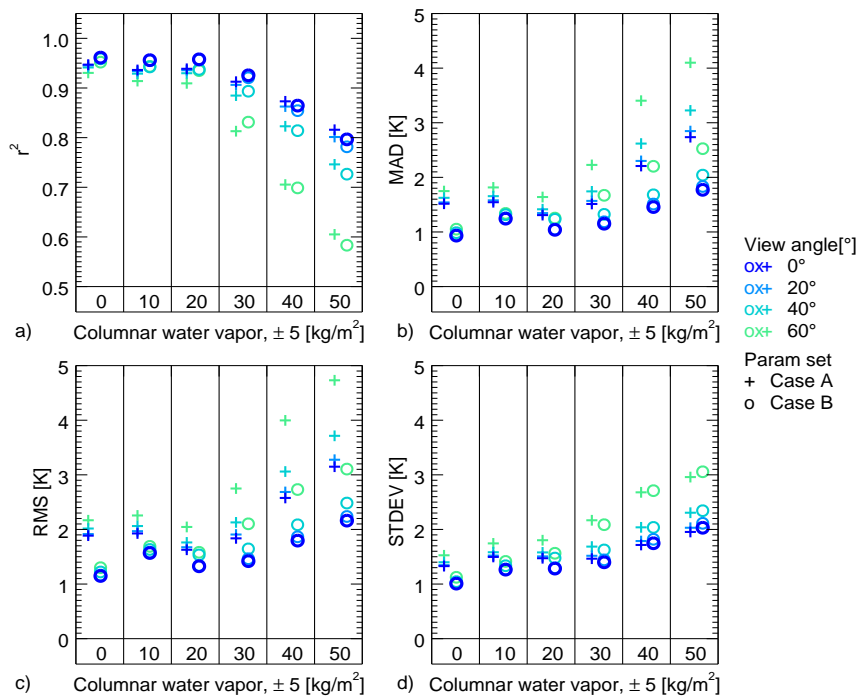
Figure 3. Performance measures (**a**) r^2; (**b**) MAD; (**c**) RMS; (**d**) STDEV for the revised Qin et al., 2001 algorithm for different CWV ranges (daytime coefficients).
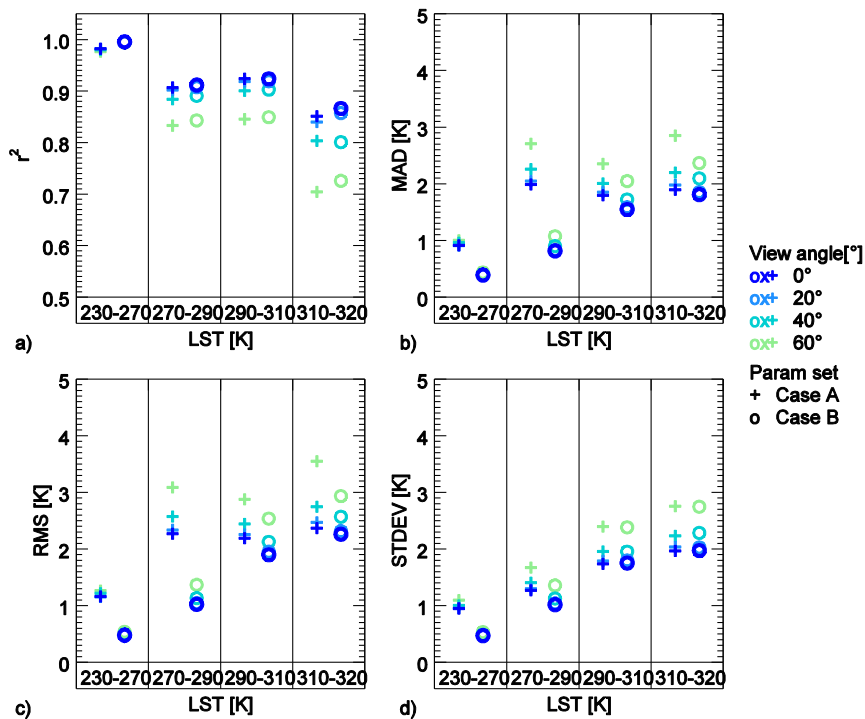

Figure 4. Performance measures (**a**) r^2; (**b**) MAD; (**c**) RMS; (**d**) STDEV for the revised Qin et al., 2001 algorithm for different LST ranges (daytime coefficients).
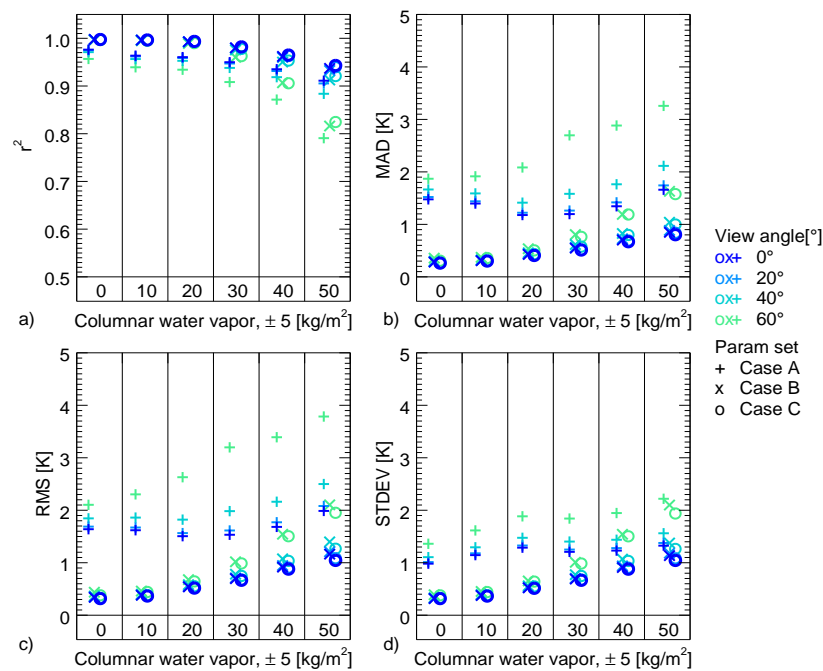
Figure 5. Performance measures (**a**) r^2; (**b**) MAD; (**c**) RMS; (**d**) STDEV for the revised Becker and Li 1990 algorithm for different CWV and LST ranges (daytime coefficients).
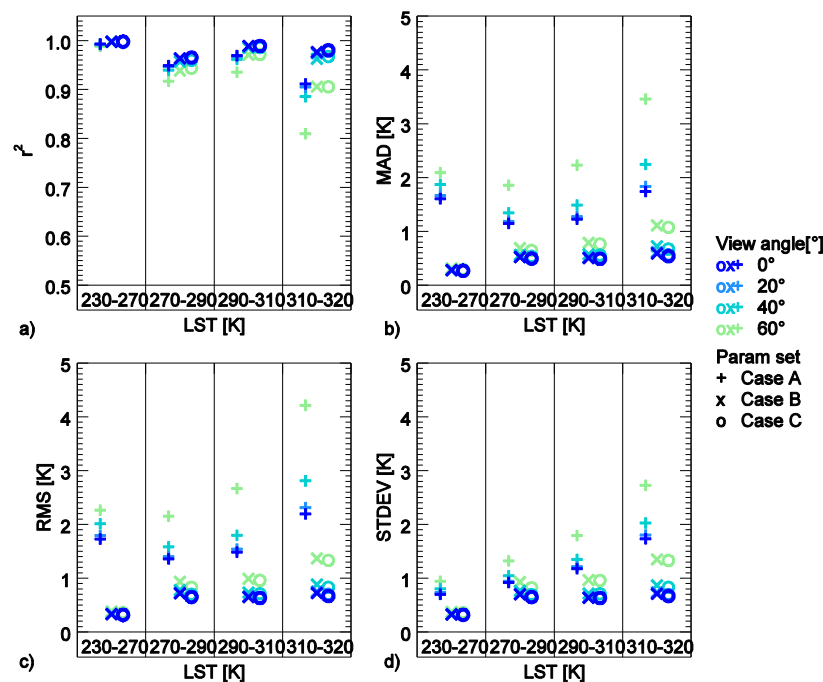
Figure 6. Performance measures (**a**) r^2; (**b**) MAD; (**c**) RMS; (**d**) STDEV for the revised Becker and Li 1990 algorithm for different CWV and LST ranges (daytime coefficients).

The performance measures of the night-time coefficients are very similar to the daytime values both in magnitude and in tendency. Figure 7 shows overviews over all methods, plotted for the different ranges of columnar water vapour, a sensor view angle of 0°, an initial guess LST range from 270 to 310 K, and separated for the day and the night coefficient sets. Where possible, the coefficient sets from case C are taken, as these were identified as the best performing coefficient sets. Most methods show the same relation with columnar water vapour classes. The methods work well in low humidity

atmospheres, but decrease in accuracy when using more humid profiles. For the split-window algorithms, the MAD stays below 0.4 K, however the mono-window algorithms have partly higher MAD values. The Qin et al., 2001 and the linear method stay below 1 K, but the Price 1983 algorithm reaches almost 2 K, while the Jiménez-Muñoz and Sobrino 2003 is even higher. The Jiménez-Muñoz and Sobrino 2003 algorithm shows low accuracies in humid conditions (mostly out of the range in Figure 7). Especially the night coefficients give lowest accuracies in very humid conditions. This is in line with the findings described in [1].
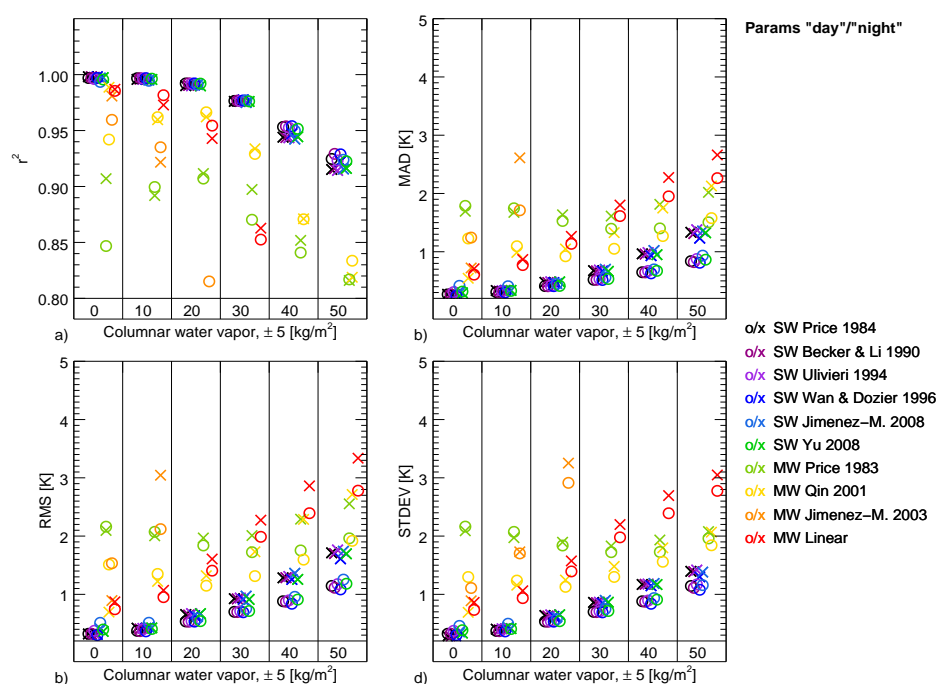

Figure 7. Performance measures (**a**) r^2; (**b**) MAD; (**c**) RMS; (**d**) STDEV of all revised algorithms, separated for different classes of CWV and an initial guess LST range from 270 to 310K. Results are retrieved from the case C coefficient set, a sensor view angle of $0°$, and daytime (given by symbol $\bigcirc$) and nighttime (given by symbol $\times$) conditions.

Figure 8 shows an extract of the statistical data for a columnar water vapour range of 0–35 kg·m^{-2}, a sensor view angle of $0°$, and separated for the initial LST classes. In addition, here, accuracy is slightly higher in very cool conditions, MAD stays below 1 K for the linear, the Qin et al., 2001 and the daytime Jiménez-Muñoz and Sobrino 2003 methods and 0.4 K all the split-window methods, except the nighttime coefficients of Jiménez-Muñoz and Sobrino 2008. At moderate temperatures, the accuracies get slightly lower again, while for higher temperature ranges, accuracies get better. These slight shifts are a result of available input data for the different ranges of columnar water vapour and initial guess LST. By narrowing the range of columnar water vapour to 0–25 kg·m^{-2}, this effect would not show in the figure anymore. Jiménez-Muñoz and Sobrino 2008 night time performs least well at very low temperatures with columnar water vapour higher than 15 kg·m^{-2}, therefore a peak value is shown in Figure 8. The Jiménez-Muñoz and Sobrino 2003 algorithm is mostly out of axis range in Figure 8.

In case of the split-window algorithms, the differences between the single methods are low; MAD and RMS differences stay below 0.2 K. Under normal atmospheric conditions the choice of algorithm does not seem to be crucial. The method of Yu et al. 2008 incorporates the cosine of the view angle. However, obviously, this expression does not have an advantage over the other methods, when using the view angle optimized coefficients (Figures 7 and 8). In case of very high view angles ($60°$) the performance of this method is even worse (data not show). In its original version (Yu et al.,

2008) the algorithm was used with coefficients regressed using all angles, the results are therefore not directly comparable.
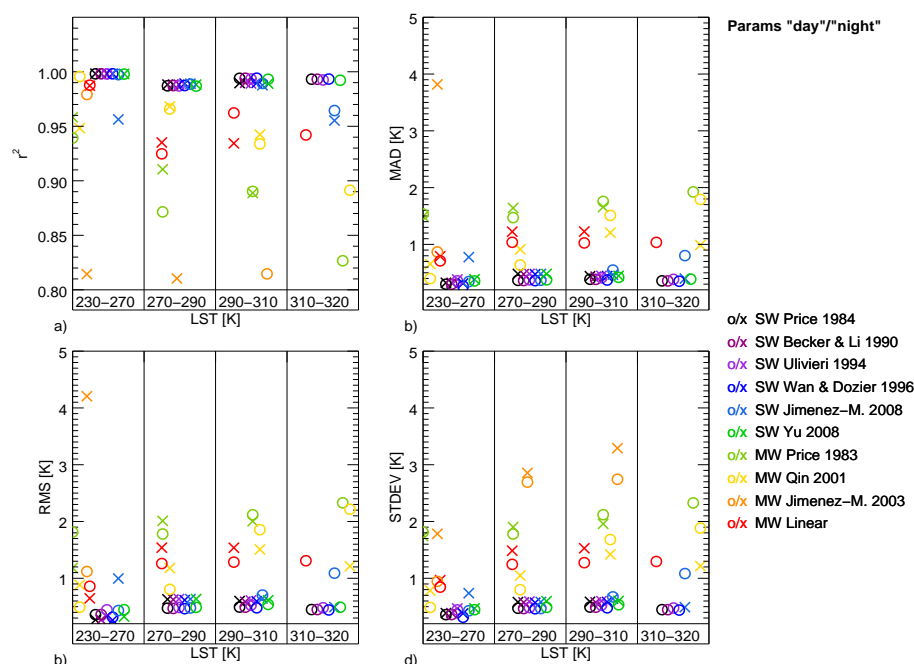

Figure 8. Performance measures (**a**) r^2; (**b**) MAD; (**c**) RMS; (**d**) STDEV of all revised algorithms, separated for different classes of initial guess LST and a columnar water vapour range of 0–35 kg·m^{-2}. Results are retrieved from the case C coefficient set, a sensor view angle of 0°, and daytime (○) and nighttime (×) conditions.

For both the mono-window and the split-window algorithms accuracy differences were found between the sensors due to their different spectral response. For most split-window equations the differences were found to be constant (difference between maximum and minimum MAD) mostly below 0.6 K, with decreasing differences in more humid conditions. The mono-window algorithms showed a higher spread with maximum differences between the sensors reaching 2 K for humid conditions. Jiménez-Muñoz and Sobrino 2003 has exceptional high deviations in humid conditions, as this algorithm is not suitable for such conditions. Table 5 shows the mean of all maximal deviations between sensors for given CWVs.

Table 5. Mean of maximal deviation of accuracy (MAD) between different sensors.

Method	CWV = 0	CWV = 10	CWV = 20	CWV = 30	CWV = 40
Mono-Window					
Price 1983	0.62	0.73	0.76	0.88	1.12
Qin et al., 2001	0.52	0.37	0.30	0.63	1.19
Jiménez-Muñoz and Sobrino 2003	0.34	1.45	7.14	19.72	36.66
Linear	0.26	0.10	0.46	1.16	1.80
Split-Window					
Price 1984	0.70	0.66	0.52	0.31	0.21
Becker and Li 1990	0.71	0.67	0.52	0.32	0.20
Ulivieri et al., 1994	0.61	0.66	0.52	0.32	0.19
Wan and Dozier 1996	0.72	0.66	0.53	0.32	0.19
Jiménez-Muñoz and Sobrino 2008	0.64	0.64	0.46	0.28	0.29
Yu et al., 2008	0.60	0.64	0.50	0.32	0.23

4.2. Sensitivity

The final accuracy of a product does not only depend on the model performance, but also on the quality of its input data. Considering the fact, that input data to LST estimation is never perfectly accurate, it is straightforward to analyse the sensitivity of each method to its input bands. In this study, sensitivity was evaluated by simple deterministic analysis. The results of an LST algorithm with first a reference input value and second with deviating input values was compared. The differences of the resulting LST from the first reference run and the other (deviating) runs show the sensitivity of an algorithm to that input variable. For this analysis, it is assumed that the deviations have a discrete uniform distribution. Figures 9–11 show the sensitivity of the Qin et al., 2001 method to columnar water vapour, emissivity of band 4 and mean atmospheric temperature in form of boxplots for daytime conditions. In case of all input variables, the error increases with increasing deviation of the input variable from its 'true' value.
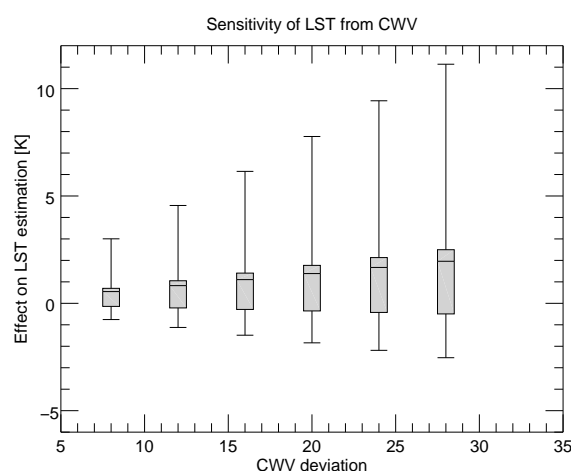

Figure 9. Sensitivity of the Qin et al., 2001 method to columnar water vapour for daytime conditions.
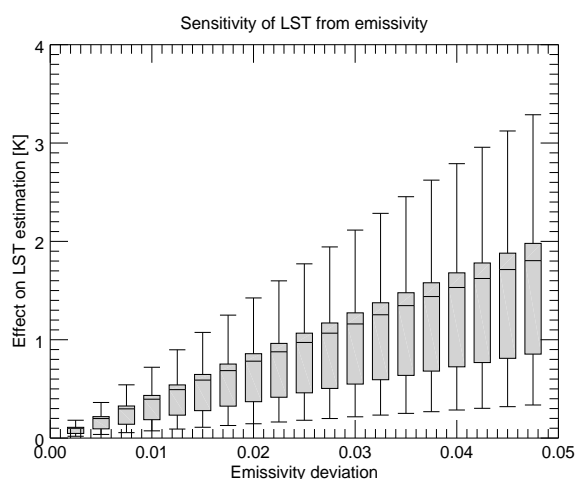

Figure 10. Sensitivity of the Qin et al., 2001 method to band emissivity for daytime conditions.

The sensitivities for columnar water vapour, emissivity, and mean atmospheric temperature show larger scattering range with higher deviations. In single cases the error occurring in the LST due to a specific error in an input dataset might be considerably higher than the average error. In case of columnar water vapour for example, the maximum error occurring at 28 kg/m^2 for columnar water vapour is 11.1 K. In case of an emissivity input error of 0.0475 a maximal error of 3.3 K would occur. However, at lower ranges of input error, which might be the majority, the resulting errors in LST are

much lower. The MAD at an input error of 8 kg/m^2 for columnar water vapour is 0.53 K and the MAD at an input error of 0.02 for emissivity is 0.65 K only. In nighttime conditions, the sensitivity is generally lower, but especially for the columnar water vapour. The MAD at an input error of 8 kg/m^2 for columnar water vapour is then 0.24 K.
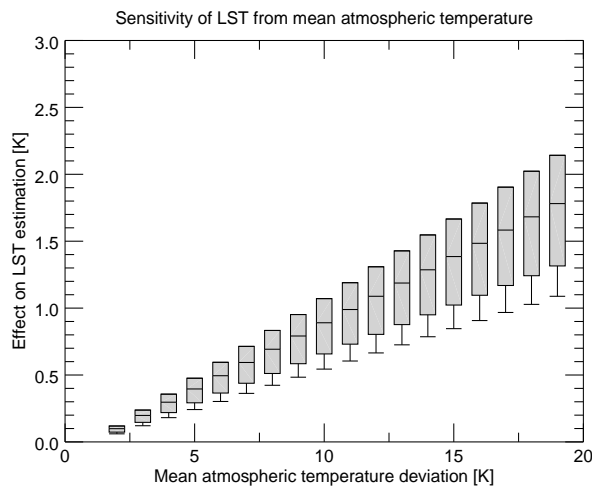

Figure 11. Sensitivity of the Qin et al., 2001 method to mean atmospheric temperature for daytime conditions.

Figures 12–14 show the sensitivity of the Becker and Li 1990 method to total columnar water vapour, mean emissivity and emissivity difference. In addition, here, the error increases with larger deviation of the input variable.
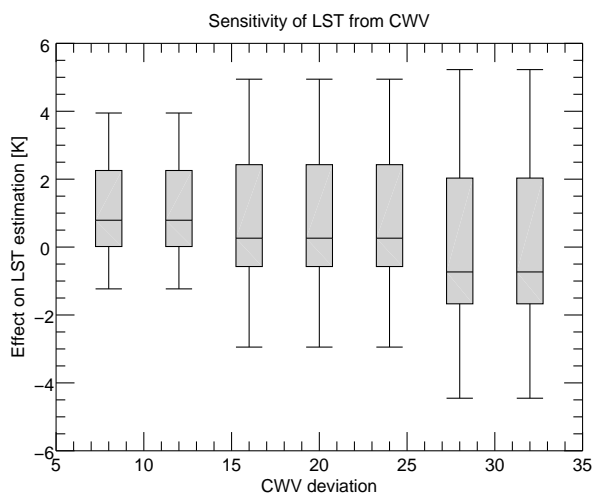

Figure 12. Sensitivity of the Becker and Li 1990 method to columnar water vapour for daytime conditions.
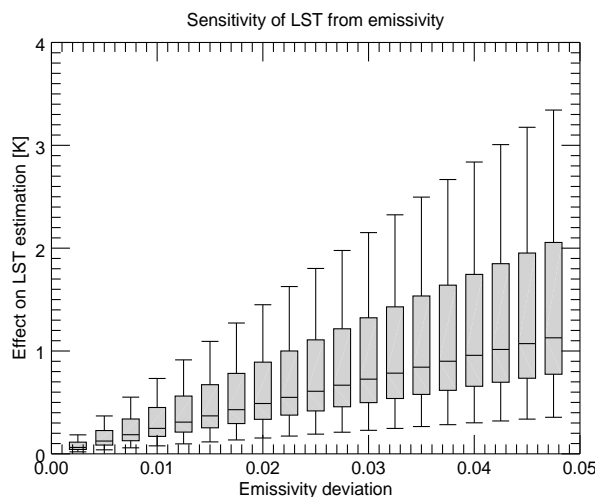
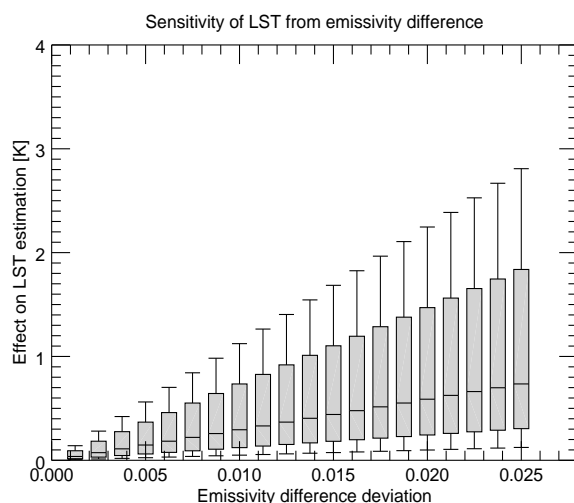**Figure 13.** Sensitivity of the Becker and Li 1990 method to mean emissivity for daytime conditions.

Figure 14. Sensitivity of the Becker and Li 1990 method to emissivity difference for daytime conditions.

Similar to the mono-window algorithm, the sensitivities for columnar water vapour, emissivity and emissivity difference show larger scattering ranges with higher deviations for the split-window algorithm Becker and Li 1990. The maximum absolute error occurring at $28 \, kg/m^2$ for columnar water vapour is 4.1 K. In case of an emissivity input error of 0.0475 also a maximal error of 3.3 K would occur. The MAD at an input error of $8 \, kg/m^2$ for columnar water vapour is 0.81 K and the MAD at an input error of 0.02 for emissivity is 0.61 K. In nighttime conditions, the sensitivity is also lower for the columnar water vapour. The MAD at an input error of $8 \, kg/m^2$ is then 0.26 K.

In case all variables are incorrect the errors sum up to a total sensitivity. This is shown at the example of an input error of total water vapour of $8 \, kg/m^2$, an emissivity error of 0.02, and—where applicable—an error of mean atmospheric temperature of 5 K and emissivity difference of 0.02. In case of the mono-window algorithms, the algorithms of Price 1983 and Qin et al. 2001 show the lowest total sensitivities. Price 1983 has an almost neglectable sensitivity to emissivity, which is mainly due to the magnitude of the first constant in the formula. Qin et al. 2001 has an extra dependence on mean atmospheric temperature, which increases the total sensitivity between 0.25 K and 0.5 K. In case of the split-window algorithms, all algorithms are in a similar range, except the Jiménez-Muñoz and Sobrino 2008 which shows a higher total sensitivity due to its sensitivity to emissivity and emissivity difference. Figure 15 shows the comparison of the sensitivities of all mono-window and split-window

algorithms. Thereby, the error of columnar water vapour was set to 8 kg/m^2, the error of mean emissivity, respective band 4 emissivity to 0.02, the error of emissivity difference to 0.02, and finally the error of mean atmospheric temperature to 5 K.

The estimation of emissivity using the VCM method requires a land use classification, which itself is prone to errors (a) due to different definitions of classes by different authors; (b) due to different classification approaches; (c) due to limited classification accuracy; and (d) due to changing surface cover over time [39]. These errors are passed to the emissivity estimation. To estimate the resulting error from misclassification, the absolute difference between the emissivity of one 'true' class and the emissivity from other 'wrong' classes was calculated for all possible values of FVC (ranging from 0 to 1). The error of a certain misclassification would then be expressed as the MAD. If, for example, a pixel has in reality the class 'woodland', but it was misclassified as 'urban area', then the MAD would be calculated as the mean of all absolute differences between the 'true' woodland emissivities and the 'wrong' urban area emissivities for the given range of FVC. Figure 16 shows the errors in emissivity for all classes for band 4 and band 5. Misclassification errors from one vegetated to another vegetated class are generally low (blue color). In addition, the misclassification from a vegetated to an urban class does not substantially change the emissivity. However, the misclassification of the classes 'bare rock/ground' do change the emissivity up to 0.06 in band 4 and 0.04 in band 5 (Figure 16 limits the plotted errors to 0.03, to enhance the low error differences). In addition, the class 'snow and ice' shows higher values up to 0.02 in band 5.
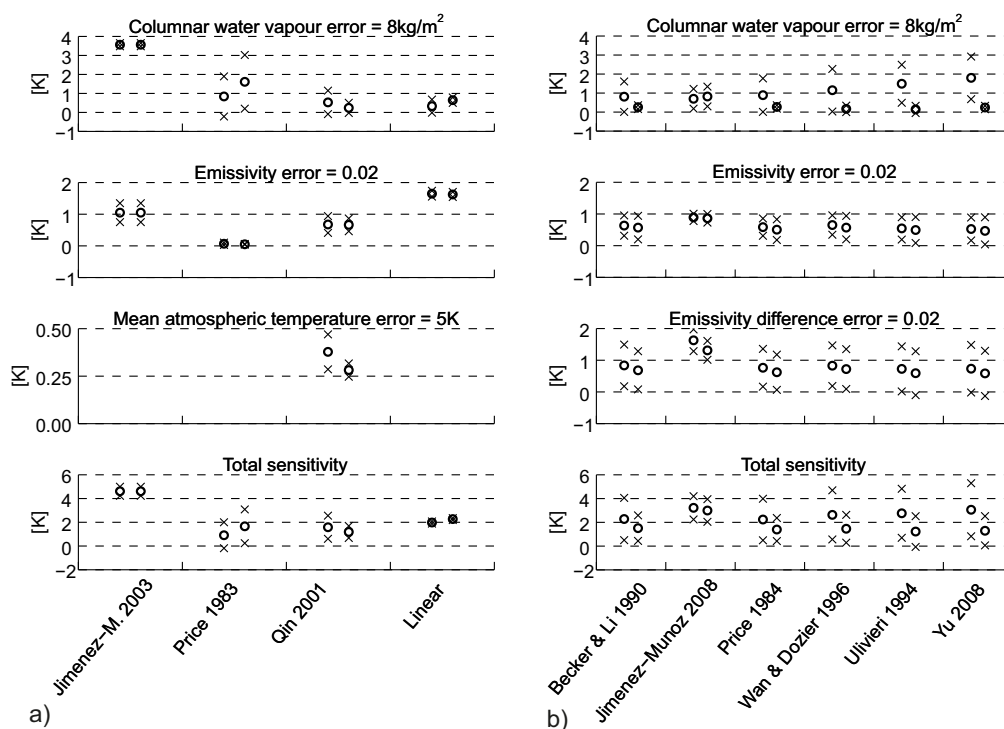

Figure 15. Overview over all (**a**) mono and (**b**) split-window methods. Circles (o) depict the MAD, × stays for the MAD ± the standard deviation (left values: daytime, right values: night time).
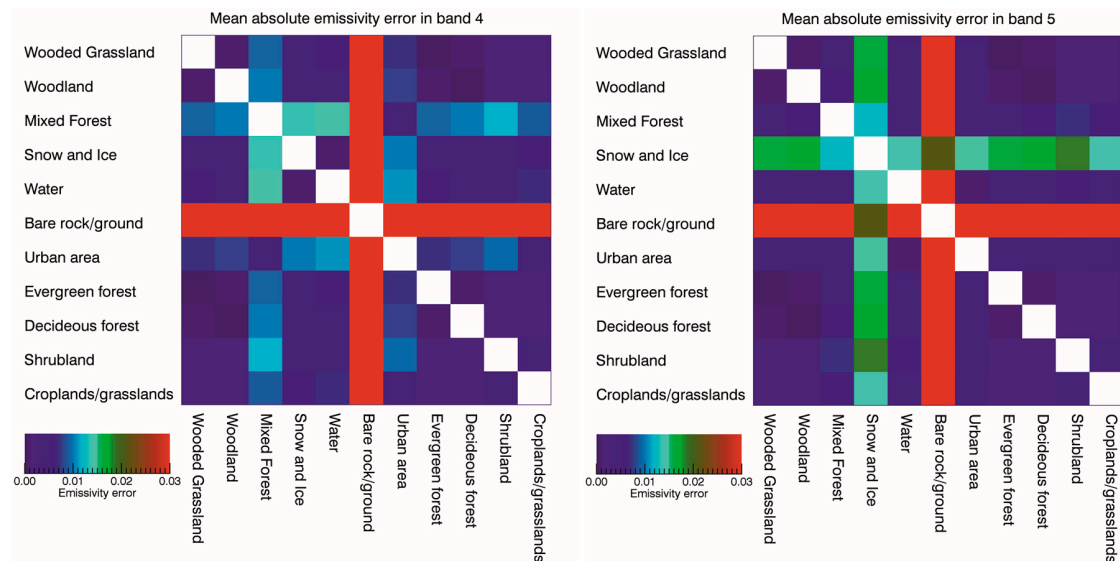
Figure 16. Error in emissivity estimation resulting from misclassification of pixels.

5. Comparison with MODIS Data

To check the validity of the approach, AVHRR LST calculated with the Becker and Li 1990 algorithm combined with the Vegetation Cover Method (VCM) from [30] was compared with the MODIS V005 LST product (MOD11A1/MYD11A1) and the results of using the original Becker and Li 1990 coefficients and the NDVI approach of [40]. The comparison was done for the year 2001 and NOAA-16. For the calculation of the new AVHRR LST, preferably Case C (consideration of ranges for view angle, columnar water vapour and first guess LST) coefficients were used. In case where no coefficients were available due to missing coefficients in the database, case B or even case A might have been used. This could happen in rare extreme situations (e.g., very cold and humid), where the database would not provide enough profiles to calculate coefficients. The direct application of the coefficients provided for discrete ranges leads to discontinuities in the resulting LST images. To counteract this, actual LST values are weighted means from different parameter classes using trilinear interpolation.

For the comparison, the data was filtered in several ways. Firstly, all pixels which were marked as cloud cover in either the AVHRR or the MODIS product were removed. Second, each cloud-free pixel was checked for its spatial homogeneity. If the standard deviation in a window of 3×3 pixels would exceed 0.5 K, the middle pixel would be removed. This resulted in a 55% reduction of the data. This filtering was necessary to reduce noise effects due to inaccurate geometric alignment. The view angle difference between AVHRR and MODIS was allowed to be maximal 20°; the difference between acquisition times was set to maximal 10 min. The minimum number of pixel per tile was set to 50. Statistical measures were then calculated for different geographical subsets as defined by the MODIS tiles. Such, overall MAD between these filtered AVHRR and MODIS is 1.8 K, overall standard deviation is 1.4 K. Using the same collection of pixels, the MAD between AVHRR and MODIS using the original approach [40] is 3.3 K.

Figure 17 shows the MAD per MODIS tile and date in relation to mean total column water vapour, mean emissivity, and mean zenith view angle. The MAD values are given for the original (grey) and the newly calculated (black) coefficients. The figures do not reveal a relation between the MAD magnitudes with any of the other variables; the error seems to be distributed randomly.

Figure 18 shows the same MADs as in Figure 17, but this time in relation to the difference of mean emissivities, mean absolute zenith angle and time difference, each between AVHRR and MODIS. To each MAD value its associated standard deviation is plotted. The MAD increase slightly with

increasing MADs of emissivity and view zenith angle. Time differences (acquisition times around noon) hardly show an influence.
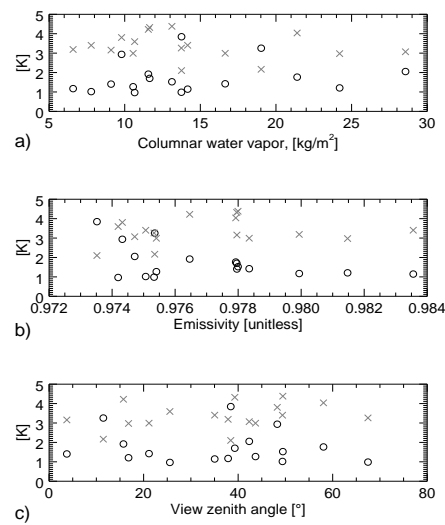

Figure 17. MAD per MODIS tile and date in relation to the (**a**) mean total column water vapour; (**b**) mean absolute emissivity difference; and (**c**) mean absolute zenith view angle difference between AVHRR and MODIS. The figures show the MAD resulting from applying the new coefficients (black circles) and the original approach (grey crosses).

Table 6 shows more statistical values in tabular form together with the number of available pixels per tile. In 13 out of 17 tiles, the MAD is below 2 K. In many cases, the AVHRR LST shows negative offset compared to the MODIS LST resulting in negative mean differences. It should be noted that the magnitude of the dataset should be enlarged for a more thorough validation. Nevertheless, the statistical values of this chapter may serve as indication of the validity the LST approach. Unfortunately, the data is, despite of application of the cloud masks and an additional cloud border buffer, not fully cleared of cloud contaminated pixels, which leads to higher MADs in some cases (e.g., Tile Nr. 16 in Table 6—data not shown).
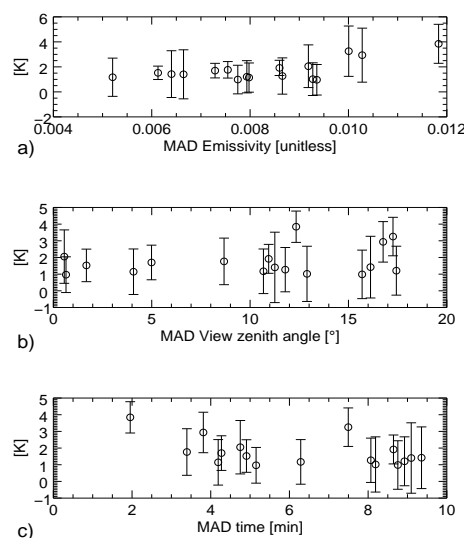

Figure 18. MAD per MODIS tile and date in relation to the (**a**) mean absolute emissivity difference; (**b**) mean absolute zenith view angle difference; and (**c**) mean absolute time difference between AVHRR and MODIS. To each MAD value (circle) the MAD ± the standard deviation is drawn.

Table 6. Statistical measures of the difference between AVHRR and MODIS of selected dates resulting from the filtering. LST AVHRR was calculated using the Becker and Li (1990) algorithm with the newly derived coefficients (LST$_{new}$) and the original approach (LST$_{orig}$).

Tile Nr.	Nr. of Pixels	MAD LST$_{new}$ (K)	MEAN LST$_{new}$ (K)	STDEV LST$_{new}$ (K)	MAD LST$_{orig}$ (K)	MAD Mean Emissivity (%)	MAD VZA (°)	Mean TCWV (kg/m^2)
1	1544	0.97	0.21	1.22	3.59	0.009	0.64	10.66
2	144	2.94	2.09	2.16	3.81	0.010	16.76	9.79
3	132	1.02	−0.89	1.31	3.40	0.009	12.89	7.79
4	3421	1.17	−0.39	1.53	3.19	0.005	10.68	6.59
5	157	1.40	−0.69	1.96	3.16	0.007	11.26	9.10
6	120	1.42	−0.05	1.87	2.99	0.006	16.12	16.64
7	3086	2.05	−1.76	1.71	3.07	0.009	0.55	28.56
8	122	1.76	1.76	0.66	4.04	0.008	8.68	21.40
9	98	0.99	−0.87	1.14	3.26	0.008	15.69	13.75
10	23,927	1.15	0.76	1.17	3.40	0.008	4.07	14.16
11	460	1.21	−0.95	1.29	2.98	0.008	17.43	24.21
12	11,541	3.25	−3.24	2.01	2.17	0.010	17.26	19.02
13	2123	1.92	1.91	0.61	4.22	0.009	10.94	11.56
14	115	1.27	−1.01	1.45	2.99	0.009	11.79	10.56
15	5136	1.70	1.69	0.58	4.32	0.007	4.99	11.65
16	313	3.84	−3.84	1.56	2.10	0.012	12.34	13.76
17	1256	1.52	1.52	0.54	4.38	0.006	1.68	13.12

6. Discussion

In this work, four mono-window and six split-window algorithms were analysed in terms of their suitability for long time series processing of AVHRR data. To make the algorithms comparable, new coefficients were calculated for all algorithms in the same simulated data framework. It was found that the performance of the split-window algorithms were very similar in all cases. The MAD stays below 0.4 K in non-humid atmospheres. From the analysis of performance all six split-window algorithms could be recommended. The performance of the mono-window algorithms differed considerably amongst the algorithms. The Qin et al. 2001 and the linear method thereby showed the best performance, while the Price 1983 and the Jiménez-Muñoz and Sobrino 2003 algorithm performed less accurate. Due to its best performance, the Qin et al. 2001 algorithm is suggested for long time series processing aiming at climate relevant trends.

Even though the coefficients were calculated for different ranges of columnar water vapour, the accuracy of the algorithms decreases with increasing columnar water vapour. Nevertheless, it is suggested to use humidity dependent coefficients, as the error rises especially in less humid atmospheres in case of coefficients that do not consider humidity dependent ranges (case A). The additional splitting of the coefficients to separate temperature ranges improved the accuracy mainly in very humid conditions. In such cases, the splitting might lower the MAD a few tenth degree. In less-humid conditions, the splitting did not improve the accuracy. The sensor view angle is considered in all cases (A, B, and C), nevertheless, accuracy is still lower in case of higher view angles. As such, view angle, humidity, and first guess LST could serve as proxy for the data quality.

Due to differing spectral response curves, it was assumed that the stability of a long time series of LST from AVHRR data would suffer in case of nonconsideration of the different sensors characteristics. It was shown that the neglecting of the influence of the spectral response curves would produce maximum differences ranging from 0.19 K to 0.71 K in case of the split-window algorithms. Maximum differences ranging from 0.10 K to 1.8 K (36.7 K for Jiménez-Muñoz and Sobrino 2003) in case of the mono-window algorithms would occur. Even if neglecting the case of Jiménez-Muñoz and Sobrino 2003, the numbers suggest using sensor dependent coefficients, as significant additional noise would be introduced to a long time series in case of nonconsideration.

For the processing of long time series data not only the performance of an algorithm is important, but also its sensitivity to its input data. The uncertainty inherent to all input data is an additional

source of noise, lowering the overall accuracy of a final LST product. As the availability of input data is non ideal and even poor in case of the early years when AVHRR was flown, algorithms with lower sensitivity to their input data are to be preferred to algorithms with a high sensitivity. It is not always straightforward to know about the associated error in an input dataset. However, some studies might indicate about the magnitude of the error of a certain variable. Ref. [41] for example compared three land surface broadband emissivity datasets. They found land cover dependent differences and RMSEs between 0.009 and 0.011. If just the error in band emissivity as input to the Qin et al., 2001 or the Becker and Li 1990 would be comparable to the findings of this study, this lead to errors between 0.1 and 0.8 K. Relating this to the GCOS requirement of accuracy better than 1 K [23], it becomes clear that errors in surface emissivity may substantially lower the quality of a product.

This study has shown that the different sensitivities amongst the split-window algorithms are more similar than the sensitivities of the different mono-window algorithms. In the presented case of Figure 15 lowest sensitivity is found for the Becker and Li 1990 and the Price 1984 algorithms with a total mean sensitivity of 1.9 K for daytime and nighttime conditions. Highest total sensitivity is found for the Jiménez-Muñoz and Sobrino 2008 algorithm with 3.1 K for same conditions. Lowest sensitivity of the mono-window algorithms is found for the Price 1983 and the Qin et al., 2001 algorithms with a total mean sensitivity of 1.3 K for daytime and nighttime conditions. Highest total sensitivity is found for the Jiménez-Muñoz and Sobrino 2003 algorithm with a total sensitivity of 4.6 K for same conditions. Against the background of long time series and the situation of imperfect input data, the Becker and Li 1990 and the Price 1984 split-window algorithms and the Price 1983 and the Qin et al., 2001 algorithms are best suited for.

The simulation of the impact of misclassification errors to estimated LST showed that mainly misclassification of bare rock and ground has a significant impact, as well as the class snow and ice in case of band 5. Any misclassification from one to another vegetation class resulted in emissivity errors of less than 0.01, which would result in LST errors of generally less than 0.3 K. To avoid large errors in LST due to misclassification it is suggested to perform an additional check for pixels, which are assigned to the classes bare rock/ground and snow based on classification maps or other procedures.

Regarding the selection of suitable algorithms for time series processing, the synopsis of the performance and the sensitivity analysis following algorithm we suggest: For the mono-window algorithm the Qin et al., 2001 is preferred due to its good performance and low sensitivity results. For the split-window algorithm the Becker and Li 1990 and the Price 1984 algorithm can be suggested showing also good performance and low sensitivity values.

A first comparison with MODIS data using the Becker and Li 1990 approach combined with the Vegetation Cover Method (VCM) from [30] showed despite of some cloud contaminated pixels a good agreement with an overall MAD between AVHRR and MODIS of 1.8 K [K], and an overall standard deviation of 1.4 [K]. This showed a clear improvement over the original AVHRR LST approach assessed in [40], which had an overall MAD of 3.3 K.

The processing has been performed by a preference of coefficients of case C (consideration of ranges for view angle, columnar water vapour and first guess LST). In case where no coefficients were available due to missing coefficients in the database, case B or even case A might have been used. This could happen in extreme situations (e.g., very cold and humid), where the database would not provide enough profiles to calculate coefficients. A selection of pixels, where case C is applied only, might improve further the comparison.

Besides the selection of suitable algorithms, the quantitative results are relevant for the assessment of global warming. The IPCC report from 2014 [42] concludes an average global surface warming of 0.85 K over the period 1880 to 2012, while the period from 1983 to 2012 was likely to be the warmest period of the last 1400 years in the Northern Hemisphere. Warming per decade varied from almost 0 K to more than 0.2 K since 1951. The assessment of performance and sensitivity of AVHRR LST shows that LST derived from AVHRR is feasible under certain conditions only as input to global warming studies. Firstly, the chosen algorithm should be applied using coefficients considering at least various levels of

columnar water vapour and view angles. Performance differences between using coefficient set Case A (only view angle considered) and Case C (view angle, total columnar water vapour, and first guess LST considered) could exceed 1 K in case of the Becker and Li 1990 algorithm—which is a multiple of the before mentioned warming per decade. Second, the algorithms should generally show good performance and low sensitivity, as presented in this study. Maximum differences between algorithms in presented total sensitivity was found to be 3.7 K for the mono-window algorithms and 1.6 K for the split-window algorithms. The general level of presented total sensitivities (except Jiménez-Muñoz and Sobrino 2003) of about 2 K is not representative of general LST retrieval, but still points, third, to the importance of the quality of the input data to the LST algorithms. Forth, only LST values carrying a low uncertainty can be used—this would exclude surfaces under very humid atmospheres or surfaces acquired by a large sensor view angle. The performance of even the best ranking algorithms in this study approaches 1 K (RMSD) in case of very humid atmospheres—in case of large view angles even 2 K.

Besides the considerations on performance and sensitivity, data from the AVHRR instruments are influenced by several other factors which complicate direct use of the data [40]. The orbital drift and the differentiation of the NOAA satellites into morning and afternoon passes for example lead to different acquisition times over the lifespan of a single sensor and the whole AVHRR time series. It follows that AVHRR LST data cannot be directly be used for climate change studies, but should be further processed to match temporal requirements of such analyses. AVHRR LST further forms valuable input as an additional source of information to studies using a multitude of sensors.

7. Conclusions

LST can be retrieved with increasing accuracy and precision, making it a suitable candidate for regional and global assessments of climate variability and change. There are attempts to use LST as a substitute for surface air temperature in areas where in situ measurements are scarce, however, it can also be used directly for change studies. AVHRR has been flown since the early 80ies on a series of platforms. The resulting long time series—35 years—is unique. Although there are some constraints to the direct use of the data (e.g., orbital drift), the resulting time series can be used to generate additional value by extending existing time series of newer sensors with AVHRR back in time or fill gaps of these time series in areas where cloud cover is predominant. As such, AVHRR LST plays an important part in improving climatological databases.

As the number of bands of the AVHRR instrument was not consistent, AVHRR/1 had only one band in the thermal domain, whereas AHVRR/2 and /3 have 2 bands, four different mono-window algorithms and six different split-window algorithms were assessed. For the comparison, new coefficients were generated in the form of a small look up table, accounting for different ranges of columnar water vapour, emissivity, emissivity difference (AVHRR/2 and /3 only) and first guess LST. Such sensor—specific parameters prevent producing artificial anomalies or wrong trends in the data, when processing long time series from AVHRR. For the generation of the coefficients, a radiative transfer model (MODTRAN V5) was used to model top-of-atmosphere brightness temperatures to given LST and the previously mentioned ranges. The least squares method was then used to generate the coefficients and corresponding statistical measures from the database. The statistical measures were used for the performance assessment of the different algorithms. Further, the sensitivity of the single algorithms to their input data was assessed, by simulating ranges of input values. The synopsis of the performance, as expressed by accuracy and precision measures, and sensitivity analysis revealed that the Qin et al., 2001 algorithm is to be preferred amongst the mono-window algorithms due to its good performance and low sensitivity results. Amongst the split-window algorithms the Becker and Li 1990 and the Price 1984 algorithm can be suggested, as both show good performance and low sensitivity values.

A comparison of the application of the Becker and Li 1990 coefficients and the Vegetation Cover Method (VCM) [30] to AVHRR with the MODIS LST product revealed a good agreement and confirmed the approach to be valid and physically sound.

Author Contributions: C.F. performed the design of the study, analyzed the data and wrote the paper. C.K. contributed to the design of the study and contributed to the manuscript. S.D. contributed to the design of the study and commented on the manuscript.

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

References

1. Li, Z.L.; Tang, B.H.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I.F.; Sobrino, J.A. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.* **2013**, *131*, 14–37. [CrossRef]

2. Mao, K.; Shi, J.; Tang, H.; Li, Z.-L.; Wang, X.; Chen, K.-S. A neural network technique for separating land surface emissivity and temperature from ASTER imagery. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*. [CrossRef]

3. Yu, Y.; Privette, J.L.; Pinheiro, A.C. Evaluation of Split-Window Land Surface Temperature Algorithms for Generating Climate Data Records. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 179–192. [CrossRef]

4. LSA SAF. Available online: http://landsaf.meteo.pt/algorithms.jsp (accessed on 7 October 2016).

5. Copernicus Land. Available online: http://land.copernicus.eu/global/products/lst (accessed on 1 June 2016).

6. NASA MODIS. Available online: https://modis.gsfc.nasa.gov/data/dataprod/dataproducts.php?MOD_NUMBER=11 (accessed on 1 June 2016).

7. Eyre, J.R. Estimation of sea-surface temperature from AVHRR data; comments on the paper by Singh et al. (1985). *Int. J. Remote Sens.* **1986**, *7*, 465–469. [CrossRef]

8. Duguay-Tetzlaff, A.; Bento, V.A.; Göttsche, F.M.; Stöckli, R.; Martins, J.P.A.; Trigo, I.; Olesen, F.; Bojanowski, J.S.; da Camara, C.; Kunz, H. Meteosat land surface temperature climate data record: Achievable accuracy and potential uncertainties. *Remote Sens.* **2015**, *7*, 13139–13156. [CrossRef]

9. Price, J.C. Estimating surface temperatures from satellite thermal infrared data—A simple formulation for the atmospheric effect. *Remote Sens. Environ.* **1983**, *13*, 353–361. [CrossRef]

10. Prata, A.J. Land surface temperature determination from satellites. *Adv. Space Res.* **1994**, *14*, 15–26. [CrossRef]

11. Qin, Z.; Karnieli, A.; Berliner, P.A. Mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *Int. J. Remote Sens.* **2001**, *22*, 3719–3746. [CrossRef]

12. Jiménez-Muñoz, J.C.; Sobrino, J.A. A generalized single-channel method for retrieving land surface temperature from remote sensing data. *J. Geophys. Res. Atmos.* **2003**, *108*, 4688. [CrossRef]

13. Anding, D.; Kauth, R. Estimation of sea surface temperature from space. *Remote Sens. Environ.* **1970**, *1*, 217–220. [CrossRef]

14. Becker, F.; Li, Z.L. Towards a local split window method over land surfaces. *Int. J. Remote Sens.* **1990**, *11*, 369–393. [CrossRef]

15. Becker, F.; Li, Z.L. Surface temperature and emissivity at various scales: Definition, measurement and related problems. *Remote Sens. Rev.* **1995**, *12*, 225–253. [CrossRef]

16. Wan, Z.; Dozier, J.A. Generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905.

17. Ulivieri, C.; Castronuovo, M.M.; Frrancioni, R.; Cardillo, A. A split window algorithm for estimating land surface temperature from satellites. *Adv. Space Res.* **1994**, *14*, 59–65. [CrossRef]

18. François, C.; Ottlé, C. Atmospheric corrections in the thermal infrared: Global and water vapor dependent split-window algorithms-applications to ATSR and AVHRR data. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 457–470. [CrossRef]

19. Tang, B.-H.; Bi, Y.; Li, Z.-L.; Xia, J. Generalized Split-Window algorithm for estimate of land surface temperature from Chinese geostationary FengYun meteorological satellite (FY-2C) data. *Sensors* **2008**, *8*, 933–951. [CrossRef] [PubMed]

20. Ouaidrari, H.; Goward, S.N.; Czajkowski, K.P.; Sobrino, J.A.; Vermote, E. Land surface temperature estimation from AVHRR thermal infrared measurements: An assessment for the AVHRR Land Pathfinder II data set. *Remote Sens. Environ.* **2002**, *81*, 114–128. [CrossRef]

21. Pozo Vázquez, D.; Olmo Reyes, F.J.; Alados Arboledas, L. A comparative study of algorithms from estimating land surface temperature from AVHRR data. *Remote Sens. Environ.* **1997**, *62*, 215–222. [CrossRef]

22. Kerr, Y.H. Land Surface Temperature Retrieval Techniques and Applications (Case of the AVHRR). 2000. Available online: http://www.cetp.ipsl.fr/1activ/21mineu/iot/publi_pdf/IRT-Book-YHK.pdf (accessed on 1 June 2016).

23. Global Climate Observing System. *Systematic Observation Requirements for Satellite-based Data Products for Climate: 2011 Update GCOS-154*; World Meteorological Organization: Geneva, Switzerland, 2011.

24. CM SAF. Available online: http://www.cmsaf.eu/EN/Home/home_node.html (accessed on 28 November 2016).

25. GlobTemperature. Available online: http://www.globtemperature.info/ (accessed on 1 June 2016).

26. MOD11A1 NASA LP DAAC. *Land Surface Temperature and Emissivity Daily L3 Global 1 km Grid SIN. V6*; USGS: Sioux Falls, SD, USA, 2015. Available online: http://e4ftl01.cr.usgs.gov/MOLT/MOD11A1.00 (accessed on 15 November 2015).

27. Bachmann, M.; Tungalagsaikhan, P.; Ruppert, T.; Dech, S. Calibration and Pre-processing of a Multi-decadal AVHRR Time Series. In *Remote Sensing Time Series—Revealing Land Surface Dynamics*; Künzer, C., Dech, S., Wagner, W., Eds.; Springer: Cham, Switzerland, 2015; pp. 43–74.

28. Borbas, E.E.; Seemann, S.W.; Hung-Lung, H.; Li, J.; Menzel, W.P. Global profile training database for satellite regression retrievals with estimates of skin temperature and emissivity. In Proceedings of the XIV International ATOVS Study Conference, Beijing, China, 25–31 May 2005; pp. 763–770.

29. Hansen, M.C.; Defries, R.S.; Townshend, J.R.G.; Sohlberg, R. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* **2000**, *21*, 1331–1364. [CrossRef]

30. Caselles, E.; Valor, E.; Abad, F.; Caselles, V. Automatic classification-based generation of thermal infrared land surface emissivity maps using AATSR data over Europe. *Remote Sens. Environ.* **2012**, *124*, 321–333. [CrossRef]

31. Kriebel, K.T.; Gesell, G.; Kästner, M.; Mannstein, H. The cloud analysis tool APOLLO: Improvements and validations. *Int. J. Remote Sens.* **2003**, *24*, 2389–2408. [CrossRef]

32. Gesell, G. An algorithm for snow and ice detection using AVHRR data An extension to the APOLLO software package. *Int. J. Remote Sens.* **1989**, *10*, 897–905. [CrossRef]

33. Jiménez-Muñoz, J.C.; Sobrino, J.A. Split-window coefficients for land surface temperature retrieval from low-resolution thermal infrared sensors. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 806–809. [CrossRef]

34. NOAA Polar Orbiter Data (POD) User's Guide. Available online: http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/podug/index.htm (accessed on 1 June 2016).

35. NOAA KLM User's Guide. Available online: http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/klm/index.htm (accessed on 7 October 2016).

36. Cho, A.R.; Choi, Y.Y.; Suh, M.S. Improvement of a COMS land surface temperature retrieval algorithm based on the temperature lapse rate and water vapor/aerosol effect. *Remote Sens.* **2015**, *7*, 1777–1797. [CrossRef]

37. García-Santos, V.; Valor, E.; Caselles, V.; Ángeles Burgos, M.; Coll, C. On the angular variation of thermal infrared emissivity of inorganic soils. *J. Geophys. Res. Atmos.* **2012**, *117*, 18. [CrossRef]

38. Hulley, G.C.; Hughes, C.G.; Hook, S.J. Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data. *J. Geophys. Res. Atmos. Clim. Dyn.* **2012**, *117*, 18. [CrossRef]

39. Kuenzer, C.; Leinenkugel, P.; Vollmuth, M.; Dech, S. Comparing global land-cover products—Implications for geoscience applications: An investigation for the trans-boundary Mekong Basin. *Int. J. Remote Sens.* **2014**, *35*, 2752–2779. [CrossRef]

40. Frey, C.M.; Kuenzer, C.; Dech, S. Quantitative comparison of the operational NOAA-AVHRR LST product of DLR and the MODIS LST product V005. *Int. J. Remote Sens.* **2012**, *33*, 7165–7183. [CrossRef]

41. Chen, J.; Liang, S.; Yao, Y.; Ren, B.; Linpeng, S.; Liu, H. A Comparative Study of Three Land Surface Broadband Emissivity Datasets from Satellite Data. *Remote Sens.* **2014**, *6*, 111–134. [CrossRef]

42. IPCC. *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Core Writing Team, Pachauri, R.K., Meyer, L.A., Eds.; Climate Change 2014: Synthesis Report; IPCC: Geneva, Switzerland, 2014; p. 151.