



NASA Contract NNH08CC96C  
AER P1342

# **Atmospheric and land surface characterization using a land surface emissivity database derived from AMSR-E and MODIS**

Final Report October 2011

Submitted to  
**National Aeronautics and Space Administration  
Goddard Space Flight Center  
Greenbelt, MD 20771**

Submitted by  
**Atmospheric and Environmental Research  
131 Hartwell Avenue  
Lexington, MA 02421**

**October 8, 2011**

**Jean-Luc Moncet**  
Principal Investigator

**Pan Liang, John Galantowicz, Alan Lipton and Gennady Uymin**

## Table of Contents

1.	SCOPE AND PROJECT OVERVIEW .....	3
2.	INTRODUCTION .....	3
3.	VERSION 1 EMISSIVITY DATABASE .....	6
4.	DATABASE IMPROVEMENTS .....	15
4.1.	Snow mask update .....	16
4.2.	RFI and instability detection .....	16
4.2.1	<i>Clear-sky emissivity-based RFI detection</i> .....	17
4.2.2	<i>Tb-based cloudy RFI detection</i> .....	23
4.2.3	<i>Comparison with DGLSP product</i> .....	24
4.3.	Dealing with spatial heterogeneities due to presence of water bodies in the AMSR-E field-of-view 26	
4.4.	Training of an AMSR-E footprint matching algorithm .....	30
5.	NEW TIME SERIES SEGMENTATION APPROACH .....	38
6.	IMPROVEMENTS IN QUALITY OF AMSR-E EMISSIVITY ESTIMATES .....	58
6.1.	Addition of Terra/MODIS for emissivity estimates in penetrating areas .....	58
6.2.	AIRS vs. NWP-based atmospheric correction .....	64
7.	APPLICATIONS .....	70
7.1.	Estimation of monthly mean diurnal cycle amplitude of land surface temperature over vegetated surfaces .....	70
7.1.1	<i>MODIS MYD11 LST product</i> .....	70
7.1.2	<i>Comparison with of microwave and infrared LST with in situ data</i> .....	73
7.2	On the potential impact of dew on microwave surface emissivity .....	97
7.3	Retrieval of LST over deserts .....	106
7.4	Global short-term surface change monitoring from R11 analysis – comparison with IR emissivities .....	116
8	CONCLUSIONS .....	122
9	REFERENCES .....	125
	SYSTEMATIC ANOMALIES IN AMSR-E DERIVED SPECTRAL EMISSIVITIES .....	129
	AUTMOMATIC R11 TIME SERIES SEGMENTATION (RTS).....	137

## 1. SCOPE AND PROJECT OVERVIEW

This Final Report describes work performed under AER Project 1342 (P1342) on “Atmospheric and land surface characterization using a land surface emissivity database derived from AMSR-E and MODIS” funded by NASA (contract NNH08CC96C). Time period: April 2008 – October 2011.

Microwave measurements from space have great potential to contribute to the detection and understanding of climate trends, due largely to the stability of their calibration and their unique ability to detect surface and atmosphere properties through clouds. Microwave products thus minimize the “cloud bias” that occurs with infrared-based analysis, in which the datasets systematically exclude cloudy areas. These biases can be particularly damaging when analyzing trends in parameters like land surface temperature and water vapor, which are strongly correlated with cloudiness. This project is the continuation of work under our former NASA grant (contract # NNH04CC43C) of developing and utilizing a microwave emissivity database with data from AMSR-E and other EOS instruments. The tasks accomplished under the current grant include (1) completing the implementation of our Version 1 AMSR-E emissivity retrieval algorithm and producing one full-year of global monthly mean emissivity data complemented by spatially and temporally matched infrared emissivities from the MODIS V.4 MYB11 daily LST product (2) assessing the potential for meeting internally set requirements for application of the database to characterization of climate sensitive parameters such as land surface emissivity and temperature, and atmospheric water vapor (3) refining emissivity retrieval procedure (to improve yield and quality of the emissivity estimates) and assessing potential impact of upgrades in external data sources on the quality of the emissivity product (4) application of the data base to estimation of land surface temperature over vegetated surfaces and deserts and validation of emissivity product. The primary goal of this study is to make the preparatory work that will lead to a second improved version of the database that is particularly well suited for retrieval/assimilation applications and routine production of accurate climate related products, and identify particular areas where future development might be needed.

## 2. INTRODUCTION

Our baseline approach to microwave surface emissivity retrieval over land has its heritage in the work of *Prigent et al.* (1997, 1999 and 2006). Estimates of microwave emissivity are derived by combining satellite microwave observations in the clear-sky with co-located infrared land surface temperature (LST) and atmospheric temperature and water vapor profiles from external sources, e.g. Numerical Weather Prediction (NWP) models or high spectral resolution sounders retrieval product. Instantaneous emissivity estimates are obtained by solving the radiative transfer equation independently for each available clear AMSR-E field-of-view and by averaging adjacent observations to the nearest grid point on a fixed Earth grid.

While atlases of microwave emissivity data are already available (*Prigent et al.*, 2006; *Karbou et al.*, 2005) the emphasis of our work is on improving the accuracy of the derived emissivities. For surface state monitoring high sensitivity is required in order to detect small changes in the surface state occurring on spatial scales of the order of or smaller than the size of the AMSR-E FOV. With respect to the use of microwave surface emissivities as a priori constraints in assimilation or in the retrieval of surface temperature and atmospheric water vapor from microwave measurements, sensitivity studies (*Moncet et al.* 2011a) have shown that uncertainties less than 0.01 (ideally of the order of 0.005) in

surface emissivity are required in order to produce a significant impact on the retrieval product. For instance, Land Surface Temperature (LST) and in particular the amplitude of the temperature diurnal cycle, is an important climate variable. In many regions of the globe (e.g. tropics, polar regions), this amplitude may be of the order of 10K or less and in order to derive useful information from the satellite observations it is desirable to keep uncertainties small compared to this number. When surface emissivity is the same for both day and night, biases in their estimates cancel out, but there are some factors (e.g. surface wetness, vegetation moisture content) that could potentially introduce systematic day/night differences. An emissivity uncertainty of 0.005 leads to an uncertainty of 1.5K in LST estimates. We believe that this number is far from being reached in a monthly averaged sense with LST products currently available from infrared or microwave sources (e.g. Moncet *et al.*, 2011b). Hence our strategy is to favor quality first, rather than yield and spatial coverage, assess the global and regional yield obtained with the highest quality data and then seek solutions for filling in the gaps, and bring all relevant information available to us from external sources to understand and flag exceptions. It is believed that this approach is necessary in order to make significant advances in the use of remotely sensed microwave data in climate applications in the long term.

Our work departs from the aforementioned work of Prigent *et al.* in the several respects:

- 1) We benefit from the good temporal co-location of the MODIS, AIRS and AMSR-E instrument on board the same Aqua satellite, a unique configuration among current satellite constellations, as well as the demonstrated good consistency between the MODIS LST product and microwave observations (e.g. Moncet *et al.* 2011b).
- 2) We focused much of our effort on developing strict quality control procedures for the combined microwave and infrared observations to filter out occurrences of daily outliers due to e.g. precipitation and sudden change in soil moisture, or occasional bad microwave or infrared data (due to e.g. to contamination by undetected clouds) which could affect the quality of the emissivity data. This strict quality control (QC) comes at the expense of a loss in yield but methods are sought to help compensating for this loss (e.g. class-based emissivity substitution – Moncet *et al.*, 2011a) and increasing the number of high quality measurements in frequently cloudy areas (see Sec.5).
- 3) We benefit from the availability of low frequency channels which are more immune to the impact of atmosphere and liquid clouds than the SSM/I frequencies used by Prigent *et al.* In fact, we rely extensively on the 11 GHz channels for detection of anomalies in the infrared (MODIS) surface temperature data and for the retrieval of LST over vegetated (non-penetrating) surfaces. In addition, we rely on the 11GHz polarization ratio,  $R11 = Tb_{11H} / Tb_{11V} \approx \varepsilon_{11H} / \varepsilon_{11V}$ , which is by construction quite insensitive to variations (or errors) in surface temperature, to monitor changes in the surface state in arid and semi-arid areas
- 4) We have shown that the impact of penetration depth over semi-arid and arid areas is not limited to sandy soils and is more widespread than originally thought at midday, i.e. the time of the Aqua overpass (Moncet *et al.* 2011a; Galantowicz *et al.* 2011). A different treatment is required over these areas to characterize surface emission (see Sec. 3 and 7.3).
- 5) We have highlighted the fact that surface emissivity is quite stable in time over many (snow free) areas. This characteristic may be exploited for comparing the quality of different external LST products (Moncet *et al.*, 2011b), by assessing their impact on the temporal standard deviation of the estimated emissivities when used as an input to our retrieval scheme. Although land emissivities tend to be higher than over ocean which tends to reduce the sensitivity of the



microwave observations to errors in input atmospheric parameters, extending the same approach to land water vapor products may prove to be valuable (e.g. Sec. 6.2).

Version 1 of our emissivity retrieval algorithm and product is summarized in Section 3. The work described in the subsequent sections focuses on refinements we intend to make to our approach in preparation for a *second improved version* of our emissivity database and on application of this database to *LST retrieval and monitoring of surface changes*. While all the refinements described in this report have been implemented and individually tested, none of them (unless mentioned explicitly) has been fully integrated into Version 2 of the emissivity algorithm at this point. One major change from Version 1, actually used in our application to LST retrievals, is the departure from our arbitrary monthly breakdown of the analysis period. Here we are examining the possibility of adopting a separate time division for each individual grid points coinciding with detected changes in the surface properties at that particular location. In areas where emissivity remains relatively stable over time periods exceeding one month, this new approach should increase the likelihood of finding clear observations, which should in turn lead to both higher yield and higher quality of our emissivity estimates. Another important area of focus is the application of the data base to LST retrieval. In particular, over vegetated surfaces where we rely on the 11 GHz measurements (mainly insensitive to errors in the specified water vapor and to residual cloud liquid water) for LST retrieval, comparisons of our LST product with in situ observations serves as an indirect means of validating several critical aspects tied to our emissivity retrieval process, i.e. the quality of the input MODIS LST estimates and the adequacy of our QC procedure.

In addition to the Version 1 description (Sec. 3), the present report addresses the following topics:

- Quality control and flagging (Sec.4)
  - Snow mask update
  - RFI flag update
  - Strategy for dealing with spatial heterogeneities due to presence of water bodies in the AMSR-E field-of-view
- Improvement in yield and spatial coverage (Sec. 5)
  - Development of a new time series segmentation algorithm exploiting long term stability of surface emissivity in some regions to increase analysis time period and number of clear-sky observations, and improve algorithm yield
- Improvement in quality of AMSR-E emissivity estimates (Sec. 6)
  - Addition of MODIS LST data from Terra for improved characterization of surface thermal forcing over highly penetrating regions – impact on quality of microwave surface emissivity estimates in arid and semi-arid areas.
  - Assessment of impact of replacing atmospheric correction derived from Numerical Weather Prediction (NWP) model fields (see Sec. 3) with standard daily Level 3 AIRS (AIRX3STD) atmospheric product on the quality of the retrieved AMSR-E emissivities
- Applications of emissivity database (Sec.7)
  - Retrieval of monthly mean land surface temperature from AMSR-E over vegetated surfaces for climate application - comparison with MODIS V4 and V5 LST product and indirect validation of the quality of the emissivity atlas
  - On the potential impact of dew on microwave surface emissivity
  - On the issue of soil temperature estimation from microwave (or combined IR and MW) observations over highly penetrating desert areas and impact of dust storms
  - Global monitoring of surface changes

As part of this study we also considered upgrading our AMSR-E Level 2A brightness temperatures. A bias analysis performed on our retrieved emissivity product highlights potential calibration issues with the latest Version distributed by NSIDC. Our findings are described in Appendix A. We have been working with Remote Sensing System (RSS) to evaluate their new version (Version 7) of AMSR-E Tb's (not yet publically available – to our knowledge) with the intent to correct these issues. Unfortunately the timing for the availability of the V7 Level 2A AMSR-E product is such that we were only able to perform a preliminary assessment of the new RSS product.

### 3. VERSION 1 EMISSIVITY DATABASE

The AMSR-E instrument is a conically scanning microwave imager with a fixed 55° Earth incidence angle (EIA) which is part of the NASA/Aqua payload (ascending and descending nominal equator crossing times (ECT) 13:30 and 01:30 local). This instrument carries V and H-polarized channels centered at 6.925, 10.65, 18.7, 23.8, 36.5, and 89 GHz (Kawanishi *et al.*, 2003). The radiometric data used to develop the current version (Version 1) of our emissivity data base is Version 09 of the AMSR-E/Aqua L2A Global Swath Spatially-Resampled Brightness Temperatures (AE\_L2A) product (Ashcroft and Wentz, 2003) distributed by the National Snow and Ice Data Center (NSIDC). In this data set, all channels have been spatially averaged to match a common spatial response pattern (Ashcroft and Wentz, 2000). For this work, we used the Resolution 2 data matched to a 51 km × 29 km footprint size (average spatial resolution = 38 km) and sampled at 10 km intervals along scan. All channels except the 6.9 GHz are available at this resolution. The spatial averaging reduces random instrument noise to < 0.15 K (*rms*) at the center of scan for all frequencies (or ~0.0005 in terms of equivalent retrieved emissivity standard deviation) before Earth gridding.

The V1 approach used to estimate the surface emissivity in the AMSR-E channels is described in details in Moncet *et al.* (2011a). Only a summary is given here.

Over moderately to highly vegetated (non-penetrating) surfaces, surface is assumed to emit at the same temperature at all frequencies (and polarizations). An estimate of the surface temperature is obtained in the clear-sky by averaging time collocated infrared MODIS LST data over the AMSR-E field-of-view (FOV). Given an atmospheric temperature and water vapor profile obtained from external sources (see below) and interpolated to the footprint location, one solves the clear-sky radiative transfer equation to derive an instantaneous surface emissivity estimate for each AMSR-E FOV. This emissivity data is mapped to a fixed 28×28 km sinusoidal Earth grid, by simply averaging together the FOV's from a same overpass falling within a 10 km radius of the center of a grid box, and temporally averaged over each calendar month to produce monthly mean day and night emissivity values for each gridded location. This method, which is referred to interchangeably in this report as 1a or non-penetration approximation (NPA), is first applied to the clearest FOV's ( $F_{clear} \geq 98\%$ , where  $F_{clear}$  represents the “clear fraction”, obtained from the antenna-pattern-weighted MODIS LST quality flag values averaged over the AMSR-E FOV – Moncet *et al.*, 2011a). It is only when the number of clearest FOV's collected during the month is insufficient for producing a reliable mean emissivity estimate that we introduce cloudy ( $F_{clear} \geq 50\%$  and  $F_{clear} \geq 20\%$ ) FOVs. A time series analysis of the R11 11 GHz polarization ratio is used to flag and exclude from our analysis outliers associated with isolated events such as changes in soil wetness over arid and semiarid areas. R11 is also used as an indicator of the temporal stability of the surface (see below). A clustering approach (Moncet *et al.*, 2011a) is finally applied to remove remaining emissivity outliers due in particular to cloud contaminated AMSR-E observations or LST estimates. The quality of our final 1a surface emissivity estimates is measured in

terms of their temporal standard deviation (ESD) evaluated over the month. Over stable regions (see below) we find that ESD is typically less than 0.01, reaching 0.005 or less over ~60% of the grid points (e.g. Figure 5), an indication of the high degree of consistency between the FOV-averaged MODIS LST's and the AMSR-E observations (e.g. Moncet *et al.*, 2011b).

The input LST data and associated quality flags used in our AMSR-E processing are from Version V004 of the 5-km Level-3 Earth gridded product from the MODIS day/night algorithm (Wan, 1999, 2008). The MODIS LST algorithm makes use of Version V004 of the MODIS cloud mask (Ackerman *et al.*, 1998, 2006). Our pre-processing uses a threshold for the local temporal LST standard deviation to filter out LST data potentially affected by residual clouds. The filtering algorithm was acquired from Wan (2008) and modified for the V1 emissivity data production to use the same threshold regardless of surface type. A 5K threshold was subjectively chosen by inspecting images of filtered LSTs and seeking to eliminate any obvious low LST values at the edges of cloud masses. The algorithm has since then been upgraded to use a surface type dependent threshold. Note that this filter tends to be less effective at locations where cloud cover is persistent and only few LST reports pass the initial quality test. In Version 1 of our emissivity product atmospheric temperature and water vapor profiles are derived from the 1° National Centers for Environmental Prediction (NCEP) Global Data Assimilation System (GDAS) analysis (Kanamitsu, 1989; Kalnay *et al.*, 1990). The 6-hourly NWP product is interpolated to the local time of the Aqua overpass and to the center of the AMSR-E footprints. The surface pressure is obtained from the GDAS analysis and adjusted using the GTOPO30 30-arc-sec topography database (available from US Geological Survey: <http://edc.usgs.gov/products/elevation/gtopo30/README.html>), spatially degraded to the AMSR-E footprint resolution. The atmospheric absorption model used in this study is from Rosenkranz (1998) (see also Liebe *et al.*, 1992).

Over arid or semi-arid surfaces, microwave sub-surface penetration causes the emission temperature at the lower frequencies to be typically colder than the surface skin temperature during the day and warmer at night. When using the NPA method to process the data at those locations, this effect results in the retrieved daytime emissivities being smaller than the nighttime estimates. The difference is the largest at the lowest (more deeply penetrating) frequencies. When a significant day/night emissivity difference is detected, the 1b or penetration time series (PTS) algorithm is applied to estimate a monthly mean emissivities valid for both day and night (Moncet *et al.*, 2011a). In the PTS algorithm, the response to temporal variations in the surface thermal forcing (described by time series of infrared observations) at some depth beneath the surface is modeled using the 1-dimensional heat flow equation (Galantowicz *et al.* 2011). Daily atmospheric corrected microwave observations covering the period analyzed (here, 1 month) are used to solve for emission temperature, depth parameter  $\alpha$  (both parameters are assumed to be independent of polarization), and emissivity, at all frequencies. More than two daily overpasses are needed to be able to solve for these parameters. In the present algorithm we also make use of DMSP F13 and F15 (nominal ascending node ECT: 18:33 and 21:05, respectively) SSM/I observations (Hollinger *et al.*, 1990; Colton and Poe, 1999) obtained from the Global Hydrology Resource Center (<http://ghrc.msfc.nasa.gov>) to complement the AMSR-E measurements. To support our analyses of SSM/I data, we use cloud data from the International Satellite Cloud Climatology Project (ISCCP), produced from multiple satellites on a 30-km grid, and available at 3-hour intervals (Rossow and Shiffer, 1999; Rossow and Garder, 1993a, b). In Version 1, the only source of surface skin temperature data used in the NPS algorithm is from the Aqua/MODIS sensor. Early attempts to use the ISCCP LST product and scale it using MODIS data produced unsatisfactory results (anomalies detected in the current ISCCP LST product over deserts are summarized in Moncet *et al.* 2011b). The 89 GHz observations have been used instead to complement the infrared observations in the determination of the parameters (amplitude and phase) that characterize

the thermal forcing. This approach is not ideal because over certain surfaces (e.g. sand) penetration may also affect the 89 GHz observations. In addition the 89 GHz channels are also highly sensitive to errors in the specification of atmospheric water vapor. As part of the current work we attempted to remedy this situation by adding LST (and cloud) information from Terra/MODIS (Sec. 6.1). In the 1b algorithm, an estimate of the surface emissivity is accepted or rejected on the basis of the goodness of fit on the combined AMSR-E and SSM/I Tb observation time series.

Many areas in the tropics (especially over rainforest or monsoon regions) are frequently cloudy. In those areas, it is common for our 1a emissivities to fail our quality control tests. In addition, many of the estimates that pass our quality tests are produced from  $20\% \leq F_{clear} < 50\%$  observations which we regard as uncertain. For these cases we attempt to substitute emissivities estimated under more favorable cloud conditions over locations with same surface type (Moncet *et al.* 2011a). This algorithm is subsequently referred to as “class-based substitution” or C-algorithm.

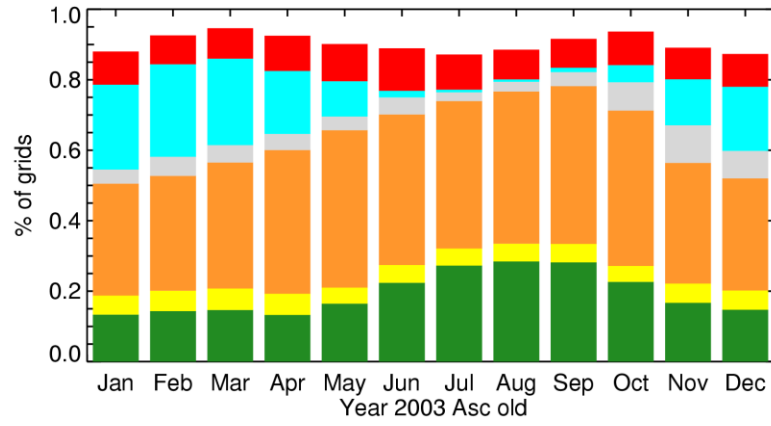
Figure 1 provides a monthly summary of the partitioning of our retrieved emissivities among the different algorithm sources and quality categories. The color code used for the database quality assessment (QA) flags throughout this report is provided in Table 1. Examples of geographical distribution of product categories for the months of January and July 2003 are given in Figs. Figure 2 and Figure 3. MODIS surface type information is provided for reference in Figure 4.

Note that 1a, 1b or C emissivities are produced for all locations except where no LST estimate is available from the MODIS day/night algorithm or, in all other areas, when the NPA algorithm fails to find a cluster for both day and night and no C-emissivity substitute can be provided (for instance, in areas where surface conditions are changing through the month). The MODIS day/night algorithm relies on day/night temperature contrast being above a certain threshold to separate surface emissivity from thermal surface emission (Wan, 2008). When these conditions are not met (i.e. Polar Regions in summer and winter seasons), an LST estimate from the split window algorithm (Wan, 1999) is provided instead but is not being used for the moment for our emissivity retrievals. Less than ideal conditions for the production of reliable surface emissivity estimates include inhomogeneous areas (e.g. coastal regions) and regions where surface emissivity varies significantly on short time scales (e.g. Sahel during African Monsoon). A mean emissivity can still be provided in those regions but emissivity on any particular day may depart significantly from its mean value hindering the use of this data to constrain inversion in such applications as retrieval or data assimilation and potentially triggering false alarms in the analysis of year-to-year surface variability. We rely primarily on the R11 temporal standard deviation,  $\sigma_{R11}$ , to flag those conditions. Throughout this report, we define as unstable any location for which  $\sigma_{R11} > 0.015$  (Table 1). Note that the “unstable” QA flag in the distributed Version 1 database is based on a somewhat stricter criterion and only applies to areas where a cluster has been found in the first place (we refer the reader to the database documentation for more details – see below). Additional flags are provided to qualify the sources of instabilities. Inhomogeneous areas are identified using the mean local spatial emissivity standard deviation LSSD (Figure 5), also referred to as SpSD. An independent database is used to flag areas where Radio Frequency Interference (RFI) is likely (Sec. 0). Note that changes in snow cover or physical characteristics of snow are not always observable in the R11 time series – the largest R11 signal often coincide with the melting phase and is probably associated for the most part with soil wetness. Snow may be highly penetrating and changes in the grain/layer structure of the snow pack may cause emissivity and penetration depth to vary on small time scales. Determining under what conditions a useful snow emissivity value can be produced is one aspect that has not been addressed in details yet. For these reasons snow conditions are flagged separately (Sec. 4.1). Also, for any given month, there

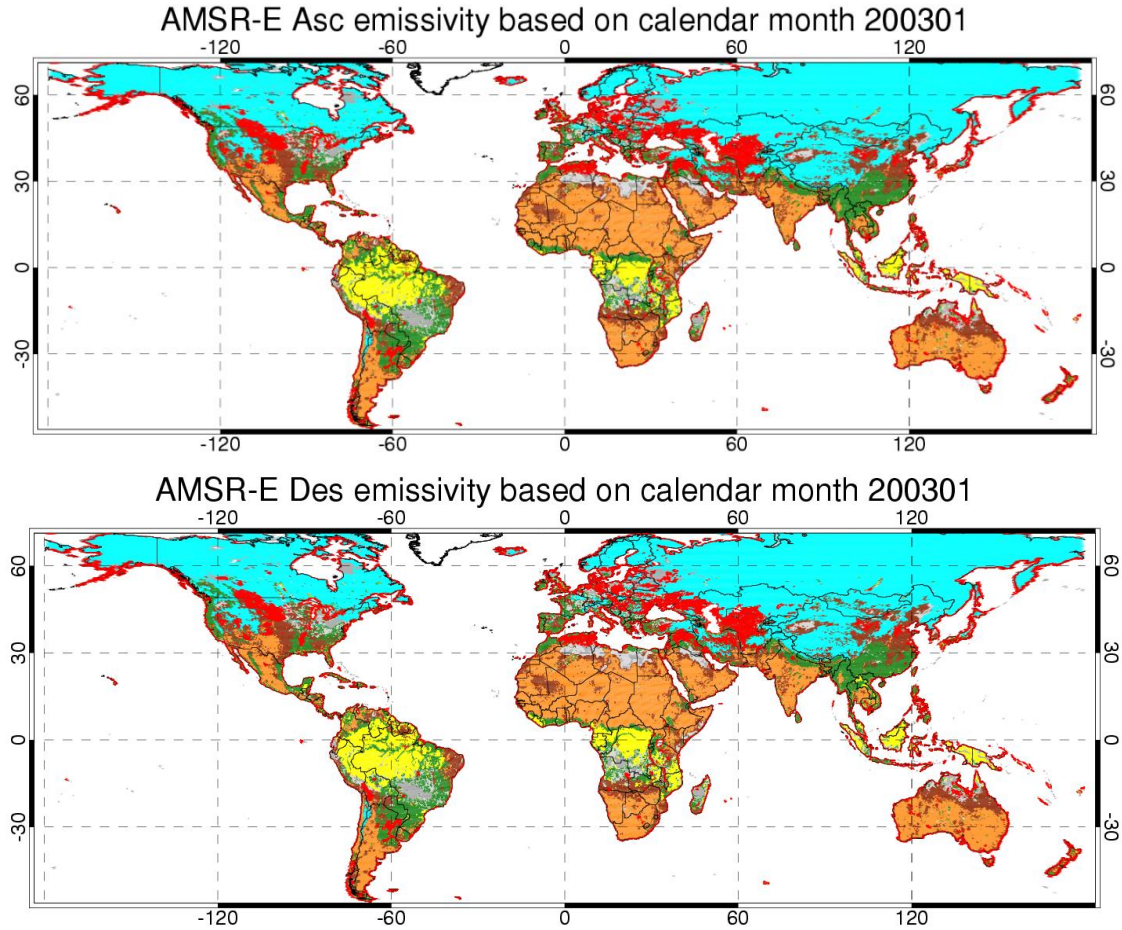
are bound to be some locations among the stable grid locations, for which the NPA algorithm cannot produce an emissivity estimate either for day or night only or for both day and night. The former are classified as “uncertain” and the latter as “missing”. These occur when cloudiness conditions are such that there are not enough observations to produce an emissivity estimate or when the estimates available do not form a tight cluster.

Color code	Explanation
1	Snow
2	Unstable or frequent 11GHz RFI
3	Data missing (no 1a nor C value could be found, or outside of area covered by the MODIS day/night LST algorithm)
4	Value found but uncertain (only a day or night time 1a estimate could be obtained)
5	High quality 1b emissivity values
6	1b case but solution does not fit input satellite data within nominal accuracy tolerance (in this case, we revert to day/night averaged 1a estimate in the merged emissivity product)
7	High quality 1a surface emissivity estimate
8	1a value found but emissivity temporal standard deviation > 0.015
9	C (class-based substitution)

**Table 1:** Emissivity database Quality Assessment flags: color code

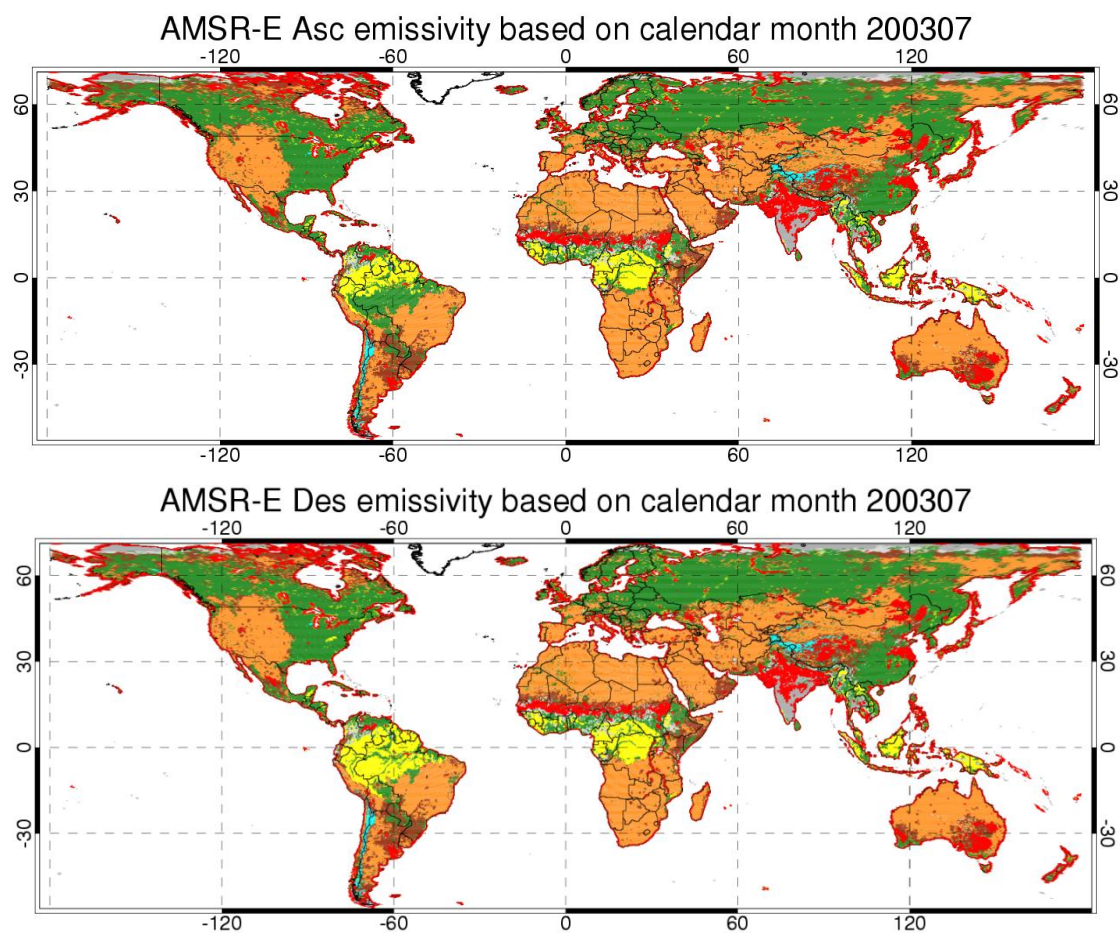


**Figure 1:** Fraction of grid points in Version 1 database for all the months of 2003 in the following categories (see Table 1): green: NPA (1a), yellow: C, orange: PTS (1b), red: unstable, and gray: uncertain (only one valid emissivity estimate for either day or night but not both). Fraction of snow covered areas is in blue. Here, Categories 7 and 8 in Table 1 are merged into a single category. Same for Categories 5 and 6. In this figure, missing values (Category 3) make up the difference between the total fraction of land points shown in color and 100%. Total number of grid points is 199,525.

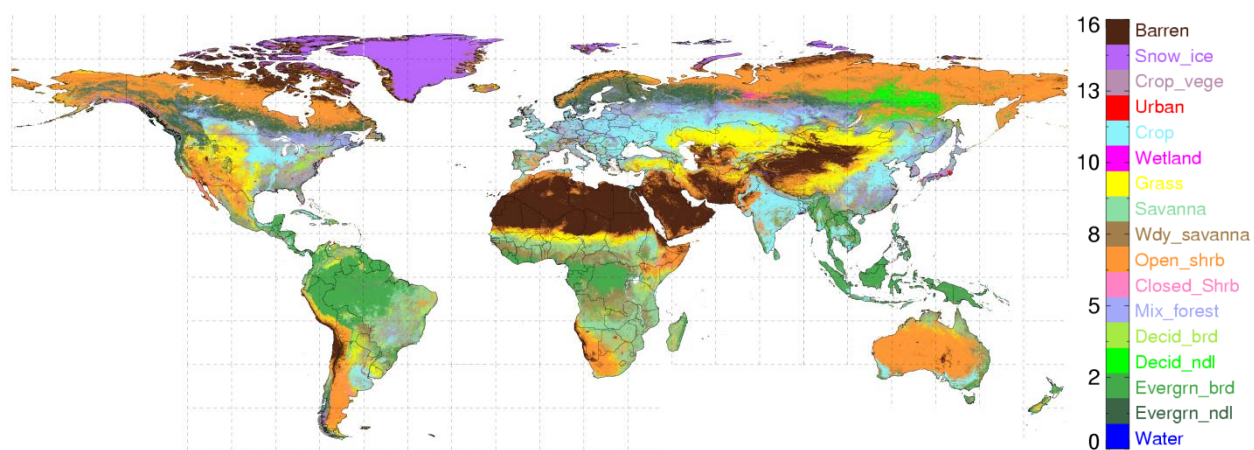


**Figure 2:** Global distribution of QA categories (see Table 1) for January 2003. In this figure, the snow mask overrides missing data in the Polar Regions.





**Figure 3:** Same as Figure 2 for July 2003



**Figure 4:** 2003 MODIS surface types

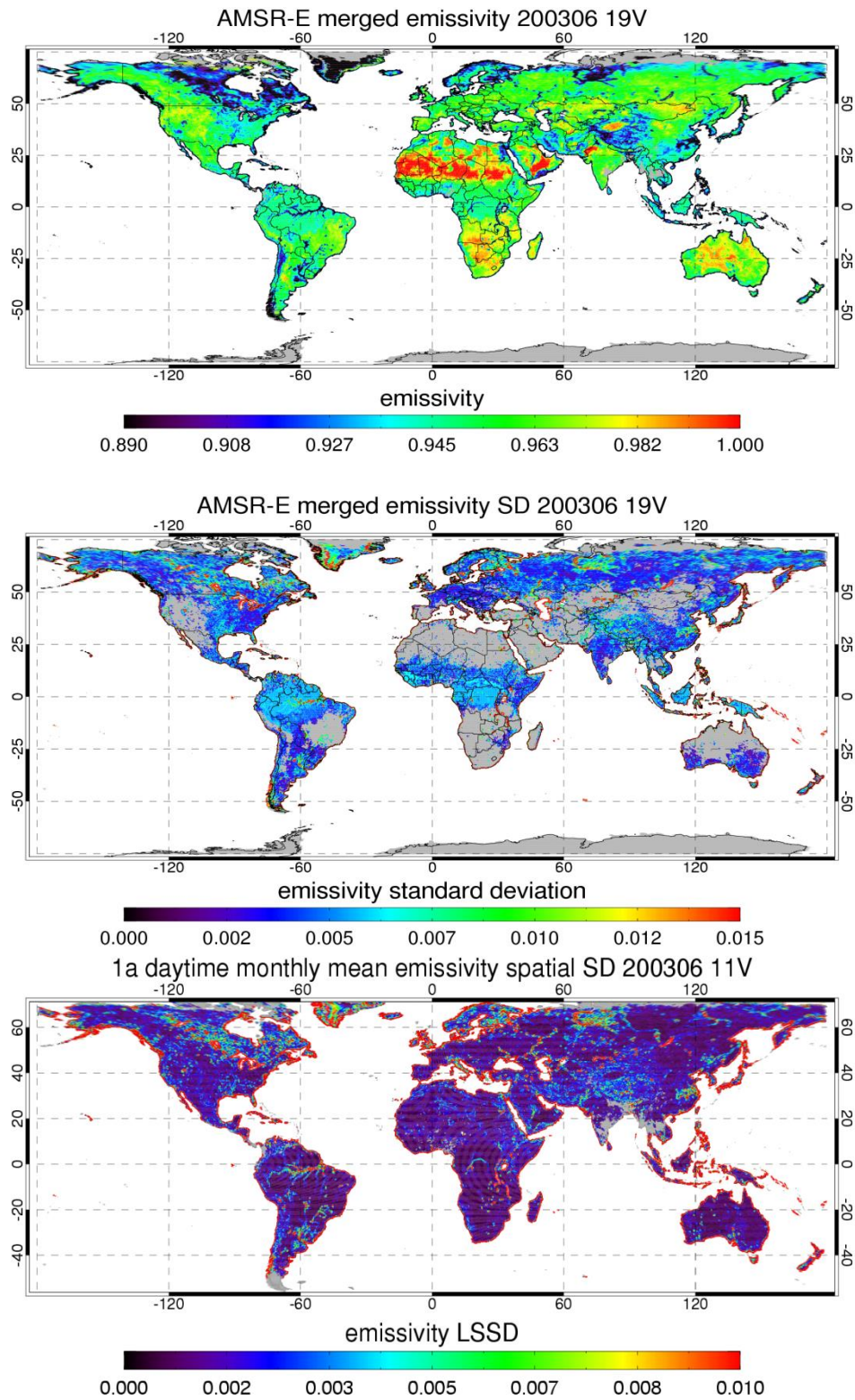
Monthly mean AMSR-E emissivity data for year 2003 is posted for open-access downloads at <http://www.aer.com/scienceResearch/mwrs/emis.html> together with documentation or may be accessed through the NOAA Center for Satellite Applications and Research (STAR) website: [http://www.star.nesdis.noaa.gov/smcd/spb/LANDEM/website/prod\\_AMSRE.php](http://www.star.nesdis.noaa.gov/smcd/spb/LANDEM/website/prod_AMSRE.php). The so-called combined database includes monthly mean daytime and nighttime 1a and C emissivities, as well as mean 1b emissivities, for all AMSR-E channels except 6.925 GHz. Auxiliary information provided with the emissivity data includes number of samples and fraction of clear samples, and 1a emissivity variance and mean spatial standard deviation. Monthly snow and RFI masks and R11-based stability index are part of the QA flags. A special version of the database delivered to Joint Center for Satellite Data Assimilation (JCSDA) also includes AMSR-E mapped infrared MODIS emissivity estimates from the MODIS LST product for concurrent analysis of infrared/microwave temporal and spatial emissivity trends. The combined database is intended for relatively sophisticated users that would like to take account of the properties of each of the data production methods in their application of the database. A merged product (Figure 5) containing the “best available” emissivity among Stage 1a, class-based and Stage1b emissivity datasets is also made available to users for whom the overriding factor is ease of data access and use, or whose applications are consistent with the criteria we used for merger (see product documentation).

Figure 6 shows a comparison between the newly produced AMSR-E emissivity maps and SSM/I derived emissivity data from Prigent *et al.* (2006) for selected months during year 2003 over Asia.

Focusing on the results for the month of November south of 42°N (snow free outside of high elevation regions), it is seen that the products from the two datasets are quite similar (although some significant local differences can be observed). Regions with the lowest emissivities generally correspond to regions where water is abundant (flooded land, e.g. Bangladesh and Yangtze or Indus River valleys, or vicinity of major rivers and lakes) but also snow covered or permafrost areas of Tibet (Figure 7). Deserts over Arabian Peninsula, China, Punjab, Turkmenistan and Uzbekistan are easily identifiable in these datasets by their high emissivity. The low reflectivity of these bare surfaces may be explained by the fact that they are observed at an angle that is close to the Brewster angle. Vegetation tends to scatter the radiation and therefore lower the emissivity (and polarization difference). In our case, emissivity estimates over the desert areas are produced by the 1b algorithm. Note that the magnitude of the emissivities over the Arabian Peninsula and Taklimakan deserts is visually the same in both datasets for the month of November.

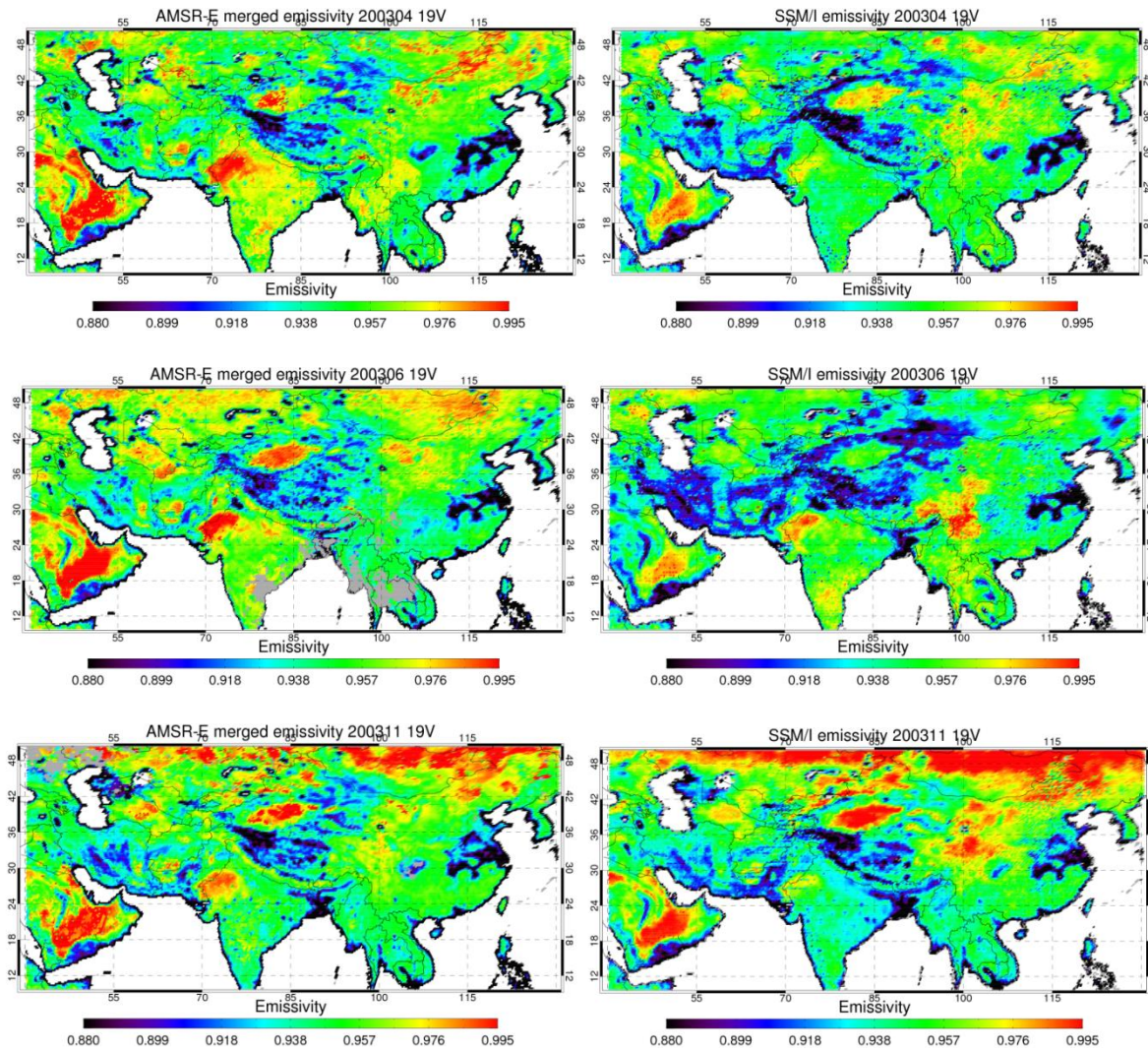
The differences between the two data sources for the months of April and June are more striking. Prigent’s database displays much lower emissivities over the most arid regions. The differences are more pronounced in June when emissivities over desert surfaces in Prigent’s database drop locally below ~0.97. Also, very extensive areas over South Central Asia (Iraq, Iran and Afghanistan) and Western China have their emissivities approaching 0.90 whereas good agreement between the two databases is maintained over highly vegetated surfaces of e.g. Eastern China, South East Asia and the Indian Peninsula at the beginning and after the wet monsoon. The patch of high emissivity in Prigent’s June map over Southern China (North of the Burmese border) could not be explained and appears somewhat suspect. By comparison, the AMSR-E maps remain steadier over the arid regions in the course of the year, which is consistent with a lack of seasonal vegetation growth cycle. The trend observed in Prigent’s data is probably artificial and is consistent with thermal contamination of the estimated SSM/I emissivities (produced with a 1a type algorithm) over the more highly penetrating surfaces, an issue that is being looked at (Prigent, *personal communication*) or seasonal biases in the





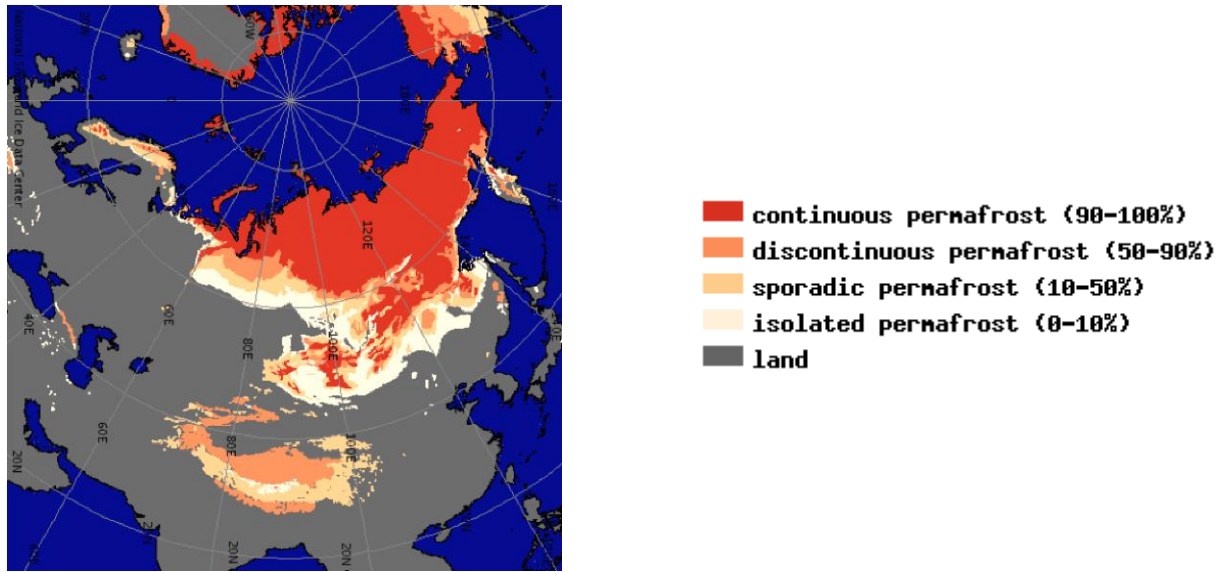
**Figure 5:** Maps of 19V emissivity (top) ESD (middle) and LSSD (bottom) from global merged product for June 2003. No ESD value is provided in areas where 1b or C emissivity product is selected.

ISCCP LST product used to generate these emissivities. Surface penetration effects are minimized in our case by detecting regions of high penetration and switching to the 1b algorithm. The lack of a pronounced seasonal variability over arid and semi-arid surfaces in our case is an indication of the skills of the 1b algorithm to separate thermal effects from surface emission properties.



**Figure 6:** Monthly mean 19V emissivity over Asia for April, June and November of 2003 from AMSR-E (left) and SSM/I (right) databases.





**Figure 7:** Permafrost extent from the Atlas of the Cryosphere (<http://nsidc.org/data/atlas/>)

#### 4. DATABASE IMPROVEMENTS

Our Version 1 surface emissivity database was designed to provide monthly climatology of surface emissivities, capturing slow seasonal trends and inter-annual variability in the surface state. One of the focuses of the present work (toward developing Version 2) is to improve both yield and accuracy of the emissivity estimates. Targeted applications of our emissivity atlas include *daily* retrieval of land surface temperatures (see Sec. 7.1 and 7.3.) and (in the future) atmospheric water vapor under cloudy conditions. Meeting internally set requirements for this type of application implies in particular departing from the arbitrary monthly time division used in the production of our initial time-averaged clear emissivity estimates and rethinking our strategy for updating emissivity values at each location (Sec. 5). For best retrieval results, we rely on the surface emissivity remaining approximately constant ( $ESD < 0.005$ ) at a given location over some period of time. Surface emissivity estimated during the clear portion of this time interval is used as a constraint in cloudy retrievals over the rest of the time period. When a change in surface state is detected at that location, a new estimate of the emissivity that is valid for the next stable time segment is produced. Note that for climate applications, unlike real time retrieval or assimilation, we have the ability to work forward and backward in time, implying that surface emissivity at time  $t$  can be estimated using observations at some future time within a same segment. Only the former climate application is considered here. As mentioned previously, we currently rely on the R11 information for detecting surface changes although the need for bringing in additional external data, especially over vegetated regions, is constantly being assessed. Our retrieval applications are currently restricted to homogeneous, RFI and snow free land areas. LST retrieval over non-penetrating (vegetated surfaces) relies primarily on the AMSR-E 10.65GHz V-polarized channel brightness temperatures because this channel is least affected by the atmosphere (and, in particular, by clouds). Obviously, LST estimates retrieved in the vicinity of sources of 11 GHz radio frequency emission are corrupted and must be flagged as such. In these areas we have the option to use the 18.7 V channel unless this frequency is itself significantly contaminated. Our ability to produce meaningful LST estimates over snow has yet to be assessed. Due to penetration effects, microwave emission

temperatures over snow covered surfaces are likely to be biased compared to infrared  $T_{skin}$  estimates which may call for the use of a technique similar to the one used over deserts (Sec. 7.3) for characterizing the surface thermal structure. However, unlike desert surfaces snow can undergo rapid metamorphism, which in many cases might seriously limit the time window over which the analysis can be performed. Also in the high latitude regions during the winter months, the diurnal temperature is weak which limits the applicability of the 1b method. Validation and refinement of the current 1a approach over snow will be the object of a separate study.

The Version 1 quality control procedure needs to be refined to accommodate the application of our emissivity atlas to production of accurate *daily* retrieval products. In this section we discuss improvements made to our snow mask and 11GHz RFI flagging. Potential approaches for better dealing with spatial heterogeneities (due to e.g. open water bodies) and avoid in particular apparent temporal mismatch between infrared and microwave derived LST (due to day-to-day variability in the AMSR-E sampling around a given grid box - see Sec. 7.1) are examined in Sec. 4.3 and 4.4.

#### **4.1. Snow mask update**

Our snow mask is based on the independent data from the *AMSR-E/Aqua L3 Global Snow Water Equivalent (SWE) EASE-Grids* product ([http://nsidc.org/data/docs/daac/ae\\_swe\\_ease-grids.gd.html](http://nsidc.org/data/docs/daac/ae_swe_ease-grids.gd.html)) distributed by NSIDC (Tedesco *et al.*, 2004a, 2004b). Three global products are available from this source: daily, 5-day maximum and monthly mean SWE on a 25km EASE-Grid. The SWE values are remapped to the emissivity database grid using a nearest-neighbor approach. In Version 1 of the database, our snow flag is determined from the monthly mean SWE value. The snow flag is set to 1 for the whole month when the monthly average SWE exceeds 20mm. While this scheme was found to be adequate for the production of our monthly mean emissivities, it would result in particular in no LST or atmospheric retrieval currently being produced for an entire month at some locations even if snow is present only for a few days. Our new strategy is to attempt to maximize the retrieval yield by producing LST data at least for the snow free portion of a given time period. As mentioned previously snow is often difficult to detect from the R11 time series, a situation that is quite common e.g. over certain areas of Siberia and Northern Canada (and temporal variability of the R11 polarization ratio is not indicative of changes occurring at the higher frequencies which are more likely to be impacted by the snow metamorphism). Independent snow data is used as a complement to the R11 index to 1) exclude clear AMSR-E observations made in the presence of snow in estimating the mean surface emissivity for the snow free days and 2) flag our daily retrieval product.

In the new framework, each day at a given location is assigned its own snow flag (set if SWE > 20 mm) from the daily SWE data. Note that the old and new snow masks may differ temporally but also geographically. However, these differences remain quite local. Note also that there are a few isolated instances (most common during the fall) of SWE being close to zero for one day, and high for some time period before and after that day. A time smoothing function is applied to remove these uncertain cases.

#### **4.2. RFI and instability detection**

X-band RFI is quite widespread over Europe, Japan and California (e.g. Figure 14). Isolated occurrence of RFI is also found worldwide usually (but not exclusively) around major populated areas

(in e.g. the rest of the US, Central America, Australia, South African and South American continents). In certain regions, the level of RFI can vary widely from day to night. In order to avoid erroneous interpretation of the surface emissivities provided in our atlas and to quality control our daily land surface temperature estimates derived from 11GHz measurements over vegetated areas (Sec.7.1), it is necessary to flag instances of significant RFI contamination.

Version 1 of our emissivity atlas uses static maps of RFI produced by Njoku *et al.* (2005) to flag potentially contaminated areas. Njoku’s approach relies on the mean and temporal standard deviations (SD) of the 11V-19V Tb slope exceeding some preset thresholds for detecting RFI contamination. These thresholds are determined from analysis of 2-D monthly histograms of mean and SD of spatially binned daily resampled AMSR-E Tb slopes.

Static maps do not reflect temporal variability in the level of RFI contamination and might arbitrarily screen out some areas where valid emissivity or LST data could be produced over some limited period of time. Moreover the RFI data used in Version 1 does not separate day and night. The Daily Global Land Surface Parameters Derived from AMSR-E (hereafter DGLSP) dataset (Jones *et al.*, 2010, 2009), distributed by NSIDC, provides daily RFI flags. This product contains flags for the 6.925, 10.65 and 18.7 channels. The DGLSP algorithm is applied to the Resolution 1 AE\_L2A AMSR-E swath data remapped to a global 25-km EASE-Grid. The 10.65 GHz RFI flags are set if either one of the two following conditions is met (L. Jones, *personal communication*):

- a)  $Tb_{11H} > Tb_{11V}$
- b)  $\Delta Tb_{11V-19V} = Tb_{11V} - Tb_{19V} > 5K$

As part of this study we assessed the adequacy of the DGLSP RFI flags for our application. In particular, we found obvious cases of RFI that were not identified as such in the DGLSP product (this issue may be in part related to the fact the DGLSP algorithm operates on lower resolution AMSR-E Tb’s) but also areas of false alarm. For these reasons, an approach that better suits our particular application has been developed. Note that for our application we are not concerned about RFI detection over coastal areas or semi-arid regions with rapidly changing surface conditions or any other areas where our algorithm currently fails to produce a reliable mean emissivity estimate (Sec. 3) since over these areas and snow covered areas alike an accurate retrieval of LST is currently not feasible.

#### 4.2.1 Clear-sky emissivity-based RFI detection

One source of difficulty with the use of global thresholds, is that thresholds should be set tight enough to avoid under-detecting RFI, which could result in significant errors (of the order of 5K, for weak RFI) in instantaneous LST estimates, yet thresholds should be kept loose enough in order to minimize false alarms due to other factors (e.g. atmospheric variability or subsurface penetration) modulating the 11-19 GHz Tb slope ( $\Delta Tb_{11V-19V}$ ). If not properly accounted for, these factors may cause RFI free spectra to be wrongly flagged by the RFI detection algorithm. In particular, over deeply penetrating arid areas, the 19 and 11 GHz frequencies naturally emit at different temperatures. Over these areas, one must be careful not to associate RFI with positive  $\Delta Tb_{11V-19V}$  values arising from the fact that emission temperature at 11 GHz is generally warmer than at 19 GHz (which senses closer to the surface) at night. False alarms can also be occasionally triggered in the daytime over areas (e.g. sandy soils) where emissivity decreases with frequency. Examples of erroneous nighttime RFI identification over arid areas in the DGLSP product are given in Figure 9 and Figure 10. Note that in the maps of Figure 9 and all subsequent maps shown in this section, the green color marks RFI areas flagged both

by the AER algorithm and in the independent database (here DGLSP) the red color, areas flagged by the AER algorithm only, and the cyan color, areas identified in the independent database as being RFI contaminated but not by the AER algorithm. In addition, we use the orange color to distinguish the areas in arid regions where the test on Tb slope used by the DGLSP algorithm gives rise to false alarms from other instances of false alarm or failed detection by the AER scheme. Gray color is used to indicate areas where the LST retrieval algorithm is currently not applied.

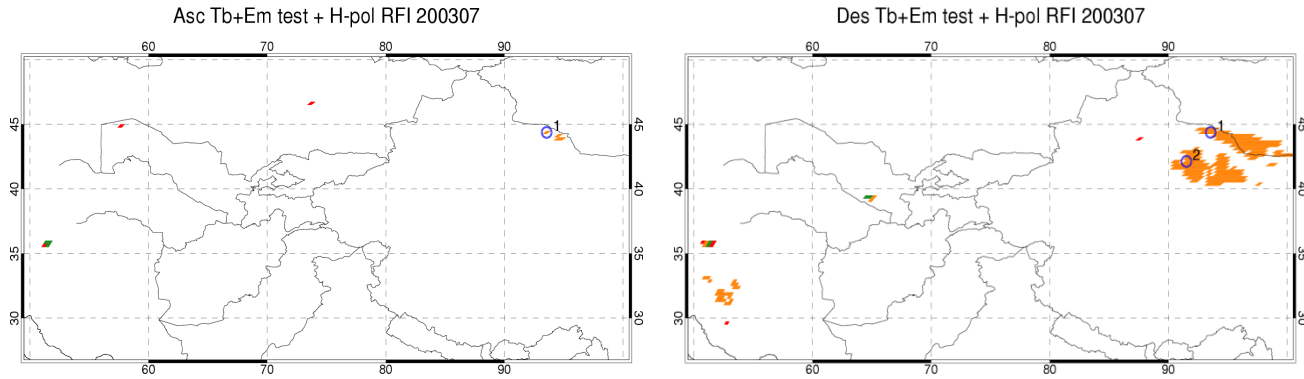
The type of false alarms described above can easily be minimized at least in the clear-sky when using the product of our 1a emissivity retrieval algorithm. As described in Sec. 3, in this retrieval process we first remove AMSR-E observations that have been flagged as cloudy (per the MODIS cloud mask) and R11 outliers, often associated with e.g. sudden changes in soil wetness over arid and semi-arid areas. In a subsequent step, a clustering scheme is applied *a posteriori* to the emissivity spectra retrieved at each overpass to screen out outliers due to e.g. X-band RFI or undetected water clouds (easily distinguishable from clear RFI contaminated spectra because of the impact of clouds on the shape of the emissivity spectrum at the higher frequencies), as well as emissivities produced from occasional bad MODIS LST data. The end product is a mean daytime and nighttime emissivity spectrum for each grid point on Earth where an emissivity cluster is found, that can be used as reference for detecting even weak intermittent RFI sources causing anomalous 11-19 GHz differences in the daily retrieved emissivities initially removed by our clustering process (e.g. Figure 10).

This approach is not designed to capture continuous emission sources since in this case even the mean of the emissivity cluster used for reference is contaminated (Figure 10). However, many of the continuous RFI sources are strong and can be identified by setting a loose threshold on the 11-19 GHz emissivity or Tb slopes. Of course, RFI weakens as we get further away from a main source and instances where the level of contamination falls below ~5K are not currently captured by our tests but another flag can be added in the future to caution about the accuracy of our product when within a certain radius of the location of the peak emission. In certain cases (e.g. Japan) it might be also helpful to rely on day/night emissivity differences.

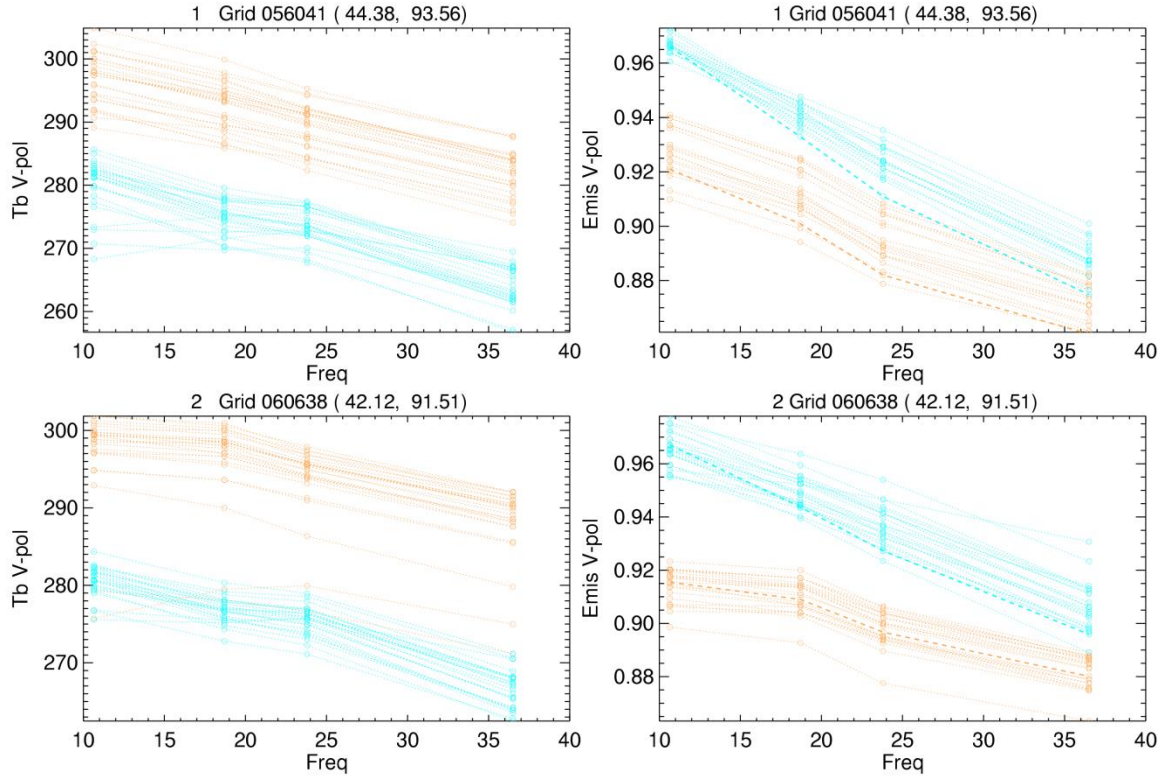
The RFI tests applied to the daily retrieved 1a emissivities are listed in Table 2. Note that the thresholds used for these tests are sensor calibration dependent and would have to be recalculated for any new version of the Level 2A Tb product. The thresholds adopted for the areas of strong penetration (1b regions) also depend on the magnitude of the sub-surface temperature gradients which reach a maximum around the time of the Aqua satellite overpass. Also note that Test 2 only applies to emissivity spectra that match the shape of our reference spectrum between 19 GHz and 37GHz which excludes emissivities derived from cloud contaminated observations, but includes emissivities obtained from bad MODIS LST data. The first order impact of LST errors is a spectrally uniform shift in the magnitude of the clear emissivities. How we deal with cloud contaminated samples (Test 3) will be explained later in Sec. 4.2.2. Note that many of the instances of RFI captured by Tests 1 and 2 can also be identified by examining the emissivity LSSD, i.e. by comparing the local standard deviation of the 11 and 19 GHz emissivities. In a small but non-negligible percentage of cases the outcome of this test is positive where Test 1 and 2 fail. For this reason we keep the LSSD test as an option.

Test #	Type	Test	Conditions
1	Emissivity slope	$\Delta\epsilon_{11V-19V} > 0.015$	1a areas
		$\Delta\epsilon_{11V-19V} > 0.02$ and $\Delta\epsilon_{11V-19V} > 2 \times \Delta\epsilon_{19V-23V}$	1b areas (daytime)
		$\Delta\epsilon_{11V-19V} > 0.035$ and $\Delta\epsilon_{11V-19V} > 2 \times \Delta\epsilon_{19V-23V}$	1b areas (nighttime)
2	Relative emissivity slope	$\Delta\epsilon_{11V-19V} - \bar{\Delta\epsilon}_{11V-19V} > 0.02$	
3	<i>Tb</i> slope	$\Delta Tb_{11V-19V} > 7.5K$ and $\Delta Tb_{11V-19V} > 2 \times \Delta Tb_{19V-23V}$	Cloudy conditions only or Clear and cloudy at C locations

**Table 2:** RFI detection tests. Test 1 and 2 are emissivity tests applied to locations where a 1a emissivity cluster is found. Test 1 is primarily designed to capture RFI contamination from continuous sources but also picks up instances of strong RFI affecting one or more clear days during the month. Test 2 captures weaker intermittent contamination not identified with Test 1. Test 3 is a *Tb* test applied to cloudy conditions or at locations where an emissivity cluster could not be found. In this table,  $\Delta\epsilon_{11V-19V}$  denotes the 11-19 GHz slope of a daily emissivity sample whereas  $\bar{\Delta\epsilon}_{11V-19V}$  represents the slope of the mean emissivity for the cluster.



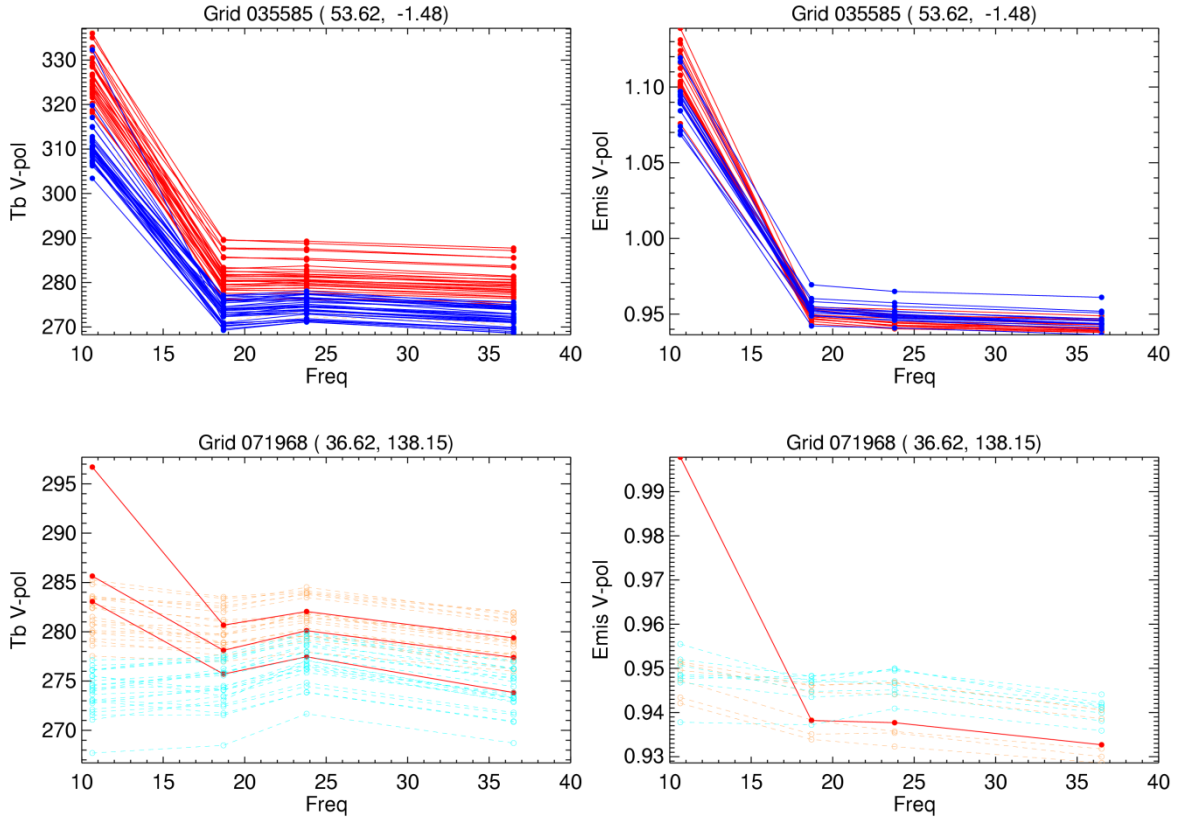
**Figure 8:** Maps showing areas where RFI has been detected at night (left) and during the day (right) by the AER and DGLSP algorithms over Central Asia in July 2003. The areas in orange are areas where the particular surface conditions triggers false alarm with the test on *Tb* slope. The complete color code is provided in the text.



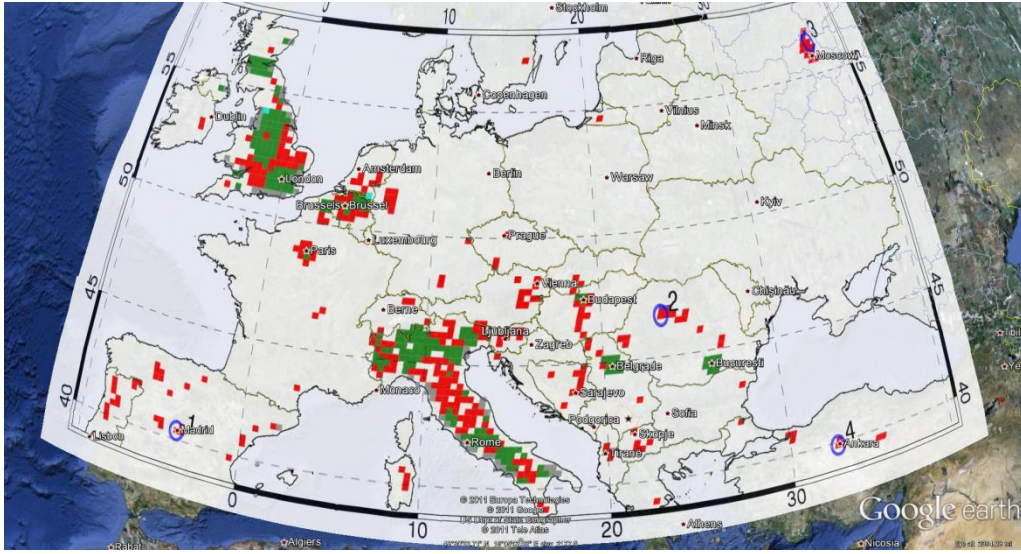
**Figure 9:** Examples of daily daytime (orange) and nighttime (cyan) Tb (left) and emissivity (right) spectra for two locations (Sites #1 and 2 in Figure 8) in North Western China where positive 11-19 GHz Tb slope is incorrectly associated with RFI (see text).

To tune our filter we used the 2003 RFI map from Njoku *et al.* as a first guess and adjusted the threshold so that all locations flagged by Njoku *et al.* are also flagged by our emissivity tests. Here we obviously excluded inhomogeneous water contaminated (coastal) locations and more generally areas where our emissivity retrieval failed to provide a cluster (e.g. frequently cloudy areas). These initial threshold values were then progressively lowered and the emissivity spectra for the locations added in this process were analyzed in order to determine whether they were potentially RFI contaminated or not. The process stops when the thresholds become so loose that they give rise to obvious false alarms. The end result is shown in Figure 11. It is apparent from this figure that our test flags more locations than the original database from Njoku *et al.* Figure 12, show examples of spectra for a few locations which are likely RFI contaminated and not part of the Njoku *et al.* static database.

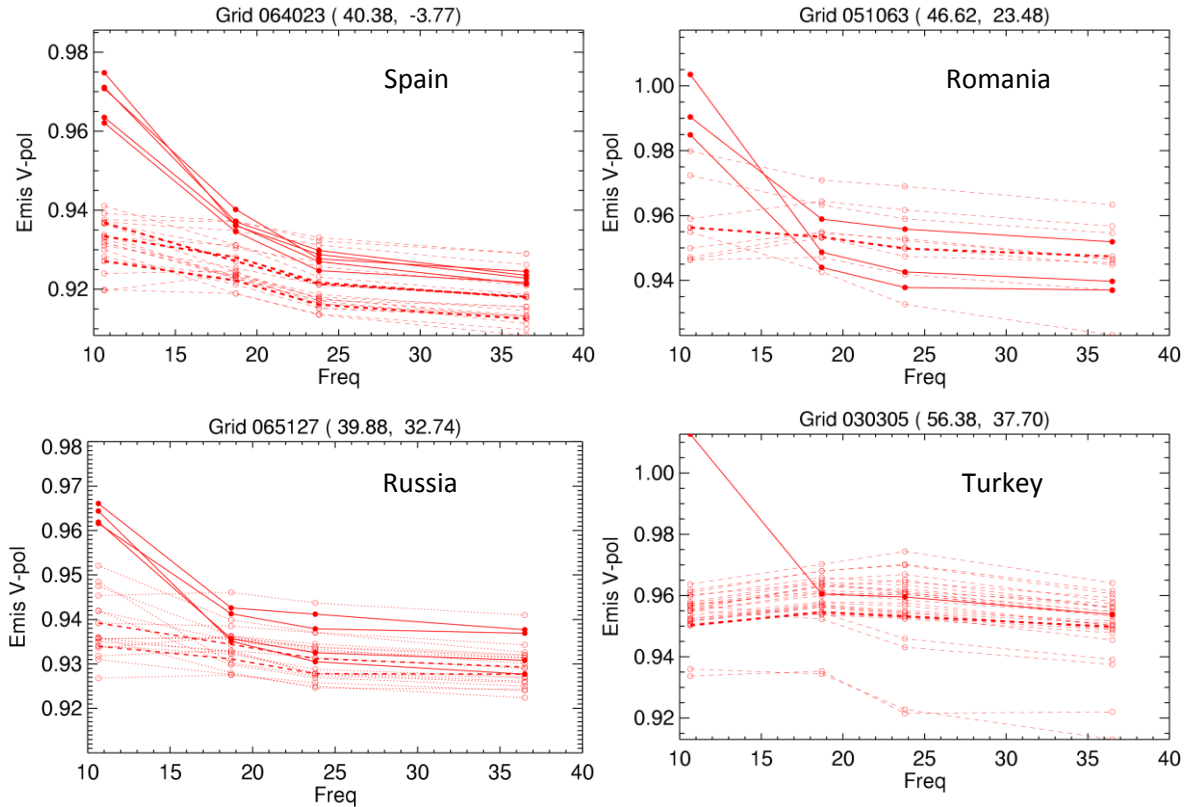




**Figure 10:** Examples of RFI contaminated instantaneous AMSR-E Tb measurements (left) and retrieved emissivity (right) spectra at two selected locations in July 2003. Observations affected by RFI are shown by solid lines (day in red, night in blue), and the non-RFI affected samples are shown by dashed lines (day in orange and night in cyan). The upper panel shows data collected around Leeds (UK), where  $Tb_{11V}$  appear to be continuously affected by RFI both in the daytime and nighttime. The lower panel data is from Nagano (Japan). In this case, only 3 daytime Tb samples appear to be significantly contaminated. Two of the contaminated Tb spectra happen to correspond to partly cloudy conditions, which are excluded from our emissivity retrieval process. The remaining sample is also filtered out by our quality control (clustering) and does not enter in the monthly mean emissivity estimation. However, LST estimates derived from the contaminated 11 GHz data using the correct mean emissivity value as a constraint would in this case display a 7-11K bias.



**Figure 11:** Map showing the areas flagged by the AER emissivity-based RFI detection algorithm in July 2003 and those flagged in Njoku *et al.* (2005) static dataset. Red and green locations on this map correspond to the locations for which at least one day has been determined to be contaminated by the AER algorithm. See text for the full color code. Note that the AER algorithm appears to be missing a few areas flagged as RFI contaminated by Njoku's mask. However, these areas are captured by the same algorithm in other months.



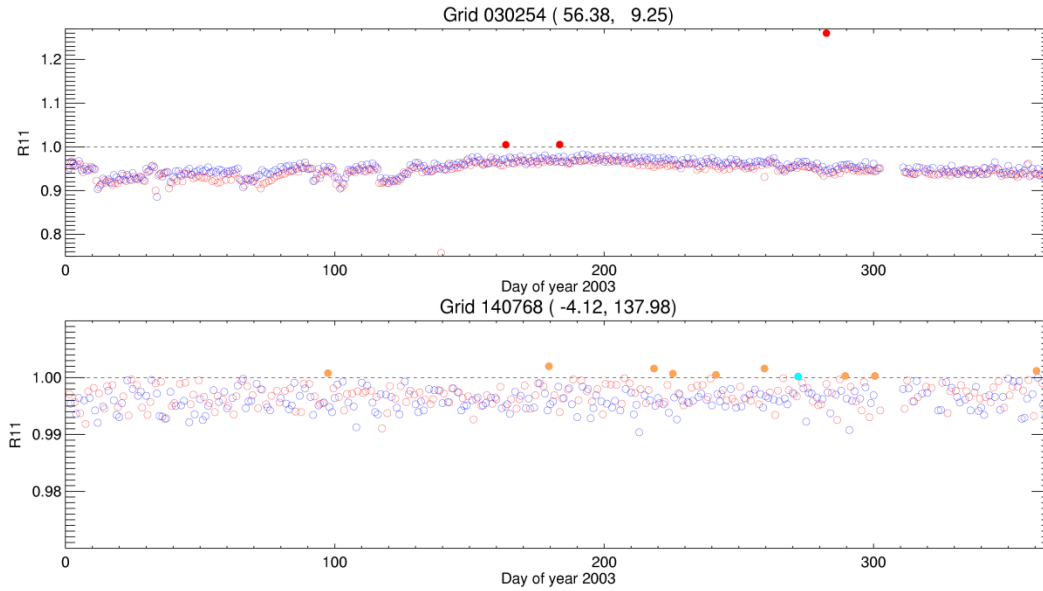
**Figure 12:** Examples of daytime emissivity spectra for locations identified by AER algorithm as RFI contaminated and not flagged in the 2003 static map available from Njoku *et al.* (2005). Same line types and colors as in Figure 10 are used to distinguish RFI contaminated from RFI-free (or uncertain) spectra.

#### 4.2.2 Tb-based cloudy RFI detection

RFI detection based on retrieved clear-sky emissivities is believed to be more robust than Tb-based detection. However by design, such a test cannot capture instances of RFI during the cloudy phases and does not apply to locations where C emissivities are used. Note that we are not concerned with the situations in which no cluster is found for both day and night and the C algorithm fails to provide a substitute since, in this case, no LST (or water vapor) retrieval can be performed with our approach. Here, the emissivity test is kept as our baseline at locations where a 1a emissivity is available, but it is complemented by an all-weather Tb test (Table 2) to attempt to identify RFI 1) on days when clouds are present or 2) in both clear and cloudy conditions when 1a algorithm failed but an emissivity estimate is provided through class-based substitution.

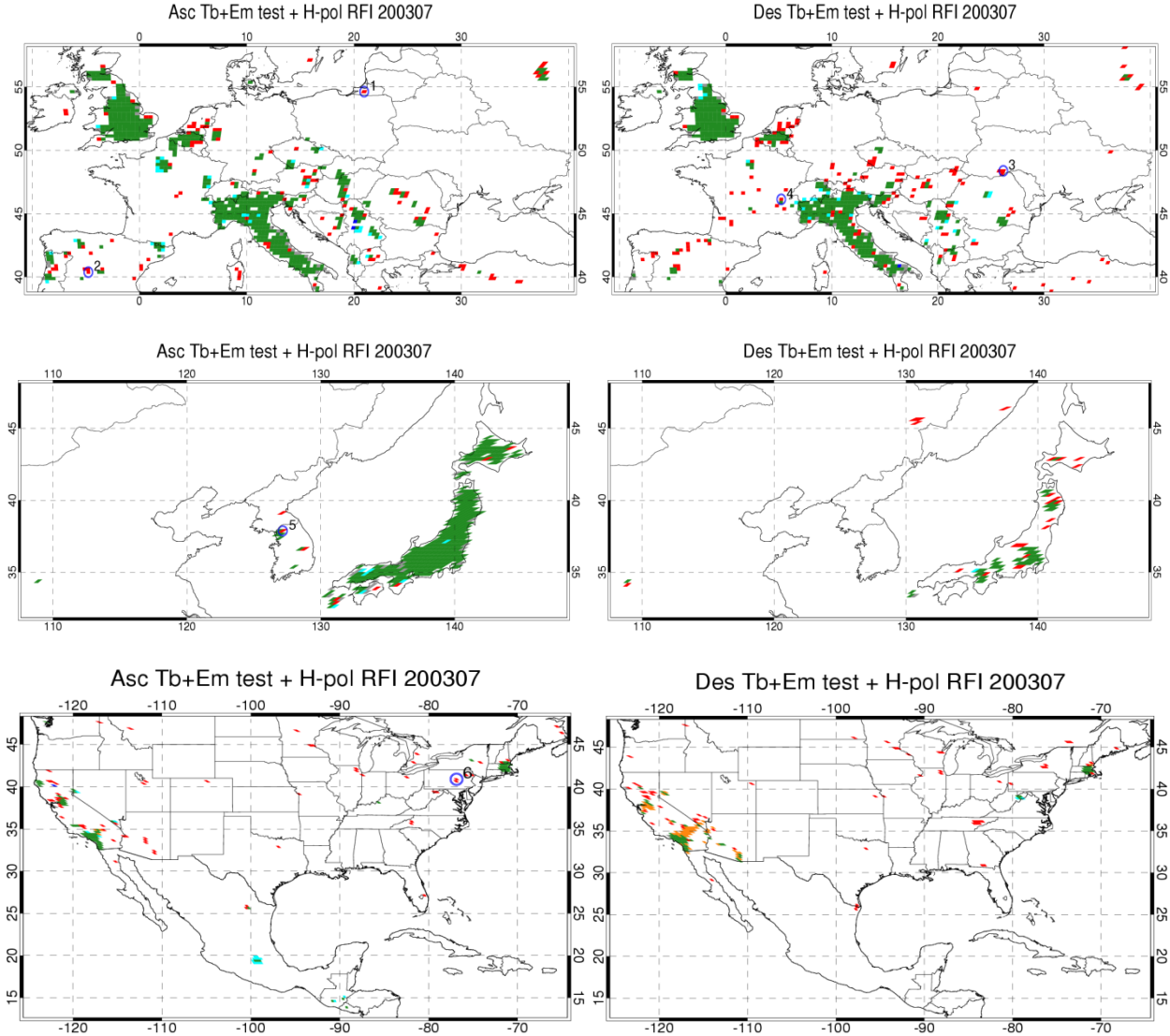
In order to determine a slope threshold for the Tb test (Test 3) we apply the test to both clear and cloudy observations in areas identified as RFI-free according to our emissivity test (Tests 1 and 2) and set the threshold to the lowest possible value producing no false alarms. Here, we are assuming that cloudy observations at some given location are RFI-free when none of the clear-sky observations at that same location are RFI contaminated. Note that some airport landing radars may operate exclusively in cloudy (low visibility) conditions. These may be located by lowering our threshold beyond the value determined with our method and removed from our training set.

Finally, as in the DGLSP scheme, we also include a test to capture strongly H-polarized RFI. However, we included a 0.003 margin to avoid false alarms due to e.g. measurement noise over weakly polarized surfaces (with  $R_{11} \sim 1$ ) (Figure 13).



**Figure 13:** Example of impact of H-polarized RFI on R11 over Denmark (top). In the top plot RFI-free R11 values are marked by open circles and instances of RFI are identified by filled circles. Bottom plot shows instances of R11 slightly exceeding 1 (over Central Irian Jaya) due most probably to instrumental noise.

### 4.2.3 Comparison with DGLSP product

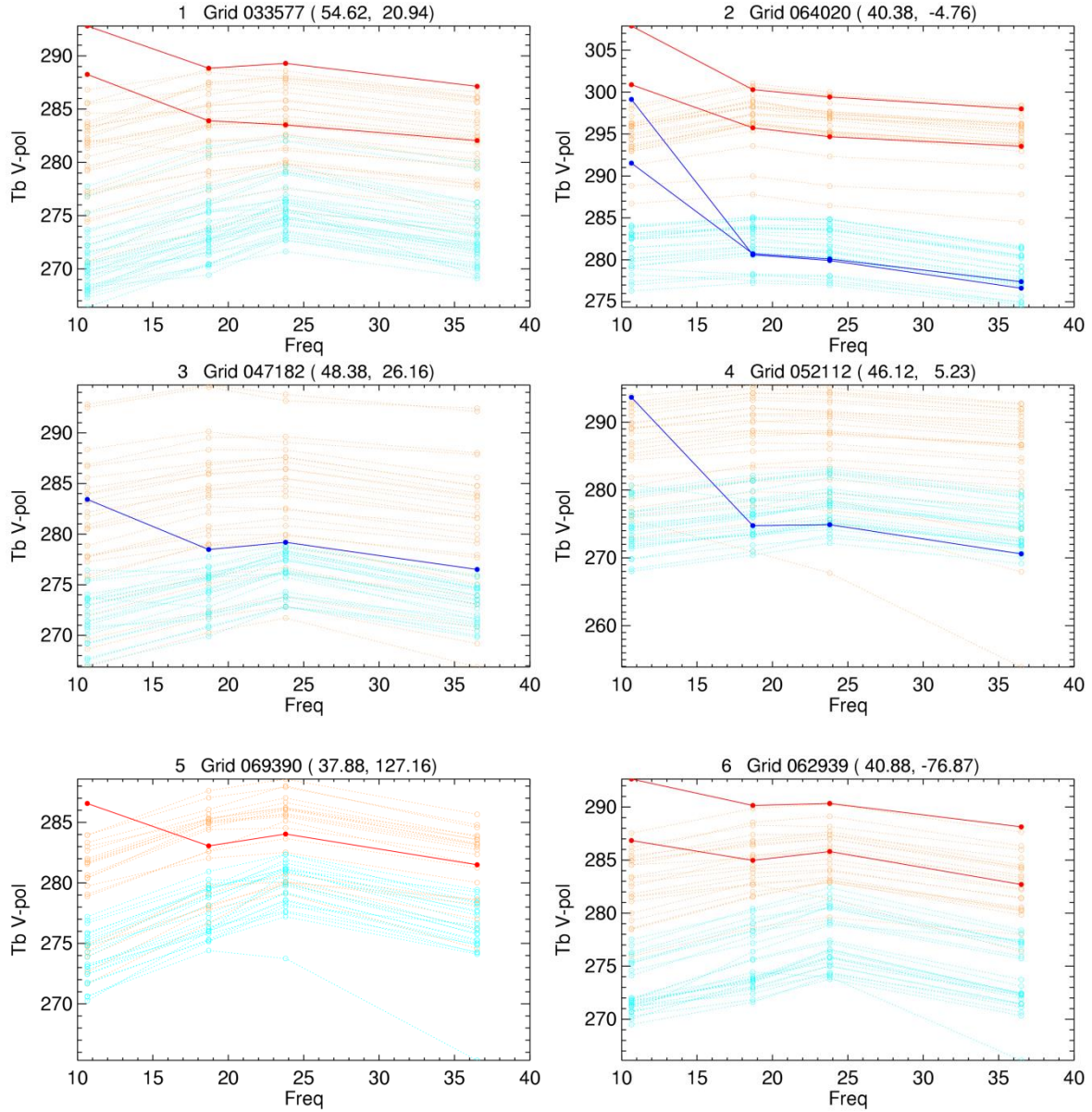


**Figure 14:** Comparison of AER RFI algorithm output with DGLSP RFI flags results for July 2003 over Europe (top), East Asia (middle) and the US (bottom). See text for color code. Note that many of the European locations - especially over Italy and England - that were identified as missing (gray color) in the map of Figure 11 now show up in green because H-polarized RFI is captured using our R11 test at these locations and we let this test override our previous mask.

Our RFI detection scheme (Test1 through 3 combined with R11 test) has been applied to a full year of global AMSR-E data. In Figure 14, we compare the output of the AER algorithm to the DGLSP RFI flags over selected regions for the month of July 2003. Out of a total of 1306 (1010 at night) locations (outside of our exception areas) identified by the AER as being RFI contaminated, globally, ~90% were detected by Tests 1 and 3 alone. When Test 2 is added, ~98% of the locations are flagged. The rest (1-2%) has only been identified by our optional LSSD test applied in emissivity and Tb space. Note that out of this total, only 886 (552 at night) were flagged in the DGLSP product. However, as we mentioned previously, the fact that DGLSP uses lower spatial resolution Level 2A product probably accounts for most of these differences. The number of instances flagged by DGLSP but removed by the AER algorithm was 88 during the day and 448 at night – most of them over arid areas. Figure 15



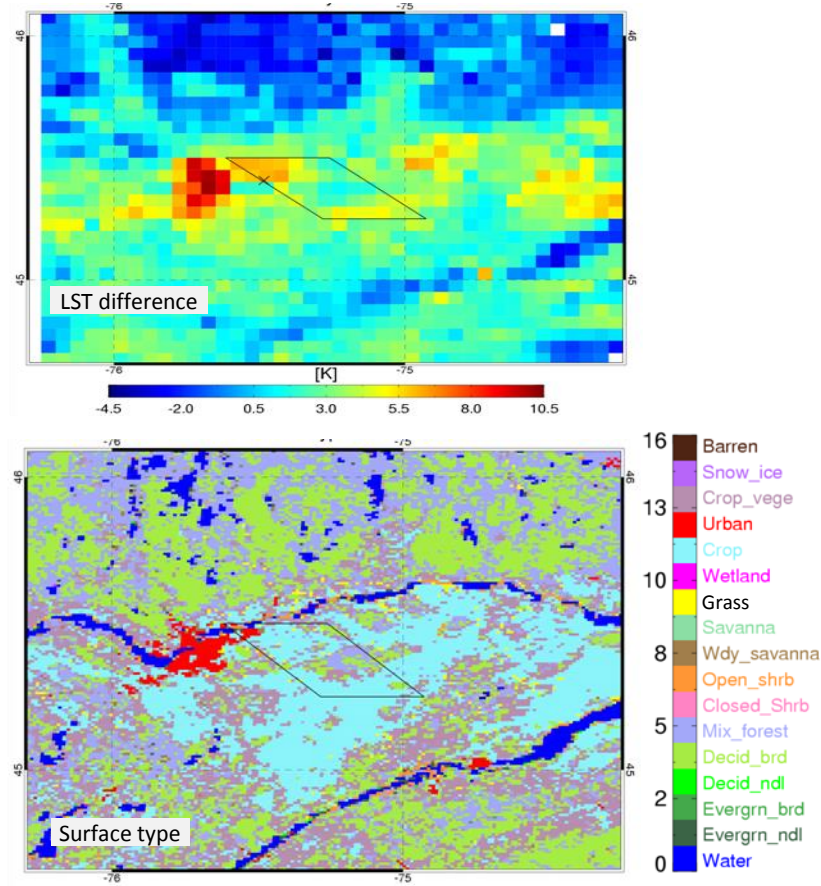
provides typical examples of Tb spectra for cases where RFI was detected by the AER algorithm and not flagged in the DGLSP product. The number of additional daytime H-polarized RFI contaminated areas captured by our R11 test is 280 (107 at night).



**Figure 15:** Examples of day/night Tb spectra for 6 locations (locations are circled in maps of Figure 14) where RFI was detected by AER algorithm and not flagged by DGLSP. In these examples there is only one or a few RFI contaminated samples in each grid box in July 2003. The non-contaminated emissivity clusters are given reference (see Figure 10 caption for explanation of colors and line types). RFI in Grid boxes 2, 4, 5 can be captured by Test 1, and in Grid boxes 1, 3 6 pass Test 2. There is one cloudy daytime sample (lower Tb in high frequency channels) in the plots of Grid box 4, although the Tb difference between 11 and 19GHz is large, it is not larger than the Tb slope between 19 and 24GHz, and it doesn't pass Test 3 for cloudy condition.

### 4.3. Dealing with spatial heterogeneities due to presence of water bodies in the AMSR-E field-of-view

On the spatial scales of our retrievals, it is not at all uncommon for there to be substantial variations in emissivity and LST within the area sampled for a single grid point. Areas with a mixture of urban land and water tend to have the largest inhomogeneities (Figure 16). These occur commonly along coastal areas and in regions of inland water bodies, including rivers, large and small lakes, ponds, and wetlands. While it is not possible to retrieve any detailed features of emissivity or LST at resolutions finer than the resolution of our data, we considered options for treatment of mixed land/water scenes, under the assumption each scene is composed of a combination of uniform land and uniform water.



**Figure 16:** Comparison of midday (13:30 LT) MODIS LST differences (top) with land surface types (bottom) for a region around Ottawa, Canada. The LST differences are relative to the *in-situ* measurement, taken at the site marked “x”, and represent composite averages of clear-sky days in July 2003. In the upper plot, the parallelogram indicates an AMSR-E analysis grid box. Ottawa is the anomalously warm urban area at the upper left edge of the grid box, and the waterways (blue) running across the middle and lower right of the plot are the Ottawa and Saint Lawrence Rivers, respectively.

A challenging factor when dealing with inhomogeneous surfaces stems from the fact that the individual satellite FOVs that are mapped to a single grid point do not all have the same weighted sampling of land and water, due to spatial offsets from the grid point center and variations in the orientation of elliptical FOVs. This issue can be addressed by footprint matching to a fixed target composite FOV for each grid point, as discussed in Sec. 4.4. Nevertheless, some variations in land/water sampling remain, because the composite FOVs do not exactly match the target composite.

Retrieval of sub-FOV information (for land and water separately) can, in principal, be done by data from multiple individual fields of view, taking advantage of their spatial offsets and their related variations in fraction of land and water coverage. This is effectively a resolution enhancement approach, which results in error amplification. An alternative approach is to use ancillary information from other sources, such as a land/water map at a resolution higher than the microwave data. When dealing with EOS data, MODIS provides such a map.

For emissivity retrieval, a MODIS land/water map can be used to estimate the land/water fraction of each grid composite FOV, using processing similar to what we do to obtain the MODIS LST for the AMSR-E FOVs. Using this map, we can compute average MODIS surface temperature separately for the land and water portions of the FOV. In this discussion, we will use the abbreviation WST to refer to the water surface temperature. The emissivity of the water portion can be computed from a model, using the WST and an estimate of the wind speed, which can be obtained from a NWP model or an analysis of surface observations. With this additional information, the emissivity of the land portion of the FOV can be retrieved.

There are a number of significant error sources in this approach to emissivity retrieval in areas with mixtures of land and water. Rather than analyzing these errors, we proceed to consider the LST retrieval problem under the ideal assumption that the emissivity retrieval in these land/water mixed areas is just as accurate as in homogeneous areas. As discussed below, even with this ideal assumption the LST retrieval is so problematic that analysis of the emissivity retrieval errors could not lead us to a satisfactory ultimate result.

We considered a simple scenario for retrieval of LST and WST from microwave data that have been footprint matched to the composite FOV of a fixed grid. The water emissivity is an additional variable, considering that it depends on WST and wind speed. As for emissivity retrieval, we take the approach of using a model with NWP or surface observations to estimate the wind speed. There may be significant errors in this estimate, particularly since the wind speed is generally quite variable near the boundaries of water bodies. With this approach, LST and WST are the only independent unknowns. For a simple error analysis, we consider retrieval using two microwave channels, so that the retrieval problem is fully determined but not over-determined. The separation of LST from WST in the retrieval depends on using channels that are sensitive to these two variables and that the ratio of the emissivities in the two channels is significantly different for land and water. The dependence of water emissivity on WST makes the problem non-linear, so we used numerical calculations for the error analysis.

A simple form of the forward model is

$$Tb_\nu = A_\nu + C_\nu [f_w \epsilon_{w\nu}(T_w) - f_w \epsilon_{lv} + \epsilon_{lv}] + B_\nu f_w \epsilon_{w\nu}(T_w) T_w + B_\nu (1 - f_w) \epsilon_{lv} T_l, \quad (1)$$

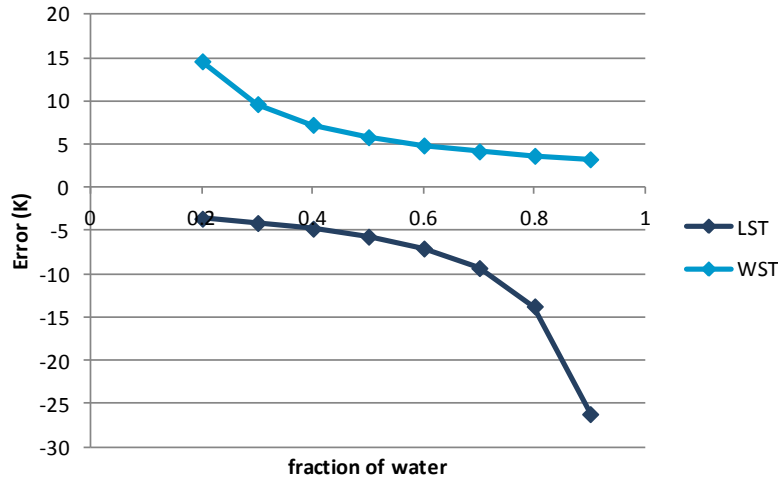
where,

$$A_\nu = \tau_\nu T_\nu^\downarrow + T_\nu^\uparrow, B_\nu = \tau_\nu, C_\nu = -\tau_\nu T_\nu^\downarrow. \quad (2)$$

$T^\uparrow$  is the upward atmospheric emission,  $T^\downarrow$  is the downward emission to the surface from the atmosphere and space,  $\tau$  is the atmospheric transmittance,  $f_w$  is the fraction of water in the FOV,  $T_w$  is the WST,  $T_l$  is the LST,  $\nu$  is the microwave frequency, and  $\varepsilon_{w\nu}(T_w)$  is the water surface emissivity. The dependence of the water emissivity on WST is explicit in this notation. The retrieval problem can be performed by writing the forward model equation for two channels ( $\nu=1$  and  $\nu=2$ ) and solving the simultaneous equations for  $T_w$  and  $T_l$ . For the error analysis, we solved a linearized form of the simultaneous equations and iterated. The linearization was made by replacing  $\varepsilon_{w\nu}(T_w)$  with  $\varepsilon_{w\nu}(T_i)$ , where  $T_i$  is the value of  $T_w$  from the prior iteration and an arbitrary value of  $T_i$  is used for the first iteration.

The error analysis was performed using the 11 GHz V and H channels. The nominal conditions were  $T_w = 290$  K and  $T_l = 300$  K,  $\varepsilon_{11V} = 0.98$ ,  $\varepsilon_{11H} = 0.94$ , with zero wind speed for the water surface. The 10-K contrast between  $T_w$  and  $T_l$  is realistic for an AMSR-E grid scale (Figure 16). The terms  $A_\nu$ ,  $B_\nu$  and  $C_\nu$  were computed with the same radiative transfer model as was used in the emissivity retrievals. Three atmospheric conditions were considered, with samples from soundings with a cold, dry environment, a mid-latitude autumn environment, and a tropical environment (with surface mixing ratios 4, 9, and 18 g/kg, respectively).

Errors in LST and WST were very sensitive to errors in  $f_w$ . For a  $-1\%$  error in  $f_w$  applied to a nominal  $f_w$  of 0.5, the LST and WST errors are  $-5.7$  K and  $+5.8$  K, respectively (Figure 17). The errors in WST increase as  $f_w$  decreases, while the errors in LST increase  $f_w$  increases. The errors were virtually identical for all the profiles tested.



**Figure 17:** Error in LST and WST for a  $-1\%$  error in  $f_w$ , as a function of the nominal value of  $f_w$ . There are no data for  $f_w = 0.1$  because the retrieval produced an unrealistically high WST, for which the water emissivity model becomes invalid.

The impact of wind speed error was significant, but considerably less than the impact of  $f_w$  errors, and the impacts were sensitive to the atmospheric condition. A positive wind speed error roughens the surface and increases the water emissivity to be erroneously high. With nominal  $f_w = 0.5$ , wind speed error of 2 m/s caused errors in LST and WST of  $-1.99$  and  $3.29$  K, respectively, for the cold, dry profile, but errors were negligible for the moist, tropical profile. The insensitivity for the tropical case



occurred because the water brightness with the roughened surface is similar to the total brightness (i.e.,  $Tb$ ) for both channels, and hence the erroneous increase in wind speed has very little effect on  $Tb_{1V}$  and  $Tb_{1H}$ .

Errors in land emissivity were also tested. In a case with  $f_w = 0.1$  (mostly land) and a land emissivity error of 0.001 was applied to both channels, the LST error was  $-0.32$  K, which is similar to what occurs when this emissivity error occurs for a retrieval for an all-land scene with an algorithm that assumes homogeneity (i.e., the standard algorithm for non-penetrating surfaces). The WST error for this case was  $0.38$  K. When the emissivity error is in consistent between the two channels, the impact on LST and WST error is larger. When the 0.001 error in land emissivity was applied only to the 11-GHz V channel, the LST and WST errors were  $-0.55$  K and  $-1.31$  K, respectively.

The scenario with land emissivity error in 11-GHz V channel was run also with  $f_w = 0.9$  (mostly water) and all other conditions the same. In this case, the LST and WST errors were  $-0.53$  K and  $0.11$  K, respectively. The relatively small error in WST for this case follows from the WST in a mostly water scene having relatively little dependence on land emissivity perturbations. It is notable that LST error in this mostly water case had a substantially lower absolute value than the WST error in the mostly land case ( $0.53$  K versus  $1.31$  K). This asymmetry follows from the fact that the nominal emissivity is lower for water than for land: the water emissivity in the 11-GHz H channel is  $0.24$  in these conditions. For illustration of the effect, consider a highly simplified, no-atmosphere radiative transfer equation:

$$Tb \approx \epsilon T_s = (\epsilon + \Delta\epsilon) \left[ \frac{\epsilon}{\epsilon + \Delta\epsilon} \right] T_s, \quad (3)$$

where  $T_s$  is the nominal surface temperature (either LST or WST) and  $\Delta\epsilon$  represents an emissivity error. The surface temperature that would be retrieved by solving this equation is

$$\left[ \frac{\epsilon}{\epsilon + \Delta\epsilon} \right] T_s, \quad (4)$$

so the emissivity error causes an error in  $T_s$  that is modulated by the factor in brackets. The lower the value of  $\epsilon$ , the more this factor differs from 1. For example,  $0.94/(0.94+0.001)=0.999$ , and  $0.24/(0.24+0.001)=0.996$ , so with a nominal  $T_s = 300$ K, the  $T_s$  errors would be  $0.3$  K and  $1.2$  K, respectively.

The error analysis indicates that the impacts of errors in  $f_w$  alone are large enough to make the retrieved LST and WST essentially useless. These assumed  $f_w$  errors are modest in relation to the 6% discrepancies against reference maps obtained where  $f_w$  was retrieved over Canada from microwave satellite data (Fily *et al.*, 2003); where the ability of Fily *et al.* to retrieve  $f_w$  depended on neglecting all other influences on FOV-average emissivity and assuming that the surface temperature of the water and land parts of the FOV are the same. In light of our results, we concluded that it is impractical to retrieve separate LST and WST in inhomogeneous composite FOVs. When a single emissivity and a single temperature are retrieved for the FOV as a whole, the results represent effective averages, where this averaging is not the same as a simple average over the covered area. Nevertheless, this approach

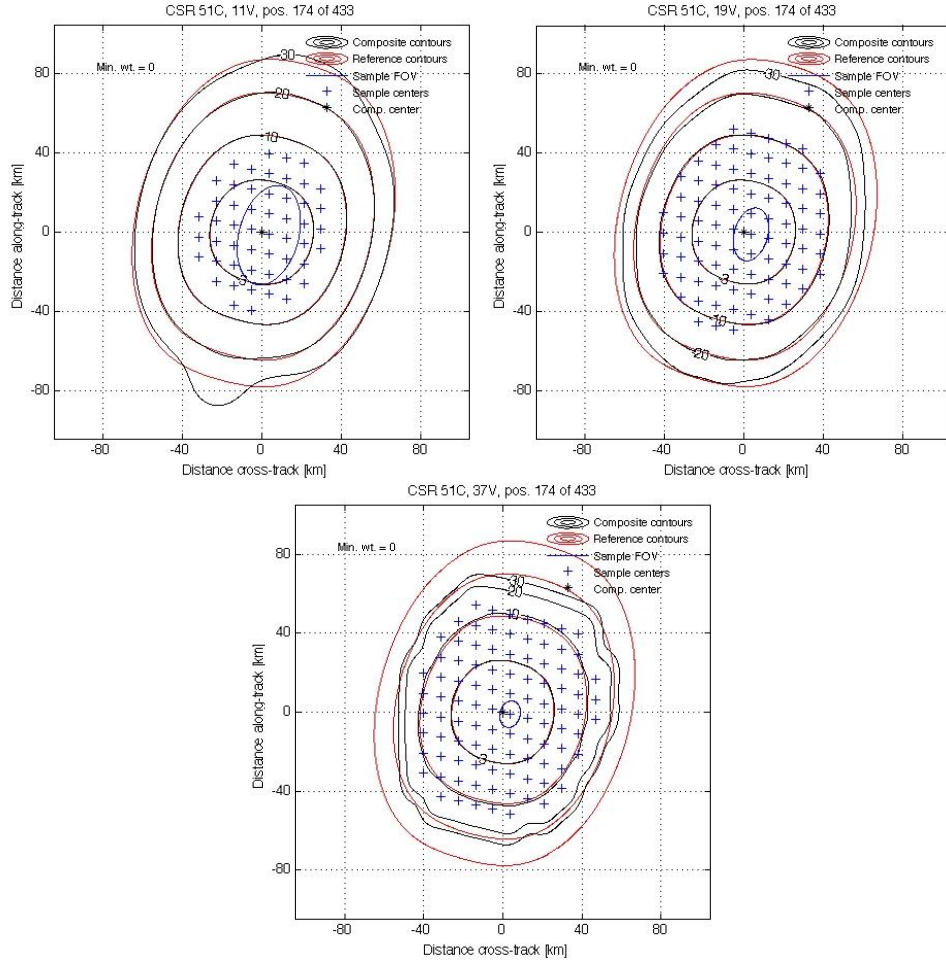
yields an average that is similar to the averaging performed by a numerical model such as a reanalysis system, and thus provides a very useful basis for comparison. In comparisons with model results, we would not expect the averaging difference to be any bigger a factor than the difference between the shape of the composite FOV and the model grid box.

#### **4.4. Training of an AMSR-E footprint matching algorithm**

Characterizing and controlling the spatial sampling properties of AMSR-E for our application is important for four reasons: isolating the slowly varying structure of the land surface cover mosaic from shorter-term (e.g., atmosphere, soil moisture, inundation, freeze/thaw, snow cover) and longer-term (e.g., vegetation seasonality) varying effects; coordinating observations from the AMSR-E channels that have different native sensor footprint sizes (not applicable to L2A product where all channels are already matched to common footprints on fixed grid in satellite reference frame); registering the AMSR-E data to land surface databases; and separating temporal and spatial variability in emissivity retrievals. We are implementing a footprint matching and regridding strategy for AMSR-E based on methods developed for the now-canceled National Polar Orbiting Environmental Satellite System (NPOESS) Conical-Scanning Microwave Imager/Sounder (CMIS) (Galantowicz *et al.*, 2003; Galantowicz, 2004).

Briefly, the footprint matching algorithm forms 2D composite footprints (herein CFOV for composite field-of-view) in the sensor reference frame from a weighed sum of neighboring native-resolution sensor samples (e.g., sensor SDRs). The weights are derived to optimize the fit of the composite footprints to a specified reference footprint shape while controlling for the composite sample noise characteristics, which are a function of the sum-of-squares of the weights. The reference footprints are carefully designed to meet several criteria: they are circularized compared to the elliptical shape of AMSR-E sensor footprints; their resolution is no smaller than that of the channel with the largest resolution to which the matching algorithm will be applied; they are the same at all scan positions; and they are highly oversampled along-track and along-scan relative to the native sensor footprints. The objective of these criteria is to provide sets of virtual sensor samples that have close to the same spatial weighting functions and are closely spaced in the sensor reference frame for a selected set of channels. These virtual samples are to be accessed by a regridding algorithm, allowing it to populate an earth-reference grid with multi-channel data that are like-sized and nearly circular, minimizing differences in spatial sampling at each grid point over time and between sensor channels.

In this CRP, the footprint matching algorithm was configured for the AMSR-E sensor and an initial set of footprint matching weights was derived and evaluated. Figure 18 shows example composite footprints and the reference footprints they are designed to match. The reference footprint type, 51C, indicates its composite spatial resolution (CSR) is 51 km and its shape is circularized (i.e., its along-scan and across-scan FOV dimensions are equivalent). To minimize hard-to-match footprint components with high spatial frequencies, the 51C reference footprint is formed by scaling and integration of the 11V pattern in the along-scan direction until the desired 51×51 km FOV size is reached. In each case in Figure 18 there is good correspondence between the reference and composite contours at the 3 and 10dB levels and varying degrees of misfit at the 20 and 30dB levels. This misfit is an acceptable trade-off to minimize the number of samples contributing to the composite and the overall complexity of the process. Footprint matching quality metrics are discussed in more detail below.



**Figure 18:** Example composite footprints at position 174 of 433 (counter-clockwise) in the 51C scan pattern. Sensor channels: 11V, 19V, and 37V. Contour levels at -3, -10, -20, and -30dB (relative to peak) are shown for the composite (black) and reference (red) footprints. Center locations of the sensor samples contributing to the composite are plotted as “+” mark and a single sensor sample FOV is plotted in blue for scale. The centroid of the composite footprint is plotted as a “\*” mark.

Footprint matching weights are tuned to optimize a cost function including goodness of fit to the reference footprint shape and noise. These factors can be controlled in two ways: by varying a tuning parameter in the weight-calculation algorithm and by varying the samples that can contribute to the composite. After the optimized weights are found, visual inspection of a sample of composite footprint contours (e.g., Figure 18) reveals obvious side-lobes or FOV mismatch. The following 2D spatial quality metrics are then calculated for selected composite scan positions to assess the overall footprint matching performance in detail:

- 1) *Absolute fit error relative to the reference footprint:*

$$F = \iint_E dA \left| G_r(\rho) - \sum_{i=1}^N a_i G_i(\rho) \right|, \quad (5)$$

where  $G_r$  is the earth-projected reference footprint pattern as a function of earth coordinate vector,  $\rho$ ,  $G_i$  are the collected of  $N$  samples contributing to the composite footprint,  $a_i$  are the

footprint matching weights, and  $E$  represents the entire area of earth upon which the footprints are projected.

- 2) *Square root of the squared error relative to the reference footprint:*

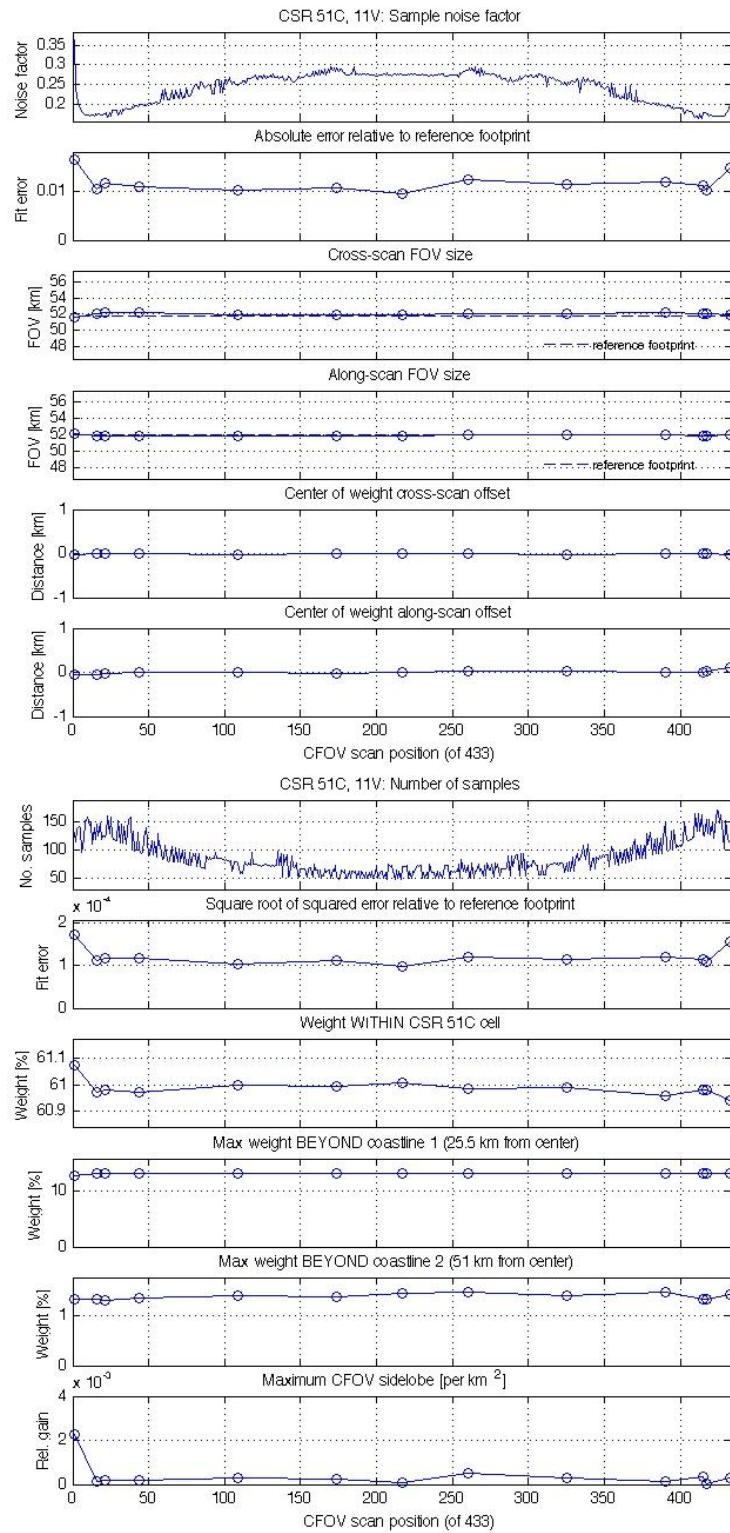
$$F_2 = \sqrt{\iint_E dA \left[ G_r(\rho) - \sum_{i=1}^N a_i G(\rho) \right]^2} \quad (6)$$

- 3) *Along-scan and cross-scan FOV size:* Calculated as the size of the smallest rectangle inscribed by the composite footprint 3dB contour.
- 4) *Center-of-weight along-scan and cross-scan offset from the reference footprint center of weight:* The center-of-weight is calculated from points within the footprint 3dB contour. The reference footprint center of weight is not at the beam peak at oblique incidence angles and the composite footprint peak is a poor measure of the center especially if undersampling results in rippling across the center of the footprint.
- 5) *Weight within 3dB contour of reference footprint ("51C cell"):* Total composite footprint weight integrated over the 3dB contour of the reference footprint, expressed as a percentage. This is an integral of the earth-projected footprint over the specified region. For comparison, the weight within the 3dB contour for a circular Gaussian reference footprint is only about 50%.
- 6) *Maximum weight beyond coastline 1:* The maximum total composite footprint weight integrated separately over four regions tangent to the reference FOV. That is, if  $Y_r$ ,  $X_r$  are the reference footprint cross-scan and along-scan FOV sizes, then the four regions are demarcated by  $y > Y_r/2$ ,  $y < Y_r/2$ ,  $x > X_r/2$ , and  $x < X_r/2$  and can be thought of as the footprint weight falling across a linear coastline. Large values indicate that either the composite FOV is much larger than the reference or that the composite has side-lobe-like spatial features. Also includes the "off-grid" weight that comes from performing all integrals over a finite region. For comparison, the weight for a circular Gaussian reference footprint is about 14%.
- 7) *Maximum weight beyond coastline 2:* Same as above only the four regions are farther from the footprint center:  $y > Y_r$ ,  $y < -Y_r$ ,  $x > X_r$ , and  $x < -X_r$ . Also includes the "off-grid" weight. For comparison, the weight for a circular Gaussian reference footprint is about 2%.
- 8) *Maximum side-lobe level:* Value of any local maxima beyond the 3dB contour in terms of gain relative to the footprint peak.

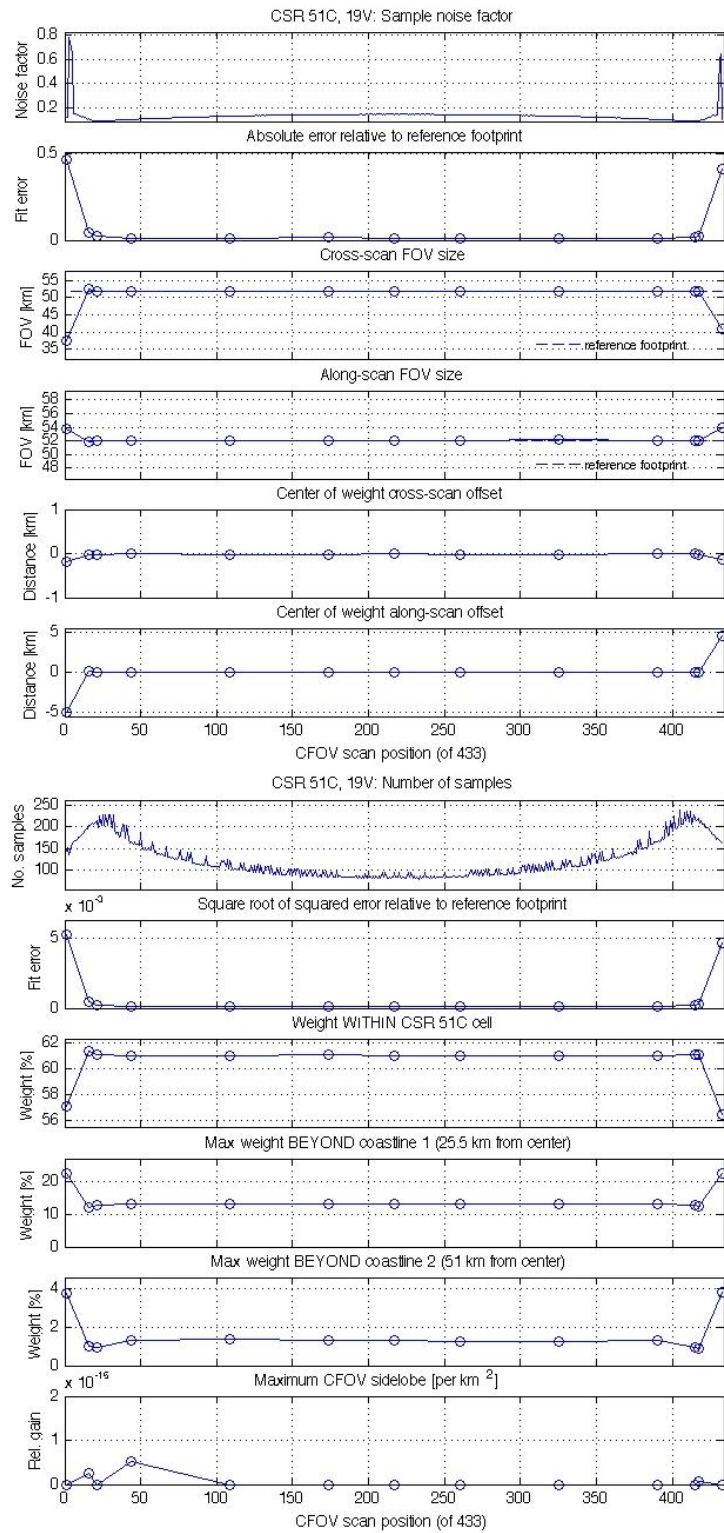
Figure 19, Figure 20, and Figure 21 give examples of these metrics for the 11V, 19V, and 37V channels matched to the 51C CSR. The figures also show the primary noise metric, the sample noise factor (SNF), which can be used to express the composite sample noise  $\sigma_c$  in terms of the sensor noise  $\sigma_i$  when  $\sigma_i$  is assumed independent per sample (which is not strictly true because of correlated noise from shared calibration samples):

$$\sigma_c = \sigma_i \sqrt{\sum_{i=1}^N a_i^2} = \sigma_i \cdot SNF \quad (7)$$

The spatial metrics indicate that the fit is good overall: in all cases, the CFOV size is about  $52 \times 52$  km, the along- and cross-scan offsets are near zero, the weight within the 3dB reference contour is about 61%, and there are no side-lobes (except near edge-of-scan). The “coastline” weight metrics are highest for 11V and decrease by degrees at 19V and 37V but in all three cases they are lower than the same metric for an idealized circular Gaussian reference footprint. By design, some scan positions near the edge of scan will be left unused due to the fact that there are insufficient samples to form a symmetric composite footprint closer to the sensor scan edge. These tests help determine the extent of the useful composite sample scan.

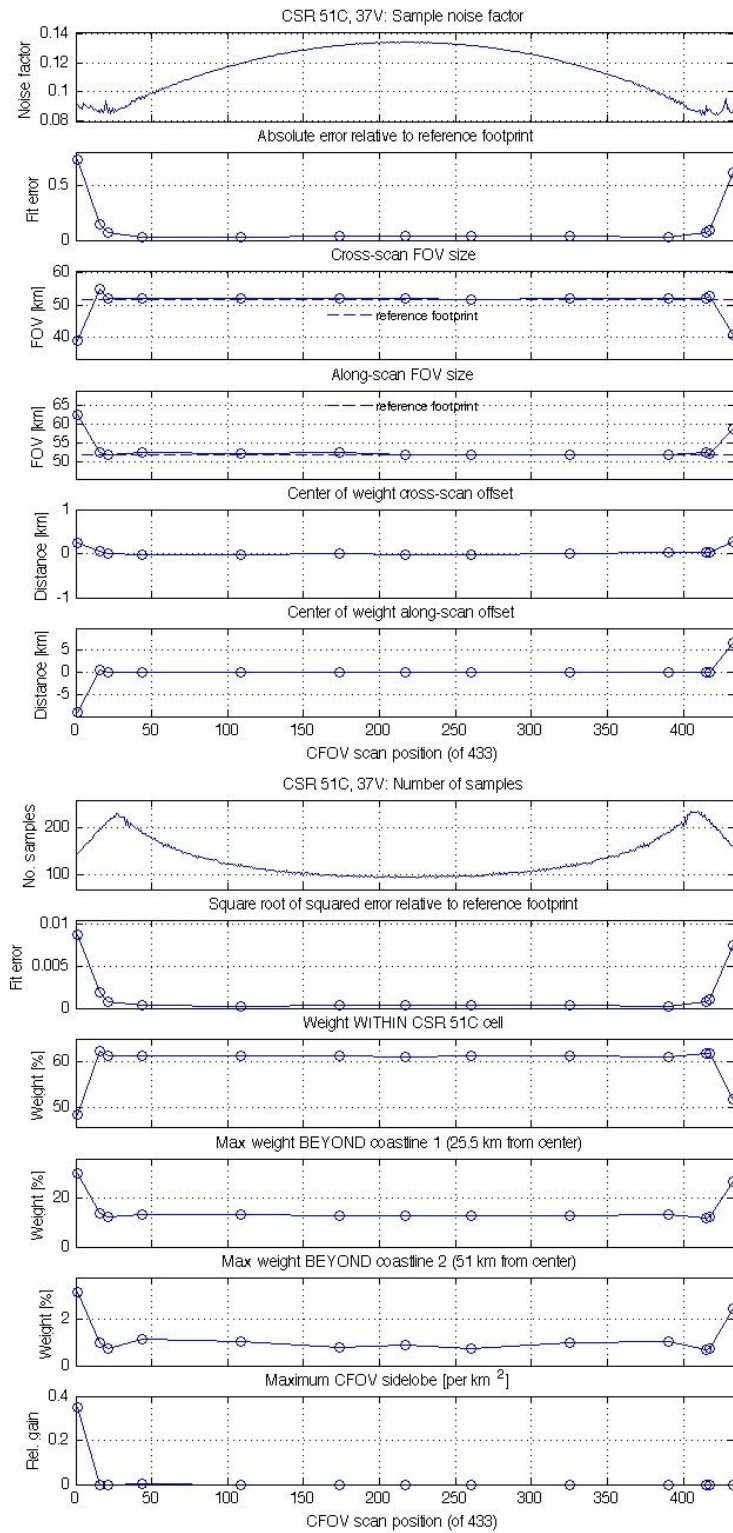


**Figure 19:** Footprint matching metrics as a function of composite scan position for 11V composites for the 51C CSR. CFOV = composite field-of-view. Sample noise factor is used to scale the sensor channel  $NE\Delta T$  to the effective noise of the composite sample.



**Figure 20:** Footprint matching metrics as a function of composite scan position for 19V composites for the 51C CSR. CFOV = composite field-of-view. Sample noise factor is used to scale the sensor channel NEAT to the effective noise of the composite sample.





**Figure 21:** Footprint matching metrics as a function of composite scan position for 37V composites for the 51C CSR. CFOV = composite field-of-view. Sample noise factor is used to scale the sensor channel NE $\Delta$ T to the effective noise of the composite sample.

As illustrated above, a baseline algorithm is now in place for calculating AMSR-E footprint matching weights with acceptable spatial and noise performance metrics. However, algorithm performance will



be reassessed when it is applied to AMSR-E data (e.g., through evaluation of variability in gridded time series data). Next, the algorithm that applies the weights to AMSR-E data and derives gridded brightness temperatures will be completed. Forms of this algorithm have already been developed for application to SSM/I and simulated CMIS datasets and these algorithms will be adapted for use with AMSR-E data.

Our Version 1 emissivity retrieval algorithm performs “instantaneous” emissivity retrieval calculations using AMSR-E Level 2A “resolution 2” brightness temperature data (e.g., resampled to match the 11 GHz native footprint size of 51×29 km). In these calculations, Earth-gridded MODIS LST data (5-km resolution) are collected for each AMSR-E footprint ellipsoid, quality-controlled via the accompanying quality flags, and averaged spatially using a truncated AMSR-E footprint pattern, providing a skin temperature estimate for the emissivity retrieval. The instantaneous emissivities are then mapped to a fixed earth grid, with all measurements falling within a 10 km radius of a given grid point assigned to that grid point. When there is more than one measurement within this radius from a single overpass, they are averaged.

An alternative processing order that takes advantage of the footprint matching described in this section proceeds as follows:

- 1) 51C footprint matching is applied to AMSR-E data in the swath reference frame with 2.5 km virtual sampling.
- 2) Footprint matched Tb’s are resampled to the earth grid by nearest-neighbor resampling (at most 1.8 km registration error) or a 2D interpolation method. Either method would produce Tb samples on the earth grid with nearly circular 51×51 km spatial weighting functions.
- 3) MODIS LST data are averaged spatially using either an idealized circularly symmetric 51×51 km spatial weighting pattern or the 51C reference footprint oriented to the local scan angle. Note that in this new scheme, the MODIS weights can be pre-calculated once and for all at all locations, which will significantly speed up our end-to-end emissivity retrieval process.
- 4) The instantaneous retrieval algorithm is performed once for this earth grid point.

The main advantages of this method over the current algorithm are: (1) it provides approximate spatial-sampling repeatability at each grid point for better separation of spatial and temporal variability and (2) it provides a gridded product with a well-defined and temporally consistent spatial extent that can be readily compared to surface type and other gridded datasets. The method may also reduce the total number of retrievals required to populate the grid but it may not be advantageous to take full advantage of this feature since it is useful to analyze a set overlapping samples to assess spatial heterogeneity in the input data and retrieved products. This analysis can still be done, e.g., by decreasing the earth grid spacing (now at 27.8 km) or by analyzing multiple composites in the vicinity of a grid point archiving a local spatial standard deviation as is currently done.

The main disadvantages of this method are: (1) the instantaneous retrieval resolution is increased from 51×29 km to 51×51 km (i.e., roughly 75% with respect to total area) and (2) there is effectively only one larger instantaneous retrieval cell per earth grid point versus up to 14 in the current method. Combined, these will reduce the overall product yield particularly near cloudy areas. One way to gain back some of that yield would be to modify the truncation threshold used when MODIS LSTs are spatially averaged to the reference footprint.

## 5. NEW TIME SERIES SEGMENTATION APPROACH

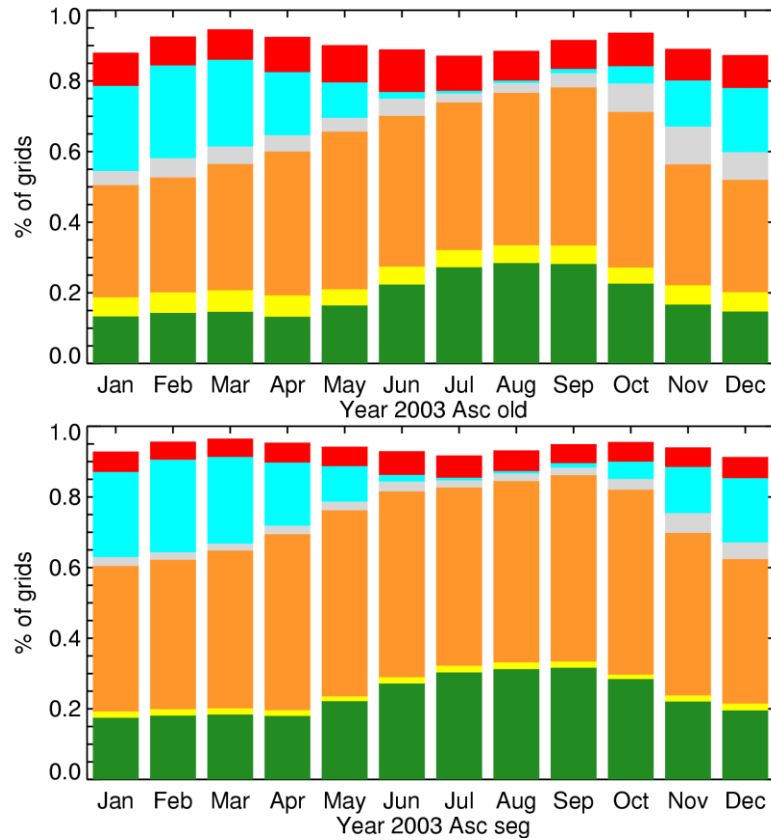
In Version 1 of our emissivity database, a new emissivity estimate is regenerated every month at each location (Sec. 3). For convenience, we label this approach Monthly Time Division (MTD).

There are periods of the year in some regions when clouds become frequent and the number of clear observations over a month long time period is scarce. The fact that we have to introduce partly cloudy observations to populate each grid box and minimize temporal gaps, introduces noise in our estimates. The class-based substitution procedure (Sec. 3) was designed to palliate this issue but this procedure has its own shortcomings. First it is only applied to highly vegetated areas. Secondly, we currently have no direct means, other than performing off line validation, of assessing the adequacy of the estimates assigned to the target location. For these reasons, we would rather attempt to increase the yield of the NPA (and PTS) algorithm and rely as much as possible on direct measurements.

The time segmentation procedure (R11-based Time Segmentation – or RTS) we are experimenting with takes advantage of the fact that there are many regions on Earth where emissivities are thought to remain relatively stable over time periods of a few months or more. These regions include deserts and highly vegetated (forested) areas. In these regions, tailoring the analysis window to the rate of change of the emissivity allows one to sometimes operate on longer (i.e. longer than one month) time segments and thereby maximize the number of clear observations used for estimating the surface emissivity. At first we rely on a time series analysis of the R11 polarization ratio performed independently at each location for detecting changes in the surface emissivity. As shown in Sec. 7.4, the R11 index is highly sensitive in particular to changes in soil moisture or state of vegetation in semi-arid regions. If it is safe to assume that a change in R11 (of magnitude exceeding the estimated noise level) is always accompanied by a change in emissivity, we recognize that the opposite is not necessarily true, especially over vegetated areas. The case discussed in Sec. 7.2 of nighttime emissivity changes over the US Corn Belt, which we attribute to dew deposition, is one example of phenomenon that has no noticeable R11 signature. Snow may also be difficult to detect by this criterion which is the reason why we make use of an independent snow mask (see Sec.4.1) in our retrieval process. Identifying other such situations is one of the objectives of our ongoing and future validation work (Sec. 7.1). More generally (although this particular enhancement has not been implemented yet), changes in surface emissivity across a segment over non-penetrating areas can be detected by examining the instantaneous highest quality 1a emissivity estimates and assessing whether a temporal trend can be identified. If an emissivity change is detected across a segment then further breakdown of that segment into shorter intervals is required to ensure that our emissivities capture those trends. This approach may fail to capture emissivity variations through a given month if no high quality estimate is available either during that particular month or during the previous or following months, but the estimate provided should still be of better quality than the one that would have been produced in Version 1 using cloud contaminated samples. Over penetrating areas, a similar approach is being used. One difference is that in this case we have to limit the width of the analysis window regardless of whether emissivity might change or not in order to accommodate the fact that the solar forcing is assumed constant through the analysis period in the 1b algorithm. We currently limit the size of the window to 30 days to make sure that the variations in the magnitude of the solar flux remains small. The algorithm is applied to different times throughout the segment. As in the non-penetrating case, the algorithm looks for a trend in the highest quality emissivity retrieval through the segment. When such trend is detected, the estimate which is the closest in time to the target month is retained. Note that the RTS method has been fully implemented but is still being evaluated and refined. The goal of this section is to provide an overview of the foreseen impact of the new RTS approach on yield and quality

of our emissivity estimates. The procedure for breaking down a full year of data into successive segments is described in Appendix C.

The impact of the RTS approach on the QA partitioning summary (see Table 1) is shown in Figure 22. To facilitate the comparison with the MTD approach this figure reproduces the plot of Figure 1. Note that in this figure, Categories 7 and 8 from Table 1 are merged into a single category. The same applies to Categories 5 and 6. Also note that the fraction of snow covered areas is kept the same in both cases. In practice, in our new approach, we take into account the fact that part of a given month may be snow free by grouping the days within that month into “snow” and “no snow” categories and by providing a separate emissivity estimate for each group. In order to avoid complicating the interpretation of our charts this detail is omitted here and if any single day is snow covered we tag the whole month as snow covered with both approaches. It is apparent from Figure 22 that the new segmentation scheme results in a significant reduction in the number of grid points classified as unstable, uncertain or missing (Categories 2, 3 and 4) as well as in those grid points for which class-based substitution is invoked (Category 9) (see Table 1 for the definition of those categories), and in a corresponding increase in fraction of grid points classified as 1a or 1b.

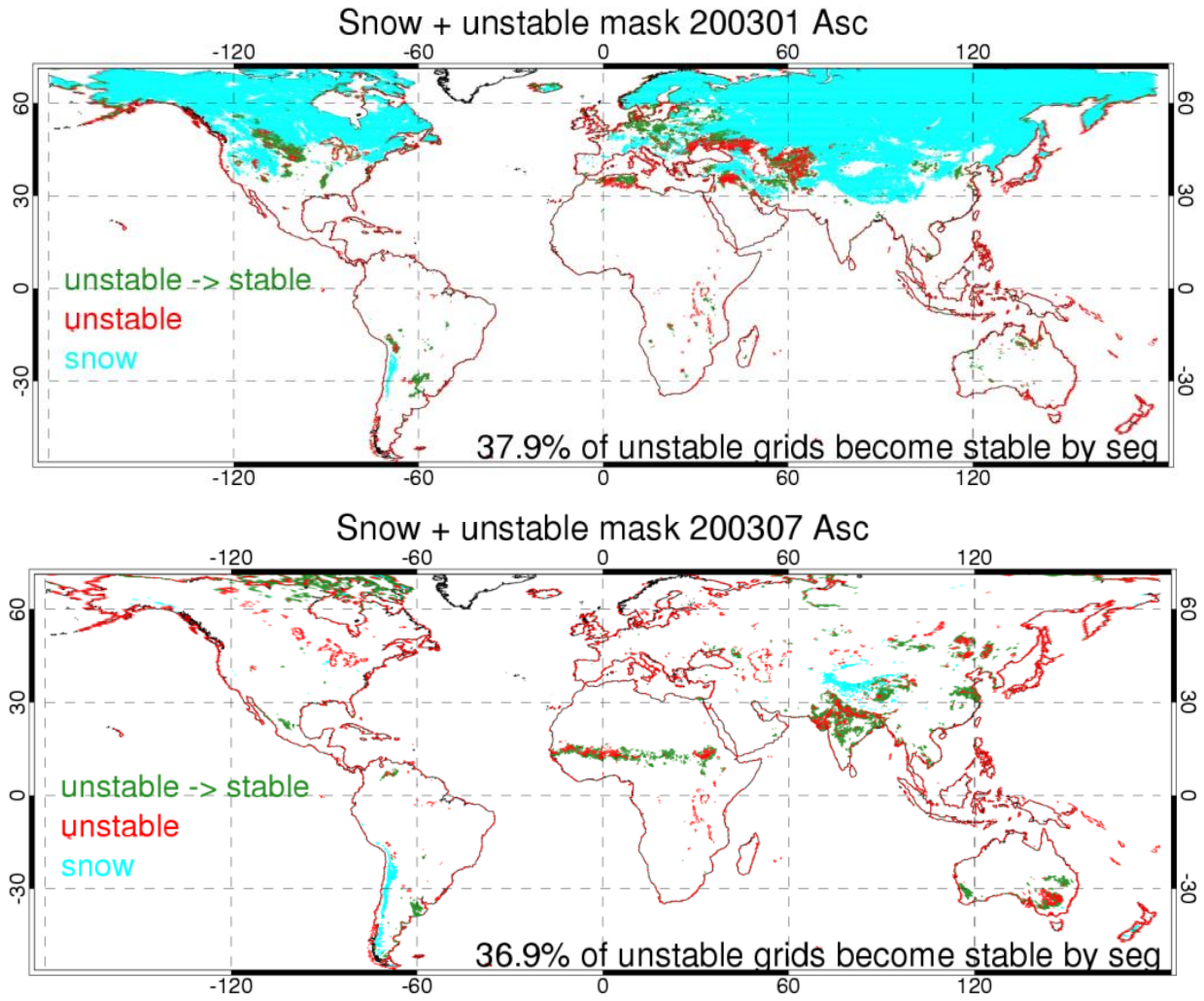


**Figure 22:** Fraction of grid points in each QA category with Version 1 MTD approach (top – reproduced from Figure 1) and with new RTS approach (bottom) for year 2003. Color code used is: green: NPA (1a), yellow: C, orange: PTS (1b), red: unstable, and gray: uncertain (only one valid emissivity estimate for either day or night but not both). Fraction of snow covered areas is in blue. . In this figure, missing values (Category 3) make up the difference between the total fraction of land points shown in color and 100%. Total number of grid points is 199,525.

In highly inhomogeneous areas (such as coastal areas) or areas where day-to-day variability is high throughout an entire month, no improvement is expected with the RTS time segmentation scheme. However, in other regions where R11 changes gradually throughout the month (due to e.g. vegetation growth) this scheme has the ability to produce two or more segments (sometimes extending beyond the first or last day of the month) in order to better capture the R11 trends, thereby lowering the R11 standard deviation in each segment (or portion of segment) within that month. Similar result is produced when a transition from stable to unstable (due to e.g. a series of precipitation events) occurs in the middle of a month. As seen in Table 3, between ~30 and 48% of the grid points that were originally classified as unstable in any given month now become stable over at least a fraction of the month when the RTS scheme is applied. In ~50% of the cases this stability period covers more than 15 days. The geographical distribution of those points is shown in Figure 23 for January and July of 2003. Note that a valid emissivity estimates cannot be produced for all the points that are newly classified as stable. This commonly occurs when the retrieval algorithm has to operate on short (of the order of a week) time periods and cloudiness is such that the number of clear observations during that period is insufficient for producing an accurate emissivity estimate. As can be inferred from Table 3:, ~15% of those points on average (maximum 30% in July) fall into this category and are classified as either uncertain or missing. The rest of those points are assigned either a 1a or 1b emissivity.

Month	%Unstable grid points (MTD scheme)	Stable grid points (in % of Column 2)	New QA categories (in % of Column 3)			Average number of stable days	
			1a	1b	Uncertain	Asc	Des
Jan	9.2	37.9	30.7	46	8.9	19.3	17.3
Feb	8.2	35	28.6	61.4	6.9	18.7	16.9
Mar	8.9	39.8	23.6	73.0	2.4	21.7	19.6
Apr	10.4	43.1	17.1	81.4	1.2	21.6	18.9
May	11.1	47.8	24.0	72.6	2.8	22.4	20.2
Jun	12.5	44.1	28.7	55.6	10.0	18.8	16.4
Jul	10.3	36.9	32.4	36.9	6.3	20.1	17.5
Aug	8.5	30.1	35.5	44.5	6.7	20.8	18.6
Sep	8.3	33.8	34.3	55.0	5.3	21.4	20.5
Oct	10.0	41.3	24.1	66.9	7.4	18.6	19.5
Nov	9.5	38.5	26.2	54.9	11.3	15.3	14.4
Dec	9.2	34.5	32.3	44.2	10.9	21.5	18.4

**Table 3:** Fraction of samples classified as unstable with Version 1 MTD approach becoming stable with new approach and monthly repartition of those samples among the various QA categories. The last two column of this stable indicates the average number for days within the month for which a valid emissivity has been produced under the new scheme.



**Figure 23:** Geographical distribution of (snow free) grid points previously unstable (with MTD approach) becoming stable and points remaining unstable with the new time segmentation approach

The second positive impact of the RTS approach results from the lengthening of the analysis window in areas that remain relatively stable over periods extending from a few months to the full year. This extension of the analysis period has the effect to increase the number of clear observations at some given location, where the algorithm previously failed to produce a valid emissivity value for individual months, allowing a same emissivity value to be produced and assigned to all the months within that window. In particular, between 80 and 92% (depending on the time of the year) of those stable points that were classified as uncertain now become either 1a or 1b (Table 4:). In this table and subsequent charts, Greenland and Antarctica are excluded from our statistics and all global percentages are based on the 178,308 remaining grid points. According to that same table, a significant fraction (~55% during the day and 15% at night) of areas where C-algorithm was invoked are now assigned a 1a value. Note that a small fraction (typically ~15% or less) are assigned a 1b value. Similarly, between 15 and 28% (depending on the time of the year) of locations that were formerly classified as 1a are now 1b (Table 4:). This is explained in some cases by the fact that 1a algorithm noise is causing the day time retrieved emissivity value to be larger than the nighttime value over mildly penetrating areas. The impact of switching from 1a to 1b algorithm over these areas should remain limited as the two algorithms produce similar solutions when penetration is small. Other reasons are provided and discussed at length at the end of this section.

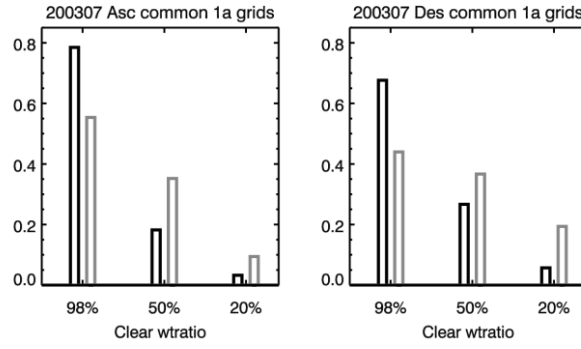
Another positive impact of our new segmentation approach is the increase in the percentage of highest quality ( $F_{clear} \geq 98\%$ ) estimates among the 1a grid points (Table 5). This impact is also accompanied by a decrease in the fraction of lower quality 1a estimates (see also Figure 24).

	1b -> 1a (%)	1a -> 1b (%)		C -> 1a (%)		C -> 1b (%)		Uncertain -> 1a (%)	Uncertain -> 1b (%)
		Asc	Des	Asc	Des	Asc	Des		
Jan	4.8	19.8	20.5	52.9	10.5	13.4	12.8	21.9	57.6
Feb	4.7	21.0	21.7	53.2	11.5	15.6	14.8	23.4	64.1
Mar	6.4	25.8	26.2	55.8	11.3	15.7	15.2	35.1	56.8
Apr	7.8	28.7	27.6	57.3	13.0	15.5	17.5	41.7	47.1
May	9.2	21.3	17.4	59.1	16.6	11.1	55.6	51.9	37.7
Jun	8.3	18.1	18.4	56.0	15.5	8.9	10.0	31.2	56.6
Jul	6.6	15.8	16.5	51.4	14.0	10.2	8.4	46.3	42.3
Aug	6.2	15.1	15.7	53.4	15.7	10.0	8.4	43.8	41.9
Sep	7.2	15.7	16.3	58.2	14.7	8.6	3.5	44.8	46.5
Oct	9.1	18.0	16.6	61.1	2.8	11.4	53.1	57.0	35.4
Nov	5.7	21.0	20.0	61.4	15.0	7.8	10.1	34.2	45.8
Dec	6.6	18.3	18.7	55.9	12.4	8.4	8.7	27.3	52.8

**Table 4:** Percentage of grid stable points classified as 1a, 1b, C or uncertain with MTD time division being assigned a 1a or 1b value with new time segmentation scheme

	$F_{clear} \geq 98\%$				$98\% > F_{clear} \geq 50\%$				$50\% > F_{clear} \geq 20\%$			
	Asc		Des		Asc		Des		Asc		Des	
	MTD	RTS	MTD	RTS	MTD	RTS	MTD	RTS	MTD	RTS	MTD	RTS
Jan	34.9	55.8	21.0	39.6	46.3	36.0	45.7	46.1	18.8	8.2	33.4	14.3
Feb	31.6	55.8	18.1	40.7	49.7	38.1	43.3	46.9	18.7	6.1	38.6	12.4
Mar	27.2	55.9	17.9	41.3	52.3	39.8	46.2	47.7	20.6	4.4	35.9	11.0
Apr	26.2	58.9	16.7	40.1	55.7	37.6	46.7	50.9	18.1	3.5	36.6	9.0
May	38.1	69.3	18.4	42.5	52.0	28.6	44.0	52.2	9.9	2.1	37.6	5.4
Jun	48.8	75.6	38.7	64.3	41.9	22.1	38.7	30.1	9.4	2.2	22.7	5.6
Jul	55.4	78.6	43.9	67.6	35.2	18.2	36.7	26.7	9.4	3.3	19.4	5.7
Aug	57.4	77.8	37.8	65.8	32.3	18.7	42.2	28.1	10.2	3.5	20.0	6.2
Sep	53.7	76.9	31.9	62.3	35.8	20.7	43.2	31.8	10.5	2.5	24.8	6.0
Oct	37.2	72.6	20.8	50.8	51.9	25.6	40.5	43.4	11.0	1.8	38.7	5.8
Nov	29.8	62.3	16.1	47.8	54.2	34.5	48.7	43.9	15.9	3.3	35.2	8.3
Dec	36.9	54.7	21.4	40.8	50.5	40.3	45.0	44.8	12.7	5.0	33.7	14.4

**Table 5:** Percentage of 1a grid points with  $F_{clear} \geq 98\%$ ,  $< 98\%$  and  $\geq 50\%$ , and  $< 50\%$  and  $\geq 20\%$ , with MTD and RTS time segmentation



**Figure 24:** Fraction of 1a locations with  $F_{clear} \geq 98\%$ ,  $50\% \geq F_{clear} < 98\%$  and  $20\% \geq F_{clear} < 50\%$  with MTD (gray) and RTS (black) methods for July 2003.

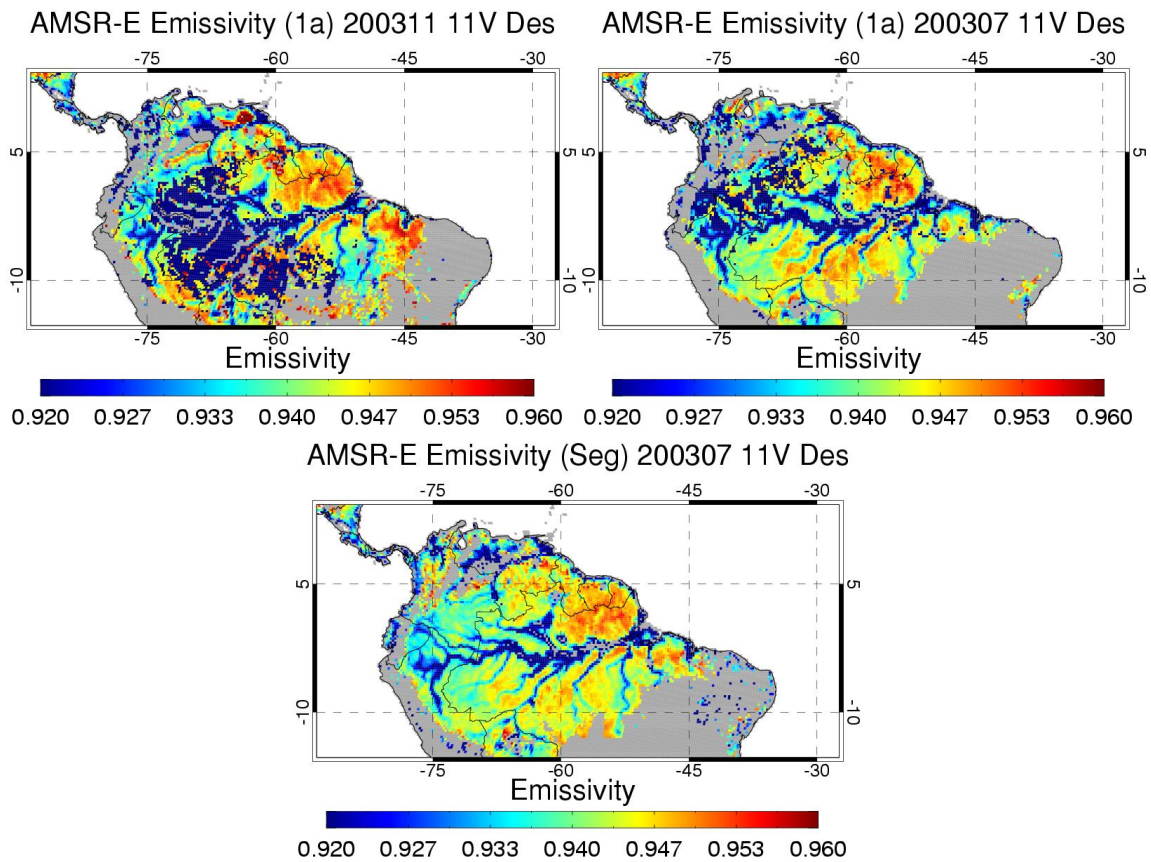
Maps comparing the geographical distribution of the QA categories listed in Table 1 obtained with the MTD and RTS approaches for some regions of interest (Amazon, African ITCZ and Sahara desert) are shown in Figure 26 - Figure 29. Global maps for the months of January and July are shown in Figure 30 through Figure 33.

In Fig. Figure 26, the increase in highest-quality 1a emissivity yield over the Amazon basin in the daytime with the RTS approach is quite striking. This region is particularly cloudy in January. In the Version 1 database, C emissivities are used over almost the entire region, with some low quality 1a estimates covering the southern edge of the basin, where the C-algorithm failed to provide a substitute (Figure 26). Whether the patch of 1b emissivities produced with the MTD approach in January over the Guyanas actually corresponds to an area of significant microwave penetration is questionable. Note that over a large fraction of the S. American tropical forest, the R11 polarization ratio remains stable through the entire year and most locations are assigned a single yearlong segment by the RTS algorithm. C emissivities with the RTS time segmentation only cover a small fraction of the Amazon region (on the Western and Northern edges of the basin), the rest being represented in large part by 1a  $F_{clear} \geq 98\%$  estimates. Note the obvious cross-talk between rivers and quality of the 1a estimates. Similar artifacts have been observed in other parts of the world and this aspect has not been investigated yet. One potential explanation is a dependence of the MODIS LST quality flags on surface background, either inherent to the LST algorithm itself or through the MODIS cloud mask, but this hypothesis has not been verified by the MODIS LST team (Wan, *personal communication*). Figure 25: shows examples of nighttime retrieved 1a emissivities with the monthly segmentation scheme over the Amazon basin for the months of July and November 2003 and corresponding emissivities retrieved using the new time segmentation approach. The emissivities produced with the monthly segmentation appear to be quite noisy spatially with large areas where they reach abnormally low values. There is also significant month-to-month variability. The new RTS emissivities are estimated using the clearest AMSR-E observations available through the year. With the RTS approach spatial variability appears to be reduced and consistency between nighttime and daytime emissivities (not shown) is also improved. Cloudiness is less extensive in the daytime and daytime emissivities are less likely to be impacted by CLW (or cloud contaminated infrared LST estimates). Note that with RTS we still revert to classification over most of this region for the nighttime emissivities but it is expected that the substituted emissivities are of better quality than those obtained with the monthly segmentation.

It is important at this stage to note that we do not rule out the fact that the new emissivities in the Amazon region might be overly stable in time and mask potential impact of e.g. leaf wetness on the surface microwave emissivity during the rainy season. The 11 GHz polarization ratio would not necessarily be impacted by such phenomenon. The clearest observations are obviously obtained during



the driest months and if such seasonal emissivity cycle existed it would not be captured by our current scheme. We provide further discussion of this aspect in Sec. 7.1. Note that if leaf wetness has indeed an impact on emissivities, whether or not we should correct for it depends on the application of the database. For instance, for surface temperature estimation a 0.005 bias in emissivity would incur a bias of 1.5K in LST which is marginally significant. In applications to rain rate retrieval, the importance of taking this effect into account is probably even less significant. Obviously, if the objective is to monitor the state of the vegetation we have to develop some ways of directly measuring this effect from satellite sensors, which in the present approach is complicated by the frequent interference of clouds and precipitation. It appears from the discussion provided at the end of this section that the level of noise in our MTD emissivities caused by this interference is probably higher than this signal itself and that even in their current state the RTS emissivities are likely to be an improvement compared to Version 1.



**Figure 25:** Nighttime (Descending pass) emissivity estimates over S. American tropical forest with MTD in November (top left) and July (top right) 2003, and with RTS (bottom). With the latter time segmentation most areas in this region are assigned the same emissivity throughout the year.

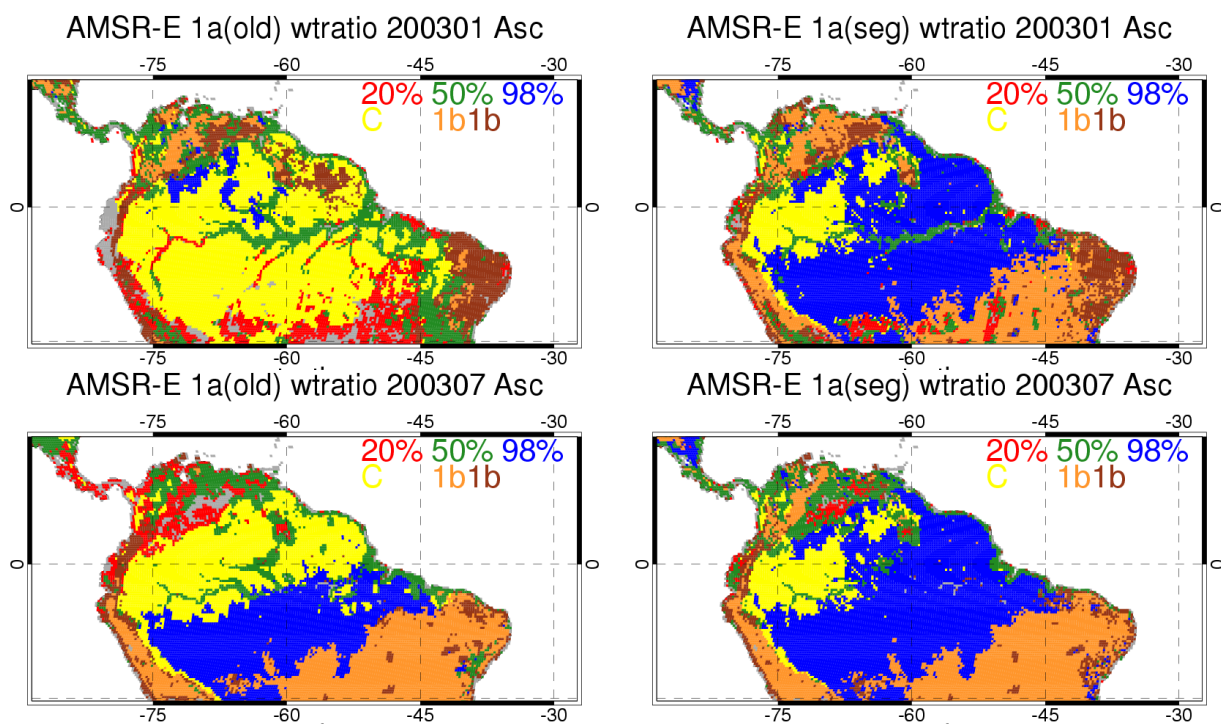
The impact of the RTS method over the African rain forest (Fig. Figure 27) is similar to what is observed over the Amazon region (i.e. increased percentage of high quality 1a emissivities and sharp reduction in C-emissivity coverage in the daytime). It is apparent in Fig. Figure 27 that the quality of the 1b estimates over Saharan regions such as Mauritania or over the African Horn improves as well, mainly because the RTS approach provides the flexibility of being able to search over longer time

periods for the set of observations that will produce the best quality emissivity estimates. The Sahel region around 15°N during the African monsoon remains a notoriously difficult region. This region is impacted by rapid vegetation growth at the onset of the monsoon, frequent cloudiness and precipitation. At the northern edge of this region, two vegetation cycles are observed, the rapid growth and senescence cycle of grass land at the time of the monsoon and a slower seasonal cycle of shrubs and trees (not shown). Analysis of surface emissivity in those regions is complicated by the fact that microwave penetration is high and both emissivity and thermal structure of the soil change on short time scales in a correlated fashion, making it difficult to separate the two effects. Moreover, cloudiness drives the temporal sampling of the observations, which when conditions are unstable can make mean nighttime and daytime emissivities appear like they are significantly different. This is one reason for screening out large isolated outliers based on analysis of the R11 time series. It is seen in Fig. Figure 27 that many of the areas affected by the African monsoon where Version 1 fails to assign an emissivity value are now covered. However the emissivity estimates provided with the RTS approach in those regions remain uncertain. This area deserves further attention and some approach for bridging spatial gaps and improving the quality of our product must be devised.

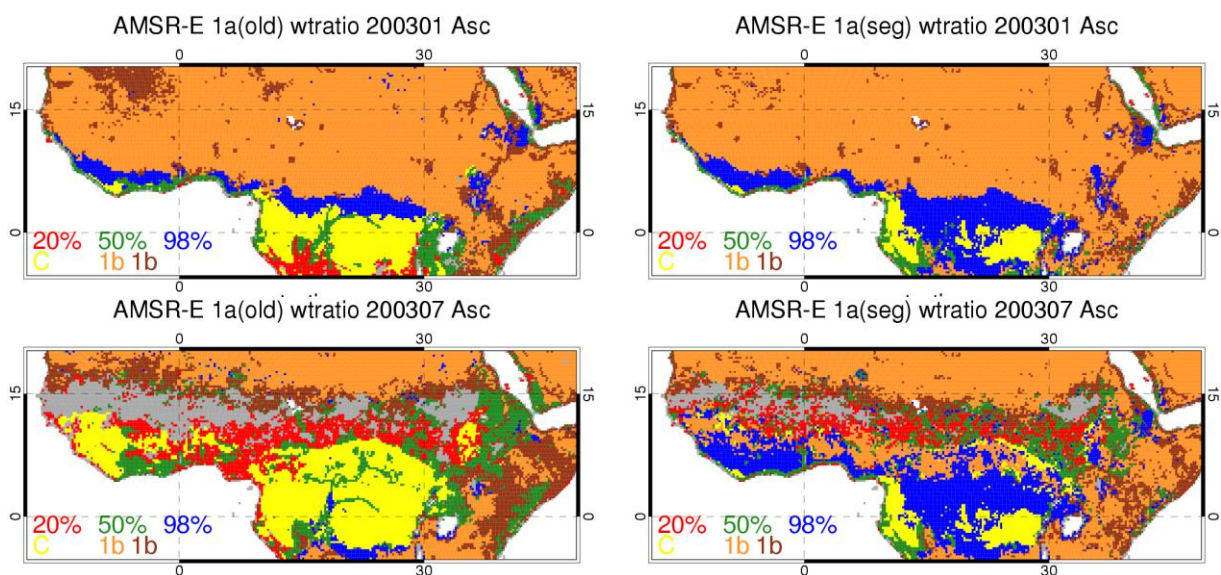
The benefits of the RTS method over the Sahara during the fall and winter season, when cloudiness and rain become more frequent in the Northern and Western parts of this region, are illustrated in Figs. Figure 28 and Figure 29. Over bare soil, the R11 polarization ratio is a good indicator of changes in surface state due to e.g. precipitation. Vegetation cycle is also easily observed in the R11 time series over semi-arid areas (e.g. Sec. 7.4). It is our experience so far that over deserts, we can trust that emissivity is stable in time when no fluctuation in R11 is detected. Note that the RTS method also provides increased coverage during the summer months when dust storms are more frequent. Dust can cause a black out in infrared data where it is the thickest. In other areas, it can cause erroneous surface temperature or atmospheric retrievals if left undetected (Sec. 7.3). We have no evidence so far of an impact of dust on microwave radiation or on the surface emissivity itself. In this case, emissivity estimates produced for a given time segment should remain valid at all times including during sand storm outbreaks.

The global plots of

Figure 30 through Figure 33 highlight other interesting features in various regions of the globe. Overall the 1b regions appear “cleaner” with the new RTS segmentation, with much fewer areas where our nominal 1b accuracy threshold is not met. Note that these maps do not show the breakdown of the 1a emissivity according to  $F_{clear}$ . As mentioned previously, a significant fraction of locations which were classified as 1a (or uncertain) with MTD are now classified as 1b with the RTS method. This is particularly obvious for the month of January in Brazil and in Africa between ~5° and 20° S, and also in North Africa, Central America and in the Sahel in July. It is also apparent in those figures that the coverage of the C-emissivities remains larger at night than in the daytime in tropical regions. We argued that the substituted emissivities might be of better quality with the RTS approach. In certain regions there may be alternate means of avoiding nighttime substitution that need to be evaluated. It also appears that many of the areas in Eastern Europe that were formerly unstable are now stable while some instabilities (probably associated with snow) are still present in the former Soviet Republics in the Caucasian and Central Asian regions. The Eastern US in the winter season is another region that remains difficult to model. This particular region is currently the focus of a detailed validation study (Sec.7.1) that we will extend in the future to the winter months.

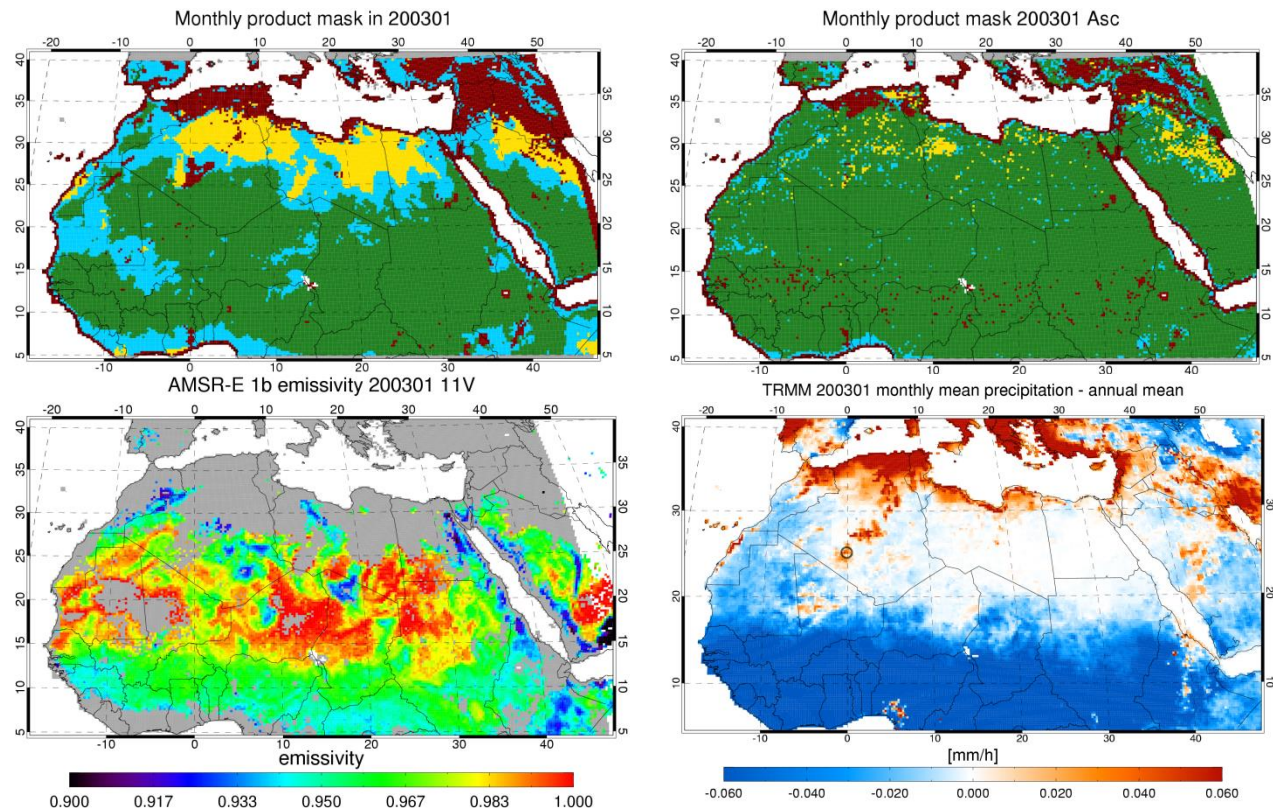


**Figure 26:** Distribution of QA categories (see Table 1) with MTD (left) and RTS (right) approaches in the daytime for January 2003 (top) and July 2003 (bottom) over northern part of South America. Color scheme used in this figure is indicated on the plots.

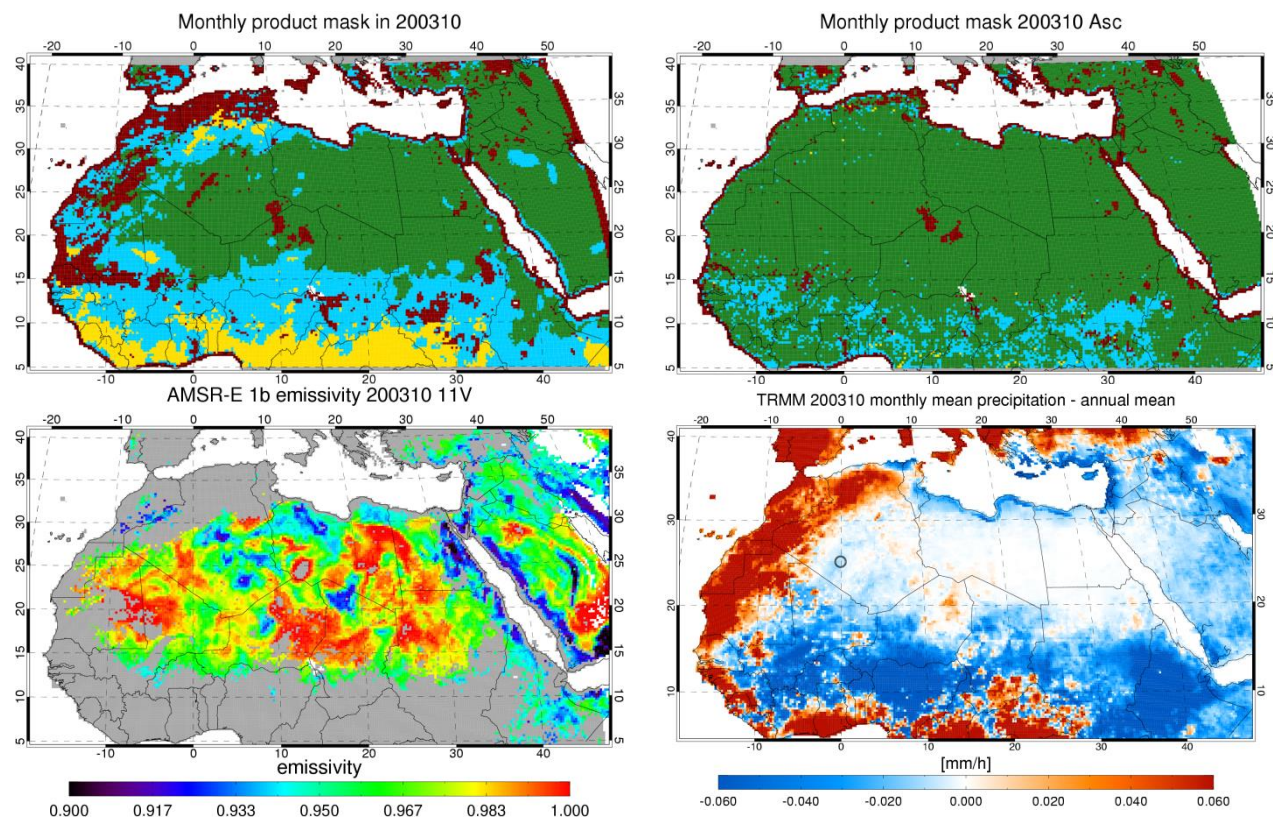


**Figure 27:** Same as Figure 26 over central part of Africa



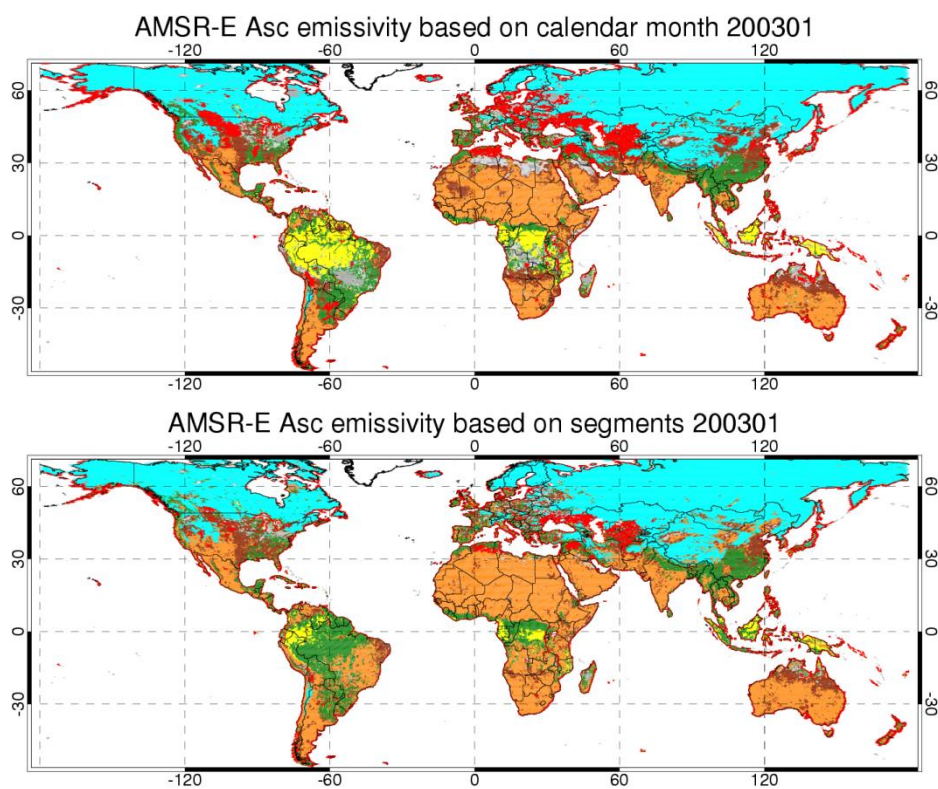


**Figure 28:** Yield estimation for MTD (top left) and RTS (top right) time segmentation approaches for January 2003. In green are the areas where the algorithm is likely to produce an emissivity estimate for the month based on frequency of clouds and temporal stability of surface properties. Blue and yellow colors indicate less favorable and unfavorable cloudiness conditions, respectively, leading to frequent cloud contamination of MODIS LST or irregular surface thermal forcing. In red are the areas where surface changes rapidly (generally correlated with TRMM precipitation – bottom right panel). The criteria applied to generate the top row panels are somewhat arbitrary. The bottom left panel (added for verification purpose) indicates a good spatial correlation between areas in green in the top left panel and areas for which the MTD 1b emissivity retrieval algorithm produces the highest quality output. In the bottom left panel, areas in grey are areas for which no reliable Version 1 monthly 1b emissivity estimate could be produced.

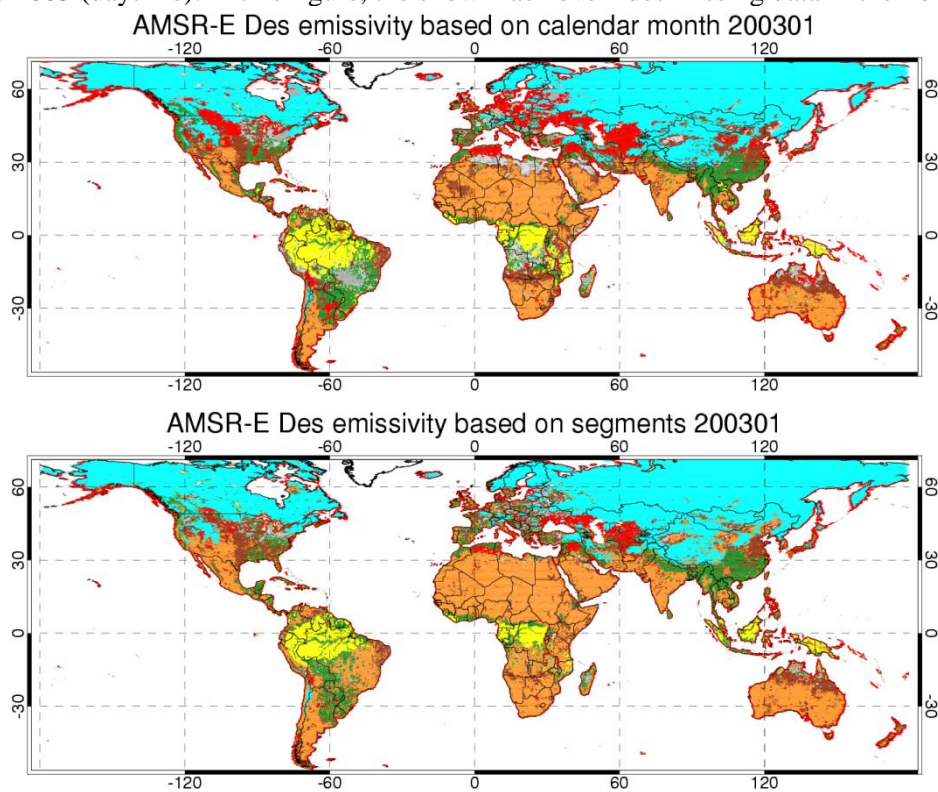


**Figure 29:** Same as Figure 28 for October 2003. Note that in the masked area south of  $\sim 10^\circ$  N, surface is highly vegetated and emissivities retrieved from the 1a algorithm (non penetrating surface) are used instead of the 1b algorithm.

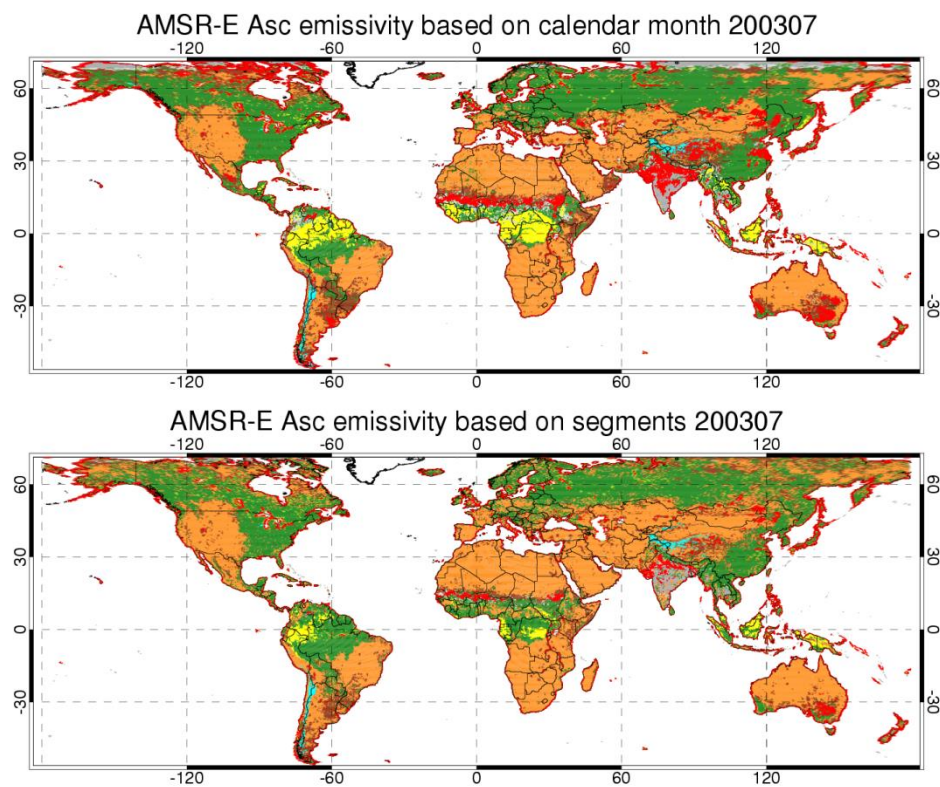




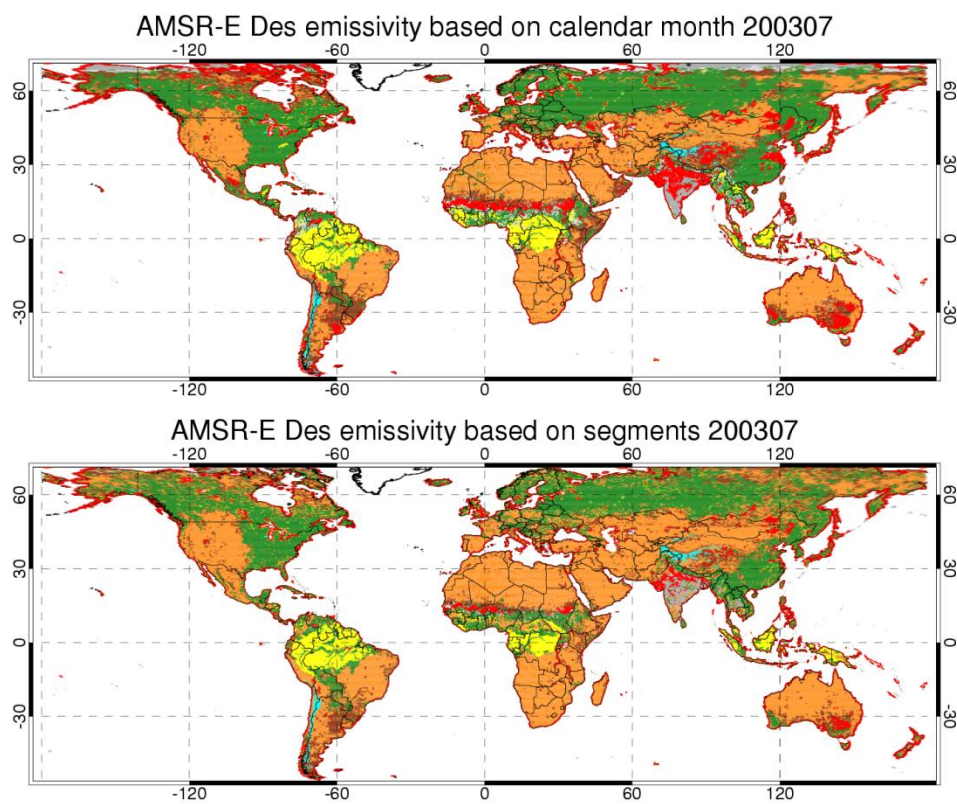
**Figure 30:** Global distribution of QA categories (see Table 1) with MTD (top) and RTS (bottom) approaches for January 2003 (daytime). In this figure, the snow mask overrides missing data in the Polar Regions.



**Figure 31:** Same as  
Figure 30 for January 2003 (nighttime)

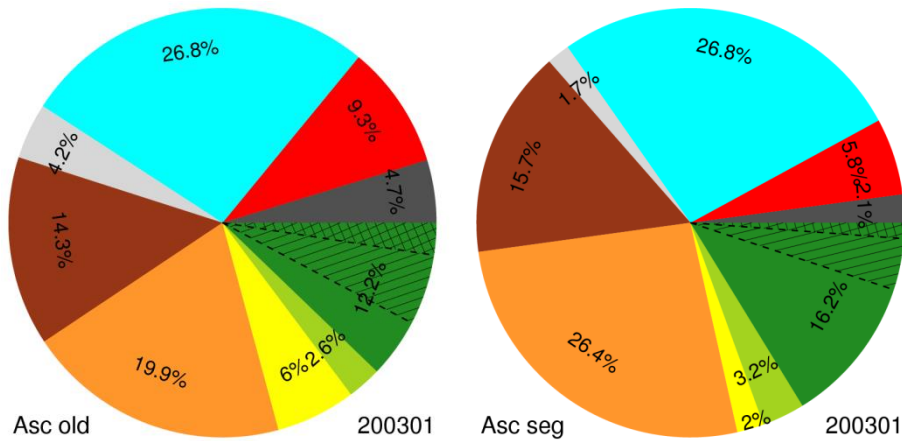


**Figure 32:** Same as  
Figure 30 for July 2003 (daytime)

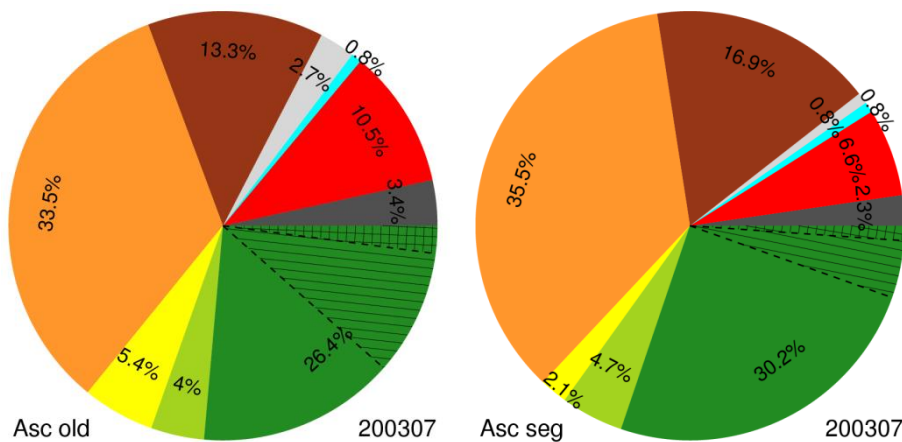


**Figure 33:** Same as  
Figure 30 for July 2003 (nighttime)

Figure 34 and Figure 35 provide a summary of the net breakdown of the 1a and 1b emissivities into different quality levels for the months of January and July 2003 for the MTD and RTS methods. The main impacts of the RTS approach, highlighted in the previous discussion, are also visible in those charts. These include a decrease in fraction of unstable or missing grid points accompanied by an increase in the number of locations assigned a 1a or 1b emissivity, as well as an overall improvement in the quality of the 1b emissivities (fewer category 6 locations) and in the quality of emissivities produced in non penetrating areas (fewer C and category 8 locations and a net shift in the remaining 1a locations toward  $F_{clear} > 98\%$  estimates). Interpretation of those charts is complicated by the fact that there are significant exchanges among the different categories (Fig. Figure 36). For instance, some of the locations that were formally unstable or missing now populate the uncertain or lower quality 1a (Category 6) or 1b (Category 8) classes making the improvements in the quality of estimates that were formerly 1a or 1b and remain 1a or 1b brought by the RTS method appear overall less significant.

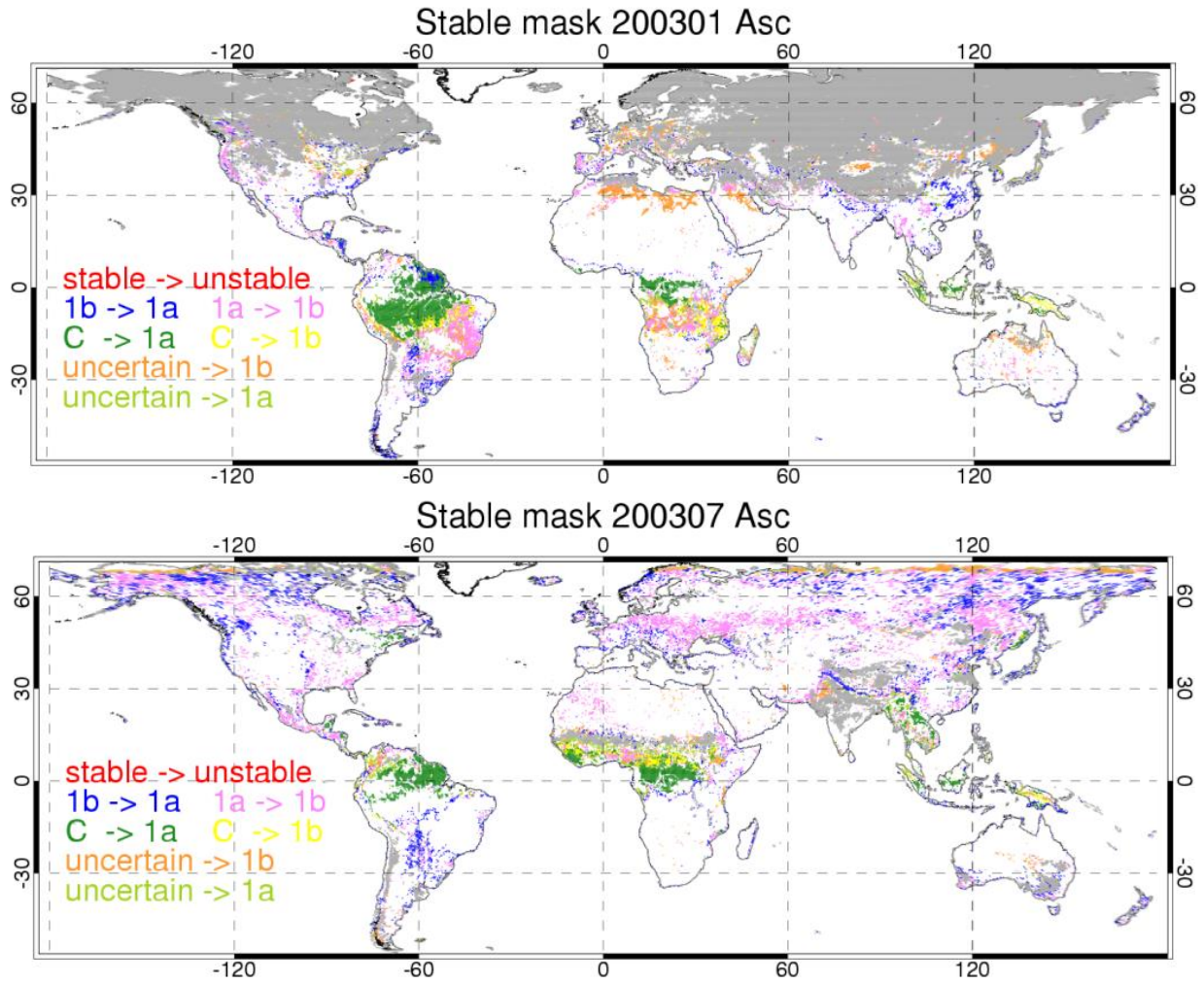


**Figure 34:** % of grid cells in each QA category with MTD (left) and RTS (right) time segmentation (daytime only). Here, the shading indicates the degree of cloud contamination –  $F_{clear} \geq 98\%$ , 50% (hatched) or 20% (cross-hatched) in the samples used to estimate. Percentages are based on the total number of grids 178,308 (excluding Greenland and Antarctica)



**Figure 35:** Same as Figure 34 for the month of July 2003





**Figure 36:** Maps highlighting the areas where changes in QA category occur with the new RTS time segmentation.

As we previously mentioned, there are a number of areas that were formerly treated as 1a, C or uncertain with the MTD segmentation and that now fall into the 1b category. In some instances the impact of cloud contamination on the retrieved instantaneous emissivities may have caused a) the monthly mean day/night emissivity differences to be of the incorrect sign (day/night emissivity differences are used for flagging instances where penetration is significant) or b) a failure to produce either a daytime or nighttime (or both) mean 1a emissivity estimate (C emissivity if available is used instead), in which cases penetration is ignored by default. In other cases, a particular location may be newly categorized as 1b with the RTS approach when the response of the state of the vegetation to e.g. precipitation cycle is too weak to be reflected in the R11 data. A change in precipitation regime may be occurring in the middle of a RTS segment spanning several months and emissivities produced from the clearest observations (hence during the drier portion of the segment) apply to the wetter portion as well.

In the three cases shown in Figure 37 (over Africa) the retrieved RTS emissivities and Version 1 MTD emissivities appear to agree reasonably well during the drier months when penetration is significant (negative day/night 1a emissivity differences). During the wetter season (shaded areas), even though

there is sometimes some hint that the sign of the daytime and nighttime emissivity differences might reverse itself (indicating a decrease in penetration effects), instantaneous emissivity retrievals may be too noisy to be able to determine which estimate (when both MTD and RTS are available) is most realistic. This is the situation encountered for cases (a) and (b) from January to April and in November and December. The RTS emissivity estimate produced for first part of the year appears to be on average no worse than the MTD estimates (note that nighttime MTD emissivities are often missing and the larger cloudy daytime emissivity values during this period may be caused by cloud contaminated MODIS LSTs – the temporal trends in those emissivities may be artificial and may not reflect the actual surface state). In November and December one can infer that the validity of the RTS emissivity in case (a) is more questionable but this situation could be flagged using e.g. NDVI information. The R11 standard deviation also increases during those months – most likely reflecting changes in soil wetness (which might also trigger the occasional low instantaneous emissivity values). In case (c), there is also good agreement between the MTD and RTS emissivities during the dry season (the MTD nighttime emissivity for the month of March looks suspect). Monthly night time and daytime 1a emissivities also appear to be close to the newly retrieved 1a emissivities in April and May. In June through October, MTD daytime emissivities, when available, are close to the night time values and both are within 0.005 of the RTS 1b emissivity. In this case, class-based substitution is invoked from April to October but the C-emissivities seem to have too much temporal variability whereas the RTS 1b emissivity – even if incorrect – seems to better represent the average instantaneous 1a emissivities.

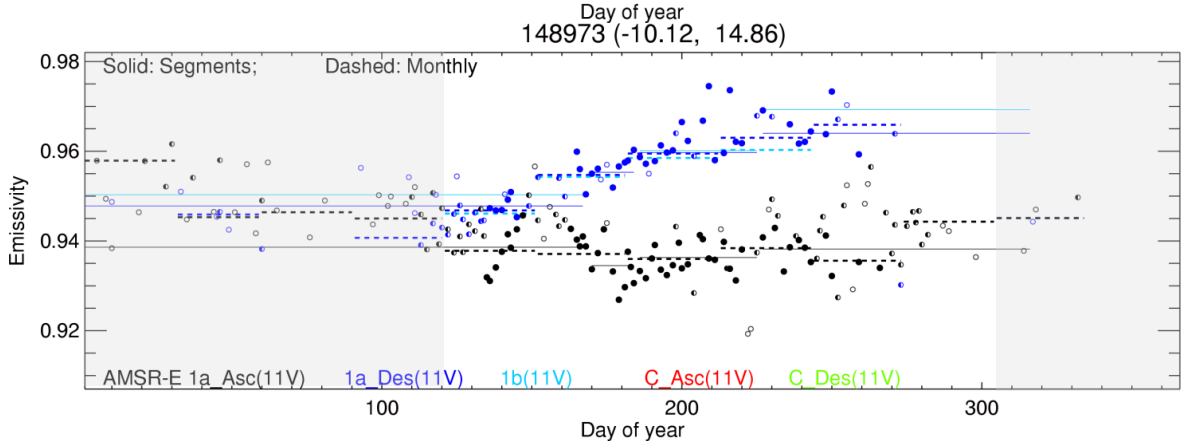
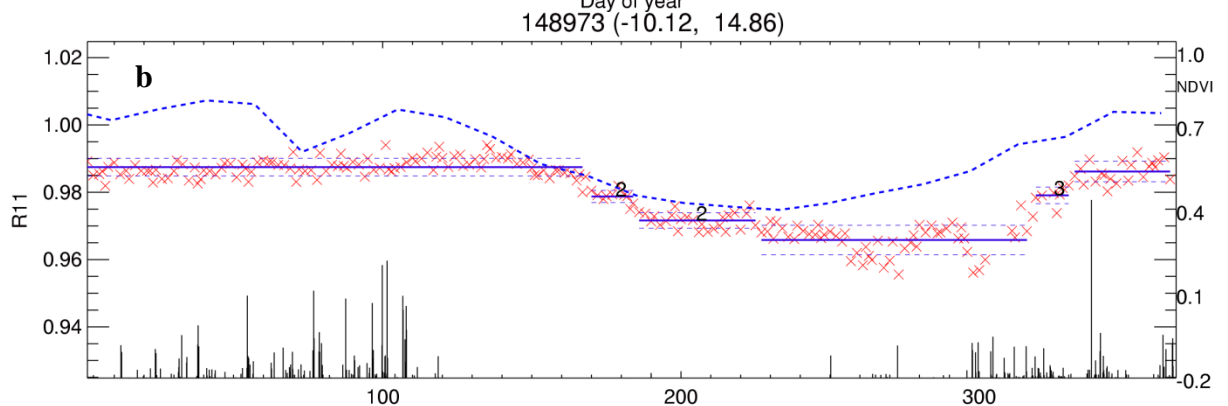
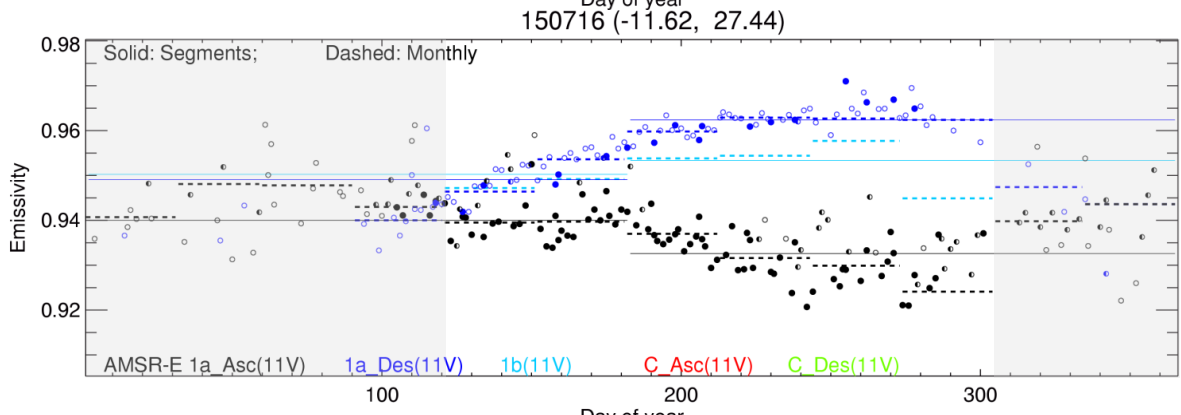
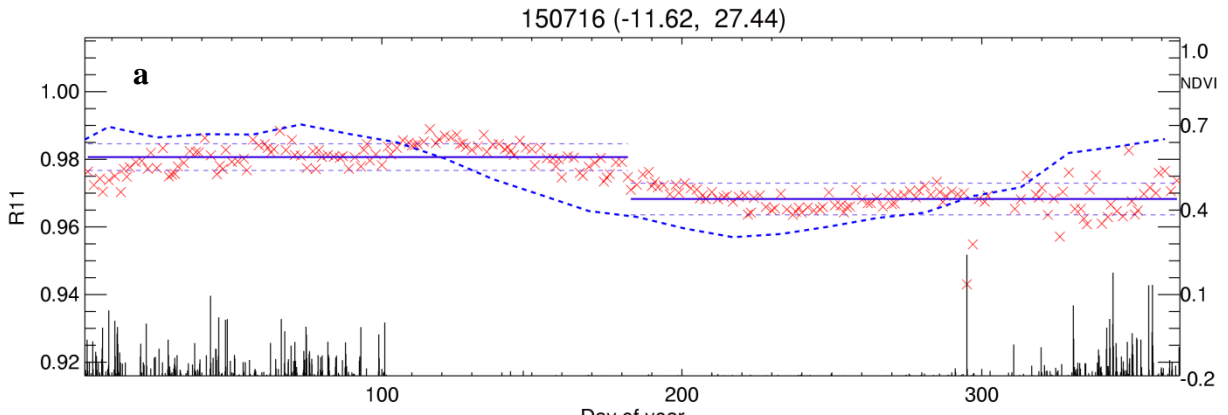
The Amazon cases shown in Figure 38 are particularly interesting as there is a clear trend in both the daytime and nighttime instantaneous emissivities (indicating an increase in the magnitude of impact of penetration) through the driest season, from April to September, that appears to reverse itself starting in September/October at the onset of the wet season. These trends seem to follow the NDVI variations and are not captured by the R11 based segmentation algorithm. The R11 temporal pattern in both cases is quite similar even though the two sites are ~800 km apart. R11 remains essentially flat through the year with a small magnitude dip below its mean value during the dry season (the reason for the artifact around days 270 - in case a - and 250 - in case b - is unknown but may be related to soil wetness). Several observations can be made:

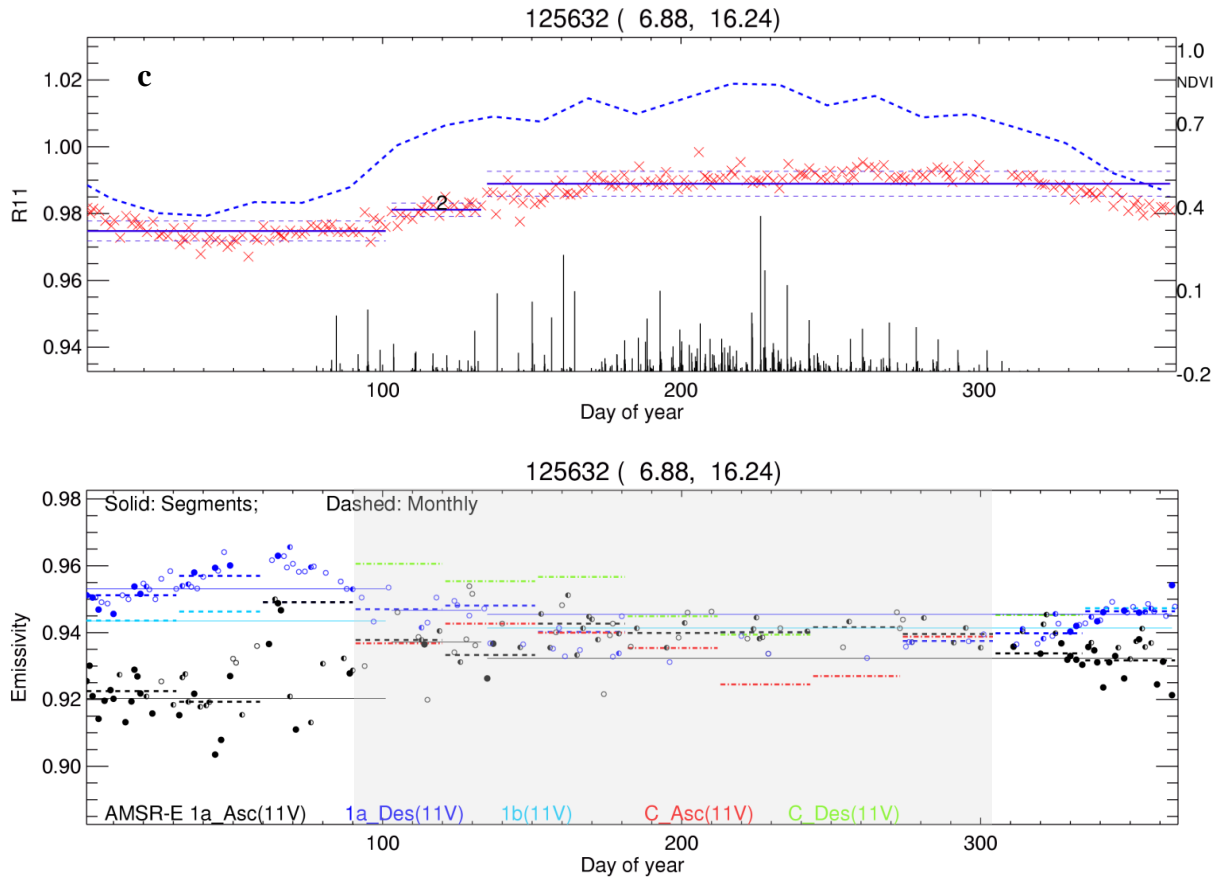
- 1) There are very few instantaneous emissivity estimates during the wet season and some of these estimates appear quite noisy (especially during the daytime where there is a clear impact of cloud contamination) which makes the quality of the MTD 1a estimates, when available, often questionable. Version1 often reverts to C-substitution for those months.
- 2) From comparison of MTD 1a and C emissivities when both estimates are available during the drier months, the C emissivities appear sometimes off by 0.01 or more and may - especially in case b - show more variability than there probably is in actuality through the wet season.
- 3) The RTS scheme in its current form is not sensitive enough to small changes in R11 in vegetated areas. This problem can easily be fixed in the dry season by analyzing the trends in instantaneous emissivity and further subdivide the initial segment to reflect those trends or by including NDVI information at the segmentation stage. In the wet season such an enhancement would not be of much help since we do not have enough clear observations to provide an emissivity estimate in the first place. The algorithm in its current form would revert to C emissivities.
- 4) During the wet season in case (a), RTS emissivities seem to agree reasonably well with the C-emissivities (differences are within 0.005). In case (b) no 1a estimate is produced with the



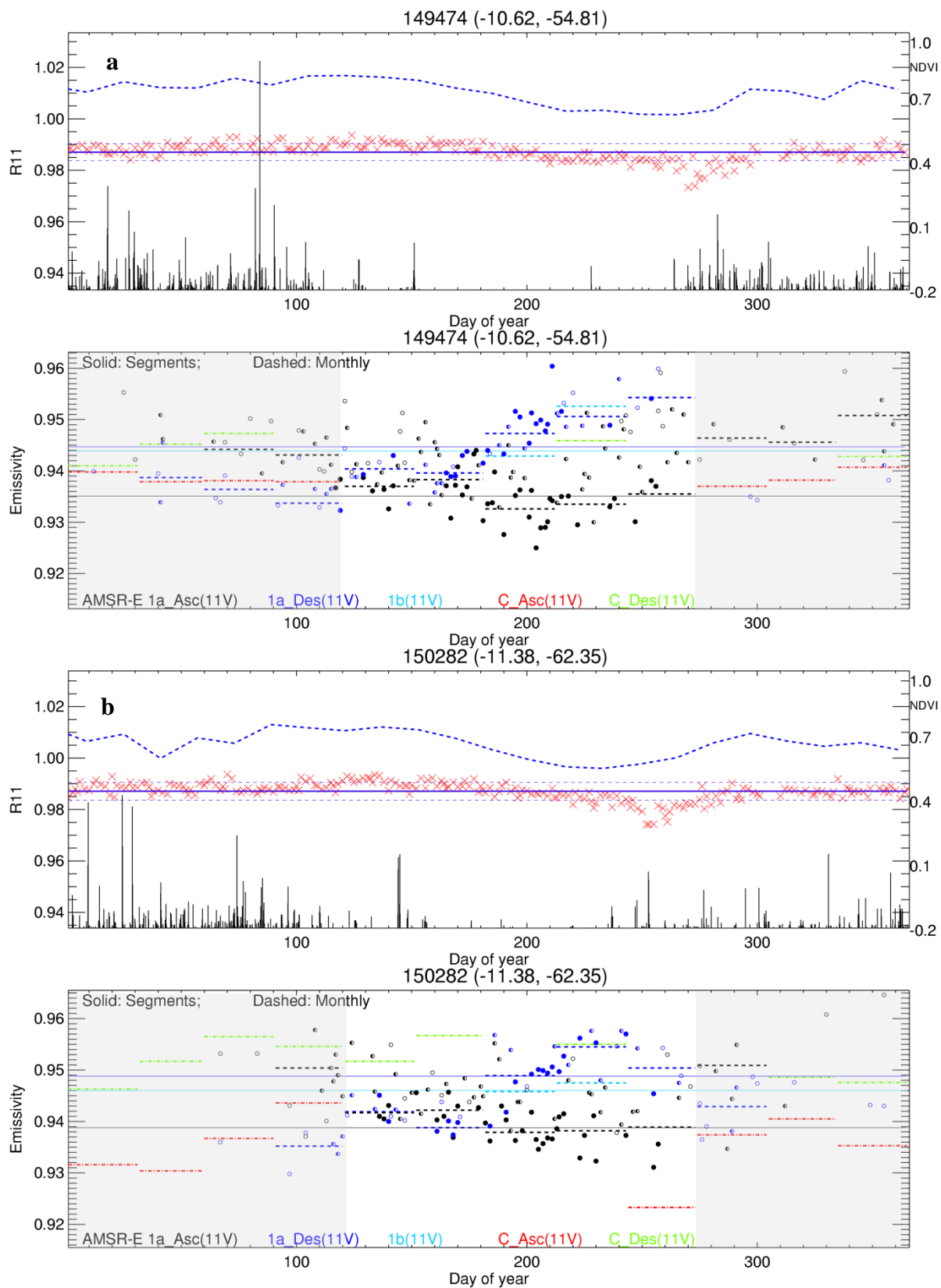
MTD approach and the C-estimates have large oscillations around the RTS value. In both cases the RTS values are probably no worse than current C emissivities (given the level of uncertainty on those estimates).

It is noteworthy that in many of the cases we examined in this study we are looking at relatively weak signals, i.e. differences between mean MTD and RTS emissivities within 0.01, and it appears that 1b estimates produced by the new RTS algorithm, despite the weaknesses identified in its current implementation, are overall no worse or even better than the Version 1 emissivities produced during the cloudy months. However, differences of the magnitude shown in the above examples are still significant for e.g. LST retrievals. Comparing the results of the RTS approach to the Version 1 data provides us with some mechanism for quickly identifying potentially problematic areas so we can focus in follow up work on explaining and resolving observed discrepancies. Based on the above examples, unless classification can be improved by taking other information (such as NDVI) into account, it is probably best for now to assume that the emissivity during the wet season is identical to the emissivity estimated e.g. at the end of that season. Of course, any impact of changes in the state of the vegetation would not be reflected in these estimates but we can for instance use NDVI for quality control. One might be able to verify the soundness of this approximation by comparing LST estimates derived from those emissivities to in situ data, as outlined in Sec. 7.1. A second issue is tied to the handling of surface penetration effects. As described in 7.3, the microwave 1DVAR algorithm used in the retrieval of surface temperature is configured differently depending on whether we are in 1a or 1b regime. So far we only have two decision levels, i.e. penetration or no penetration, regardless of which frequency might be affected. In the case of mild penetration, as illustrated above, we cannot ignore the impact at 11 GHz (as this would translate into biases of as much as 3K in the peak of the dry season, i.e.  $\sim 0.01$  in emissivity). However, we have the option of dropping the 11 GHz channel and operating in 1a mode using the higher frequencies (19 GHz and above). Surface emissivity may also be more stable at 19 GHz than at 11 GHz, which would partially alleviate concerns with the approach outlined above for supplying an emissivity estimate during the wet season.





**Figure 37:** Time series of emissivities retrieved with new RTS scheme (lower panel - solid lines) versus monthly Version 1 emissivities (dashed) for 3 locations in Angola (a), Democratic Republic of Congo (b) and Central African Republic (c) in year 2003. Instantaneous daytime and nighttime NPA (1a) emissivities are represented by black and blue circles, respectively. Filled circles represent mostly clear samples, half filled circles, partly cloudy and open symbols, mostly cloudy. Time series of R11 (red crosses) and NDVI (dashed line) are shown in the upper panel. The results of the automated R11 segmentation are superposed to the R11 time series. Daily normalized TRMM precipitation from the closest reporting grid box is shown by the vertical lines. Gray shading in emissivity plots is used to indicate the rainy months, during which instantaneous emissivity retrievals are sparse or missing.



**Figure 38:** Same as Figure 37 for two locations in Mato Grosso (a) and Rondônia (b) states (Brazil) over deforested areas on the Southern edge of the Amazon basin

## 6. IMPROVEMENTS IN QUALITY OF AMSR-E EMISSIVITY ESTIMATES

### 6.1. Addition of Terra/MODIS for emissivity estimates in penetrating areas

The effective emission depth parameter retrieval algorithm (Galantowicz *et al.*, 2011) or 1b algorithm uses a thermal model that relates the time-of-day variation of subsurface temperatures to that of the surface. The model includes 5 parameters that describe the surface temperature variation: phase ( $\phi_n$ ) and amplitude ( $A_n$ ) parameters for two sinusoidal functions with 24- and 12-hour periods, and a diurnal mean temperature parameter ( $T_0$ ):

$$T_s(t) = T(0, t) = T_0 + \sum_n A_n \cos(n\omega_0 t + \phi_n) \quad (8)$$

The baseline solution uses two datasets to derive these parameters in two steps. First, SSM/I and AMSR-E 87/89 GHz V-pol. Earth-leaving Tb's, aggregated over a period of days, are used to derive the two surface sinusoid phase parameters. Second, Aqua MODIS LST is used to solve for the amplitude and mean diurnal temperature parameters, with the phase parameters held constant.

The 87-89 GHz data are useful for phase information because: (a) 87-89 GHz penetration depth is small so the phase parameters of the effective emission temperature diurnal cycle should be close to those of the surface temperature and (b) there are six sampling times per day from the combination of two SSM/I sensors and AMSR-E. However, several factors may introduce errors, including microwave emissivities less than one, atmospheric contamination, and non-zero emission depths at 87-89 GHz. Therefore, the purpose of this analysis is to investigate whether the combination of Terra and Aqua MODIS LST time-series can be used to solve for all five surface temperature model parameters simultaneously and, if so, whether the solution produces higher quality products than the baseline method.

A new method was developed to solve for the five surface temperature model parameters using only MODIS Terra and Aqua LST arrayed by time-of-day, with up to four samples per day over a minimum of 10 days of observations. The new method solves for all five surface temperature model parameters simultaneously<sup>1</sup> using the same algorithm as the baseline method, albeit with MODIS LSTs replacing 87-89 GHz Tb's. The methods differ in the subsequent step: the baseline method replaces the amplitude and diurnal mean temperature parameters by fitting a three-parameter model to MODIS-Aqua LSTs while the phase parameters are held constant; the new method skips this step entirely. That is, the revised method derives the surface phase, amplitude, and mean temperature parameters in one step from MODIS Aqua and Terra LSTs, while the baseline method derives them in two steps from 87-89 GHz Tbs and MODIS Aqua LSTs.

One important note is that the MODIS Terra LSTs can be filtered for cloudiness independently of the MODIS Aqua LST cloud filtering tests. For the results below, grid-averaged MODIS Terra LSTs were excluded when cloudiness reduced the yield to less than 20% of MODIS LST pixels over the grid point averaging area. The 20% threshold is consistent with the MODIS Aqua cloud filtering threshold but MODIS Aqua cloud detection and quality control is done over each AMSR-E footprint prior to resampling from footprints to the earth-grid.

---

<sup>1</sup> Note that MODIS observation time at a given location is not the same from day to day providing effectively observations at more than 4 time-of-day over a 10 day analysis period, which is sufficient for solving for the 5 parameters in Eq. 8.



For control purposes, the baseline method code was run for a set of 22,014 earth-grid points within the region 15°N-41°N latitude and 18°W-58°E longitude and including the Sahara (see figures below). The time series analysis period was July 2003 (31 days).

Figure 39 and Figure 40 compare the 19V effective temperature model RMS residual error for the baseline and revised methods. This metric is used as an overall thermal and emission model figure of merit because it includes data from AMSR-E and SSM/I (i.e., up to six observations per day) and 19V Tb's are less susceptible to atmospheric contamination than 37 and 89 GHz.

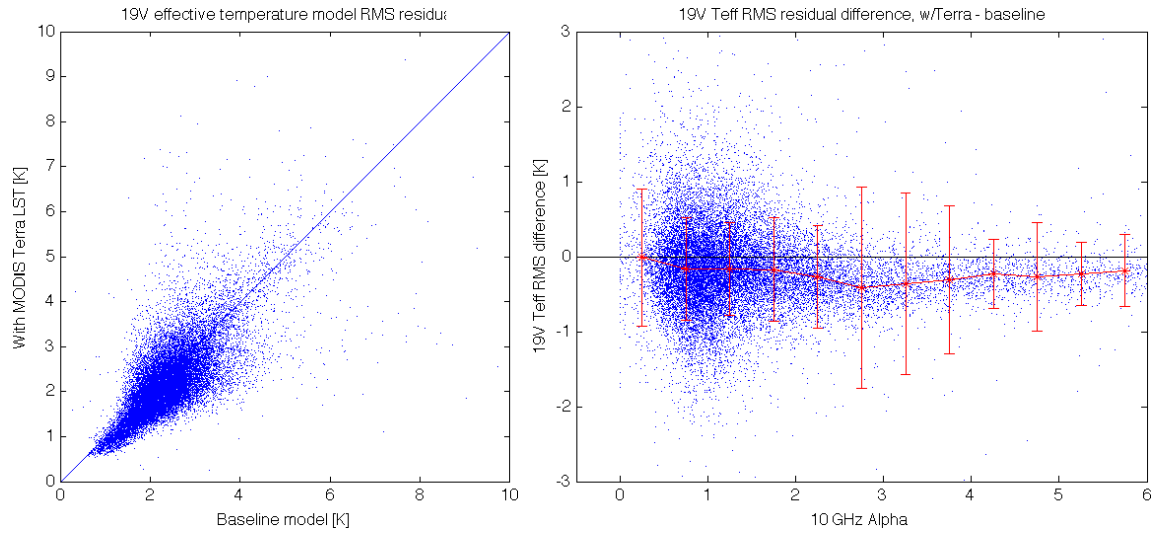
Figure 39 shows that the Terra LST method ("w/Terra") yields lower fit errors on average at all 10 GHz alpha-parameter bins – see Fig. 88 caption ( $\alpha_{10}$ , calculated from 0 to 6 in 0.5 increments). This suggests that using the Terra LSTs reduces some sources of noise in the model's temperature representation relative to the effective temperature at 19 GHz. Figure 40 shows that many low-error areas are larger in the w/Terra case than the baseline (i.e., darker blue areas in the Sahara and Arabian deserts). Some of the marginal areas where 19V effective emission temperature ( $T_{eff}$ ) errors improve straddle gradients between low and high  $\alpha_{10}$  (see Figure 46). It is possible that the elimination of microwave data in the surface temperature solution improves the accuracy of the solution where emission depth spatial heterogeneity is most significant. It might be possible to improve results with the microwave-only solution (baseline case) by improving the repeatability over time of spatial registration and gridding, e.g., through footprint matching to circularized virtual microwave footprints that are oversampled in the sensor swath reference frame.

Figure 41 and Figure 42 compare the first surface temperature model amplitude parameter,  $A_1$  (corresponding to the sinusoid with 24-hour period);

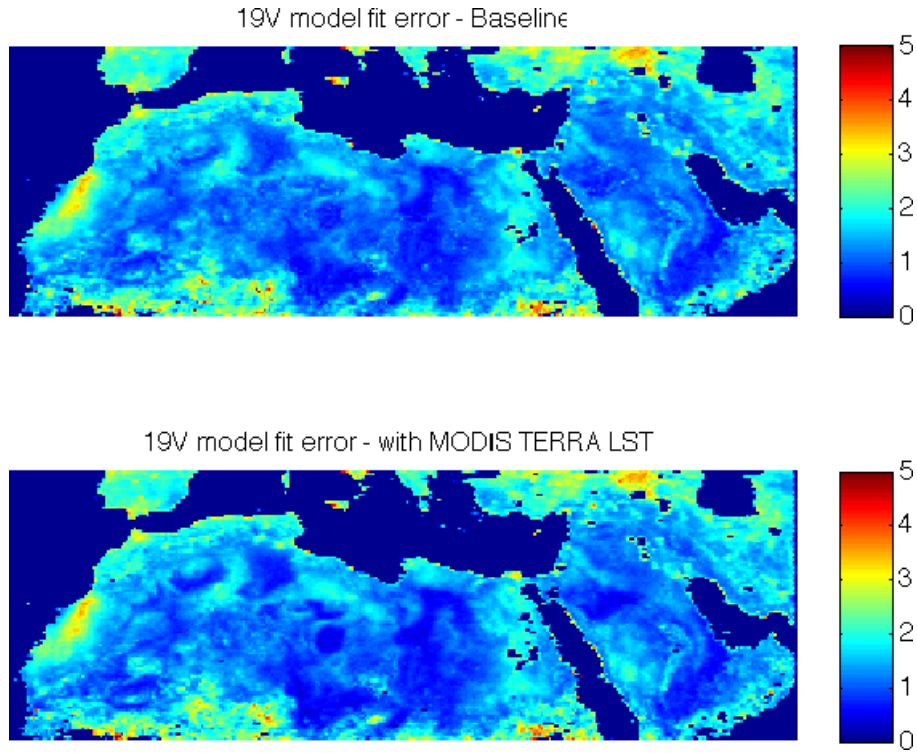
Figure 43 and Figure 44 compare the corresponding phase parameter,  $\phi_1$ .

Figure 41 shows that there is good agreement in  $A_1$  between the baseline and w/Terra model relative to  $A_1$ 's range of variability (~5 to 20 K): mean w/Terra vs. baseline differences are at most 1 K per  $\alpha_{10}$  bin. However,  $A_1$  w/Terra is consistently less than  $A_1$  baseline for all  $\alpha_{10} > 1.5$  bins. Furthermore, Figure 42 shows that there are several areas of significant w/Terra vs. baseline differences: (1) spatial noise along coastlines is lower w/Terra, (2) spatial noise in the eastern Sahara desert appears to be lower w/Terra, and (3)  $A_1$  values in the SW Sahara and central Arabian deserts are noticeably lower.

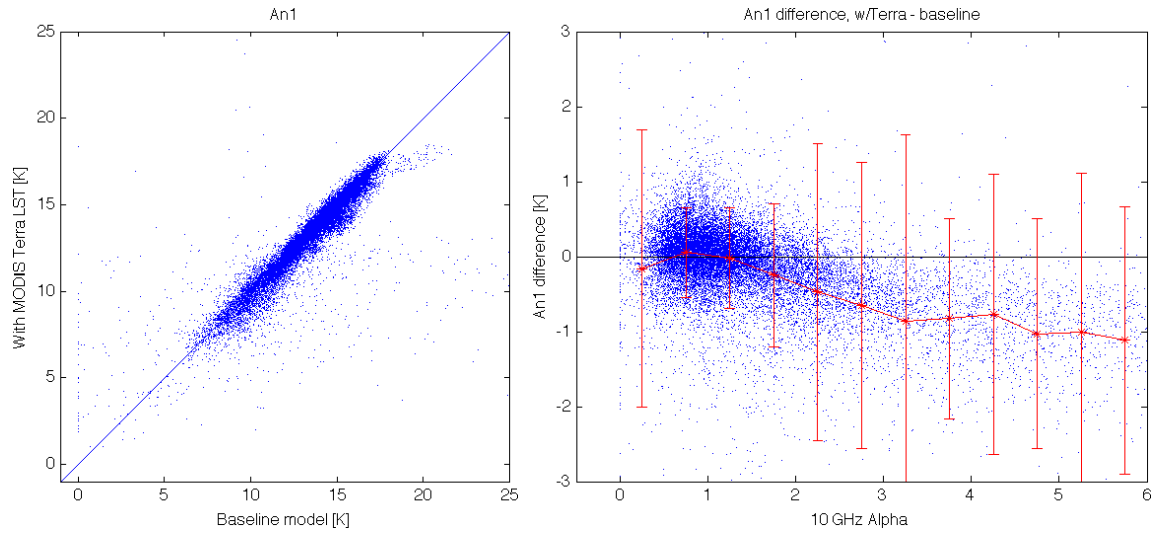
Figure 43 and Figure 44 show that despite the relative consistency in  $A_1$ ,  $\phi_1$  differs significantly (relative to its range) between the baseline and w/Terra cases. The differences are significant overall and within each  $\alpha_{10}$  bin. The overall  $\phi_1$  standard deviations (in radians) are 0.18 w/Terra and 0.21 for the baseline. These results indicate that time-of-day sampling differences between the w/Terra model (up to four MODIS LSTs per day) and baseline model (up to 6 microwave observations per day) produces a significantly different diurnal surface temperature model solution. Because the w/Terra model produces overall lower 19 GHz effective temperature model RMS residuals, there is reason to believe that the w/Terra model is more accurate. However, further tests (e.g., with LST data from geosynchronous satellite data) are needed to assess the accuracy of the surface temperature model itself over the full diurnal cycle.



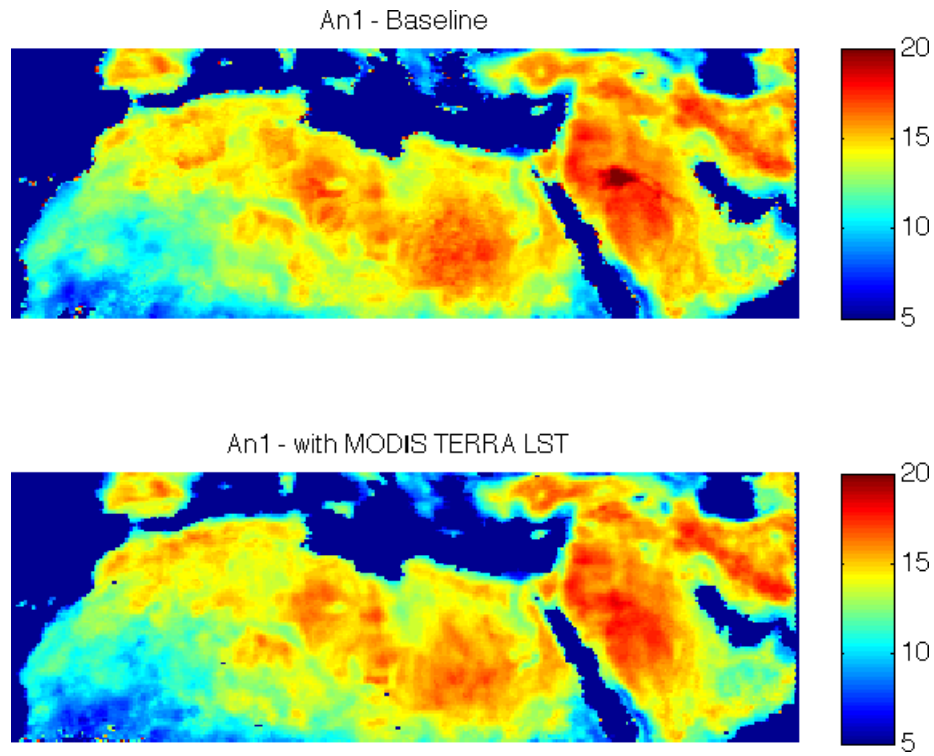
**Figure 39:** 19V effective temperature model fit error. Left: “With Terra” vs. baseline model. Right: “With Terra” – baseline difference vs.  $\alpha_{10}$  from w/Terra model. Line shows mean and standard deviation for 0.5  $\alpha_{10}$  bins.



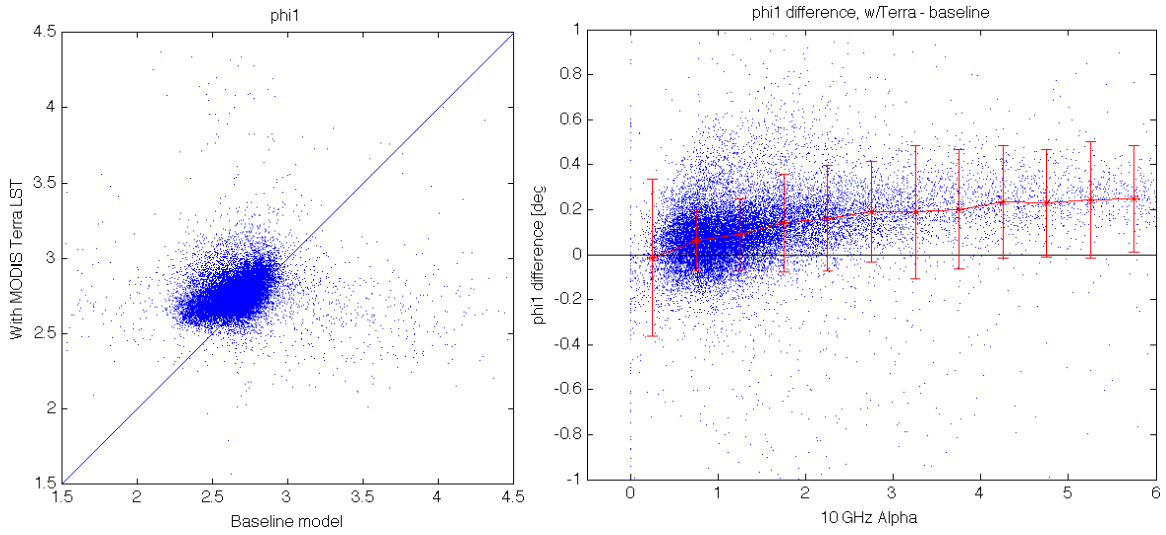
**Figure 40:** 19V model fit error (in K). Top: baseline. Bottom: with MODIS Terra LST



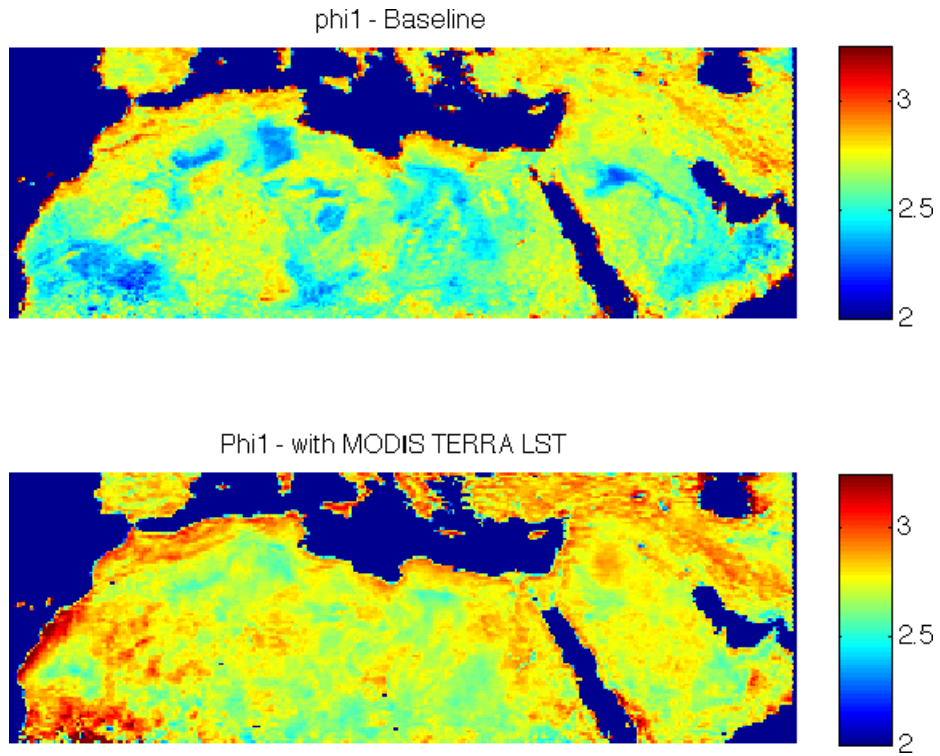
**Figure 41:** Surface model 24-hour sinusoid amplitude,  $A_1$  (in K). Left: “With Terra” vs. baseline model. Right: “With Terra” – baseline difference vs.  $\alpha_{10}$  from w/Terra model. Line shows mean and standard deviation for 0.5  $\alpha_{10}$  bins.



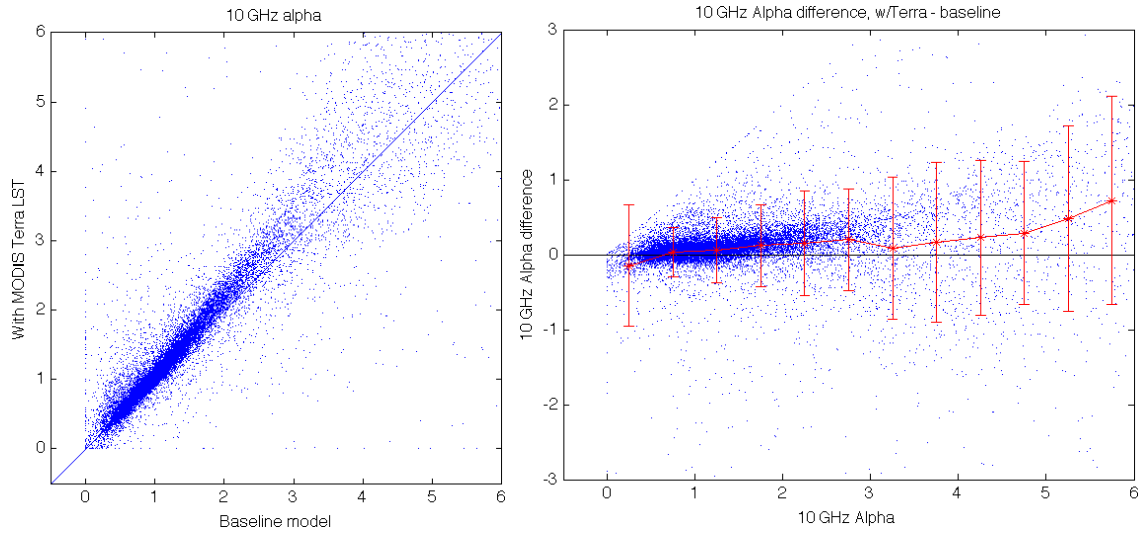
**Figure 42:** Surface model 24-hour sinusoid amplitude,  $A_1$ (in K). Top: baseline. Bottom: with MODIS Terra LST



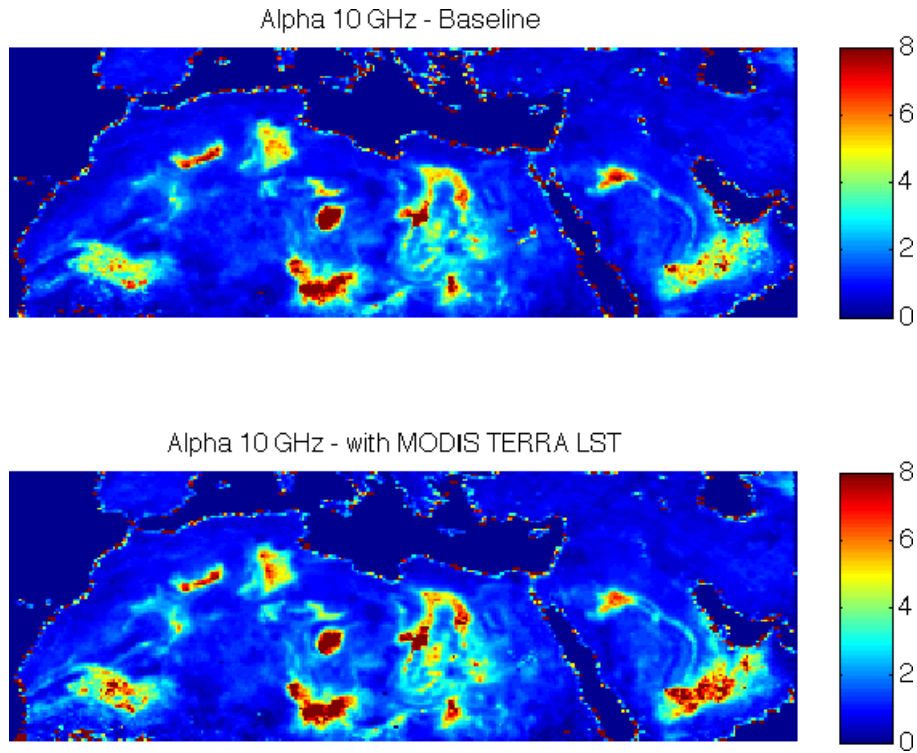
**Figure 43:** Surface model 24-hour sinusoid phase,  $\phi_1$  (in radians). Left: “With Terra” vs. baseline model. Right: “With Terra” – baseline difference vs.  $\alpha_{10}$  from w/Terra model. Line shows mean and standard deviation for 0.5  $\alpha_{10}$  bins.



**Figure 44:** Surface model 24-hour sinusoid phase,  $\phi_1$  (in radians). Top: baseline. Bottom: with MODIS Terra LST



**Figure 45:** 10 GHz emission depth parameter,  $\alpha_{10}$ . Left: “With Terra” vs. baseline model. Right: “With Terra” – baseline difference vs.  $\alpha_{10}$  from w/Terra model. Line shows mean and standard deviation for 0.5  $\alpha_{10}$  bins.



**Figure 46:** 10 GHz emission depth parameter,  $\alpha_{10}$ . Top: baseline. Bottom: with MODIS Terra LST

Figure 45 and Figure 46 compare the 10 GHz emission depth parameter,  $\alpha_{10}$ , derived with the w/Terra and baseline models. Alpha skews higher w/Terra in all but the lowest  $\alpha_{10}$  bins. Since increasing alpha has the combined effect of reducing the temperature cycle amplitude and shifting its phase relative to the surface temperature model, it is possible that differences in either the surface

temperature amplitude or phase may be responsible for alpha changes. The  $\alpha_{10}$  maps (Figure 46) show some regional  $\alpha_{10}$  increases (e.g., southern Arabian Desert). Alpha also notably decreases in the northern part of the Arabian Desert where An1 decreased. There is no noticeable difference in  $\alpha_{10}$  spatial variability in the w/Terra case.

## 6.2. AIRS vs. NWP-based atmospheric correction

In Version 1 of our emissivity retrieval algorithm, the clear-sky atmospheric correction uses temperature and water vapor profiles from the 6-hourly  $1^\circ$  resolution NCEP/GDAS analysis interpolated to the local time of the Aqua overpass and to the center of the AMSR-E footprints (Sec. 3). Surface pressure from the NCEP model is corrected using terrain height information derived from a high resolution topographic database. Prior to its use in our emissivity retrieval process, the spatial resolution of terrain data is degraded to that of the AMSR-E observations and mapped onto a fixed Nyquist sampled Earth grid.

Regional and temporal biases in the NWP model fields are obviously reflected in our retrieved emissivities (e.g. Figure 48 and Figure 49). When using these emissivities (and associated error covariance matrix) as a constraint in atmospheric (e.g. column water vapor) retrievals or in assimilation, the biases also manifest themselves in the final results, limiting the improvement that can be expected (in a time-average sense) from analyzing the microwave observations in the clear sky. Such retrieval product is still useful in that it more accurately reproduces the actual day-to-day variability in atmospheric state. Most importantly, (although the sensitivity of the microwave observations to water vapor may be impacted by liquid water) it provides valuable information under cloudy conditions and in particular in situations where models are often deficient (e.g. convection).

In addition to the bias issue, errors in the modeled water vapor generally cause significant apparent temporal variability in the 1a emissivities, especially in the 24 and 89 GHz channels and act in such a way as to decorrelate the emissivity in those channels from the emissivity in the least water vapor sensitive channels. When the emissivity covariance is determined empirically from the emissivity measurements themselves, it results a reduction in the magnitude of the off-diagonal elements of the covariance matrix and hence in the effectiveness of the surface emissivity constraint in retrieval applications. Alternate ways of estimating the covariance matrix elements can be devised to partially alleviate this issue but this topic is beyond the scope of the present study.

As part of this study, we had planned on assessing the impact on the quality of our AMSR-E emissivity estimates of basing our atmospheric correction on the AIRS retrieval product (instead of the NCEP data). The AIRS product is typically more accurate than the NCEP analysis. Also, like the MODIS LST product, it has the advantage of its excellent time co-location with the AMSR-E observation, thereby limiting the impact of errors due to temporal interpolation. In this study we use the Version 5 AIRS L3 Standard daily product (AIRX3STD) distributed by GES DISC (Goddard Earth Science Data and Information Service Center).

AIRS Standard L3 temperature profiles are specified on a fixed 24-level pressure grid from 1 to 1000mb, and water vapor profiles on 12 levels from 100 to 1000mb. The L3 data is derived from the best quality L2 products (retrieved from combined AIRS and AMSU measurements, without HSB) and

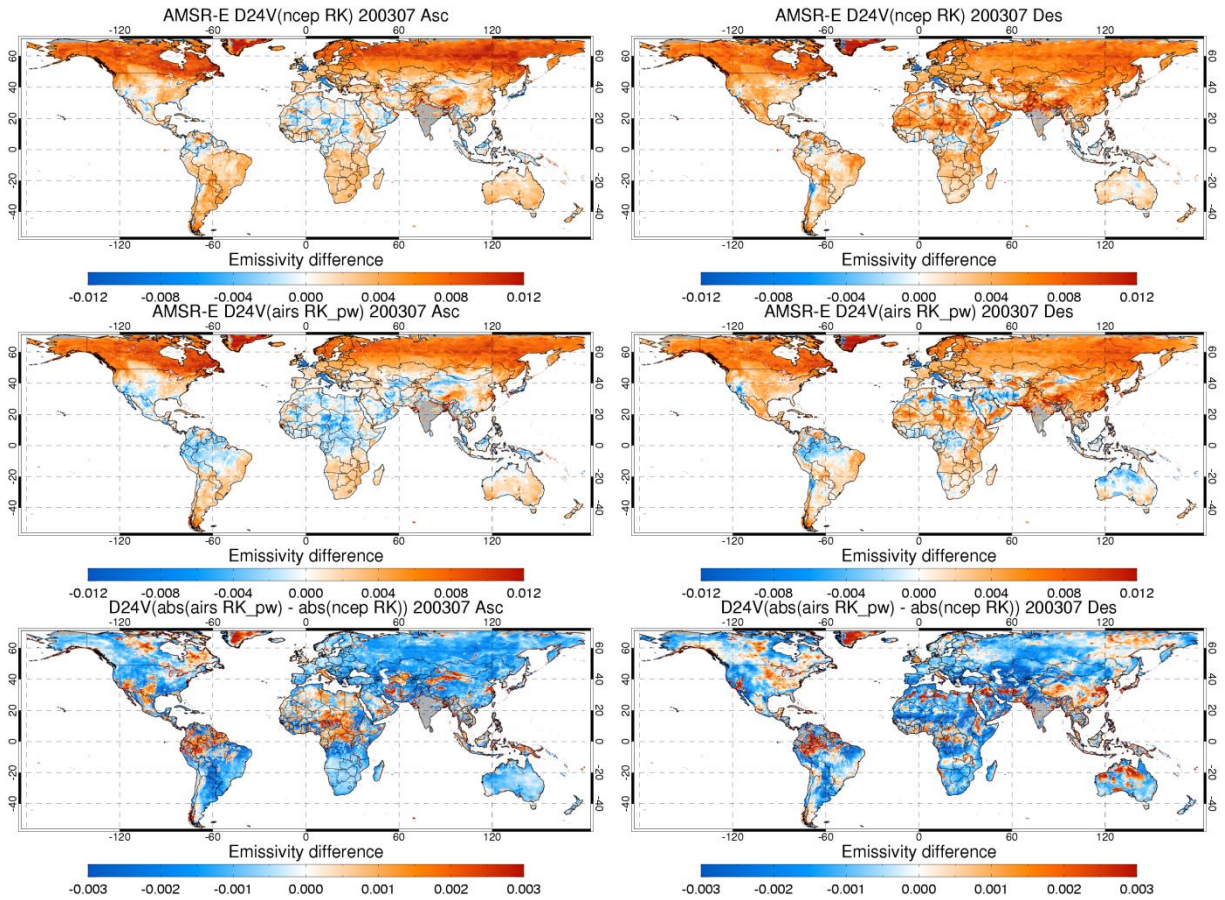


mapped on a fixed 1° degree-resolution Earth grid. Here, all valid data are used. Profiles containing missing levels (after spatial interpolation) are rejected.

To provide some assessment of the quality of the AIRS and NCEP atmospheric corrections we rely on an analysis of their relative impact on the shape and temporal stability of the retrieved 1a AMSR-E land surface emissivities. For instance, the data set that best reproduces the temporal variability in the actual atmospheric state should result, when used as an input into our retrieval process, in minimum variability in the AMSR-E emissivities estimates. Such criterion has already been used over stable regions (identified using our R11 polarization ratio) to compare the quality of different LST sources, namely the MODIS and ISCCP datasets (Moncet *et al.*, 2011b, Jimenez *et al.* 2011). In the work we refer to we showed that ESD significantly increases when replacing MODIS LSTs with ISCCP's, an indication that the quality of the ISCCP LST data set is questionable in some regions. The same approach applies here in the comparison of NCEP and AIRS atmospheric fields, except that in this case, one cannot examine the ESD of the water vapor sensitive channels (e.g. 24 GHz) in isolation of the other channels. In particular, errors in input surface temperature affect all the channels to a similar degree. In order to isolate variability in the 24 GHz channel that is caused mainly by water vapor it is desirable to remove the component due to LST. This can be achieved in some approximate fashion for instance by fitting the emissivities in the least water vapor sensitive neighboring frequencies, i.e. 11, 19 and 37 GHz, with e.g. a second degree polynomial and interpolating the function to the 24 GHz frequency. The first order impact of variability in the LST is removed by taking the difference between the actual retrieved 24 GHz emissivity value and the value of the quadratic function at 24 GHz. With this simple technique we can also identify potential biases in the water vapor fields. Note that in this process, we do not mean to imply that emissivity varies quadratically across the frequency range but rather that it varies smoothly with frequency, a valid assumption over forest and other densely vegetated surfaces. Issues might arise over deserts with this approach, since in this case, our 1a emissivities contain the signature of the subsurface thermal gradients (Sec.3), only removed by the 1b (PTS) algorithm. These areas should be treated with caution. However, so long as sub-surface temperature also varies smoothly, the above model still applies.

Note that one cannot examine the issue of the quality of the input atmospheric product, without also taking into account the Radiative Transfer Model (RTM) accuracy and the AMSR-E Tb calibration. All these factors can influence the results of our analysis in a similar way. The RTM used in our emissivity retrievals is based on Rosenkranz' (RK) absorption model (Sec. 3). Differences between this model, the model employed by Remote Sensing Systems (RSS) in their calibration work, and the AER MonoRTM model (Clough *et al.*, 1992, 2005) are detailed in Appendix B. Some previous studies had revealed calibration problems with the V.09 AE\_L2A AMSR-E Tb product (see Appendix A) currently used to generate our emissivities (Sec. 3). In particular, the 19 GHz channel emissivities appear to be too high when compared to the other frequencies. The AMSR-E 89 GHz emissivities are also significantly larger than the SSM/I 87GHz emissivities. The fact that issues were identified with these particular channels is indicative of some calibration problems which may also affect other channels. RSS has recently performed some recalibration of the AMSR-E Tbs. A cursory analysis of the new dataset is provided in Appendix A. However, because of the timing of the (non-official) release of this new dataset we have not been able to use it as originally planned for the AIRS/NCEP comparison study. For this reason, and also because the current Tb version may soon be obsolete, we simply illustrate here the technique employed for comparing the quality of AIRS and NCEP atmospheric products. In order to avoid potential issues due the artifact at 19 GHz, the 19 GHz channels were excluded from our analysis and we revert to a simple linear fit between 11 and 37 GHz channel to remove surface temperature effects from the 24 GHz emissivity. The  $D_{24}$  index used here is defined in Appendix A.

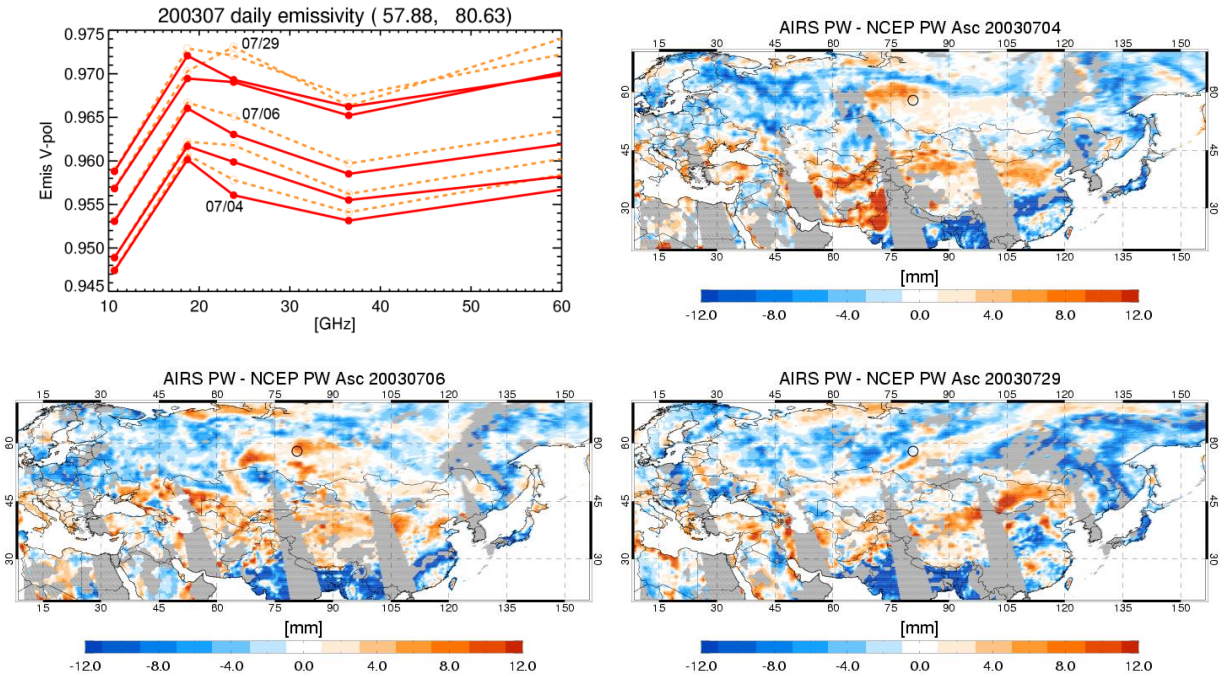
Figure 47 compares the July monthly mean 1a emissivities obtained with NCEP and AIRS atmospheric correction. Here, both the NCEP and AIRS profiles were horizontally interpolated to our 28 km grid and vertically interpolated on a 40-level pressure grid for the Radiative Transfer (RT) calculations. AIRS profiles containing missing levels (after interpolation) are rejected. Remaining water vapor profiles were scaled to match the precipitable water (PW) values provided as part the AIRS L3 Standard product dataset. Note that in Figure 47 there is a reversal of the sign of  $D_{24}$  occurring at night over sandy desert areas (e.g. Saharan ergs, Arabian Peninsula, Karakum and Taklimakan deserts). This feature is due to microwave penetration and originates from changes in the curvature of the subsurface emission temperature profile from day (negative) to night (positive). The negative values over England and Italy are caused by 11GHz RFI. Another interesting area is this region of large positive  $D_{24}$  extending across the N. Hemisphere around 60°N, most pronounced on the ASC pass. This particular anomaly had already been identified and this issue is addressed in more details in Appendix A. It is interesting to note that it is still present when using AIRS atmospheric profiles (although the magnitude of the  $D_{24}$  bias appears to be slightly smaller with AIRS) – which might discount biases in NCEP product as a potential cause. It is apparent from Figure 47 that the AIRS based atmospheric correction produces lower  $D_{24}$  values over a large fraction of the globe, which in regions where NCEP  $D_{24}$  is positive goes in the direction of making the emissivity spectra smoother. In other areas (e.g. Amazon, central Africa - over the Equatorial forest and the Sahel regions around 15°N, and New Guinea) the AIRS correction causes the 24 GHz emissivity to become more negative (e.g. Figure 49) which translates into an increase in the absolute magnitude of  $D_{24}$ . However, as noted in the caption of Figure 49, this increase could also be an artifact of calibration errors. Results are expected to differ significantly when new RSS calibration is used. The tropical regions listed above tend to be highly cloudy at this time of the year. It is possible that the impact of clouds on the quality of the AIRS product also plays a role, although as shown in Figure 50, the standard deviation of  $D_{24}$  improves over many of these areas when AIRS correction is used. Note that these results were obtained using our original monthly time division approach (Sec. 5). The new RTS time segmentation gives us more flexibility for selecting the clearest conditions, which should help avoiding instances where AIRS cloud-clearing does not perform well.



**Figure 47:** July 2003 mean D24 with NCEP (top) and AIRS (middle) atmospheric correction for ascending (left) and descending (right) passes. Bottom plot shows the (AIRS -NCEP) differences.

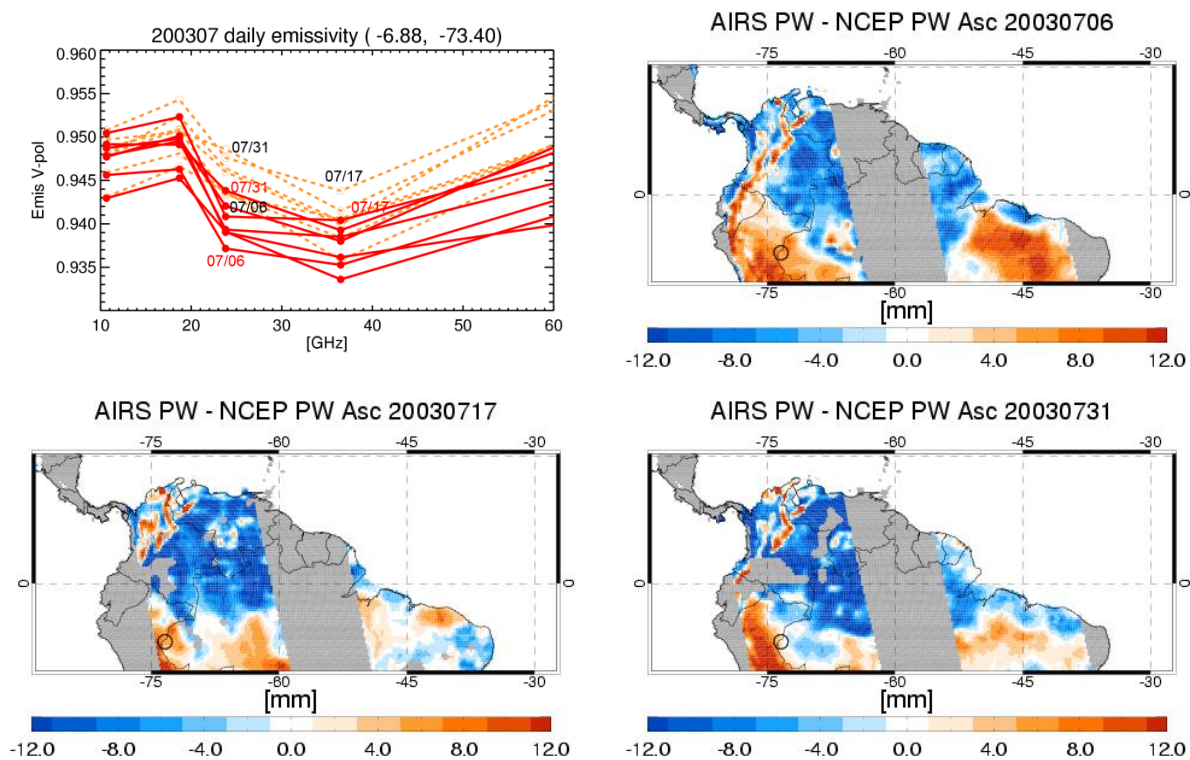
It is apparent from Figure 50 (daytime map) that there is no consistent improvement in  $D_{24}$  standard deviation across the globe when using AIRS. Again, some caution should be exercised in interpreting these results as some of our conclusions are likely to change when applying the same analysis to the recalibrated AMSR-E Tbs. Using the AIRS atmospheric product appears to improve the temporal stability of our 24 GHz emissivities in some regions, e.g. over Russia, Western Canada, and South Eastern US. The largest reduction in  $D_{24}$  standard deviation in the daytime is found north of the Amazon basin, in parts of the African equatorial forest and in the Sahel region. It seems somehow neutral over much of the Southern Hemisphere and Western Europe, but results in increased  $D_{24}$  standard deviation over Central Asia (e.g. Iran, Tibet and Himalayas), over many parts of the US and Mexico, as well as around the Southern edge of the Amazon basin and in the Northern Andes. There are also differences between the day time and night time results with the nighttime results appearing better than the daytime ones in some areas of South America, parts of the US, and over Central Asia, but generally getting slightly worse over the rest of the world. Note that over deserts, for the reasons we gave previously, the method can produce uncertain results (some areas of degradation probably linked to day time dust storm activity over Western Sahara are recognizable on the maps of Fig. 50). Over penetrating areas it might be best to look at the impact on the 1b algorithm fitting errors. The area of high nighttime  $D_{24}$  standard deviation over Finland and Northwestern Russia (Fig. 48) should be ignored. This particular artifact results from a mismatch in the time stamps of the AMSR-E data and gridded AIRS L3 product when crossing the midnight line at edge of swath, an issue that will be fixed in the future. We have not attempted to explain the day-night differences in Figure 50, in the other

regions listed above. This aspect needs to be addressed more carefully when analyzing the new RSS Tbs. Some of the observed differences might be linked to cloudiness. How much cloudiness impacts the current yield when using the AIRS based atmospheric correction for emissivity retrievals is another issue we need to worry about. Note also that the trends shown in Figs. Figure 47 and Figure 50 are similar whether we use the RK model or a hybrid model which combines, our water vapor spectroscopy with RK O<sub>2</sub> absorption parameterization (although some differences are observed when switching models). However, the AER model that is based on LBLRTM (Clough et al., 1992) has evolved since these experiments were made (see Appendix B). Ideally it would be best to eliminate RT models as source of differences in this type of analysis and ensure that the RTM used to perform emissivity retrievals is the same as the one used to calibrate the microwave instrument.

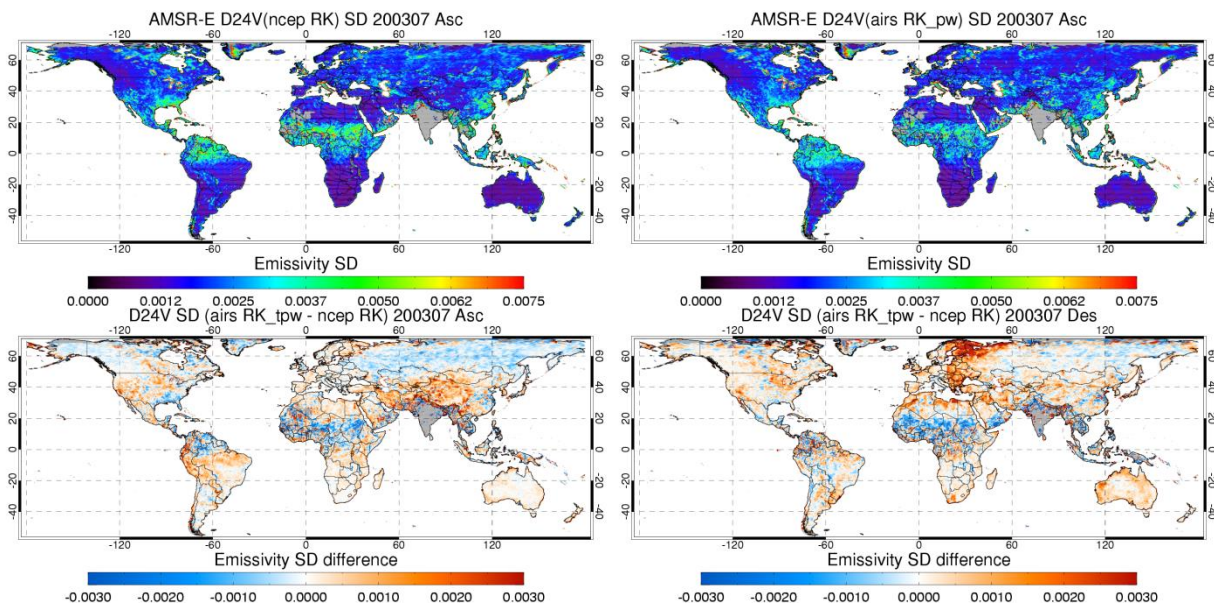


**Figure 48:** Example of maps of AIRS and NCEP PW differences over Asia for July 4, 6 and 29, 2003(daytime) and corresponding retrieved emissivity spectra for a location over Russian boreal forest marked by a circle in the PW maps. In the top left plot, red solid lines show emissivity spectra derived with AIRS atmospheric correction, and the orange dashed lines shows emissivities retrieve using NCEP. In this example, the AIRS correction tends to make  $D_{24}$  smaller and closer to zero than with NCEP. Note that the odd appearance of these emissivity spectra is due to calibration issues with the current L2A AMSR-E Tbs. With the new RSS calibration, the emissivity trend in the 11-37GHz frequency range is more linear. However, as one can infer from Appendix A (Fig. A4), the magnitude of the impact of AIRS at 24 GHz shown here is over Russia unlikely to bring this frequency in alignment with the other frequencies and  $D_{24}$  should remain positively biased.





**Figure 49:** Same as Figure 48. Top left panel is for a location in the Amazonian forest. The data used in this figure is from July 6, 17 and 31. In this example, AIRS  $D_{24}$  is lower than NCEP, and tends to be negative over most of the Amazonian forest. However, it is apparent from the plots of Appendix A (Fig. A3, A8) that with the RSS new calibration, the AIRS correction could instead result in a better alignment of the 11, 19, 24 and 37 GHz emissivities over most vegetated regions.



**Figure 50:**  $D_{24}$  monthly standard deviation with NCEP (left) and AIRS (right) atmospheric correction (top). Only the results for the ascending passes are shown. AIRS-NCEP  $D_{24}$  standard deviations differences (bottom) for ascending (left) and descending passes (right).



## 7. APPLICATIONS

### 7.1. Estimation of monthly mean diurnal cycle amplitude of land surface temperature over vegetated surfaces

In the absence of independent measurements of land surface emissivity on the scale of the AMSR-E field-of-view one has to rely on indirect validation to assess the quality of our emissivity product. Here we use the atlas of emissivity produced by the Version 1 algorithm described in Sec.3 (with the new time segmentation scheme incorporated), under clear condition only, as a constraint in the estimation of land surface temperatures over homogeneous (snow free) vegetated surfaces in all weather conditions. Here the 11V channel, which is the least sensitive to atmospheric contribution, is our primary LST retrieval channel. The retrieved microwave LSTs are compared to independent in situ air temperature measurements from surface stations. Surface temperature retrieval over arid and semi-arid areas is addressed in Sec. 7.3. In these areas, validation sources are scarce and air and skin temperature differences may be large at the time of the Aqua satellite overpass so that the former cannot be used as reference in our comparison.

Validating LSTs produced using the 11V channel, is a means of assessing the quality of the MODIS LST product (and cloud mask) entering in our emissivity generation process and to validate our QC approach while keeping errors in the input water vapor and temperature profiles from interfering with this analysis. Errors in the clear-sky atmospheric correction affect mainly the accuracy of the higher frequency emissivities. Ensuring that the impact of these errors is small is particularly important for applications of the emissivity product to precipitable water and cloud liquid water retrieval over land - but this aspect can be addressed separately (see Sec. 6.2).

It is important to bear in mind through this comparison that in situ surface air measurements do not constitute a truth. These measurements have their own noise sources. Moreover, air temperature measured a few meters above the surface may differ significantly from surface emission temperature, although these differences are generally smaller over dense vegetation than over bare soil. Finally, in situ measurements are point measurements and do not generally represent the average conditions across the microwave footprint. Before describing regional results, we discuss the results of comparisons between retrieved AMSR-E LST and in situ air temperature measurements at 18 selected sites from the FLUXNET network. These sites represent a variety of vegetation types including deciduous, evergreen and mixed forests, grassland and shrub land, across Northern Europe, the US and Canada, and in the tropics (see

Table 7). This type of detailed analysis is useful for providing some benchmark in assessing how closely the microwave derived LST can reproduce the in situ measurements. As part of this study we also compare monthly mean infrared (IR) LST from the V4 and V5 LST MYD11 products (Wan, 2007; Wan, 2006) with both microwave and in situ measurements. The results discussed here are for the month of July of 2003.

#### 7.1.1 MODIS MYD11 LST product

Our AMSR-E surface emissivities are retrieved using the V4 daily MYD11B product from the day/night LST algorithm available on a 5 km sinusoidal grid. As described in Sec.3, this product makes use of the Version 4 MODIS cloud mask. As mentioned in this same section, an additional filter is applied to exclude *a posteriori* potentially cloud contaminated LST estimates not removed by the

MODIS cloud mask. The daily and monthly MYD11C products used in this study originate from the same source as the MYD11B daily product. The main differences are that the MYD11C products are provided on a  $0.05^\circ$  regular Lat-Lon Climate Modeling Grid (CMG) and that the data has already been processed through the external cloud filter so the data can be used as is, without any further post processing. Each  $0.05^\circ$  pixel has its own quality flag which provides an assessment of the nominal accuracy of the LST data (

Table 6). Note that the Version 5 LST product was also considered as part of our inter-comparison study allowing us to assess the impact of switching to this new LST dataset for our AMSR-E emissivity database generation. The Version 5 LST algorithm makes use of the newest version (V5) of the MODIS cloud mask. It also fully integrates the split window and day night algorithms (Wan, 2008). Some authors have warned against the use of V5 MODIS LST product over deserts (Hulley and Hook, 2009) where surface IR emissivities were found to be significantly biased. However, the quality of the V4 and V5 LSTs over vegetated surfaces is believed to be similar (Wan and Li, 2008; Coll et al., 2009). One of the objectives addressed in the development of Version 5 was to increase the yield of the LST algorithm (Wan, 2008). The spatial coverage of MYD11 V4 and V5 for different quality levels are compared in Figure 52. Note that for the purpose of comparing the MODIS product with our retrieved AMSR-E LST, the former is averaged to the same 28 km resolution sinusoidal grid used for reporting the AMSR-E results (Sec. 3). In the present study, two levels of quality control (QC) are considered. The first level (denoted thereafter by QA=0) only includes the highest quality MODIS LST estimates. The QA=0 data available on a particular day is first spatially averaged over a given AMSR-E 28 km grid cell and the resulting daily LST estimates are temporally averaged to produce a 28 km resolution monthly mean LST estimate for that cell. In order to increase spatial and temporal coverage we relax our quality requirement by bringing in the next quality level, QA=01, on days when no QA=00 sample is found, to produce an LST estimate for these days. It is only when no QA = 01 data is found that we include the next QA = 10 level. This second QC level derived from this procedure is denoted subsequently as QA  $\geq$  0. Note that the QA =11 data is not used in this study. The number of these samples is relatively small and their quality is often significantly degraded. Note also that the order in which the temporal and spatial averaging operations are applied matters. Slightly different results are obtained over vegetated surfaces<sup>2</sup> when performing the temporal averaging of the  $0.05^\circ$  data first and then spatially averaging the high resolution monthly product but the differences are not significant enough to affect our conclusions. In order to be able to compare daily microwave and infrared products, it is necessary to adopt the first approach.

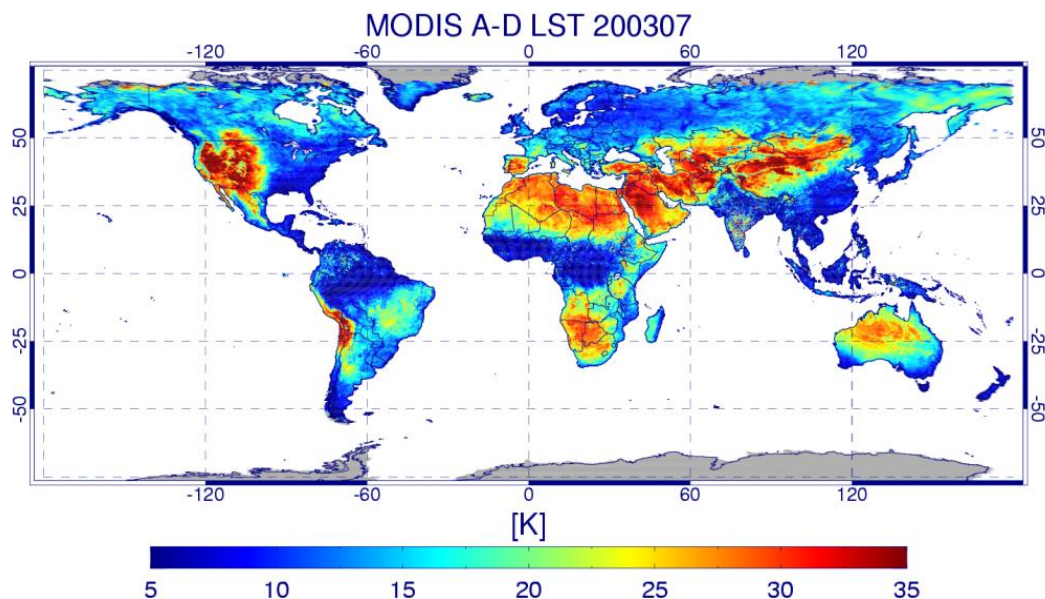
It is apparent from Figure 52 that the spatial coverage obtained with the highest quality (QA=0) V4 data is quite poor. The highest quality V5 data provide almost complete global coverage, a coverage only achieved with V4 by merging together QA=00, 01 and 10 data.

LST QA flags	Explanation
00	Estimated error < 1K
01	Estimated error between 1 and 2K
10	Estimated error between 2 and 3K
11	Estimated error > 3K
NA	LST estimate not available

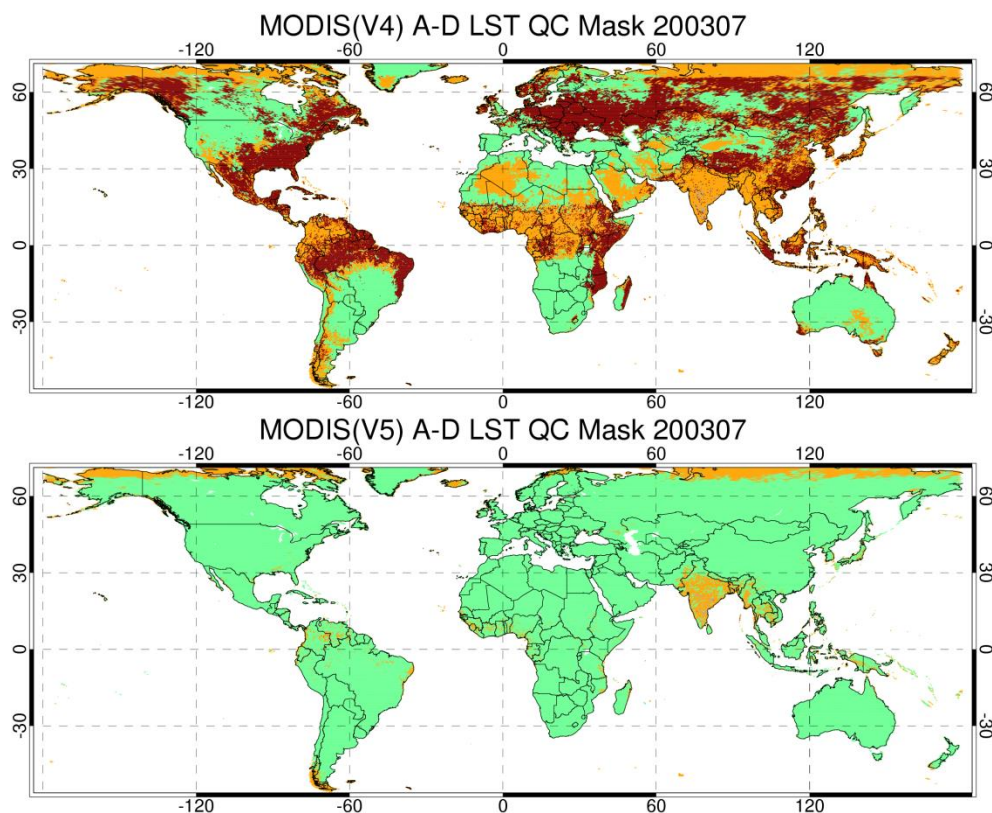
---

<sup>2</sup> Differences between the two approaches may be larger over predominantly clear (small day-to-day variability in solar forcing) thermally inhomogeneous desert areas but this consideration falls outside of the topic covered in this section

**Table 6:** MODIS MYD11 LST QA flags. Note that data assigned a QA value = 11 are not considered in this study and are treated as NA.



**Figure 51:** 28 km-resolution global map of mean Ascending-Descending (A-D) MODIS LST for the month of July 2003. This map was derived using the V5 MODIS product. Missing areas are in gray.

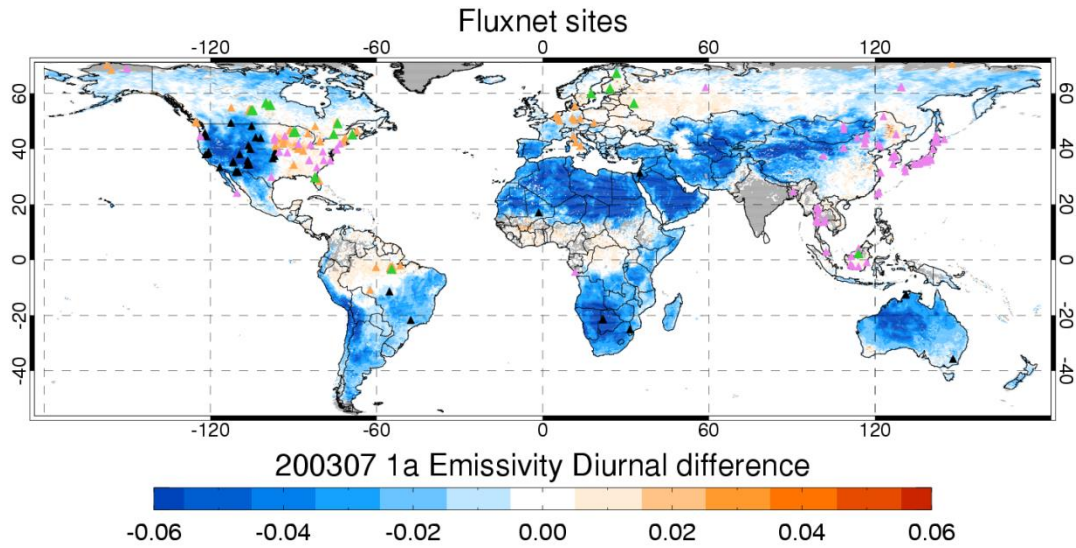


**Figure 52:** July 2003 spatial coverage of the V4 and V5 MODIS LST products. Colors indicate the quality level of the LST data. Green: QA =00, Orange: QA=00-01 and Brown: QA=00-10.

## 7.1.2 Comparison with of microwave and infrared LST with in situ data

### 7.1.2.1 FLUXNET in situ sites – overview of comparison results

Figure 53 shows the locations of the FLUXNET sites for which air temperature records are generally available as well as the sites selected for this particular study. The latter are listed in Table 7. Figure 54 shows one example (for the SK-HJP site – site #9) of mix of surface types typically encountered around some of the selected FLUXNET sites. This figure provides some idea of the order of magnitude of the impact of spatial heterogeneities on the differences between 1) air surface temperature reported at the in situ site, 2) MODIS 0.05° LST product averaged over the AMSR-E 28 km grid box surrounding the site, and 3) LSTs estimated from AMSR-E observation on a 51 km × 29 km quasi-elliptical FOV with center nearest to the measurement tower.



**Figure 53:** Location of sites from FLUXNET network for which air temperature from tower measurements are available. The sites used to analyze year 2003 data are in green. The black symbols represent the sites located in arid and semi-arid areas (ignored for this study), the orange symbols, the sites having incomplete temperature data records or sites that are located in unstable (per R11 standard deviation) or highly inhomogeneous areas. The sites from which data is available for years other than 2003 are shown in pink.

The results of the comparison of the 28km resolution MODIS (V4 and V5) and AMSR-E monthly mean Ascending-Descending overpass (A-D) LST differences with the in situ data for the 18 sites are shown in Figure 55. Here, the in situ data is reported every 30 minutes and interpolated to the exact times of the MODIS and AMSR-E observations. Only the highest quality level (QA=0) MODIS data is included in this comparison. On days when no 0.05° sample is available within the 28km grid, no estimate is produced for that day. The number of days within the month for which a 28km V4 or V5 estimate is available is summarized in Table 8. The impact of adding the QA=01 and 10 data is discussed later in this section. Examples of daily time series for few selected sites are shown in Figure 56. As shown in this figure, the daily temperature trends are well reproduced by the microwave data, which illustrates the robustness of our approach. It is apparent from Figure 55 that the IR consistently overestimates the mean diurnal cycle amplitude in the midlatitudes, with V5 being often slightly closer to the in situ estimates than V4. The maximum discrepancy between IR and in situ surface temperature is ~5 and 6K for V5 and V4, respectively. Due to the persistent cloudiness at the lower latitude sites in

	Station Code	Lat., Lon.	Location	Meas. height	Vegetation type (canopy height)
1	FISod	67.36, 26.64	Finland	NA	Evergreen coniferous, ENF (10-18m)
2	FIHyy	61.85, 24.29	Finland	4, 8, 16, 33, 50, 67 m	Evergreen coniferous, ENF (18m)
3	SENor	60.09, 17.48	Sweden	NA	Evergreen coniferous, ENF, scot pines
4	RUFyo	56.46, 32.93	Russia	NA	Evergreen coniferous, ENF, old spruce
5	CANS4	55.91, -98.38	Manitoba	10m	Boreal, ENF (5-6m)
6	CANS5	55.86, -98.49		9m	Boreal, ENF (4-5m)
7	CANS7	56.64, -99.95		6m	Open Shrublands
8	SK-OBS**	53.99, -105.12	Sask. (Can.)	1m	Old mature Black Spruce overstory, feather moss ground cover (7m)
9	SK-HJP/OJP	53.94, -104.65		1m	Jack pine seedlings, grasses, herbs, lichen, dead wood
10	QC-HBS00	49.27, -74.04	Quebec	1.6 m	Regenerating Black Spruce, Vaccinium spp., Ledum, moss
11	ON-EPMB	45.41, -75.48	Ontario	3m	Ericaceous shrubs, deciduous shrubs (0.2-0.3m)
12	USSyv	46.24, -89.35	W. Mich.	36m	Mixed forest (22m)
13	USLos	46.08, -89.98	N. Wisc.	10.2m	Deciduous Broad-leaf Forest
14	USHo1	45.20, -68.74	Maine	26.5m	Mixed Forest (20m)
15	USSP1	29.74, -82.22	N. Florida	32m	Evergreen Broad-leaf Forest (20m)
16	USSP2/3	29.76, -82.24	N. Florida	18.5m	Slash pine plantation, woody Savannas (<10m)
17	ID-Pag	2.35, 114.04	Palangkaraya Indonesia	Above canopy	Tropical forest (26m)
18	Santarem-Km77	-3.26, -54.54	Brazil	11.3m	Tropical forest (28m)

**Table 7:** List of selected FLUXNET sites sorted by region and latitude (ENF = Evergreen Needleleaf forest)

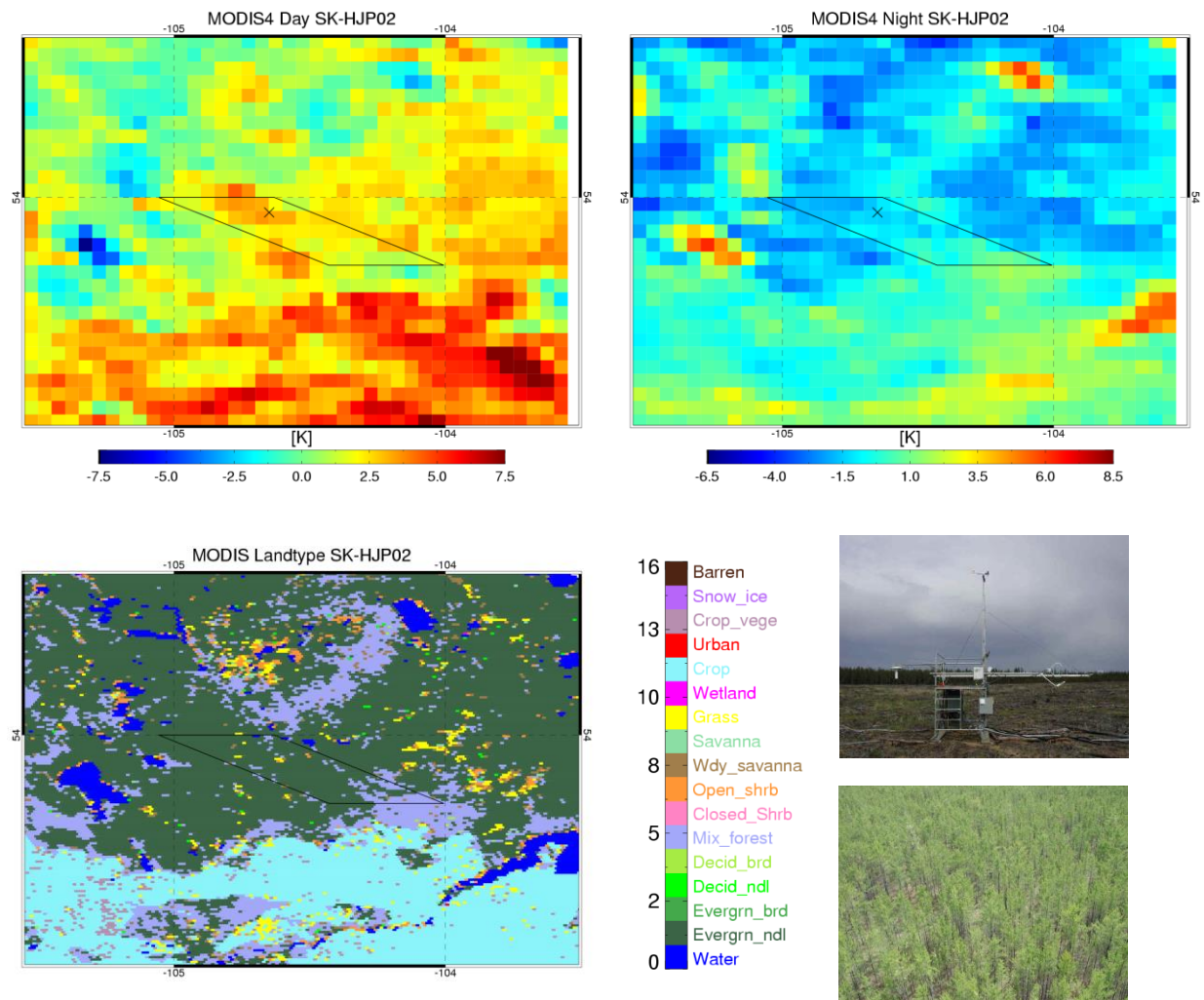
Sites		V4 QA=0		V4 QA≥0		V5 QA=0		V5 QA≥0	
		Day	Night	Day	Night	Day	Night	Day	Night
1	FISod	13	21	28	30	22	20	27	28
2	FIHyy	19	23	25	30	24	23	25	29
3	SENor	12	12	26	30	26	19	26	23
4	RUFyo	11	16	26	28	23	23	27	27
5	CANS4	13	11	28	31	26	20	29	27
6	CANS5	16	13	28	29	27	18	28	28
7	CANS7	21	18	29	31	24	23	29	25
8	SK-OBS	18	18	27	30	25	25	28	29
9	SK-HJP02	13	17	25	29	24	26	26	29
10	QC-HBS00	5	7	21	31	13	13	21	22
11	ON-EPMB	11	10	23	27	19	17	22	21
12	USSyv	17	16	23	30	22	21	23	26
13	USLos	12	14	24	28	23	20	25	24
14	USHo1	7	10	23	30	19	17	22	22
15	USSP1	1	1	16	23	12	14	18	20
16	USSP2	0	3	16	24	11	11	17	17
17	ID-Pag	11	0	21	5	19	3	21	6
18	Santarem-Km77	11	0	23	10	22	15	23	20

**Table 8:** Number of daily 28 km resolution MODIS LST estimates available for the month of July 2003 at each of the FLUXNET sites listed in

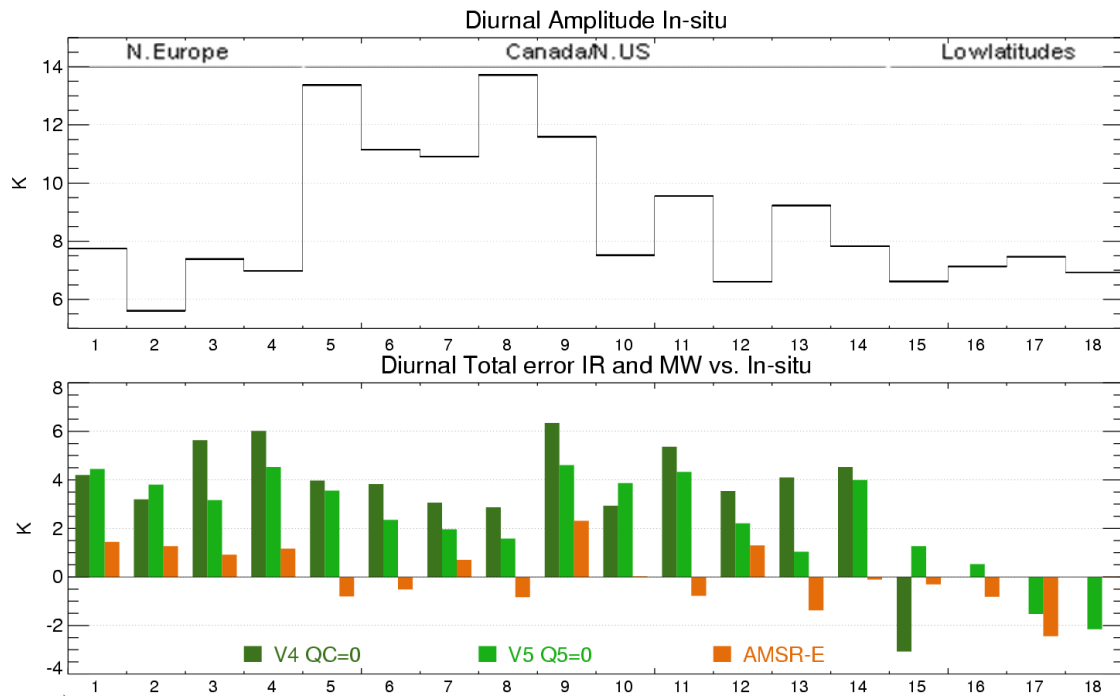
**Table 7** for V4 and V5 and, for each MODIS version, for our two quality levels: QA=0 and ≥0



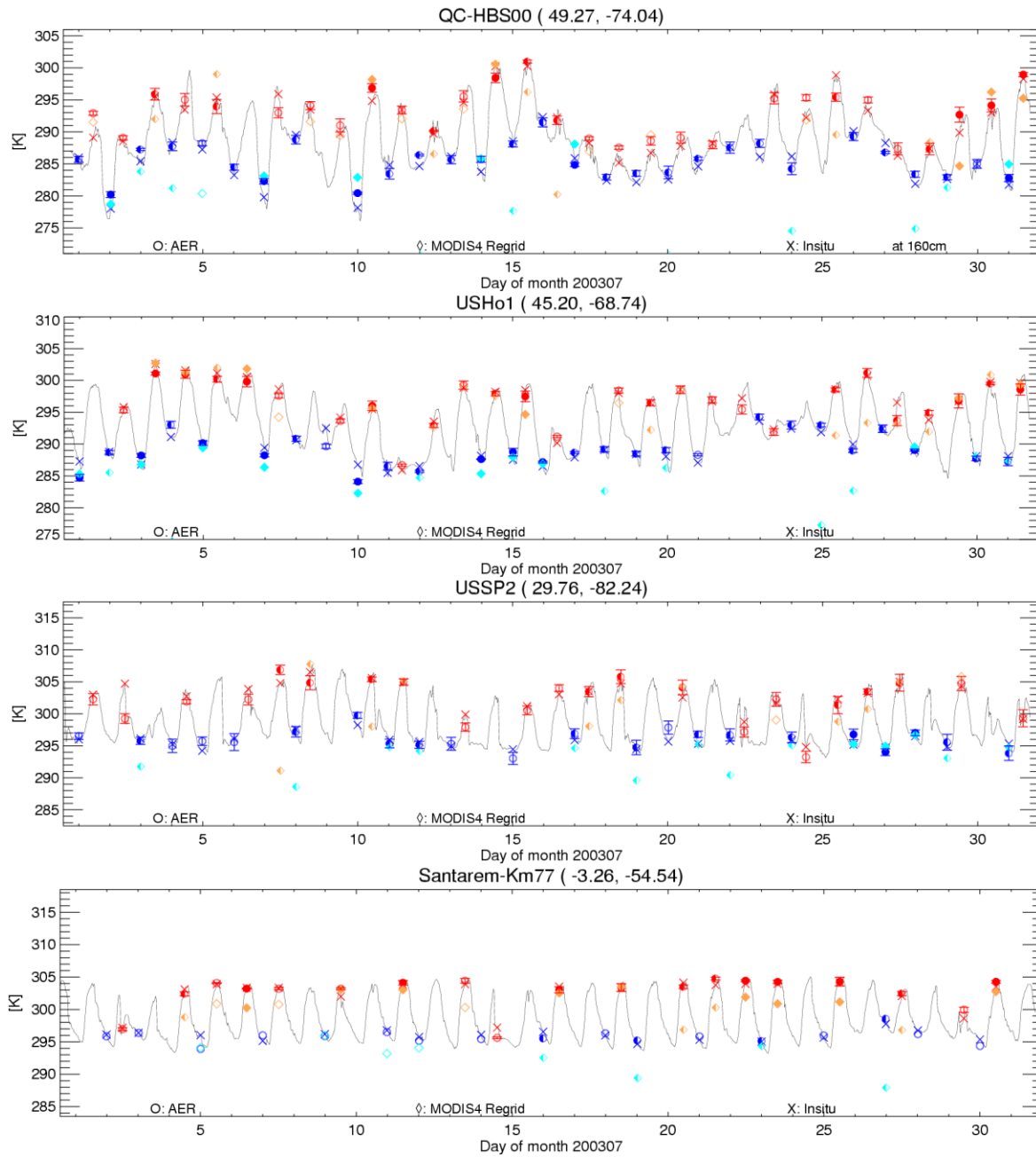
July, no V4 QA=0 A-D LST estimate is produced for those sites. By comparison, the mean microwave derived LSTs are significantly closer to the in situ measurements. In this case, the mean difference with in situ is generally within  $\pm 1.5\text{K}$  and slightly exceeding  $2\text{K}$  in only two instances. One of those cases (site #17) is a particular instance where emissivities are for some yet unknown reason inconsistent with the surface type (tropical forest). This case deserves to be examined in more details.



**Figure 54:** Surface thermal heterogeneity due to variations in surface type around the SK-HJP site (bottom left MODIS surface type map). The 28 km AMSR-E grid box nearest to the in situ site (marked as “X” on the maps) is outlined in black. The dominant surface type in this region is Evergreen forest but the tower itself is located on cleared land (bottom right pictures). The top two plots represent the monthly mean surface temperature differences between MODIS day time (left) and nighttime (right) LST and time collocated tower measurements. The crop land south of the site appears significantly warmer than the surrounding forested areas. Lakes are also clearly visible in the MODIS temperature anomaly fields. They appear colder during the day and warmer at night.



**Figure 55:** Monthly mean A-D surface temperature differences from in situ measurements at the 18 FLUXNET sites (upper panel) and differences between A-D temperatures from MODIS (V4 and V5) and AMSR-E, and in situ data (lower panel). Here only the highest quality MODIS data is included. No high quality V4 data is available at sites 16 (daytime), and 17 and 18 (nighttime) in July 2003.



**Figure 56:** Examples of daily time series for FLUXNET sites in Quebec (QC-HBS00) (a), Maine (USHo1) (b), Florida (USSP2) (c), and in the Amazon basin (Santarem-Km77) (d). The in situ measurements are represented by the continuous line. The red crosses and red circles represent the daytime in situ measurements (at the time of the AMSR-E overpass) and the AMSR-E LST estimates, respectively. The corresponding nighttime data is shown in blue. The nighttime and daytime MODIS V4 data are represented by the orange and light blue diamonds. Filled symbols indicate that the conditions across the 28 km grid were clear at the time of the observations. Open symbols are used for overcast conditions and half-filled symbols for partly cloudy.

### 7.1.2.2 *MODIS versus in situ surface temperatures: detailed analysis*

More insight into the causes for the discrepancies between infrared and in situ data shown in Figure 55 can be gained by examining the mean daytime and nighttime temperature differences and by separating errors due to temporal sampling from those due to the combined algorithm/measurement noise, and representativeness of the data. The total monthly mean difference between test and reference data (in this case, infrared and in situ, respectively), can be express as the sum of two terms,

$$\delta \bar{T}_{ref} = \delta \bar{T}_{ref}^* + \mathcal{G}_{ref}, \quad (9)$$

where,

$$\delta \bar{T}_{ref}^* = \frac{1}{L} \sum_{l=1}^L T'_l - \frac{1}{L} \sum_{l=1}^L T_{ref,l}, \quad (10)$$

and,

$$\mathcal{G}_{ref} = \frac{1}{L} \sum_{l=1}^L T_{ref,l} - \frac{1}{N} \sum_{n=1}^N T_{ref,n}. \quad (11)$$

In the above equations,  $N$  is the number of days for which a report is available from the reference data source (generally, for in situ observations, the number of days in the month) and  $L$ , the number of days for which coincident LST data is available from the MODIS MYD11 product. This number depends on the selected QA level (Table 8). Here, in situ and infrared surface temperatures are denoted by  $T_{ref}$  and  $T'$ , respectively.

The daytime and nighttime  $\delta \bar{T}_{insitu}^*$  and  $\mathcal{G}_{insitu}$  mean difference terms for all 18 sites are shown in Figs. Figure 57 and Figure 58. In these figures, we added the differences obtained with MODIS V4 QA $\geq 0$ . The LST bias resulting from the sparse temporal sampling of the infrared surface observations ( $\mathcal{G}_{insitu}$ ) can make up a significant portion of the total mean surface temperature error. As expected, this term is usually positive during the day and negative at night. The magnitude of  $\mathcal{G}_{insitu}$  is larger in the midlatitudes than in the tropics. Many locations in the tropics show little day-to-day variability (e.g. Santarem site in Figure 56). In this case, fewer samples are required to get a reasonably accurate estimate of the mean day and nighttime temperatures. The magnitude of the sampling bias, is generally between +1 and +3K in the daytime for the midlatitude sites. It tends to be lower at night, since the thermal contrast between clear and cloudy phases is often weaker, but can reach below -3K at some sites. The example of time series provided in Fig. 6 for the QC-HBS00 site (site #10) illustrates well the impact of sparse temporal sampling on mean daytime and nighttime IR surface temperature estimates. In this case, daytime and nighttime temperatures are highly modulated by the presence of clouds. Day/night differences may reach over 10K during clear outbreaks and may be as low as to  $\sim 3$ K under cloudy conditions. In the case of QC-HBS00, MODIS misses a long stretch of quasi-overcast conditions lasting from day 16 to day 24 and other isolated cloudy occurrences where nighttime temperatures are nearly as high as the daytime temperatures. These cloudy gaps explain the significant positive and negative sampling biases obtained in this case with the MODIS V4 QA=0 product in the daytime and nighttime, respectively. Note that the sign of the  $\mathcal{G}_{insitu}$  term can sometimes be opposite to what one might generally expect. These exceptions occur for instance when the temporal

representation is poor, i.e. only a few samples are available through the month and at least one sample happens to coincide with a cloudy day (not flagged by the cloud mask). This is the situation encountered for instance with MODIS V4 at the USSP1 site (site #15).

Obviously the magnitude of sampling biases tends to be smaller with V5 and with V4  $QA \geq 0$  than with V4 ( $QA=0$ ) since both datasets have a higher temporal coverage (Table 8). Note that at the Florida sites (sites # 15 and 16) and at the ID-Pag and Santarem sites (# 17 and 18), which are typical of equatorial regions, the infrared sampling is poor due to the high frequency of clouds. In particular, in the ID-Pag and Santarem cases, there are no  $0.05^\circ$  high quality MODIS V4 samples available at night during the whole month of July, making it impossible to provide an estimate of the mean night time LST and diurnal cycle amplitude for that month. A similar situation occurs for the USSP2 site during the day.

The behavior of the  $\delta\bar{T}_{insitu}^*$  term is less systematic and more difficult to explain than the temporal sampling bias as representativeness errors may contribute significantly to this term and depend on local characteristics of each site as well as weather conditions. Representativeness errors include differences between skin and surface air temperature, which tend to be positive during the day and negative at night. Although the magnitude of these differences is expected to be smaller over vegetated surfaces than over bare soil, it may still be significant and is likely to vary with local atmospheric conditions and in particular with the amount of mixing occurring in the BL. Spatial heterogeneities also play a significant role. Presence of lakes, urban areas, variability in the vegetation cover are all factors that contribute to the differences between grid-cell averaged infrared surface temperature and the temperature measured at the tower site. The sign of this effect may vary from site to site. For instance, the 28-km grid box around the ON-EPMB site (site #11) includes the Ottawa urban development which contributes to make the MODIS LST warmer on average than the in situ measurements both at night and during daytime (Fig 16). On the other hand, the impact of lakes (e.g. Figure 54) tends to be positive and night and negative during the day, while surfaces where vegetation is sparse, which respond more quickly to changes in the solar forcing and cool more rapidly at night, will show an opposite behavior. Noise in the MODIS day/night LST algorithm could also play a significant role, especially when the number of samples available is small.

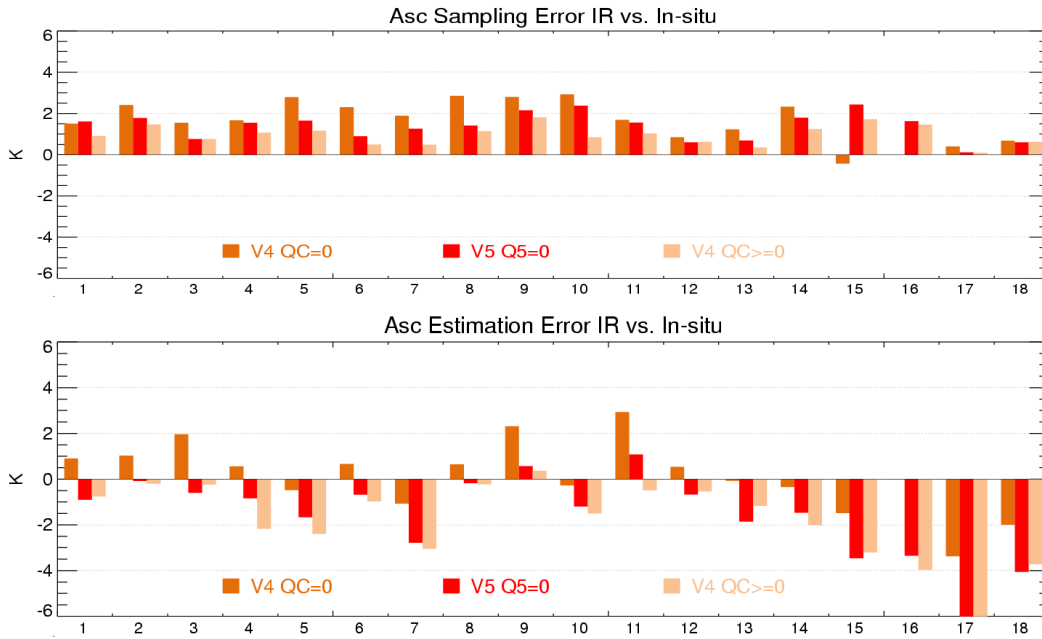
Another major contributor to the  $\delta\bar{T}_{insitu}^*$  term is residual cloud contamination which usually causes negative errors in the IR LST estimates both during daytime and nighttime. It is apparent from examination of Figs.

**Figure 57** and **Figure 58** that the  $\delta\bar{T}_{insitu}^*$  term is predominantly negative for MODIS V5 whereas for V4, as expected based on the previous discussion, it does not have a clear preferential sign. With V5, in the daytime,  $\delta\bar{T}_{insitu}^*$  is most strongly negative in the lower latitudes (sites # 15-18) where cloudiness is persistent. At night this term is negative at all sites. As shown below, the difference between V4 and V5 products most likely results from the impact of cloud contamination of the V5 data.

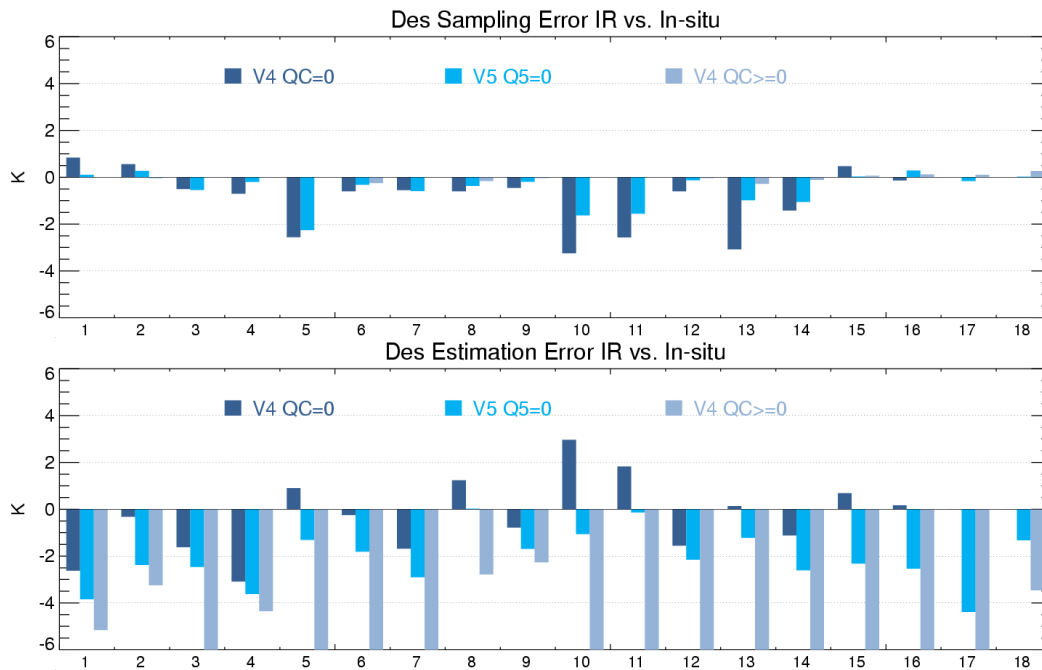
Figure 60 shows a daytime and nighttime comparison of the  $QA=0$  V4 and V5  $T_{skin}$  minus in-situ air temperature differences averaged over days when both MODIS products are available (Case A), with the V5 minus in-situ surface temperature measurements averaged over days when only V5 produces a  $QA=0$  estimates (Case B). In the latter case, a V4 LST estimate is either non-available or flagged as lower quality. Note that in Case A, V5 LSTs tend to be colder than V4, which may be due to differences inherent to the methodologies employed to produce V4 and V5 LSTs. However, it is clear from this figure that this negative difference increases significantly in Case B. Based on a careful



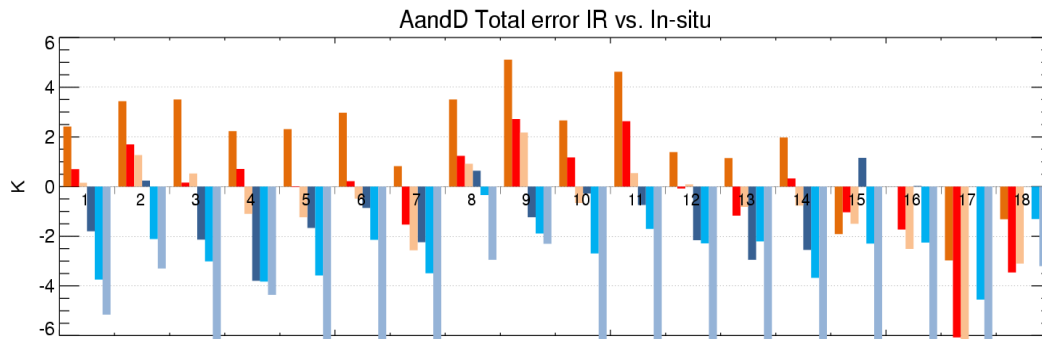
examination of the multi-spectral cloud imagery covering the 18 selected sites for July 2003, it appears that a significant fraction of the V5 highest quality  $0.05^\circ$  LST estimates not coinciding spatially with a V4 QA=0 estimate are cloud contaminated, an indication that the increase in yield obtained with V5 is often achieved at the expense of a larger degree of cloud contamination. This aspect is illustrated by Figure 61, Figure 62 and Figure 63. The two effects combined (reduced sampling error and cloud contamination compensating for daytime positive bias in V4 LST estimates) result in the mean daytime surface temperature difference with in situ measurement being generally smaller (and in some cases slightly negative) with V5 than with V4 QA=0 in the midlatitudes (Figure 59). At night (or in the daytime at the lower latitude sites) the magnitude of the negative differences between infrared and in situ LST increases with V5. As a net result, the mean A-D LSTs computed from the V5 QA=0 product July 2003 is closer than V4 QA=0 to the in situ estimate (Figure 55), but this improvement appears to be somewhat fortuitous as cancellation of errors due to residual cloud contamination at night and during the day seems to play a significant role.



**Figure 57:** Daytime  $\mathcal{G}_{insitu}$  and  $\delta\bar{T}_{insitu}^*$  computed at 18 FLUXNET sites for July 2003 for MODIS V4 QA=0 (orange), V5 (red) and V4 QA $\geq$ 0 (light brown)



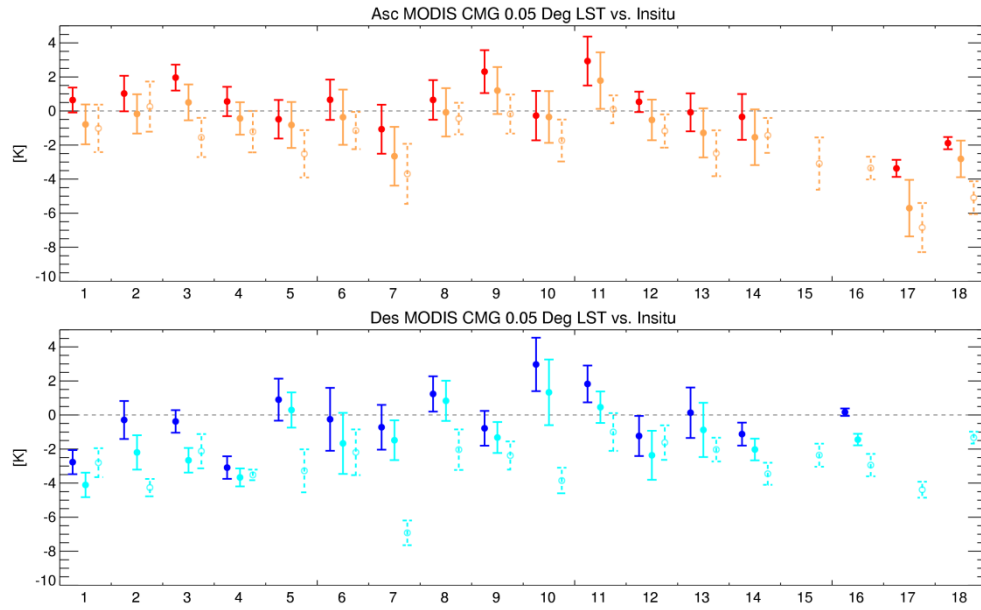
**Figure 58:** Same as  
Figure 57 for nighttime



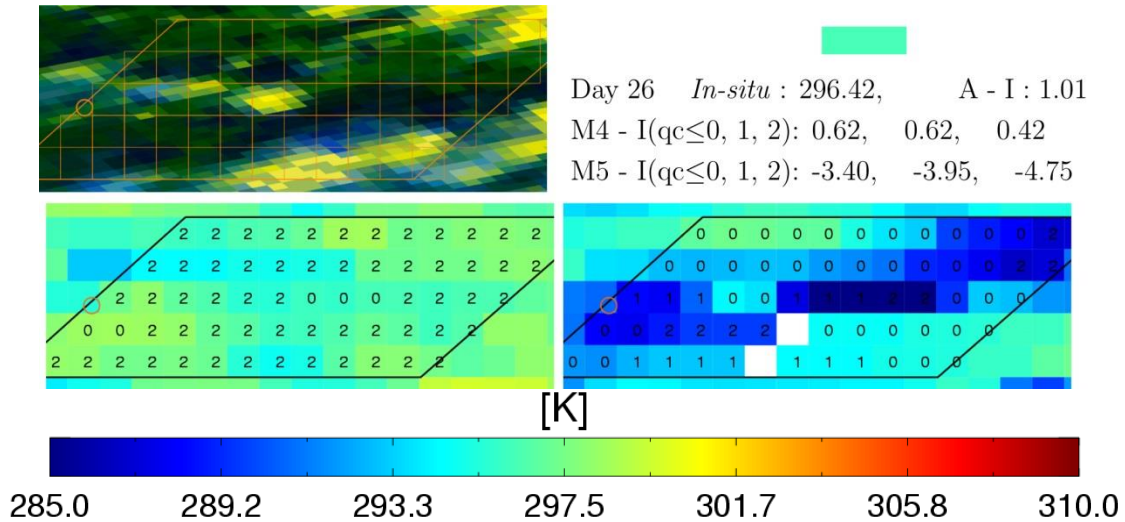
**Figure 59:** Monthly mean daytime and nighttime differences between MODIS LST and in situ air temperature at 18 FLUXNET sites in July 1003 for MODIS V4 QA=0, V5 and V4 QA $\geq$ 0. Color code is identical to the one used in

Figure 57 (day) and Figure 58 (night).

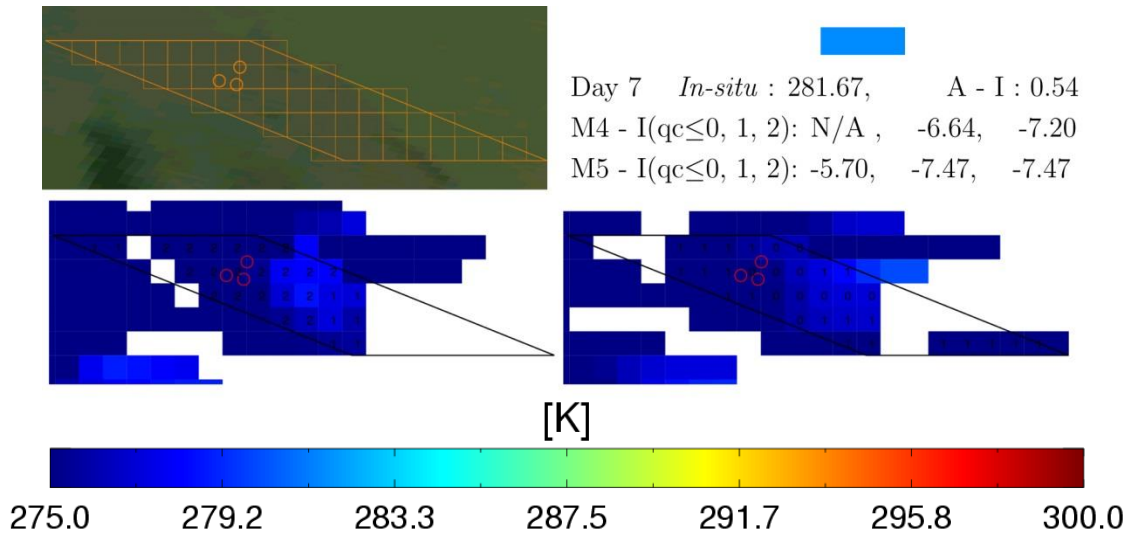
It is seen from the examples of Figure 63 that the impact of cloud contamination on the MODIS V4 (QA=01 and 10) LST estimates can be large. This impact is more pronounced at night and cause large negative bias in the mean nighttime MODIS V4 QA $\geq$ 0 LST estimates exceeding the range of the y-axis in Figure 58. For this reason, it is best to switch to the V5 version for increased spatial coverage rather than attempt to improve the yield of the MODIS V4 product by relaxing the QA threshold.



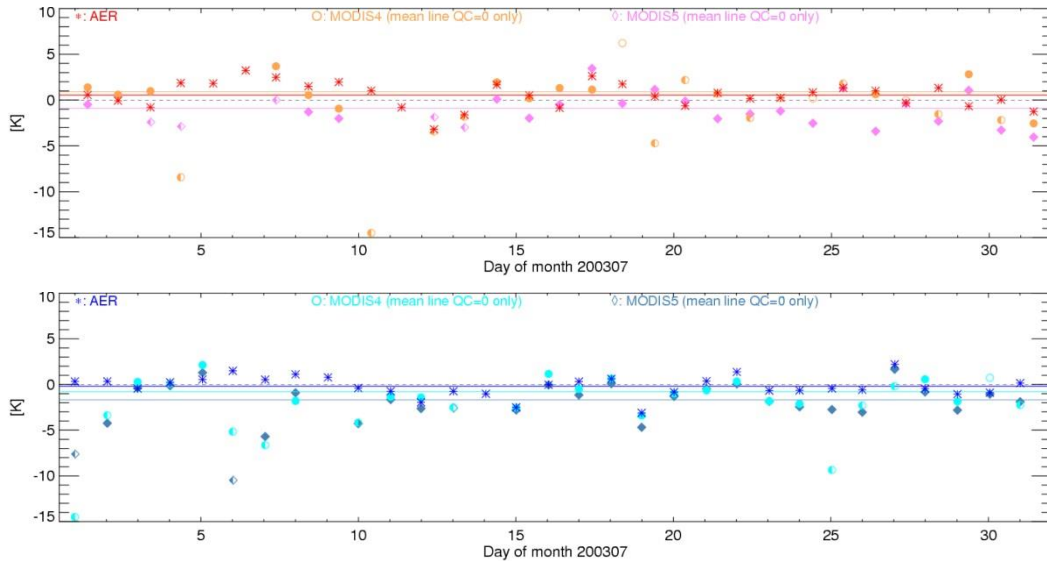
**Figure 60:** Monthly mean (and standard deviation) daytime LST difference (top) between V4 QA=0 LST (solid red) and in situ air temperatures for all 18 FLUXNET sites. V5 QA=0 differences averaged over days when a QA=0 estimate is available (Case A) are represented by the solid orange lines. Dashed orange lines show the V5 QA=0 differences averaged over days when no coincident V4 QA=0 estimate is available (Case B). Bottom plot show the nighttime results. Mean V4, V5 (Case A) and V5 (Case B) LST differences are shown in solid blue, solid cyan and dashed cyan, respectively.



**Figure 61:** Example of MODIS daytime multi-spectral cloud imagery (top) and V4 (bottom left) and V5 (bottom right) LST QA flags for a 28 km AMSR-E grid cell at FISod site (July 26 2003). Circles mark the tower location. Small squares are the 0.05° CMG grid boxes on which the MODIS MYD11 product is reported. QA levels (0=00, 1=01 and 2=10) are indicated on the bottom maps. Invalid data are left blank. Measured in situ temperature on that day is 296.42 (top right) and difference with AMSR-E retrieved LST is 1.01K. It is seen in this figure that only few V4 QA=0 estimates are produced and they coincide with the clearest part of the scene. However some V5 estimates falling in cloudy areas (e.g. bottom right corner of the scene) are flagged as high quality. Resulting bias in averaged V5 QA=0 LST for this cell (compared to in situ) is -3.4K (top right) versus 0.62K for V4.



**Figure 62:** Example of MODIS nighttime multi-spectral cloud imagery (top) and V4 (bottom left) and V5 (bottom right) LST QA flags for a 28 km AMSR-E grid cell at SK-HJP02 site (July 7 2003). Circles mark the tower locations. Small squares are the  $0.05^\circ$  CMG grid boxes on which the MODIS MYD11 product is reported. QA levels (0=00, 1=01 and 2=10) are indicated on the bottom maps. Invalid data are left blank. Measured nighttime in situ temperature on that day is 281.67 (top right) and difference with AMSR-E retrieved LST is 0.54K. It is seen in this figure that no V4 QA=0 estimate is produced for that day. However many V5 estimates are flagged as high quality even though conditions are overcast. Resulting bias in averaged V5 QA=0 LST for this cell (compared to in situ) is -5.7K (top right).



**Figure 63:** Daily time series of differences between satellite-derived LST and in situ air temperature measurements for FISod site (Site #1) in the daytime (top) and SK-HJP02 (Site #9) at night (bottom). In these plots, the grid averaged MODIS V4 data is represented by circles and MODIS V5 by diamonds. Filled symbols are used for QA=0 estimates, and half-filled for QA  $\geq 0$  (open symbols represent the lowest quality data – not used in this study). AMSAR-E estimates are marked as “\*”. Horizontal lines show the monthly mean differences between the QA=0 infrared satellite products and the in situ measurements (each mean value is represented a single point on the plots of Figure 60). Mean AMSAR-E LST differences are shown by the red (day) and dark blue (night) lines. This figure shows that QA=0 V5 estimates coinciding in time with a QA=0 V4 estimate tend to be biased low. Mean differences between V5 QA=0 and V4 QA=0 estimates is  $\sim -2$ K for the first case and  $-1$ K in the second case. It is also seen that the level of cloud contamination of the V4 QA $\geq 0$

estimates can be large, with differences reaching -15K occasionally (largely exceeding the nominal uncertainty values provided with the product - Table 6). Note that the examples shown in this figures do not constitute the worst cases (see Figure 60).

### 7.1.2.3 *Microwave versus infrared surface temperatures*

The temporal sampling bias with the AMSR-E data is obviously mostly insignificant since LST estimates are provided in both clear and cloudy conditions. The only missing days are due to gaps in the satellite coverage or to outliers removed by our R11 filtering. To a first order it can be assumed that the mean differences between the AMSR-E LST and in situ measurements (Figure 67) are entirely explained by representation errors. In this respect the microwave data is bound to inherit some of the characteristics of the infrared data since, by construction, the microwave emissivities (and in particular the 11V emissivities used here for our LST retrieval) are effectively such that the microwave derived surface temperatures match the clearest coincident V4 infrared estimates in an average sense.

In principle, microwave and infrared LSTs should be in almost perfect agreement when averaging the two products over “common” days (i.e. days for which coincident microwave and infrared data are available), instrumental and LST algorithm noise being the only factors accounting for the differences. However, several other factors must be accounted for:

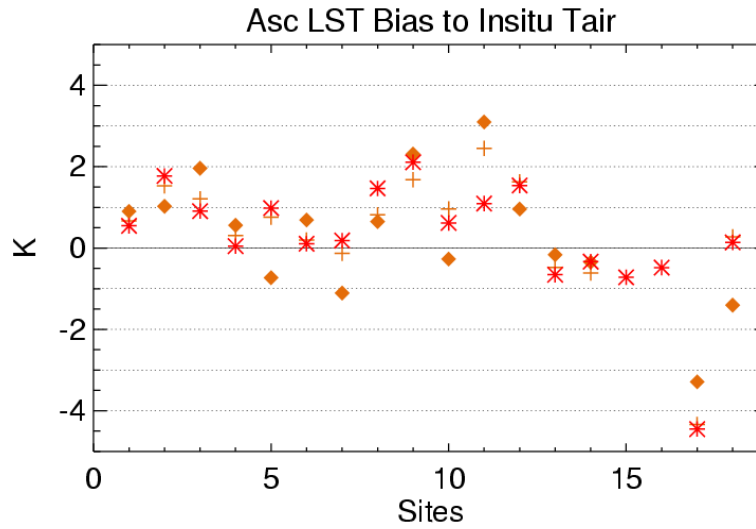
- 1) Surface inhomogeneities: in the current validation framework, the MW derived LSTs correspond to the average LST values over the MW footprint while the infrared temperature it is compared to, is an average over the nearest grid box. Surface inhomogeneities may cause significant differences between the LST estimated from the two data sources. These differences may vary in time because the microwave does not sample the exact same locations from day to day. We are looking at minimizing this issue in the future by mapping the microwave channels to a common effective footprint on a fixed Earth grid and by weighing the high resolution IR LST by the effective antenna pattern, thereby ensuring a coherent and consistent treatment of IR and MW data sources.
- 2) Under-analyzed cloud cover: (most likely to arise in partly cloudy conditions) might cause abnormally low MODIS LSTs even after spatial averaging. Note that these outliers are generally captured and filtered out by our QC procedure in the process of estimating our microwave surface emissivities but, in this validation experiment, all MODIS estimates are used.
- 3) Clear bias in partly cloudy conditions: mean IR QA=0 LSTs used in this validation experiment are constructed by averaging together only the clear pixels, with no restriction on the minimum number of clear samples. Surface temperatures at those clear locations tend to be warmer during the day and colder at night than the average temperature of the scene which may cause significant differences between the AMSR-E and in situ estimates in mostly cloudy conditions. For AMSR-E emissivity estimation we select the FOVs with the largest fraction of clear measurements, thereby minimizing the clear bias.

It is apparent from Figure 64 and Figure 65 that mean daytime and nighttime AMSR-E LST estimates are generally closer to the in situ measurements than the MODIS LSTs (a comparison of mean A-D LST estimates is provided in Figure 66). Exceptions include Sites #5, #8, but even in these cases the mean AMSR-E and in situ LST estimates don't differ by more than ~2K. It is seen in these figures that

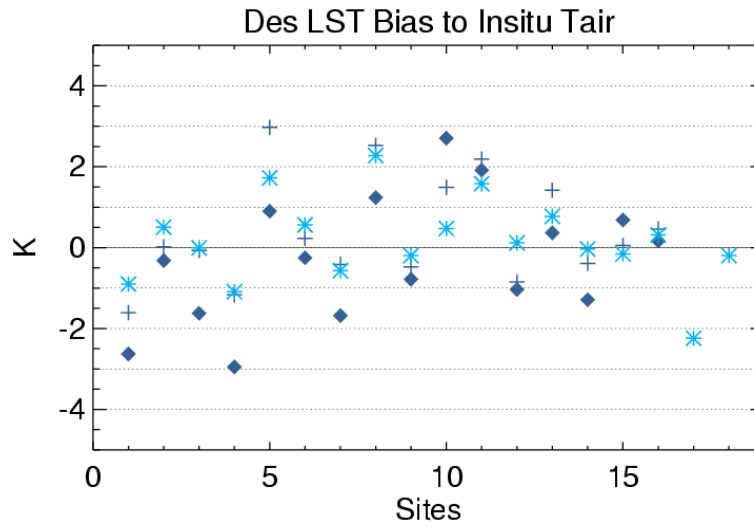


differences between mean microwave and in situ temperatures tend to decrease when averaging the daily differences over the entire month as opposed to averaging only over the days when coincident MODIS and AMSR-E observations are available. Obviously one expects the impact of random differences to be minimized by averaging over a longer time period, but the fact that the magnitude of both thermal inhomogeneities and representativeness errors is smaller in overcast conditions than in the clear-sky (when IR data is produced) is another factor.

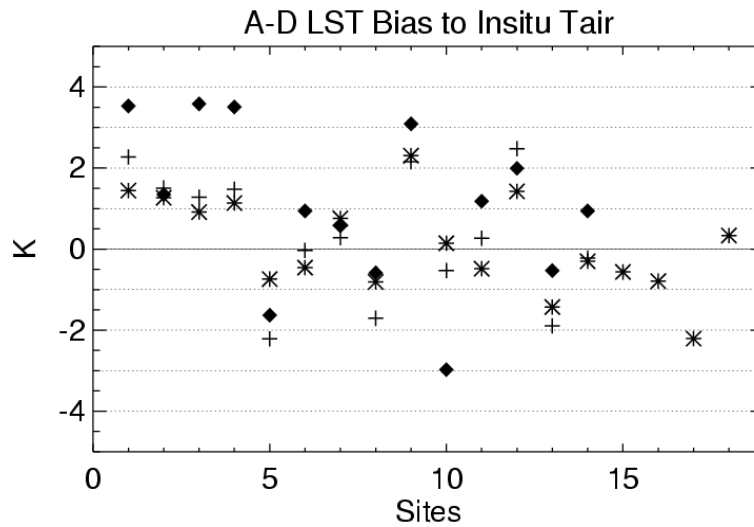
It is apparent in Figure 67 that the total monthly mean AMSR-E minus in situ temperature differences are usually within 1.5K both in the daytime and nighttime, approaching or slightly exceeding 2K only at a few sites (Site #2 and 9 in the daytime; Sites #5, 8 and 11 at night). This performance is to be compared with the plot of Figure 59. It is clear by examining all 18 sites, that the MW retrievals are more robust than the IR retrievals. In particular, the MW technique does not suffer from the sampling problem that negatively impacts the infrared (i.e. large sampling errors resulting from low yield – as with V4 QA=0 – or significant impact of cloud contamination when increasing the yield – as in V5) and leads to discrepancies between IR and in situ measurements commonly exceeding 2K and approaching or exceeding 4K in over a fair number of sites. Note that site #17 (ID-Pag) is problematic for both infrared and microwave.



**Figure 64:** Average daytime MODIS V4 QA=0 LST minus in situ temperature differences (i.e.  $\text{MODIS } \overline{\delta T_{\text{insitu}}^*}$ ) (diamonds) vs. AMSR-E minus in situ differences averaged over days for which a MODIS estimate is available (“+”) vs. total AMSR-E minus in situ differences ( $\text{AMSR-E } \overline{\delta T_{\text{insitu}}^*}$ ) (“\*”) for all 18 FLUXNET sites for July 2003.

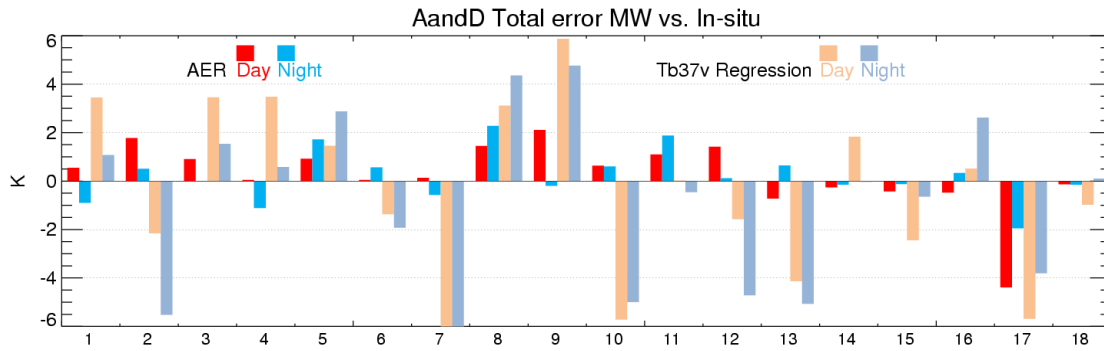


**Figure 65:** Same as Figure 64 for nighttime



**Figure 66:** Same as Figure 64 for A-D LST

For the sake of completeness we added to this comparison the results of applying the global single channel 37V regression developed by Holmes *et al.* (2009) for the retrieval of land surface temperature from AMSR-E. Clearly, this algorithm does not perform nearly as well as our 11V retrieval when compared to the independent in situ observations at the 18 FLUXNET sites. The proposed global regression technique does not consistently improve on the MODIS performance, despite the fact that the latter is more impacted by clouds. This is particularly true for the higher latitude sites where the performance of the regression is worse than both V4 and V5 MODIS (at night and during the day) over ~50 % of the cases. This result illustrates the necessity to account for the spatial variability in surface emissivity and to update these emissivities as needed, as is done with the AER method, in order to maintain consistency over time with the clear MODIS measurements and improve on the performance of infrared techniques in cloudy areas.



**Figure 67:** Monthly mean daytime (red) and nighttime (blue) AER AMSR-E minus in situ surface temperature differences at the 18 FLUXNET sites for July 2003. Also shown in this figure is the daytime and nighttime performance of the Holmes *et al.* (2009) single channel global regression in light brown and light blue, respectively.

### 7.1.3 Regional scales analysis

Although the main objective of the analysis described in the previous section was to validate our microwave surface emissivities, this type of analysis is also useful for providing an idea of the performance achievable with infrared and microwave surface temperature retrieval techniques.

As part of the above study we highlighted significant discrepancies between the V4 and V5 MODIS products and described the impact of including lower quality V4 data to fill in the gap in spatial and temporal MODIS V4 LST coverage. Based on the results of Sec.7.1.2.2, it was decided not to upgrade to MODIS V5 for the production of our surface emissivities. The data analyzed in the previous section, indicates that V5 uses a looser criterion than V4 for flagging data as high quality. Having too many even mildly cloud-contaminated LST estimates makes our clustering algorithm less efficient at removing outliers. Some authors also cautioned against using the V5 product over deserts (Hulley and Hook, 2009) and against mixing V5 and V4 data. We also showed that our simple 11V based microwave LST estimation technique performs well when compared to in situ data over vegetated surfaces, a somewhat surprising result when considering the potential impact of representativeness errors. The success of the microwave retrieval technique is mainly attributed to its weak sensitivity to (non-precipitating) clouds and to the effectiveness of the QC procedure, used in the process of constructing our emissivity atlas, at screening out potentially bad data. However, the analysis described above remains limited by the short time period analyzed (we have plans to perform multiyear analysis in the future and cover more seasons) and the small number of FLUXNET sites available worldwide.

Our broader objective is to extend our performance analysis to larger regional scales. It is important to note that we are not looking to measure the performance of the microwave derived LST's by comparing its product to some "truth" (for this type of validation study it is preferable to e.g. compare high spatial resolution infrared measurements to ground based interferometric measurements) but simply ensuring that the microwave algorithm performs *as expected* when compared to infrared LST products on similar spatial scales. Note that the results presented in this section are still preliminary. In particular, up to this point we have ignored precipitation in our LST retrievals and assumed that in the regions analyzed its effect is sporadic and would not significantly affect our monthly statistics. In the future, precipitation can be identified by running the full 1D-VAR inversion (with constraint on surface emissivity derived from clear-sky analysis) on multi-channel observations. The expectation is that the direct impact of hydrometeors on the microwave measurements (not taken into account in a

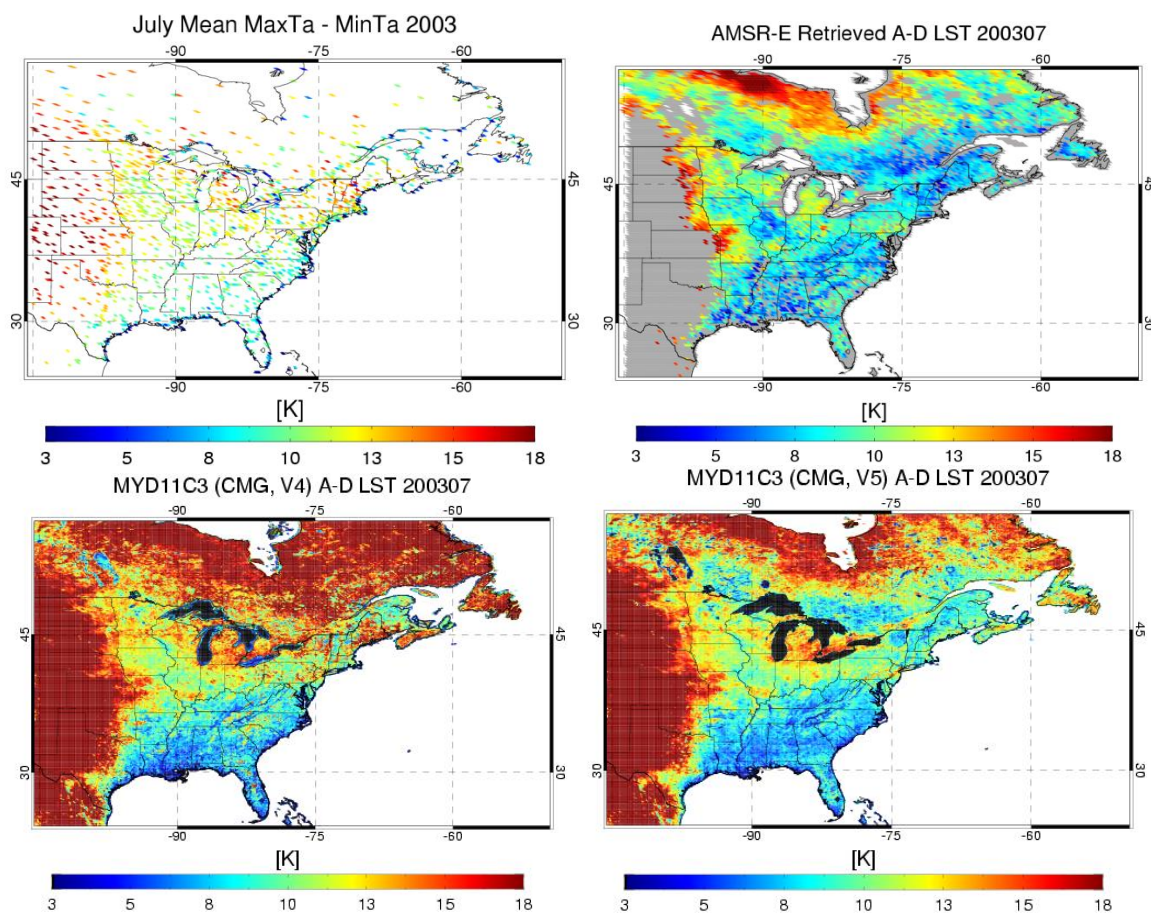
non-scattering RT model) combined with the impact of surface wetness on its emissivity would prevent the retrieval scheme from converging.

It is noteworthy that the main differences between microwave and infrared LST estimates highlighted in the above study are also reflected in the global comparison between the two sources of LST (not shown). However, for a detailed validation it is best to work on regional scales. The planned approach is to divide the globe into a number of regions (including Eastern US and Canada, Europe, China and the tropical belt, e.g. Amazon basin, African rainforest and South East Asia), identify anomalies in the MW or IR LST difference fields and attempt to correlate those with surface features or known surface/atmospheric dynamics. In such comparisons, it is useful to bring as many surface observations as possible. Although as mentioned before one cannot always rely on those observations as a source of truth, this information can help identify possible causes for the observed discrepancies in the remotely sensed data.

Figure 68 shows an example of monthly mean A-D LSTs (a proxy for diurnal surface temperature cycle amplitude) derived from AMSR-E and MODIS V4 and V5 (from the 0.05° MYD11C product) over the Eastern US and Canada for July 2003. For reference, we added maps of mean daily maximum – minimum surface air temperatures reported by weather stations. In this figure, the Western part of the US is masked out in the AMSR-E map because this region corresponds to high penetration surfaces treated with the 1b algorithm. Note that our AMSR-E emissivities are retrieved using the daily MODIS V4 MYD11B product. A map of monthly mean MODIS V4 (QA=00) A-D LST from the MYD11B product and differences with mean AMSR-E A-D LST is included in Figure 69 for comparison.

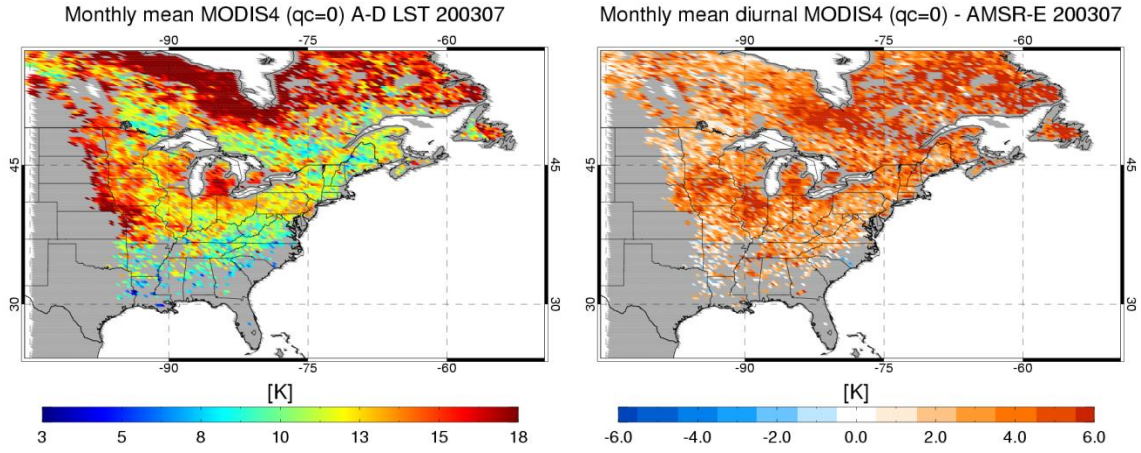
Maps of Figure 68 and Figure 69 highlight similar features in the MODIS V4 mean LST fields. Here, the MYD11B QA=0 data is remapped to the AMSR-E 28 km grid prior to performing the temporal averaging. The MYD11C product is constructed as a monthly average of all quality levels, which explains its higher spatial coverage (e.g. Southeastern United States which is frequently cloudy at this time of the year and mostly missing in Figure 69). It is apparent from these figures that MODIS V4 produces systematically larger A-D LSTs than AMSR-E over Quebec and Northern Ontario, Michigan and the Corn Belt region. As shown in Fig. Figure 70, the  $\delta T_{MW}^*$  term is relatively small both at night and during the day. The largest negative values occur mostly at night over Canada and may be related to the high density of lakes present in those regions (note that grid points located in the vicinity of large water bodies are screened out by our quality control – Sec. 4.3 describes our plans for dealing with this aspect and improve the algorithm yield in Version 2). However, there is no coherent large scale pattern in this region that would indicate a clear correlation between the  $\delta T^*$  anomalies and the biases identified above. Our analysis of the Northern US and Canada reveals that sampling errors (Fig. Figure 70) explain much of the observed differences between MODIS and AMSR-E mean A-D LST. As expected, MODIS sampling errors are generally positive during the daytime and negative at night, which results in a positive bias in the MODIS A-D LST estimates exceeding 5K locally. The largest daytime positive differences are observed in Eastern Canada, and in Southern Michigan, while the largest negative nighttime differences are found in a band stretching eastward from Northern Ontario to Newfoundland and in the Corn Belt region (this latter region is analyzed separately later in this section and in Sec. 7.2). Note that these generally correspond to regions where the number of clear daytime or nighttime observations in July 2003 is less than 15 (Figure 70). However, a perfect correlation between frequency of clear observations and magnitude of the sampling error cannot be expected as other factors (such as cloud type, vegetation, humidity...etc) influence the response of surface temperature to cloudiness.

We also examined the sampling and estimation errors for the MODIS V5 LST product. These are shown in Figure 71 for the region of interest. It is apparent when comparing this figure with Figure 70 that the magnitude of the V5 sampling errors (using AMSR-E as reference) is significantly reduced compared to V4. This result is consistent with our previous analysis (Sec. 7.1.2.2) and simply reflects the fact that V5 produces a larger number of observations flagged as high quality at most locations. However in the case of V5, the  $\delta T^*$  term is largely negative over the entire region, which if we refer to the analysis performed at the FLUXNET sites is attributed in part to intrinsic differences between the clear V4 and V5 products but also the larger degree of residual cloud contamination accompanying the increase in number of observations (Sec. 7.1.2.2). The systematic differences between MODIS V5 and AMSR-E mean daytime and nighttime LST estimates are significant but they partially cancel out in the computation of the mean A-D LST which results in a V5 estimates being overall closer to AMSR-E product than V4 (Figure 72).



**Figure 68:** Monthly mean AMSR-E (top right) and MODIS MYD11C3 (V4 and V5) (bottom) A-D LST compared to average daily maximum – minimum reported air temperature from surface stations (top left) across the Eastern US and Canada for July 2003.





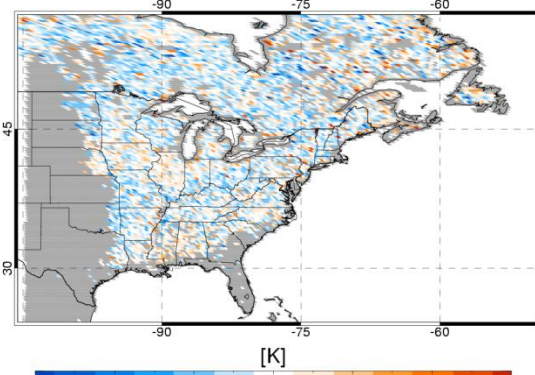
**Figure 69:** Monthly mean MODIS V4 (QA=0) A-D LST from MYD11B product remapped to AMSR-E 28 km grid (left) and difference with mean AMSR-E A-D LST (right).

As part of this validation study, we performed a cursory comparison between satellite products surface air temperature reports from ground stations. Characterizing the differences between those two sources of data may be tedious and often inconclusive. Over data dense areas one cannot possibly examine all the ground/satellite observation matchups and we instead use an automated analysis to isolate obvious anomalies in the temperature difference fields which then allows us to focus on a few selected sites to help explain the observed discrepancies. Here we only provide a high level summary of our conclusions. In Figure 73 we compare the infrared sampling errors computed using AMSR-E as our reference,  $\mathcal{G}_{MW}$ , to those computed based on hourly surface station reports,  $\mathcal{G}_{insitu}$ , for both ascending and descending overpasses. It is apparent from this figure that  $\mathcal{G}_{MW}$  and  $\mathcal{G}_{insitu}$  correlate well spatially, which gives us reasonable confidence in the AMSR-E-based results. We also examined the mean daily differences between AMSR-E retrieved LST and surface air temperatures reported by the surface stations. Because we expect the biases between infrared surface skin temperature measurements (used in the derivation of our AMSR-E surface emissivities) and surface air temperature to be partially reflected in the microwave AMSR-E LST estimates, and because surface skin-air temperature contrast tends to be more pronounced in the clear-sky than under cloudy conditions, we consider two terms separately:

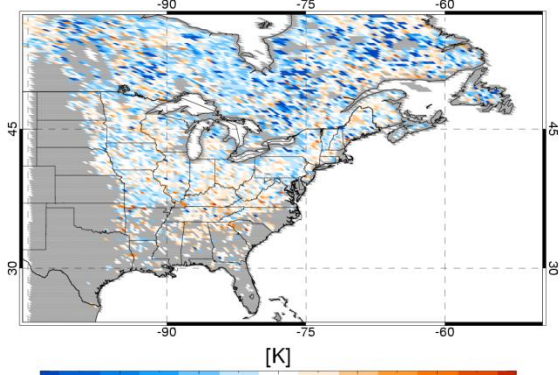
- $\overline{\delta T_{MW-insitu}^{clr}}$ , obtained by averaging the daily temperature differences over the clear days only (per MODIS cloud mask), and,
- $\overline{\delta T_{MW-insitu}^{cld}}$ , obtained by averaging the differences over the remaining cloudy days.

As shown in Figure 74, day time  $\overline{\delta T_{MW-insitu}^{clr}}$  is generally positive, i.e. clear-sky AMSR-E LST estimates are warmer than the in situ surface air temperatures. Differences commonly reach over +3K over moderately to highly vegetated areas (East of  $\sim 95^\circ\text{W}$ ). If the surface air temperature measured at those sites was closer to the actual microwave emission temperature than the MODIS skin temperatures (i.e. microwave emissivities computed using MODIS LST were too low), those positive biases would persist during the cloudy phase. The fact that they tend to vanish in cloudy atmospheres at many locations indicates that the AMSR-E surface emissivities derived at those locations using MODIS  $T_{skin}$  as an approximation of the microwave emission temperature are probably correct. At night, as expected, the  $\overline{\delta T_{MW-insitu}^{clr}}$  values are generally negative or neutral with magnitude seldom exceeding  $\sim 1.5\text{K}$  over the region of interest.

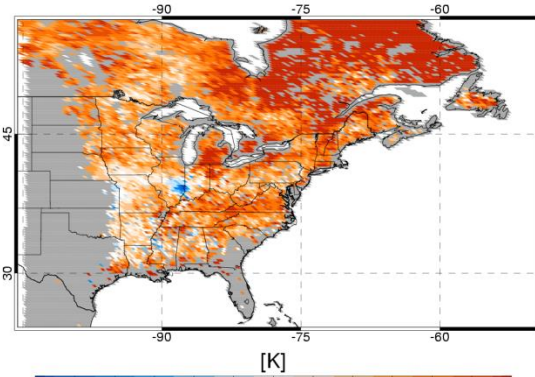
Asc commonday mean MODIS4 (qc=0) - AMSR-E 200307



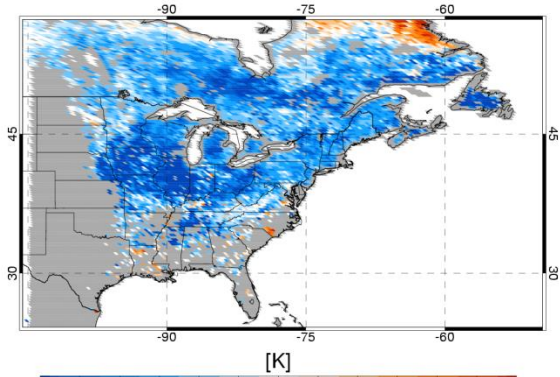
Des commonday mean MODIS4 (qc=0) - AMSR-E 200307



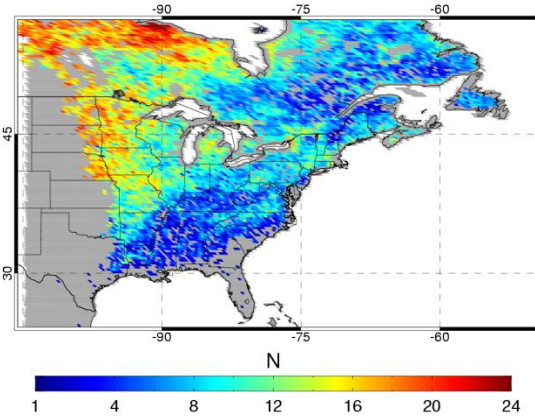
Asc monthly mean MODIS4 sampling error 200307



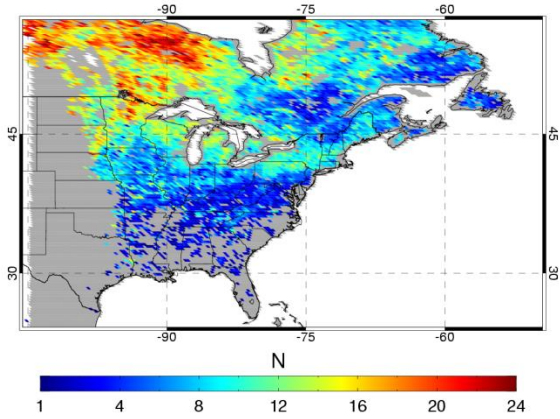
Des monthly mean MODIS4 sampling error 200307



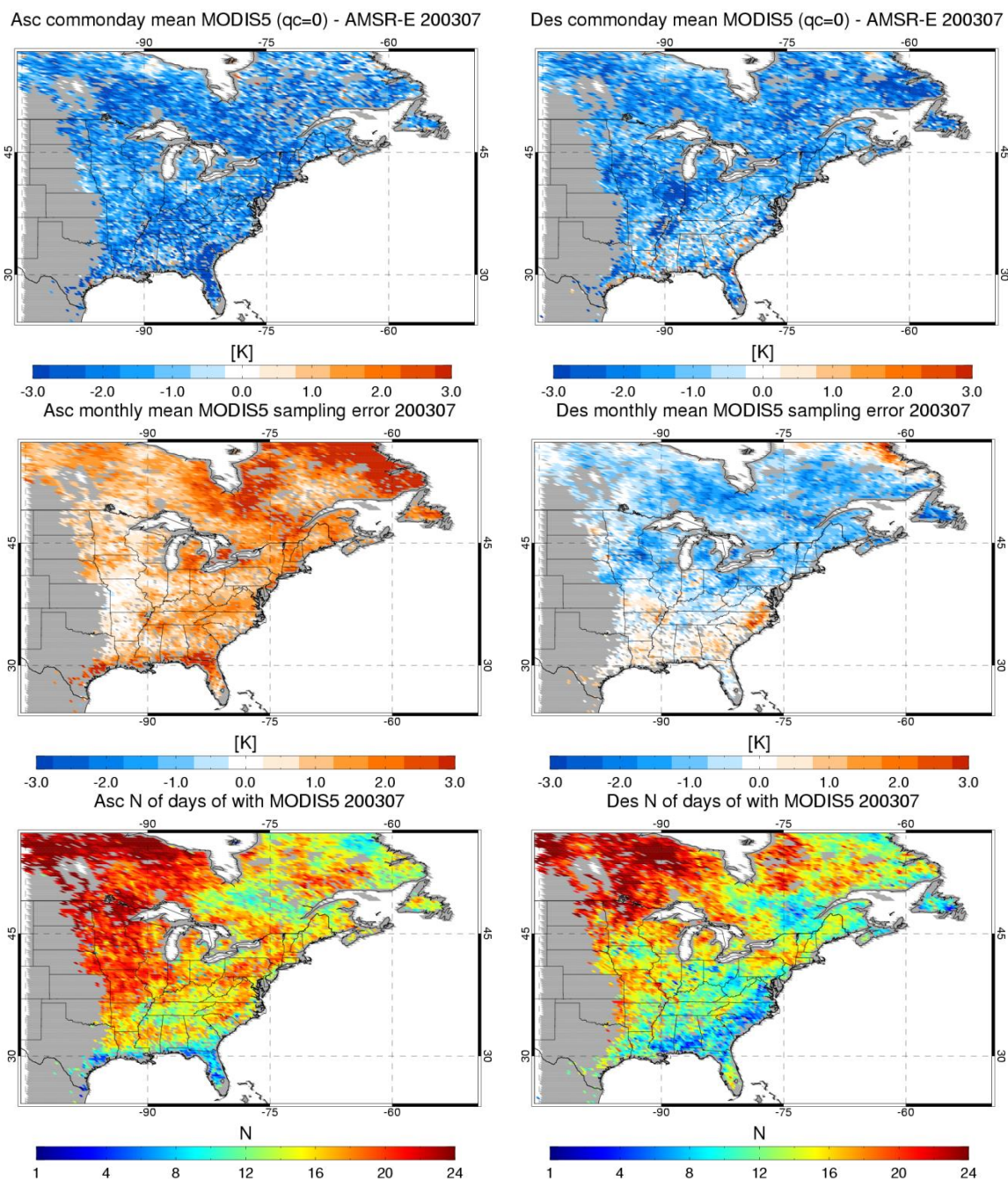
Asc N of days of with MODIS4 200307



Des N of days of with MODIS4 200307

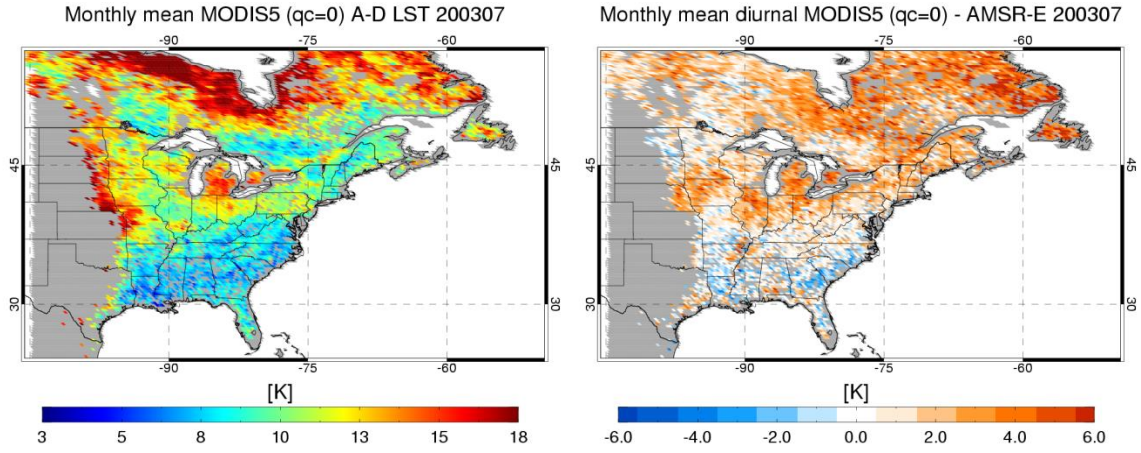


**Figure 70:** MODIS daytime (left) and nighttime (right) sampling bias (top) and  $\delta T^*$  (middle) over Eastern US and Canada evaluated using AMSR-E as a benchmark for July 2003. Bottom plots show the number of days in the month for which MODIS (QA=00) data is available.

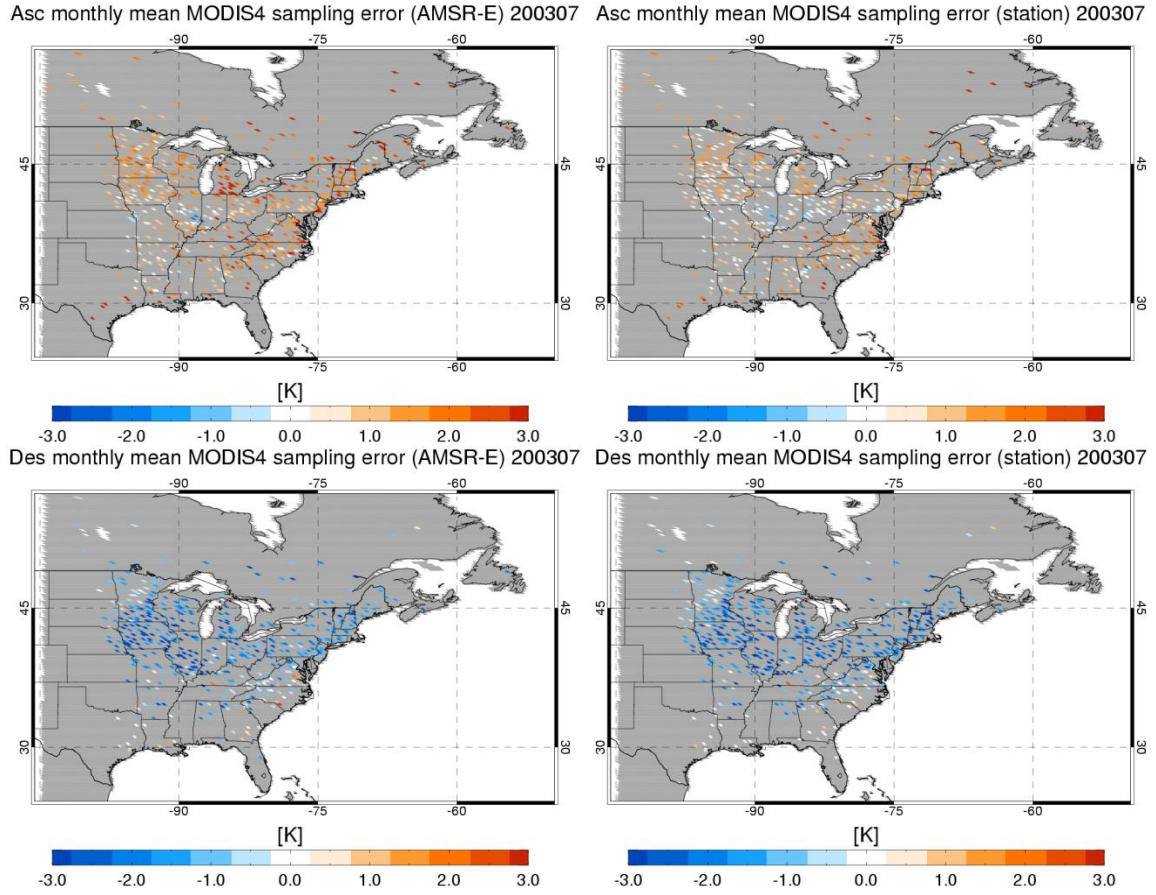


**Figure 71:** Same as Figure 70 for MODIS V5 LST product





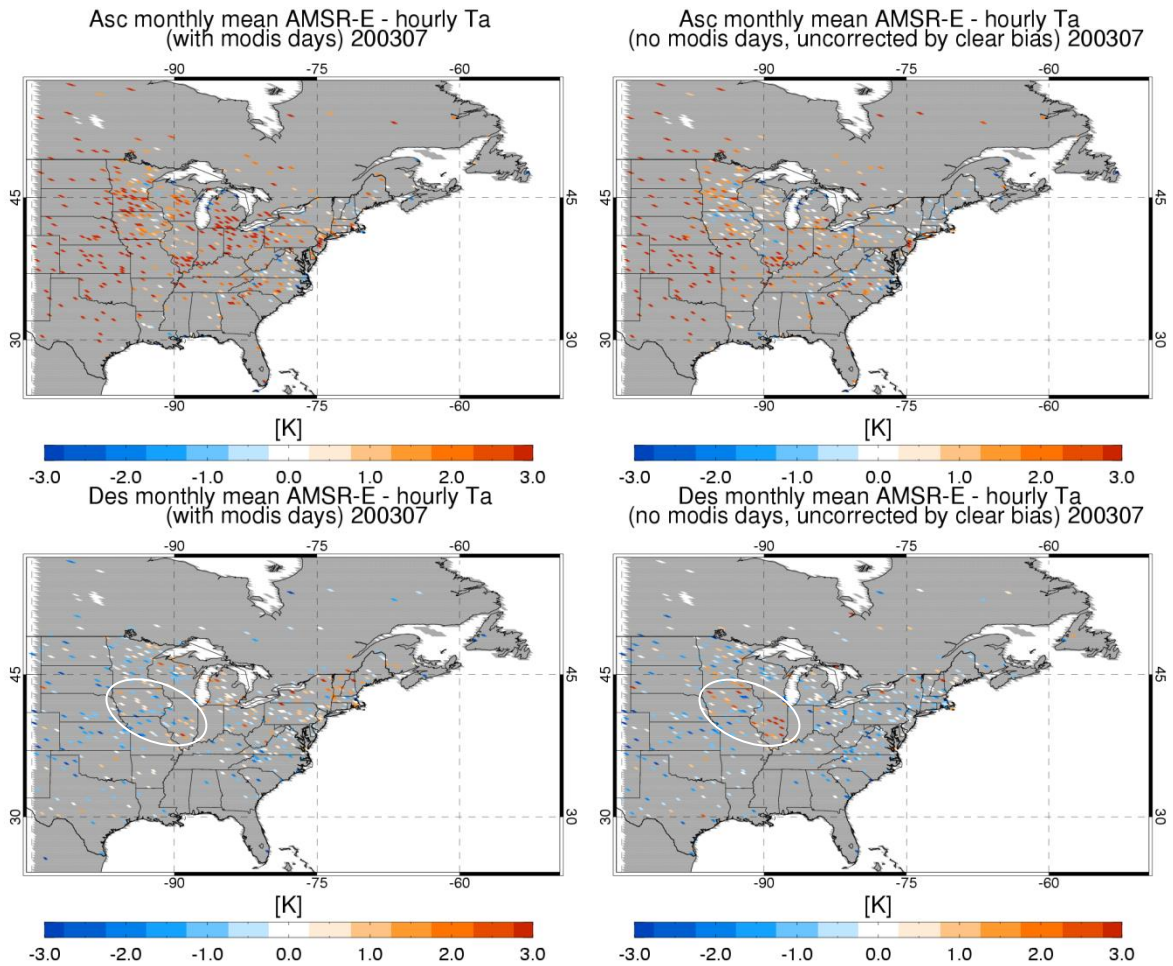
**Figure 72:** Same as Figure 69 for MODIS V5 LST product



**Figure 73:** Daytime (top) and nighttime (bottom) MODIS V4 Q=0 sampling error computed using AMSR-E (left) and surface air temperature data (right) as reference for July 2003.

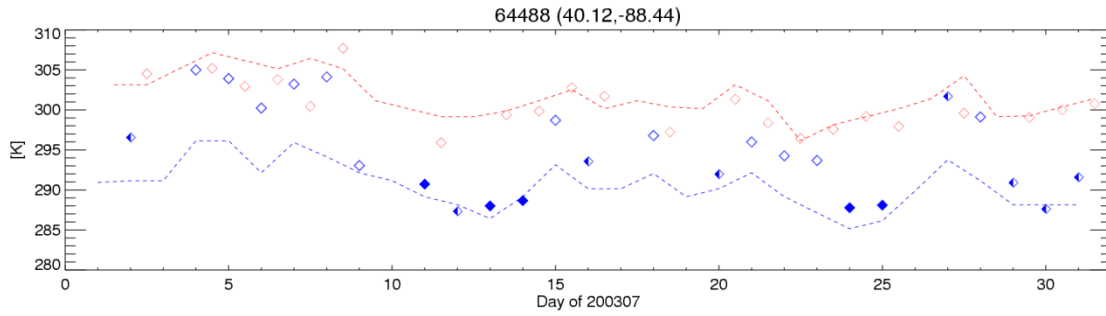
Note that over a number of sites within the Corn Belt region (see circled region in Figure 74) the  $\delta \bar{T}_{MW-insitu}$  term goes from being slightly negative or positive in the clear-sky to significantly positive under cloudy conditions. This particular anomaly is not visible in the daytime plots and coincides with a region of maximum day-night retrieved AMSR-E emissivity differences (Sec.7.2). In our approach,

microwave surface emissivities are evaluated using clear-sky observations and we assume that emissivity does not change under cloudy conditions (unless precipitation occurs). Comparison between the daily AMSR-E LST estimates and the max/min air temperature records at one of those sites near Champaign, IL, is shown in Figure 75. It is seen in this figure that while the clear AMSR-E LST estimates agree reasonably well with the minimum air temperature reports during the clearest nights (and the daytime LSTs are consistent with the reported daily air temperature maxima), retrieved cloudy nighttime LST appear to be too warm and close to the daytime temperatures. This particular behavior is indicative of a change in surface emissivity coinciding with the transitions between clear and cloudy phases. This particular artifact is analyzed in more details in the next section and appears to correlate with nighttime dew deposition on corn fields.



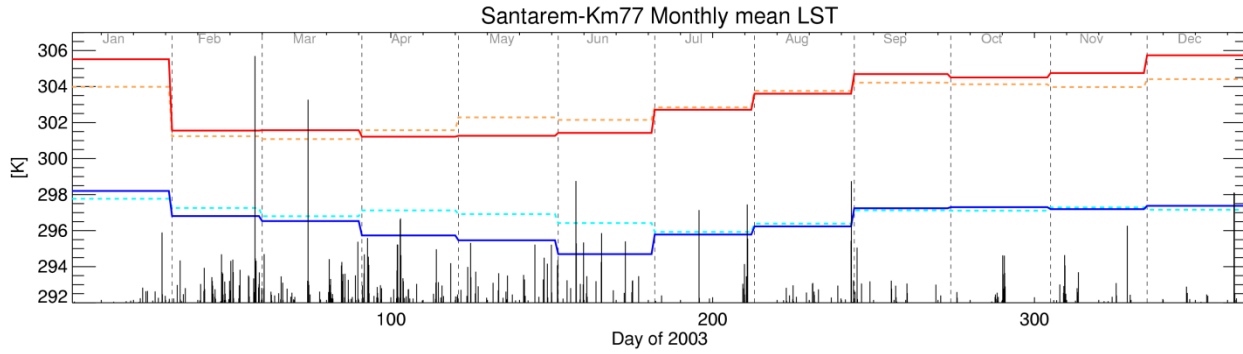
**Figure 74:** Maps of AMSR-E LST minus station air temperature data averaged over clear days (left) and cloudy days (right) for day time (top) and nighttime (bottom) in July 2003



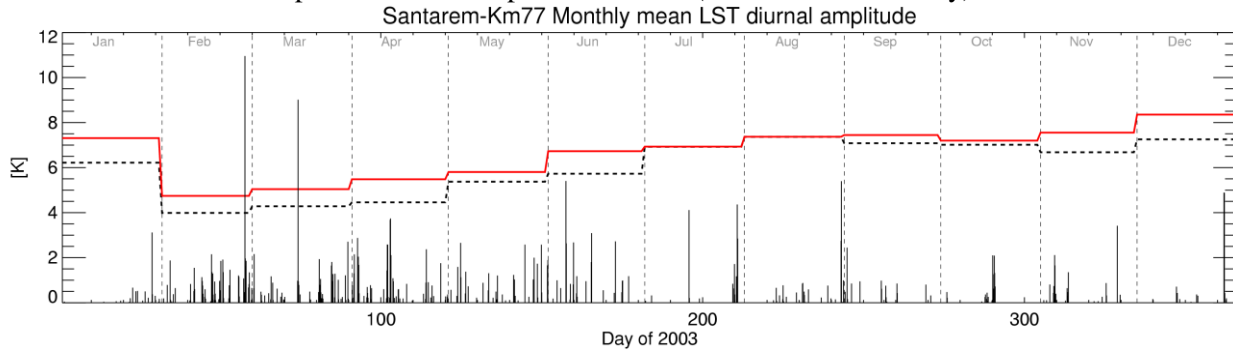


**Figure 75:** Time series of AMSR-E daytime (red diamonds) and nighttime (blue diamonds) LST near Champaign, Illinois. Filled nighttime symbols correspond to clear conditions, open symbols, to overcast and half-filled symbols to partly cloudy conditions. Red and blue dotted lines represent daily maximum and minimum temperature reports from nearest surface station. There is a good correspondence between daytime AMSR-E retrievals and maximum temperature data. At night, the agreement between AMSR-E LST and minimum surface air temperature is also good in the clear-sky but AMSR-E LSTs are generally biased high (and closer to the daily maxima) in cloudy conditions, indicating that 11GHz emissivity may be different under clear and cloudy skies.

We alluded earlier in Sec. 5 (in the context of our discussion of the performance of the R11-based RTS time segmentation approach over tropical forests), to the fact that our new RTS approach, which emphasizes the use of observations from the driest (hence clearest) season in deriving an emissivity estimate that is representative of the whole year (no appreciable variation in R11 is observed in those areas over that entire period), might mask subtle seasonal emissivity variations linked to the precipitation cycle. In order to verify this hypothesis we extracted the daily LSTs retrieved from the AMSR-E 11V measurements (by fixing the 11V emissivity to the value obtained with the RTS approach) at the Santarem site in the Amazon basin, and compared the monthly mean daytime and nighttime values to the mean 11.3 m in-situ air temperature reported at this site. The results of this comparison are shown in Fig. Figure 76. It is seen in this figure that we have almost perfect agreement between the two datasets at night for the months of January through March and July through December. From April to June, the mean nighttime LSTs from AMSR-E appear to underestimate the in situ measurements by 1.5-1.7 K. In the daytime, the results are only slightly different. In this case, AMSR-E LSTs appears to be higher than the in situ values during the N. Hemisphere fall and winter months, with positive differences reaching a maximum of 1.5 K in January. A negative daytime bias coincides in time with the negative nighttime bias but only reaches a maximum of -1K in May. While it appears that there might be a trend in the mean LST differences, the magnitude of this trend according to this figure does not exceed 2-2.5K (peak to peak). Assuming that one could neglect the contribution of other sources of discrepancies between the two data sets (which is somewhat unrealistic), the impact shown in these figures would correspond to a variation in the actual surface emissivity over the entire year not exceeding 0.0085 in the daytime and 0.0065 at night. However, one also has to account for measurements errors, as well as other factors such as the differences between the temperature of the top of the canopy measured by the infrared sensor and the air temperature within the canopy which is probably the largest during the daytime and during the drier season. For instance, the latter could contribute to the positive daytime biases in the AMSR-E LSTs observed in November, December and January. Note that all sources of error considered, the microwave estimates reproduces well the seasonal variations in monthly A-D temperatures and differs from the in situ estimates by no more than 1K ( Figure 77) which is quite good compared to our expected performance in other regions. We are obviously planning on conducting more in depth examination of our hypothesis in the future but based on the cases analyzed to this date, we have no elements suggesting a need to refine our current R11-based time segmentation approach in the context of LST retrievals over forested areas.



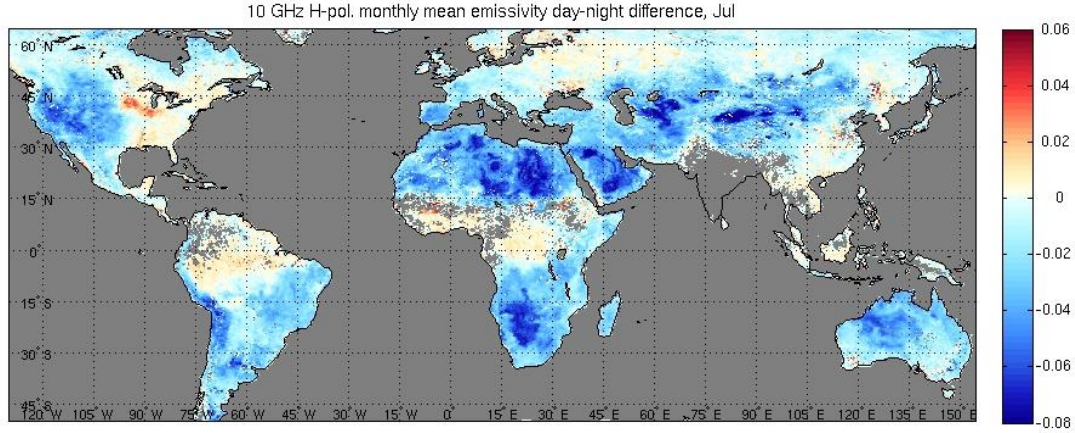
**Figure 76:** Monthly mean daytime (red) and nighttime (blue) AMSR-E LST (solid) versus 11.3m air temperature measurements (dashed) at the Santarem (Brazil) site for year 2003. Normalized rainfall from TRMM is overlaid to the plot. Maximum reported rain rate (at the end of February) is 20mm/h.



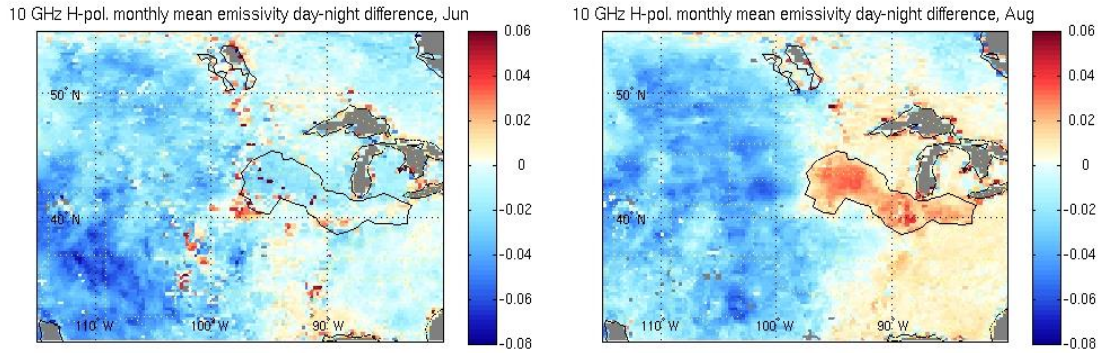
**Figure 77:** Monthly mean D-A surface temperature from AMSR-E (solid) and in situ measurements (dashed) for year 2003.

## 7.2 On the potential impact of dew on microwave surface emissivity

As part of our emissivity quality analysis, we have investigated observed positive day-night emissivity differences over Midwestern US in the summer months. A few large regions with persistent systematic positive day-night emissivity differences - i.e., daytime 11 GHz V-pol. emissivity more than 0.01 greater than nighttime on a month-average basis - occur in the 2003 dataset (Figure 78). One particularly distinct region occurs in North America only in July and August and stretches roughly from southern Minnesota to Ohio (Figure 79).



**Figure 78:** July 2003 day-night 11 GHz H-pol. month-average emissivity differences.



**Figure 79:** June (left) and August 2003 day-night 11 GHz H-pol. month-average emissivity differences. The region where large positive differences occur in July and August (but not in June or September) is outlined. Large negative differences are due to microwave penetration over deserts (Moncet et al., 2011.; Galantowicz et al., 2011).

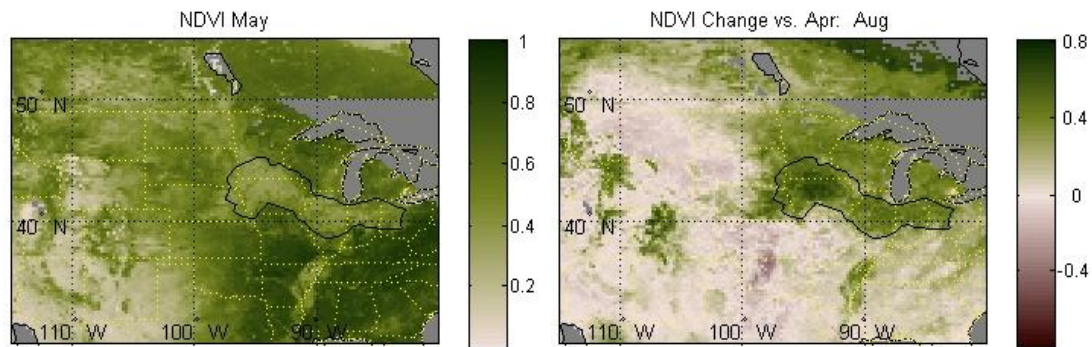
These regions are of interest because they are likely to indicate either an emissivity product quality control issue or a phenomenon that systematically alters emissivity between day and night. Possible sources of emissivity retrieval error include: (a) A combination of nighttime fog and temperature inversion that results in invalid nighttime MODIS LSTs (warm-biased), (b) daytime residual cloud contamination that results in invalid daytime MODIS LSTs (cold bias), (c) effective emitting temperature biases due to the differences between the MODIS and AMSR-E viewing geometries, and (d) AMSR-E day-night calibration errors or EIA effects. Possible sources of emissivity change include: (a) frequent nighttime dew on vegetation that depresses emissivity vs. drier daytime, (b)

intercepted precipitation on vegetation that occurs more frequently at night and depresses emissivities, and (c) wetter nighttime soils that depress emissivity. A determination of the source or sources of the observations is necessary to resolve questions about how to best use the dataset in particular for surface temperature retrievals, including: Can day and night emissivities be taken as equivalent but noisy samples of the same true emissivity? Is the day-night average the best estimate of the true emissivity at any time? Are day or night emissivities demonstrably superior? Should only one type be used as the best estimate at all times of day? What are the time scales of phenomenon affecting emissivities and what type of filtering methodology should be used to predict the most probably emissivity at a given time?

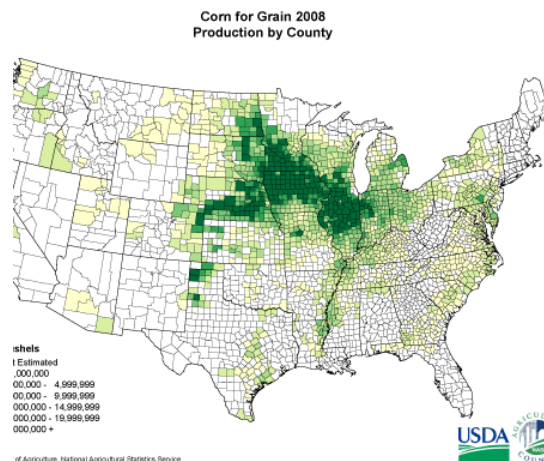
To explore these questions, we have investigated the spatial extent of the day-night emissivity differences over time and compared them to other parameters such as NDVI, crop type, and precipitation, and examined time-series of emissivities and other parameters for grid points typifying positive, negative, and near-zero emissivity difference. Several key findings have helped narrow the range of probable effects:

1. The day-night differences in July and August (hereafter the “day-night anomaly”) are highly correlated to the spatial extent of corn crops, NDVI, and the change in NDVI over the growing season. Early in the growing season (May NDVI, Figure 80) the day-night anomaly region has lower NDVI than surrounding areas while at the peak of the growing season (August NDVI, Figure 80) NDVI in the region has increased more than other areas. The high concentration of corn crops in this region (Figure 81) explains the large NDVI change. Although it is clear that neither corn crops, high NDVI, or NDVI change are the sole cause of the day-night anomalies, the high correlation to corn crops suggest that we can use corn as a model land surface type in further day-night anomaly analyses.
2. We have found no correlation between the day-night anomaly and precipitation patterns. For example, Figure 82 shows days with precipitation greater than 1 mm for July and August, 2003. Furthermore, the number of days with significant precipitation is too low in many areas to produce a persistent day-night anomaly throughout both months. Nevertheless, precipitation remains an important source of shorter-term emissivity change that may more frequently affect nighttime observations than daytime.
3. Data time-series have been examined for selected grid points exhibiting the day-night anomaly (“red” points in Figure 79) neutral day-night conditions (“white”), and negative day-night emissivity differences (“blue”). (Note that “blue” points have been attributed to diurnally varying thermal gradients and large emission depths in dry areas and are being treated using a thermal modeling solution incorporated in our baseline processing as the “1b” algorithm.) Figure 83 and Figure 84 are a set of example time series for Benton County, Iowa, which is typical of “red” points with positive day-night emissivity differences. First, the day-night emissivity change is nearly identical at V and H polarization and there is little systematic day-night difference in either the emissivity or brightness temperature polarization ratio relative to the gradual changes in polarization ratio likely attributable to vegetation growth. This suggests that the cause of the day-night anomaly is neither a polarizer (e.g., a change in the soil surface dielectric constant) nor a factor that strongly reduces polarization (e.g., absorption by the plant canopy that obscures the surface). Second, the *daytime* emissivity is more stable (around 0.95 at V-pol.) over time than *nighttime*. Again, the more gradual change in emissivity particularly at H-pol. is likely due to vegetation growth. This also suggests that daytime cold-biasing of LST by undetected clouds is unlikely. Furthermore, V-pol. emissivities at nearby “neutral”

grid points (near zero day-night difference) with similar NDVI are also close to 0.95, which suggests that the (a) higher daytime emissivity value is correct and (b) variation in the surface media dielectric constants due to temperature changes are not a primary factor. Third, nighttime emissivity anomalies are predominantly positive—bringing them closer to the daytime emissivities—and are associated with positive nighttime LST anomalies but not variations in AMSR-E observation time or earth incidence angle. This suggests that (a) nighttime emissivity estimates are responding to surface changes related to higher LST or (b) a separate phenomenon is increasing LST estimates and emissivities simultaneously. Finally, similar time-series at 19 and 37 GHz (not shown) suggest that the causes of the day-night anomaly and positive nighttime emissivity anomalies are frequency neutral (e.g., not caused by surface dielectric changes).

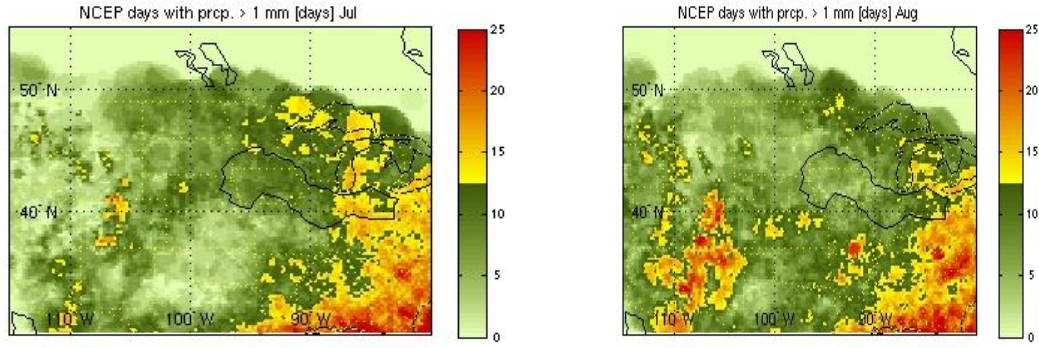


**Figure 80:** May NDVI (left) and August NDVI anomaly relative to April.

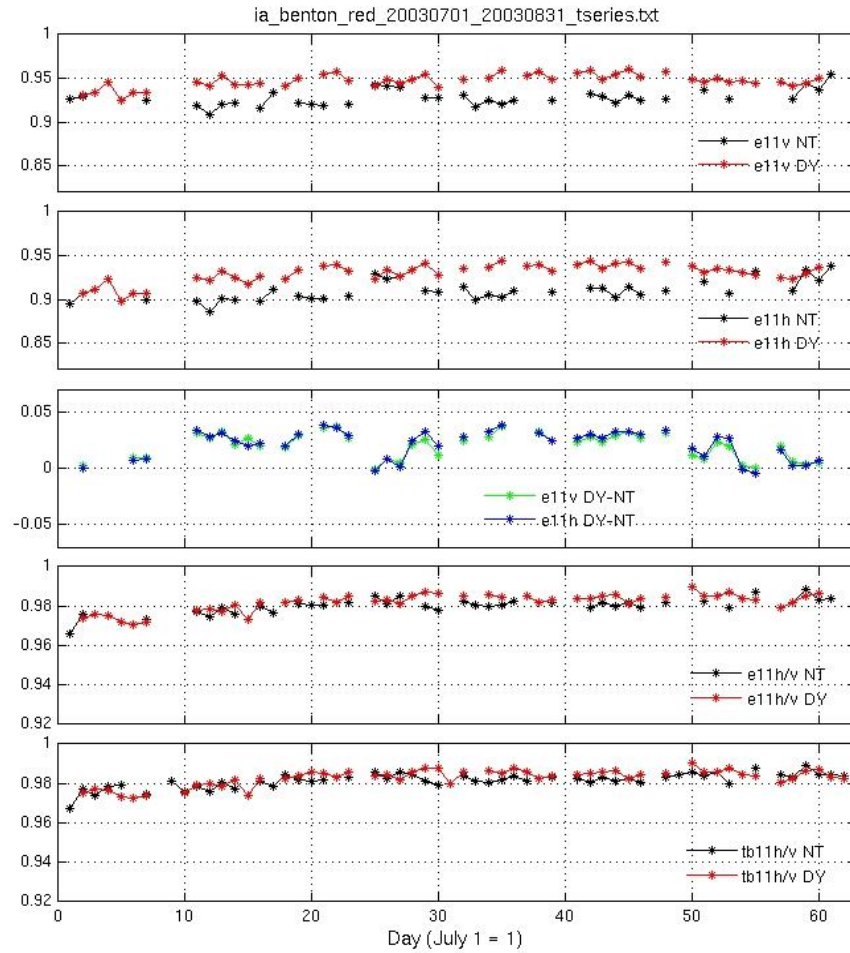


**Figure 81:** Corn yield by county for 2008. Note that yield is not normalized by county size.

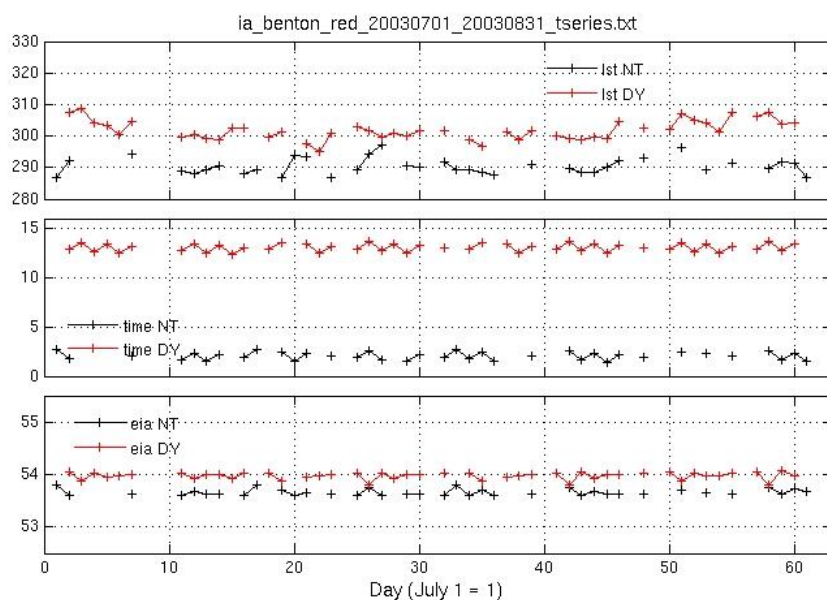




**Figure 82:** July (left) and August days with precipitation > 1 mm (from NCEP analyses).



**Figure 83:** Time-series over July and August for Benton County, Iowa. From top: 11 GHz V-pol. day (DY) and night (NT) emissivities; 11 GHz H-pol. day and night emissivities; 11 GHz V and H-pol. day-night emissivity differences; 11 GHz day and night emissivity polarization ratio ( $\varepsilon_{11H}/\varepsilon_{11V}$ ); and 11 GHz day and night brightness temperature polarization ratio (H/V).



**Figure 84:** Time-series over July and August for Benton County, Iowa: From top: MODIS land surface temperature [K] day (DY) and night (NT) averaged to AMSR-E resolution; time of day of AMSR-E observations; and AMSR-E earth incidence angle.

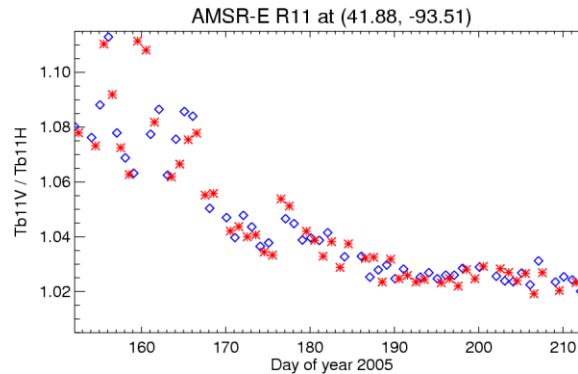
Candidate hypotheses for the dominant cause of the day-night emissivity anomaly include dew, fog, and thermal heterogeneity. Scene thermal heterogeneity would manifest itself as a bias between the effective microwave and infrared emission temperature due to the differing AMSR-E and MODIS viewing geometries and microwave penetration through the canopy. Although plausible, daytime thermal heterogeneity would typically be higher than nighttime so this hypothesis does not explain why nighttime emissivities are more strongly affected. Thermal effects are also unlikely to explain day-night emissivity differences as high as 0.05, which suggests a total temperature effect near 15 K, or 7.5 K night and day MODIS-AMSR-E biases.

Dew and/or fog effects have been found to be consistent with the data observed so far and we are in the process of further investigating these phenomena. Dew effects on microwave emission have been investigated by other researchers and Kabela *et al.* (2009) have shown that it can be persistent at AMSR-E nighttime observation times in the region of interest. We hypothesize that nighttime dew reduces emissivity by increasing scattering in the vegetation canopy relative to daytime. Dew has been found to reduce emissivities at 1.4 GHz (Hornbuckle *et al.* 2007) so it is plausible that this mechanism could occur at the higher AMSR-E frequencies, which should be more susceptible to scattering. When dew is not present, nighttime emissivities would increase toward the daytime mean. Nighttime emissivities increase during days 25-27 in Figure 83, and meteorological data from this period indicate that nighttime wind speeds and temperatures were high relative to the rest of the month, which would suggest a reduced chance of dew formation.

Nighttime fog and atmospheric inversion could also depress the nighttime emissivities by inducing a warm bias in MODIS LST's relative to microwave effective emission temperatures. Favorable conditions for dew and fog formation are similar (still air, high relative humidity, temperature inversion) so it is difficult to separate these effects and in fact they may both be significant factors at times. Further investigation is needed, first, to determine if day-night emissivity differences are widely correlated to conditions favorable for dew and fog formation (e.g., using NCEP North American

Regional Reanalysis data) and, second, if fog can be detected in nighttime satellite imagery (e.g., using MODIS or GOES 3.7 and 11  $\mu\text{m}$  bands).

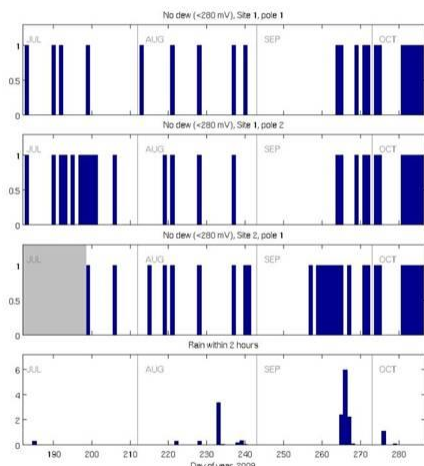
To test the dew hypothesis, we have collaborated with Dr. Brian Hornbuckle at Iowa State University who has made summertime field measurements of dew amount in the region of interest. Two datasets are available, one from 2005 and one from 2009. The 2005 dataset included extensive observations over six field sites from early June to early July. We processed the Aqua data for this period to derive emissivities and ancillary products including the 11 GHz polarization ratio, R11 (V-pol/H-pol; Figure 85). Like 2003 we found that persistent day-night emissivity differences did not occur until July, yielding too few days over which to thoroughly evaluate the dew hypothesis. However, the SMEX05 data do indicate that nighttime dew is likely to occur most nights at Aqua overpass times, confirming part of the hypothesis. Also, we observed that R11 declined during June and appeared to stabilize in July, suggesting that the effect of vegetation growth at 11 GHz had reached a limiting condition at that point and that this condition must be met before persistent day-night emissivity differences can be observed.



**Figure 85:** Time series of AMSR-E R11 polarization ratio during the SMEX05 field campaign. July 1 is day 182.

The 2009 field data included automated dew observations from a four sites with sensors at two canopy heights. Like the SMEX05 data, the 2009 data also indicate persistent dew most nights at the Aqua overpass time. Figure 86 shows two slides discussing analysis. The first slide shows (left) the nights WITHOUT dew at 3 sensors from July through mid-October and (right) the agreement between MODIS LST and *in situ* air temperatures on clear-sky nights. To reconcile the sensor differences, an *ad hoc* no-dew test was defined as any of the three sensors reporting less than 280 mV (where 274 mV is the manufacturer's threshold for dew detection). The air-LST agreement (about 0.8 K rmse) on both dew and no-dew nights (slide 1, right) is further evidence that fog (which should bias the MODIS LSTs) is not a major factor in the observed day-night differences and it allows us to replace MODIS LSTs with air temperature at night in emissivity retrievals at this site to add more points of cloudy-sky data to our analysis.

## Preliminary analysis with 2009 (SMEX09) Iowa dew field measurements\*

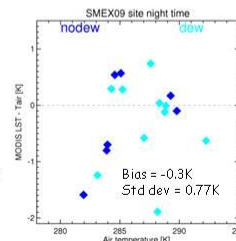


\*Experiment conducted by Brian

- 48 - Hornbuckle from U. of Iowa

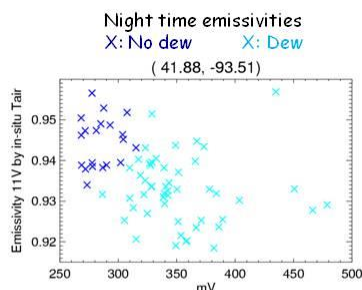
- Nighttime AMSR-E overpass times *without* detected dew
  - 3 automatic dew sensors (mV output)
- Sensor disagreement suggests light dew amount
- Ad hoc "no-dew" algorithm:
  - Any of 3 sensors reporting <280 mV

Reasonably good agreement between MODIS LST and in situ air temperatures in the clear-sky (night time)

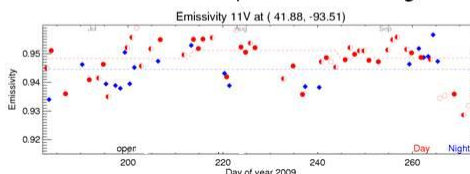


## Results

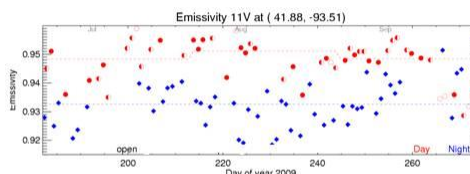
- Preliminary analysis indicate correlation between  $\Delta$ DN anomalies and occurrence of dew deposition on corn leaves



AMSR-E emissivities derived using in situ air temperatures at night



(a) No dew (see previous slide)



(b) Dew

- 49 -

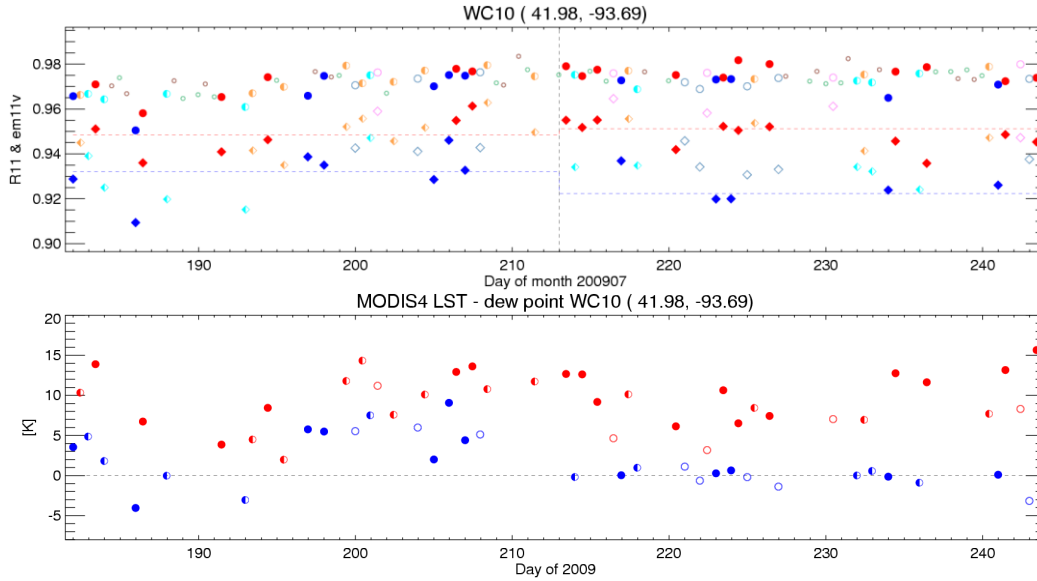
**Figure 86:** Slides discussing recent dew analysis results.

The second slide in Figure 86 shows nighttime air-temperature based emissivities for cases flagged as dew and no-dew present. Some day-to-day noise in the data is to be expected because of the presence of a reservoir about 20 km from the field site and the retrieval grid point, which will variously affect the gridded emissivities depending on the orientation of the AMSR-E footprints. The left plot shows emissivity as a function of the *average* dew sensor voltage—which is another measure of total dew amount - and the no-dew cases (flagged with the *ad hoc* algorithm) have distinctly higher emissivity than those with dew. The right plots show the time series of daytime (red) and nighttime emissivities



(blue) for no-dew and dew cases. The no-dew nighttime cases (avg. 0.945) are consistent with the daytime emissivities (0.95) while the dew cases are lower (0.932) by a significant amount relative to the day-to-day noise.

For the experiment described above, we made use of in situ measurements of leaf wetness to help with dew detection but such measurements are not commonly available and have their own uncertainties. In particular there is some level of arbitrariness in the way we set the voltage thresholds above which we can decide that dew is indeed present. To complement this experiment and in order to extend our analysis to larger spatial scales we have been examining the correlation between the 11GHz emissivity changes and meteorological factors influencing nighttime dew formation.



**Figure 87:** R11 (circles) and retrieved  $\varepsilon_{11V}$  (diamonds) time series at SMEX09 site during July and August time period (top plot). The bottom plot shows the difference between MODIS surface temperature and dew point from in situ measurements. In both plots, daytime data is in red and nighttime in blue. Filled symbols indicate clear conditions, empty symbols overcast, and half-filled partly cloudy.

Dew is difficult to predict. There are factors known to favor dew formation, including strong surface cooling (i.e. absence strong emission from clouds), little mixing (low wind speed) and high relative humidity, but they are difficult to use in practice to accurately predict dew formation (e.g., Kabela et al., 2009). In particular, dew formation depends strongly on type of surface and vegetation. For instance, dense broadleaf vegetation (such as corn) is more apt to induce dew formation than short grass, since the surface of the canopy is more thermally isolated from the ground, the main source of heat at night, and therefore can reach lower temperatures by radiation cooling. In this case, one should compare surface skin temperature (as measured to an IR sensor) to the dew point temperature of the air at the top of the canopy. The results of such a comparison for the SMEX09 campaign are shown in Figure 87. It is seen in this figure (right half of bottom plot, blue symbols) that MODIS V4 nighttime infrared temperatures are very close to dew point measured by the in situ sensor throughout the month of August. During the month of August, the nighttime AMSR-E emissivities depart from their daytime value almost every night (top plot, diamond symbols). There is a two week time interval from day 195 to 208 when dew was not detected and night time emissivities are close to their day time values. It is



apparent from Figure 87 that the MODIS temperatures for those days exceeded the dew point temperature by ~5K.

For our global nighttime LST (or water vapor) retrievals, wherever we get a significant positive day/night differences in our emissivity estimates we need an index we can rely on to help decide on a day-to-day basis which value of the 11 GHz emissivity to use as a constraint at night. If indeed dew is responsible for the observed drop in AMSR-E emissivities, then surface dew point temperature is an obvious choice. Unfortunately, surface dew point (or relative humidity) information provided by global NWP models is not sufficiently accurate for this type of application. In addition we previously relied on the infrared skin temperature dropping below the measured surface dew point temperature as an index for dew formation but infrared measurements are not available in overcast conditions. One option that has not been investigated in depth yet is to first use our daytime 11 GHz emissivity estimates for our LST retrievals and compare the results to the NWP model surface dew point temperature. The fact that the daytime emissivity values are higher than the actual observed surface emissivity when dew is present (causing a negative bias of 3-6K in retrieved LSTs) combined with the fact that actual observed LST should drop below the dew point, should result in an estimated LST that is significantly lower than the model dew point temperature, in which case the retrieval can be repeated using the estimated nighttime emissivity value. This negative bias in retrieved LST may be sufficient to overcome the uncertainties inherent to the model. As an alternative we can also rely on MODIS cloud cover to switch between estimated daytime and nighttime emissivities. We have shown in Section 7.1 (Figure 75) that instances where the observed nighttime 11GHz emissivity rise to their daytime value often coincide with overcast conditions, when nighttime surface cooling is minimal. In this case, we would have to make use of auxiliary cloud type or optical depth information to help with the decision. Thin cirrus clouds, for instance, do not have as much impact on the net nighttime radiation flux as most water clouds and would not be as effective as preventing dew formation from occurring. One issue with this approach relates to the difficulty of detecting water clouds at night underneath ice clouds. In conclusion, there is no known index we can rely on to make a correct decision 100% of the time. Short of performing a simultaneous retrieval of emissivity and surface temperature, which has its own limitations too, one can only hope at this point that occasional errors in predicted daily surface emissivities only have a limited impact on monthly mean temperature estimates.

Although not 100% conclusive due to the small case study size, the in situ tests conducted at the SMEX site are supportive of the dew hypothesis. So is the correlation between emissivity and cloudiness discussed in Section 7.1. A recent study (Hunt et al., 2011) has demonstrated that diurnal cycle in water content in corn has a measurable effect on the transmissivity of mobile telephone signal. However, the radiometric effect of a nighttime increase in vegetation water content would be opposite to the effect described here. Also there is no obvious reason why water content cycle would systematically be impacted by cloudiness. Of all the hypotheses examined here, dew is the only phenomenon that seems to fit with all the pieces of evidence collected so far.

### 7.3 Retrieval of LST over deserts

Land surface temperature retrieval over bare soil (in desert areas) is more difficult than over vegetated surfaces because of the necessity to take microwave sub-surface penetration (Moncet *et al.*, 2011a, Galantowicz *et al.*, 2011) into account. Over vegetated areas surface temperature can be estimated from the 11 GHz channels alone (except under precipitating conditions – Sec. 7.1). In arid areas, lower frequencies sense deeper into the soil than the higher frequencies (Figure 88) and emit at a different temperature (Fig. Figure 89). Also the 89 GHz channels that provide the near surface temperature information are highly sensitive to water vapor (and cloud liquid water) implying that biases that might originate from errors in the externally specified water vapor field (or from the presence of water clouds) have to be corrected.

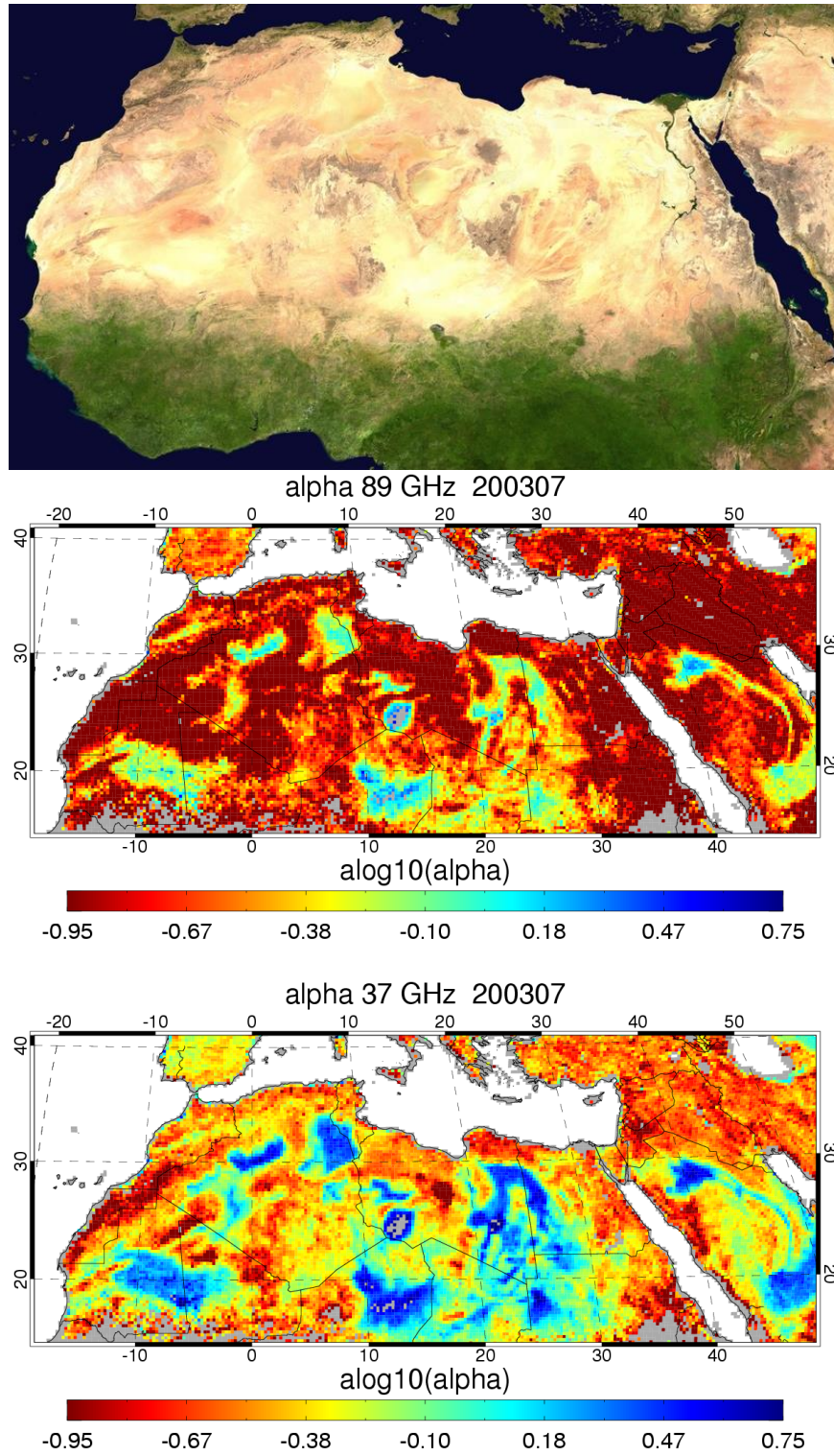
Here, surface temperature estimation over deserts is handled in the context of a simultaneous surface and atmospheric parameter retrieval using 1D variational approach. In this approach, surface properties are characterized by holding emissivities fixed to their estimated 1b values (representative of a given time period over which R11 remains stable), and by retrieving a polarization independent surface emission temperature profile from the multi-frequency observations. Assuming that atmospheric temperature can be constrained using e.g. the NWP model fields as background, we still need to retrieve a water vapor profile, cloud liquid water and cloud pressure (or emission temperature), in addition to a single soil temperature for each frequency, from the available radiometric information. The main difficulty over deserts stems from the fact that CLW both strongly impacts the 89 GHz absorption and surface emission temperature, making it difficult to detect small liquid water amounts unless near surface (89 GHz) temperature is strongly constrained. Using a coupled atmosphere land surface model to constrain the solution could be helpful but in this case, one still would need to account for the impact of the residence time of the cloud on the surface temperature profile. Information from a single satellite is not sufficient for that purpose. Over weakly polarized surfaces which only represent a few isolated areas over the Sahara desert (see Fig. Figure 91), more degrees of freedom than pieces of information available from the radiometric measurements are required to separate both effects, which is likely to result in significant crosstalk between CLW and  $T_{89}$ , even if water vapor is assumed perfectly known. This cross talk can be reduced over highly polarized surfaces since cloud liquid water and water vapor (constrained by the 22 GHz observations) are the only variables impacting the  $T_b$  polarization. In the context of variational retrievals, the covariance matrix describing the uncertainties (and their correlation) in the background soil temperature can be derived from actual retrievals in the clear-sky and eventually adjusted empirically or from land surface models.

Note that most microwave retrieval schemes over land ignore the impact of penetration and attempt to solve for V and H-polarized emissivities and a single surface temperature,  $T_s$ , using instantaneous observations. Such an approach has the disadvantages over the one proposed here that 1) it requires more parameters to characterize the surface emission ( $N_f$  frequencies times 2, for V and H-polarized emissivities, plus  $T_s$ , as opposed to  $N_f$  soil temperatures in our case) and 2) the surface parameters it produces are unphysical, making this type of surface parameterization inappropriate for the coupling with a land surface model in the context of assimilation.

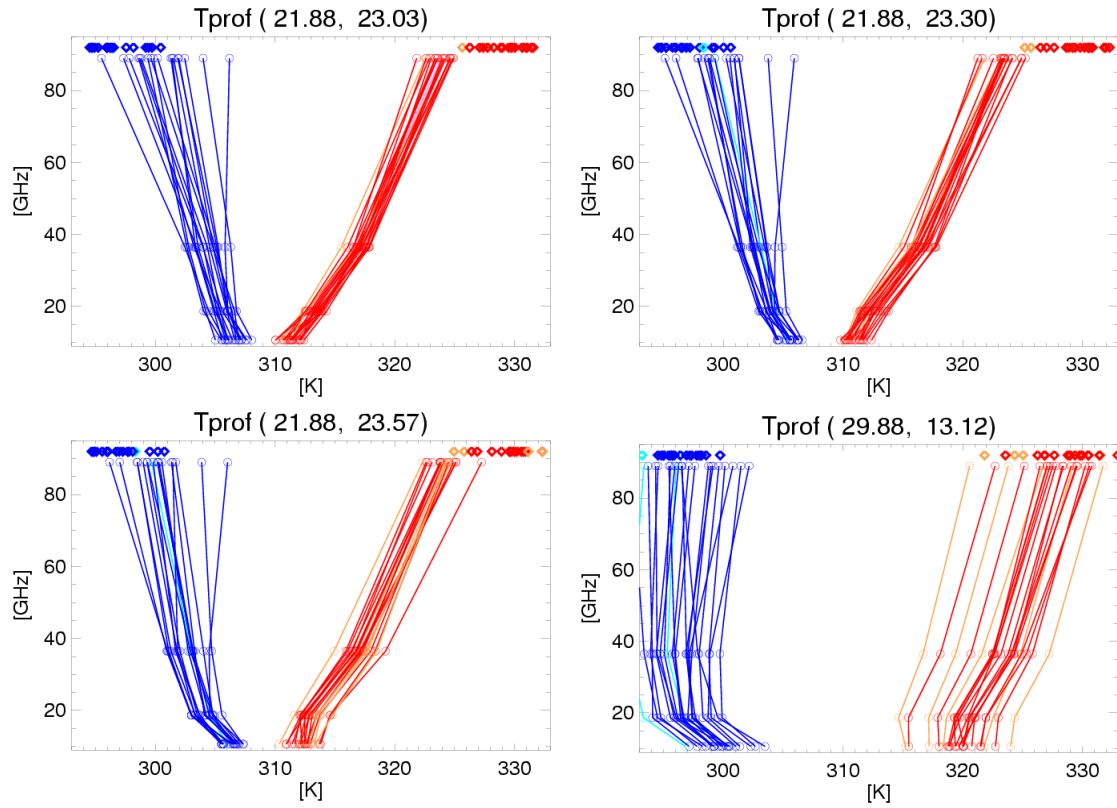
Figure 89 shows a few examples of retrieved emission temperature profiles over the Libyan Desert (southeastern Libya) and near Tabaqah (northwestern Libya). The first location is mostly composed of siliceous minerals. The subsurface penetration is fairly large over this area during the day, at the time of the Aqua satellite overpass, leading to positive temperature differences of 10-12K between the 89 GHz and the most deeply penetrating 11 GHz measurements. MODIS skin temperatures are also significantly higher than the 89 GHz temperature, suggesting that penetration at 89 GHz is significant

over this area. The daytime measurements appear to be very regular from day-to-day. At night the vertical gradients tend to be less pronounced (and differences between MODIS and 89 GHz emission temperature are also generally smaller) but we find more variability in the soil thermal structure at night than during the day for the cases shown here. Atmospheric water vapor may be a significant modulating factor infrared radiation fluxes at night. The conditions over the second site appear more isothermal and, in particular, the daytime contrast between 89 GHz emission temperatures and MODIS skin temperatures tend to be much smaller than over the first site. Day-to-day surface temperature variations are also more important (also noticeable in MODIS LST) and tend to more readily propagate down to the lower frequencies (especially noticeable at nighttime). Note that over the Tabaqah site, temperature is warmer at night at the surface than deeper into the soil. We attribute this artifact to the warm biases in the AMSR-E EA-L2A 89 Tbs (Appendix A).

Figure 90 shows examples of daytime and nighttime maps of emission temperature differences between 89 GHz and the lower frequencies (37, 19 and 11 GHz) over the Sahara desert and Arabian Peninsula. These maps show high positive temperature differences between 89 GHz and 11 GHz over most of the Sahara desert during the day and somewhat smaller negative differences at night, another indication that impact of penetration is quite widespread and not restricted to sand deserts. As expected, the temperature contrast with 89 GHz tends to decrease with increasing frequency during the day. At night, the areas of largest negative 89-37 (or 19) GHz temperature differences clearly coincide with sandy deserts and regions where siliceous rocks dominate. Temperature differences between 89 and 37 (or 19) GHz become smaller (and even positive) in rocky areas, a result which is consistent with the sample temperature profiles shown in Fig. Figure 89. As mentioned previously we attribute the fact the 89 GHz emission temperature exceeds the 37 or 19 GHz temperatures in some (less deeply penetrating) areas to calibration errors. The observed features in the maps of Figure 90 appear to correlate with the surface lithology.

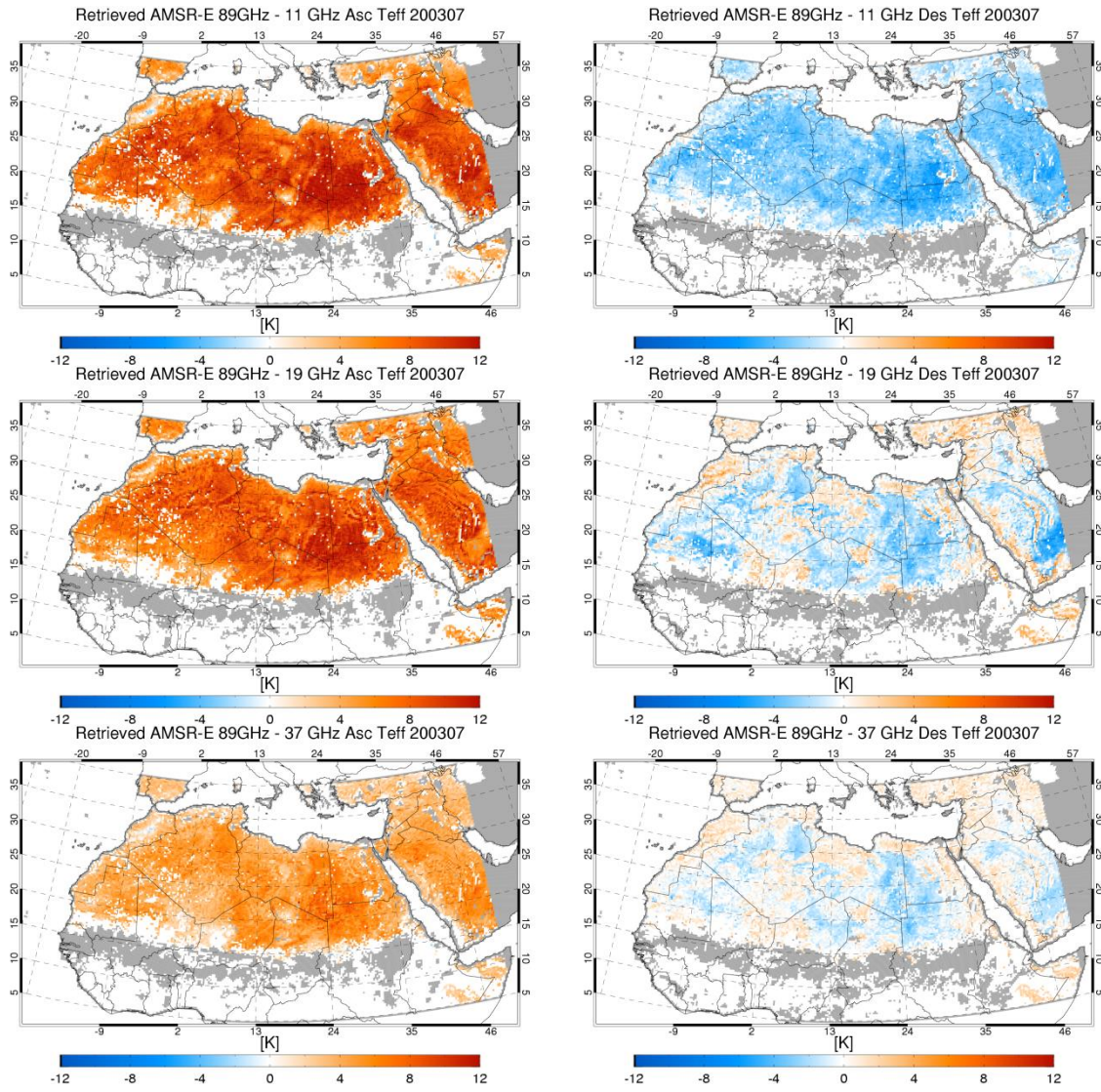


**Figure 88:** Maps of the emission depth parameter  $\alpha$  (see Moncet et al. 2011a; Galantowicz et al., 2011) retrieved from our 1b algorithm at 89GHz (middle) and 37 GHz (bottom). Note that  $\alpha$  measures the attenuation of the amplitude of the surface thermal wave as sensed by microwave channels (the deeper the channels senses the larger the attenuation) and not the physical depth in the medium. Sand deserts (bright surfaces in top visible image) are easily recognizable in both maps from the fact that they are highly penetrating (largest  $\alpha$  values)

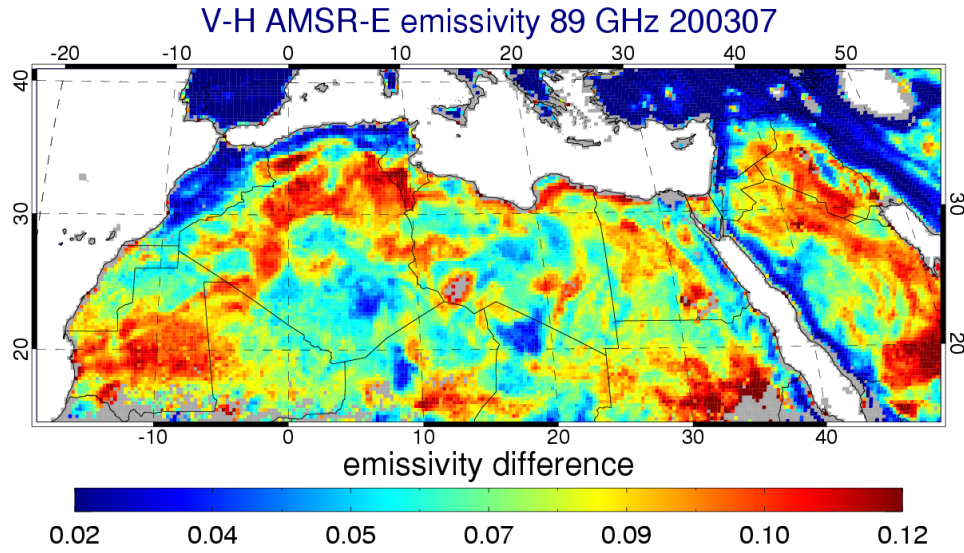


**Figure 89:** Examples of emission temperature profiles (nighttime in blue and day time in red) retrieved over Libyan Desert (southeastern Libya) – siliceous rocks - and near Tabaqah (northwestern Libya) – mixed carbonate, siliceous and evaporite rocks in July 2003. MODIS skin temperature measurements are indicated by the diamonds. Note that we use frequency as our y-axis coordinate which makes emission depth appear to be the same for both sites. In fact all that can be inferred from the microwave observations is the amplitude of the diurnal thermal cycle at the some unknown depth which is likely to vary with soil type but not the physical depth itself. A thermal conductivity model of the soil would be required to infer the latter.



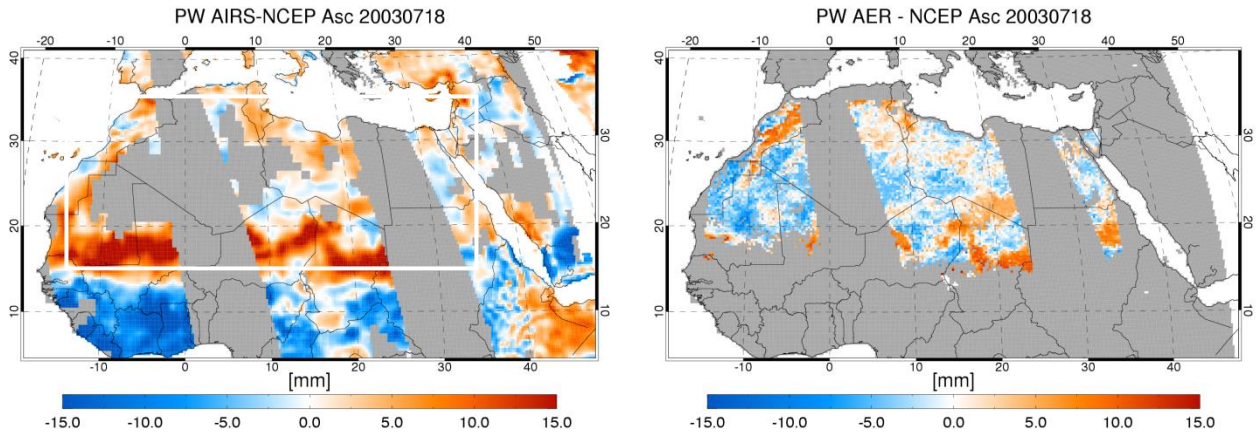


**Figure 90:** July 2003 month-average effective emitting temperature ( $T_{eff}$ ) retrieval differences relative to 89 GHz, i.e.,  $T_{eff}(89 \text{ GHz}) - T_{eff}(11 \text{ GHz})$  (top), 89-19 (middle), and 89-37 (bottom). Left: Daytime. Right: Nighttime.



**Figure 91:** Map of 89GHz polarization difference

Figure 92 compares our retrieved column water vapor with AIRS PW from the AIRX3STD product. Note that the AMSR-E retrievals were performed using the V09 AE\_L2A Tb product. The quality of the retrieval results is obviously significantly impacted by the Tb biases discussed in Appendix A. However, it is shown in Figure 92 that our AMSR-E based water vapor correction goes in the right direction. Agreement with AIRS is expected to improve when using the new RSS V7 L2A Tbs and when using AIRS-based correction in deriving our AMSR-E emissivities (Sec. 6.2). Note that some significant anomalies in the retrieved 89 GHz temperature had been observed in areas impacted by dust storms when using the NCEP atmospheric correction without any further adjustment. It is not believed that dust itself has a direct impact on the 89 GHz since the size of the airborne particles is only of the order of a few tens of microns (e.g. Goodie and Middleton, 2001). However, the dust storms over Western Sahara are often associated with humid air incursions from the moist subtropical regions south of 15°N. In these areas, the NCEP precipitable water may have significant errors and time interpolation errors can also be large. In the cases analyzed, the NCEP model often underestimates the actual atmospheric moisture behind the dust front when compared to AIRS. Unfortunately, the AIRS observations are affected by the dust and no data is generally produced within the dust storm itself. When AIRS retrieval product is available in those areas its quality is questionable. The fact that our AMSR-E water vapor correction agrees reasonably well with AIRS and that the  $T_{eff}$  biases observed around dust storms can be removed by adjusting the column water vapor based on the AMSR-E Tbs suggests that the observed anomalies might be related to water vapor errors.

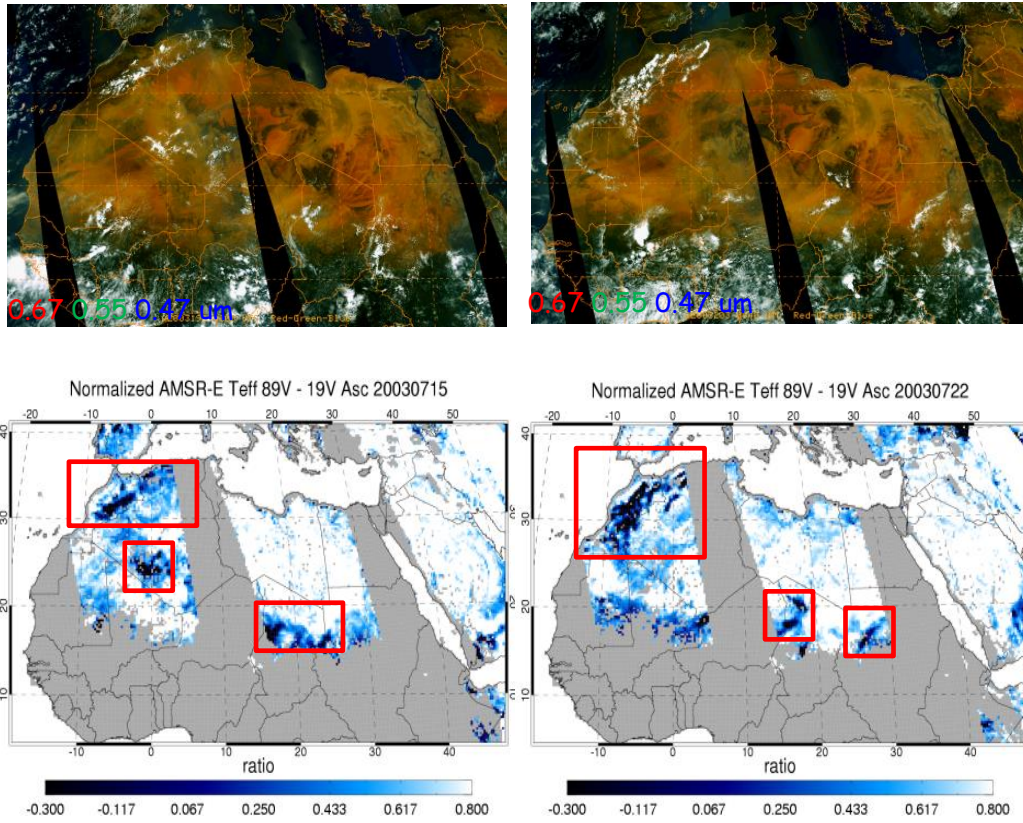


**Figure 92:** AIRS minus NCEP precipitable water (left) and AMSR-E precipitable correction (right) over Sahara desert for July 18, 2003. Areas in gray in the left panels are areas where an AIRS PW estimate is not available. Although there are significant local differences in the magnitude of the AMSR-E and AIRS retrieved PW, there is overall good agreement between areas of positive and negative PW correction.

Short of retrieving liquid water amount (and cloud altitude) from AMSR-E Tbs, which as we argued earlier may be difficult at least over certain surfaces, we can at least attempt to detect the presence of clouds and flag our retrieval product accordingly. Because of the low frequency of occurrence of water clouds over Sahara, at least at this time of the year, screening out data produced from cloud water contaminated AMSR-observations is expected to only have a limited impact on monthly average LST or PW estimates. Cloudiness is more persistent during the fall and winter months in parts of the Sahara desert (e.g. Figure 28 and Figure 29) but in this case one might be able to take advantage of the fact that surface conditions are closer to being isothermal to reduce the number of degrees of freedom in the retrieval. Note that the impact of cirrus clouds on the microwave frequencies up to 89 GHz can be ignored. Similarly we are working under the hypothesis that dust does not affect the microwave observations.

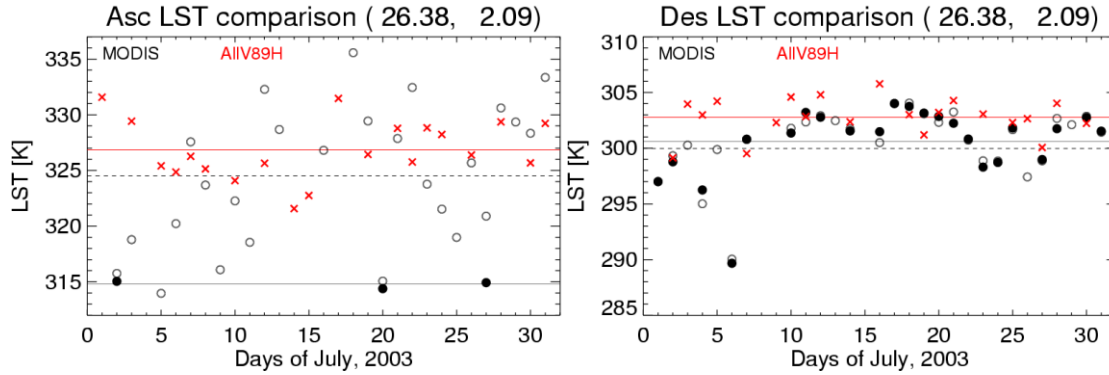
MODIS cloud analysis is used to identify clouds but also provides useful phase information. However it only identifies the phase of the cloud particles at the top of the cloud (which may be misleading in e.g. convective environment) and it cannot be relied upon for liquid water detection underneath thick cirrus clouds. Also, our variational retrieval problem being under constrained (when using background information from climatology), convergence is always reached, i.e. absence of convergence cannot be used to indicate the presence of (non-precipitating) water clouds in the field of view.





**Figure 93:** Enhanced multi-spectral MODIS imagery (top) showing clouds (in white) over Sahara desert. The bottom panels display the deviation of the 89GHz – 19 GHz emission temperature differences from its local maximum for the month of July. The data shown in the left and right columns are for July 15 and July 22, respectively. It is apparent from this figure that areas with the strongest negative anomalies in retrieved  $T_{eff, 89V} - T_{eff, 19V}$  are associated with the presence of clouds.

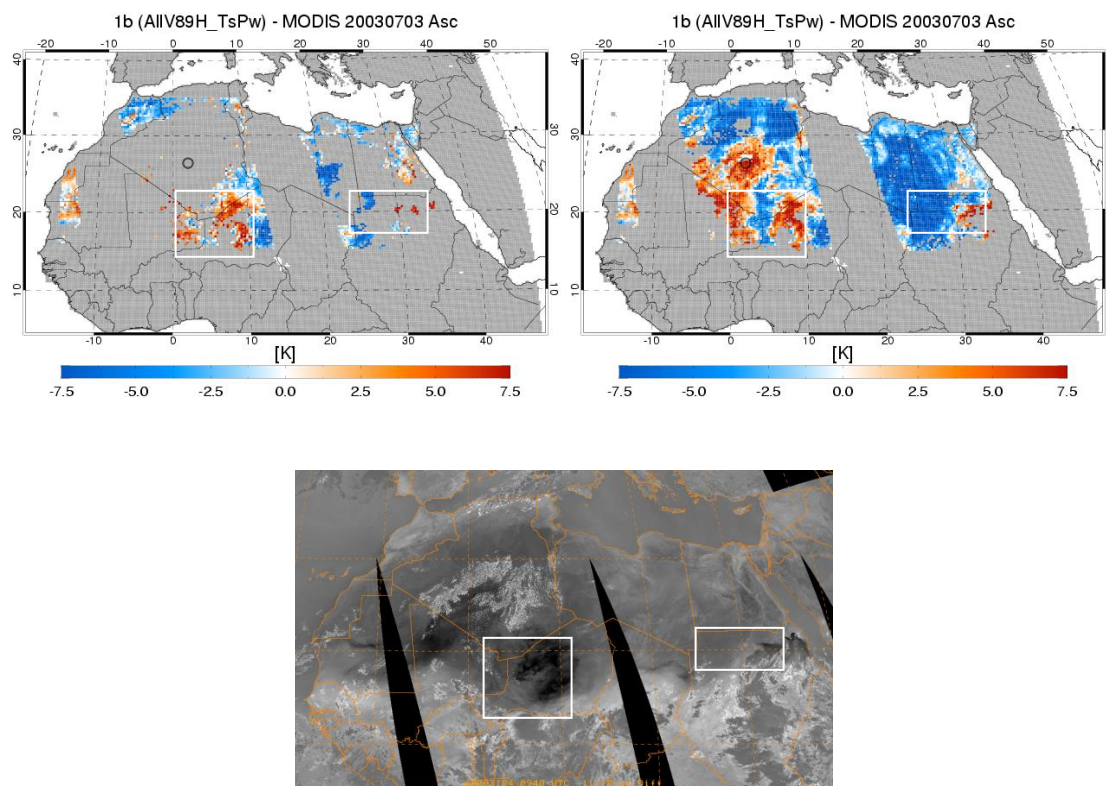
In the short term, we will explore the possibility of relying on temporal analysis of the microwave Tbs (e.g. 89 GHz polarization information), or on an analysis of surface temperature profiles retrieved under the clear-sky assumption, for the detection of moderate to high liquid water amounts and flagging of surface temperature and water vapor estimates produced under those conditions. As mentioned previously, liquid water absorption is larger at 89 GHz than at the lower frequencies and causes a decrease in the 89 GHz Tb's when the cloud is above a certain altitude and an increase for lower clouds (the altitude at which the change of sign occurs is a function of the contrast between atmospheric emission and surface brightness, which depends in particular on surface emissivity). The effect of scattering by large ice particles is also stronger at 89 GHz than at 19 GHz (note that precipitation can also be identified through the impact of soil wetness on the R11 polarization ratio - e.g. Sec. 7.4). Both scattering and liquid water absorption contribute to lower the 89 GHz Tb polarization difference. The presence of water clouds also causes the 89 GHz surface emission temperature to be colder than in the clear-sky during the day and warmer at night. The effect of clouds on temperature deeper into the soil depends on clouds residence time, microwave penetration and thermal conductivity of the soil. Figure 93 provide examples of the impact of clouds on the retrieved surface temperature profiles in the daytime. This figure shows how one can identify the presence of clouds in each grid cell by monitoring the departure of the daily retrieved 89-19GHz  $T_{eff}$  slope from the local maximum slope value (assumed to reflect clear conditions) for the month. Because surface properties are highly inhomogeneous, decision thresholds may vary with location.



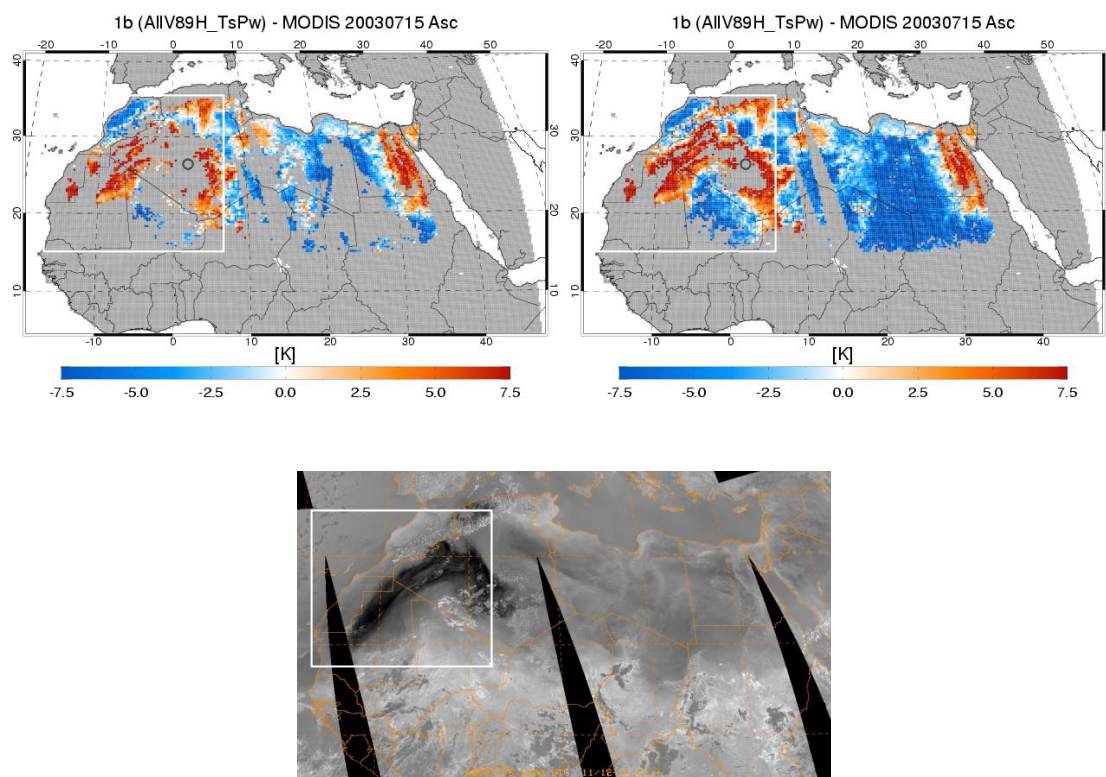
**Figure 94:** Time series of LST estimates from MODIS (in black) with strict QC (filled circles) and relaxed QC (open circles) and AMSR-E (red) for daytime (left) and nighttime (right) for grid point located at 26.38°N, 2.09°E (Western Sahara) in July 2003. Monthly mean LST are indicated by the horizontal black lines for MODIS (solid: strict QC; dashed: relaxed QC) and red line for AMSR-E. Our working assumption is that the low LSTs produced by MODIS in the daytime and the high variability in the MODIS LSTs may be due to under-analyzed dust (and clouds?). Over these areas, we expect the MODIS clear-sky LSTs to be warmer than AMSR-E LSTs in the daytime (the opposite is observed in this case) and colder at night. The better agreement between MODIS and AMSR-E at night could be explained by the diurnal cycle of wind-borne dust in this region.

Because clouds are quite infrequent over many desert areas, the clear bias in infrared monthly mean surface temperature estimates is expected to be smaller over these areas than over vegetated surfaces. Microwave derived LST estimates might still be useful for the quality control of the IR  $T_{skin}$  data. Figure 94 shows an example of time series of MODIS (strict and relaxed QC) and AMSR-E 89 GHz LST over Western Sahara. This particular location (marked by a circle in the maps of Figure 95 and Figure 96) was selected because it situated in an area where AMSR-E mean A-D temperature estimates for July were larger than MODIS mean A-D  $T_{skin}$ . Over desert regions, one would expect to see differences of the opposite sign. It is apparent from this figure that there are only three daytime MODIS LST estimates flagged as high quality during the month of July at this location and these LSTs are colder by more than 10K than the mean retrieved AMSR-E temperature. In this case, relaxing the quality control on the MODIS data reduces the mean LST difference with AMSR-E, but this improvement appears to be somewhat fortuitous given the high daily variability in the IR LSTs. One source of problems in this region is residual contamination by dust (not screened by the Version 4 MODIS mask) of the MODIS LST estimates. Figure 95 and Figure 96 show maps of AMSR-E minus MODIS LST differences for July 3 and July 15. The areas of large positive differences between AMSR-E and MODIS LSTs (relaxed QC) clearly correlate well with locations where dust/clouds are likely to be present based on images of 11-12  $\mu\text{m}$  MODIS brightness temperature. Using stricter MODIS QC masks most of these areas but not all of them. The fact that fewer MODIS estimates pass the stricter QC tests and that some of them can still be highly contaminated explains why improving the QC in this example degrades rather than improves the MODIS monthly mean LST estimates. This shows how AMSR-E surface temperature retrievals could potentially be exploited over deserts as an additional test to detect residual dust/cloud contamination not screened out by the MODIS mask. Note that this results were produced using the Version 4 MODIS LST product. Version 5 includes improved cloud flags from the Version 5 MODIS mask but we have not tested it over deserts yet.





**Figure 95:** Map of differences between AMSR-E and MODIS LSTs (left: strict MODIS QC; right relaxed MODIS QC) between 15 and 35°N over Sahara desert for July 3, 2003. The gray areas in this region correspond to areas where MODIS data does not satisfy the QC criteria. An image of the 11-12  $\mu\text{m}$  MODIS brightness temperatures is shown at the bottom. The dark areas in this image correspond to areas of probable dust and the bright areas correspond to clouds. There is a good spatial correlation between these areas and the areas of large positive LST differences between AMSR-E and MODIS in the top panels.



**Figure 96:** Same as Fig. A2 for July 15, 2003.

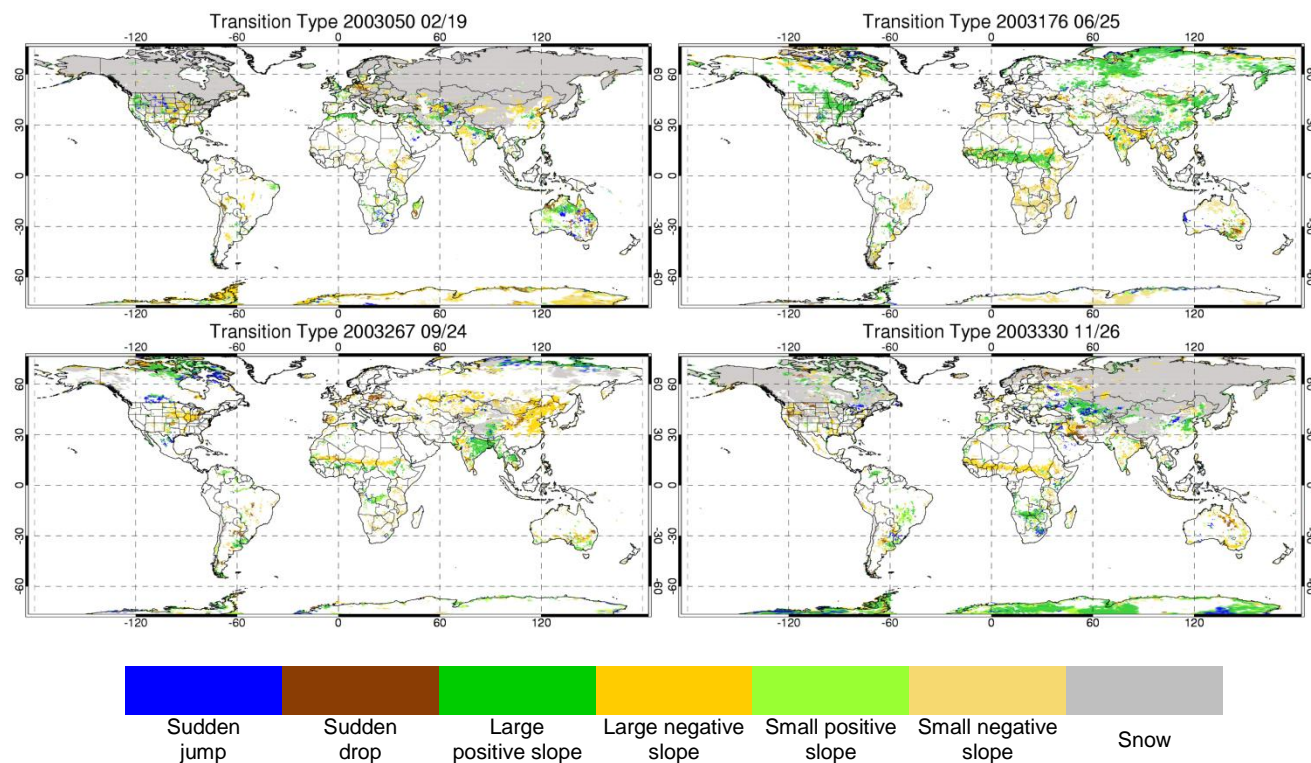
#### 7.4 Global short-term surface change monitoring from R11 analysis – comparison with IR emissivities

As previously noted, the 11 GHz measurements are less sensitive to the atmosphere than the higher frequency measurements, and the polarization ratio is little impacted by variations in surface temperature. Because the 11 GHz radiation penetrates more deeply within the vegetation canopy it is also more affected by changes in the soil properties than the higher frequencies. For these reasons, the R11 polarization ratio is a useful indicator of changes in e.g. soil moisture or state of vegetation affecting the surface emissivity over arid and semi-arid areas.

An algorithm for identifying time periods (segments) during which surface properties are assumed to be stable, from analysis of R11 time series has been described in Sec. 5 and Appendix C. The algorithm is currently applied to a full year of data at all gridded locations over land. Combined AMSR-E and MODIS observations collected within each time segment at a given location are grouped together to produce an emissivity estimate at that location that is valid for the entire interval. A by-product of this process is the flagging of outliers, and transitions in emissivity state occurring over the course of the year. This information is useful for quality control of e.g. LST or PW retrieval products but also for the monitoring surface dynamics, i.e. for tracking isolated events (e.g. precipitation, flood) producing significant emissivity changes on time scales of a few days and slower changes due to e.g. vegetation growth or harvesting.

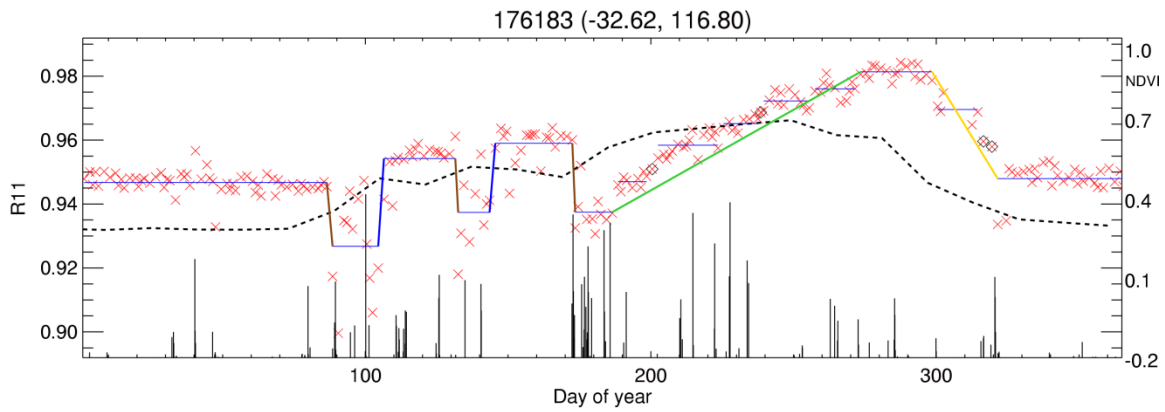
The output of the R11 time series analysis is used to produce global weekly transition maps (Figure 97) that provide an overview of temporal surface changes occurring worldwide. Each color in the map represents a different type of transition (see legend of Figure 97). Sudden transitions of R11 in snow free areas are usually associated with impact of precipitation on soil wetness and subsequent dry downs, smaller positive slopes often correspond to vegetation growth, leaf greening, and negative slopes in R11 correlate with crop harvesting or senescence. For example, the corn growing and harvesting seasons in the US Midwest are easily identifiable from the June and September maps, respectively. The seasonal vegetation cycle linked to the African monsoon in the Sahel region is also apparent on those maps.

An example of R11 time series for a location in the Australian South Western wheat belt is provided in Figure 98. It seen in this figure that surfaces is quite stable during the Southern hemisphere summer. Impact of precipitation gives rise to high R11 variability in the spring. Changes in soil properties are more easily observed when vegetation is sparse (and vegetation water content is low). The rise in R11 from the May to August coincides with the wheat growing season. It is apparent from this figure that the R11 response to the vegetation growth lags behind NDVI. Both indices decline again in the fall (October through November) during the harvesting. The maps of Figure 99 give a broader view of the surface dynamics on continental scales. In the northern part of Australia, grass grows during the monsoon season from December to March, and dries quickly after the end of the wet season. The growth and harvesting season in the Southwestern and Southeastern grain belts are quite apparent on those maps. MODIS NDVI maps in Figure 100 show similar cycles.

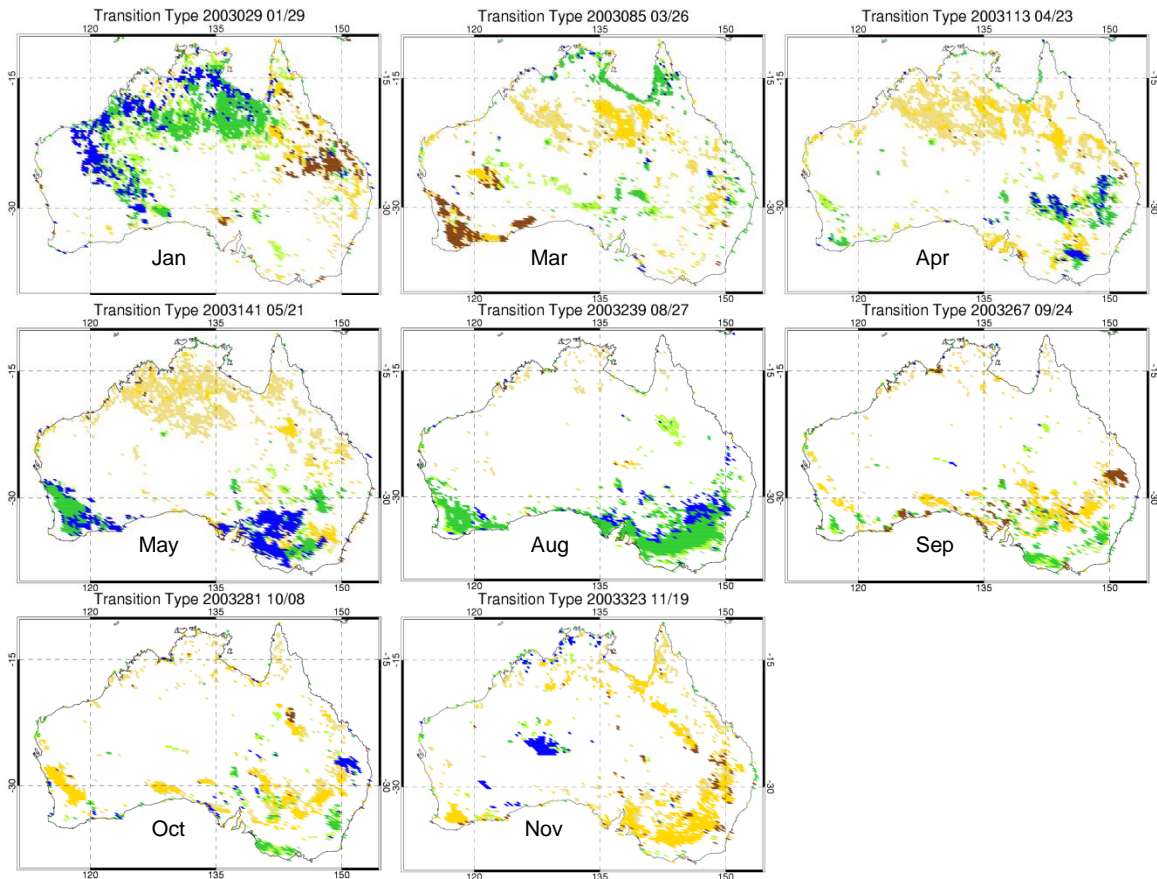


**Figure 97:** Examples of weekly R11 transition maps. In this figure white areas indicate areas where no change in R11 is detected through the time period. The signification of the other colors is indicated in the legend.

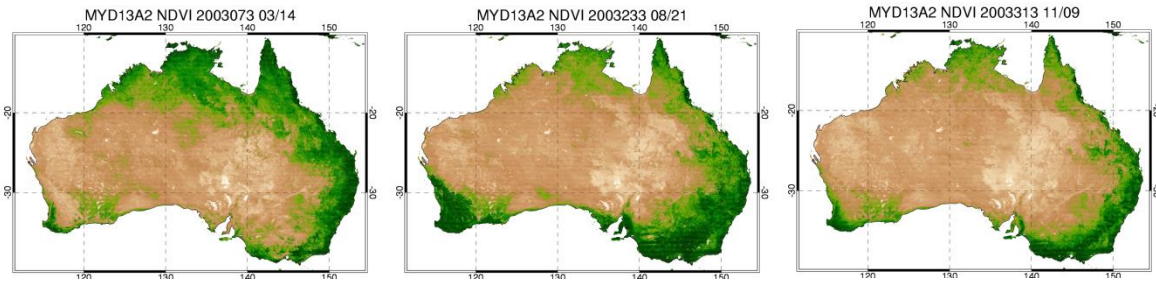




**Figure 98:** Example of R11 time series for a grid box located in the wheat growing region of Southwestern Australia. Daily R11 values are marked by the red crosses, and the segments are shown by solid black lines. Precipitation amounts from Tropical Rainfall Measuring Mission (TRMM) interpolated to the locations of the gridded AMSR-E measurements are overlaid on the plot. Dotted line represents the Normalized Difference Vegetative Index (NDVI) from the 16-day MYD13A2 MODIS product. The solid lines correspond to the R11 time series analysis output. The transition types are distinguished by different colors (color code is given in Figure 97).



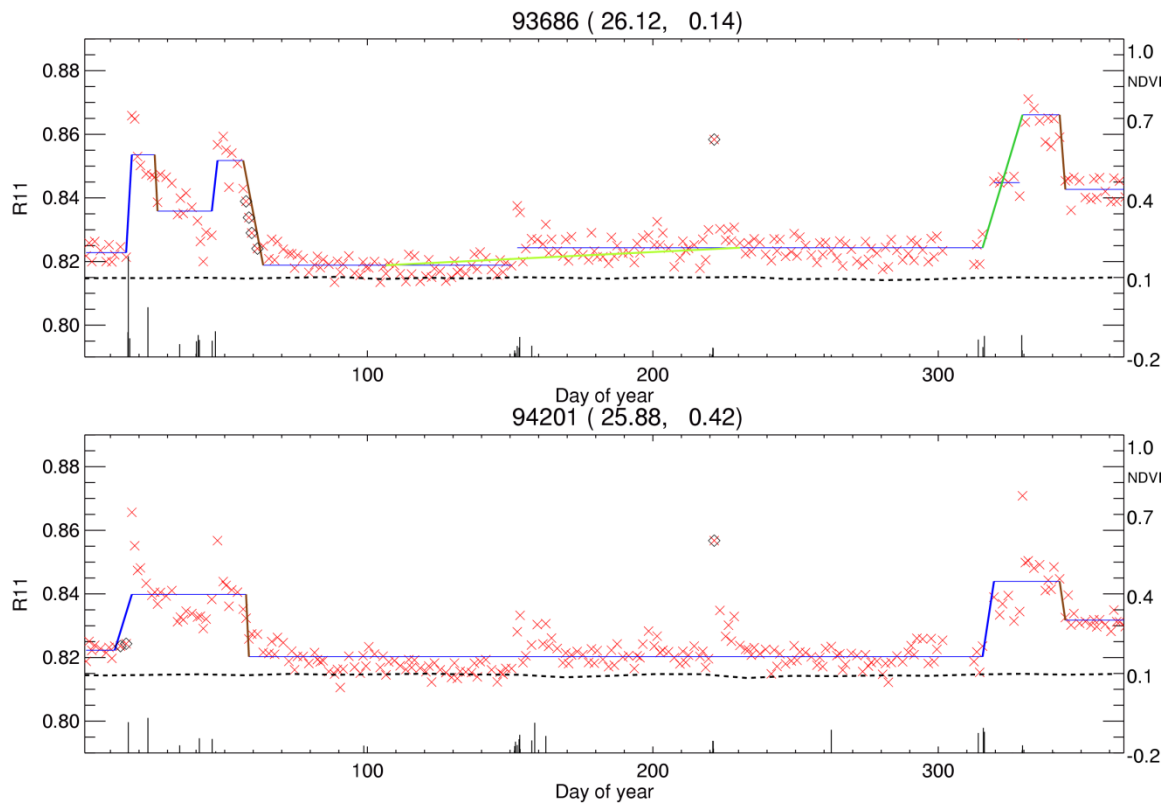
**Figure 99:** Examples of Weekly R11 transition maps over Australia in 2003.



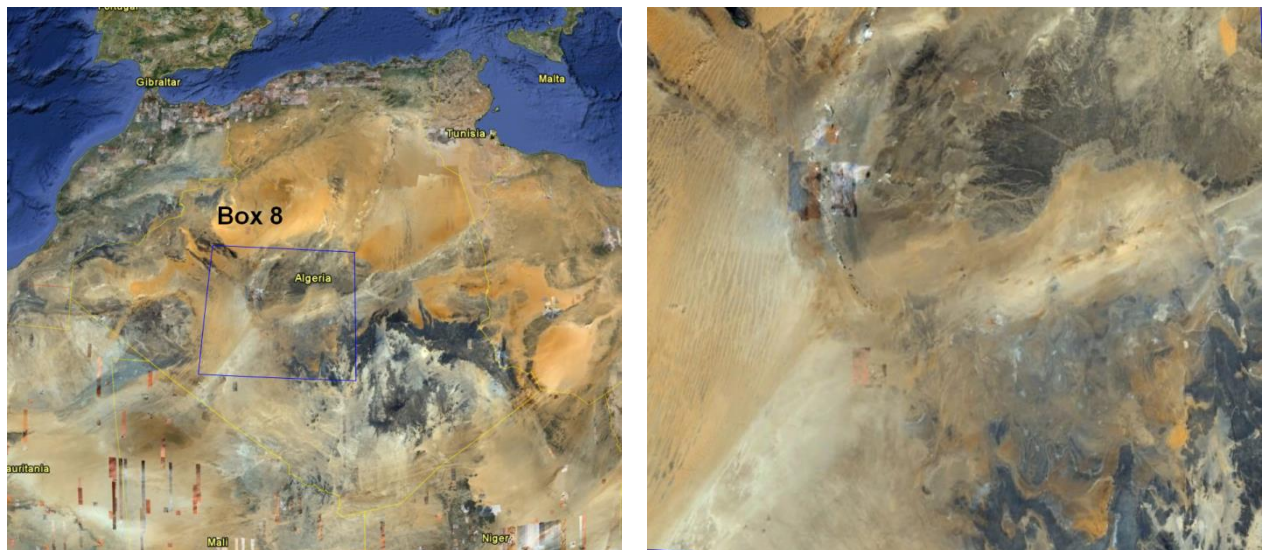
**Figure 100:** MODIS NDVI maps over Australia for March 14, June 21 and November 9.

Examples of impact of soil wetness on the R11 time series over Sahara desert are provided in Figure 101. Emissivities over the Sahara desert tend to remain relatively stable all year long. However, sharp R11 transitions associated with sporadic rainfall or flooding events followed by periods of dry down lasting a week or more are not uncommon (Figure 101). Over deserts, the R11 response to wetness can be in the form of positive or negative jumps, the latter being attributed to lower H/V polarization ratio of wet surface. Increase of R11 is shown to be linked to flood but not exclusively. The longer lasting events do not necessarily occur in the areas of most intense rainfall. Significant run off may occur over bare saturated soil or rocks and water tends to accumulate in the low-lying areas. Figure 102 shows the area (Box 8) which includes the two locations from which the R11 time series in Figure 101 were extracted. Figure 103 overlays local map of R11 transitions for the 2<sup>nd</sup> and 3<sup>rd</sup> weeks of January of 2003 to flood potential map produced from surface elevation and soil type. The locations where R11 jumps are observed transitions form a distinctive triangle. From this figure it is apparent that no change in R11 is observed in areas of low flood potential (black or transparent) while there is a spatial correlation between the areas of higher flood potential and the areas where R11 jumps occur. From an examination of the precipitation maps (Figure 104) in and around Box 8 for the period of time extending from 1/13 to 1/16 it is apparent that the areas where R11 transition occurs do not coincide with the areas of observed rainfall, reinforcing the hypothesis that the observed jump in R11 could be due to flooding.

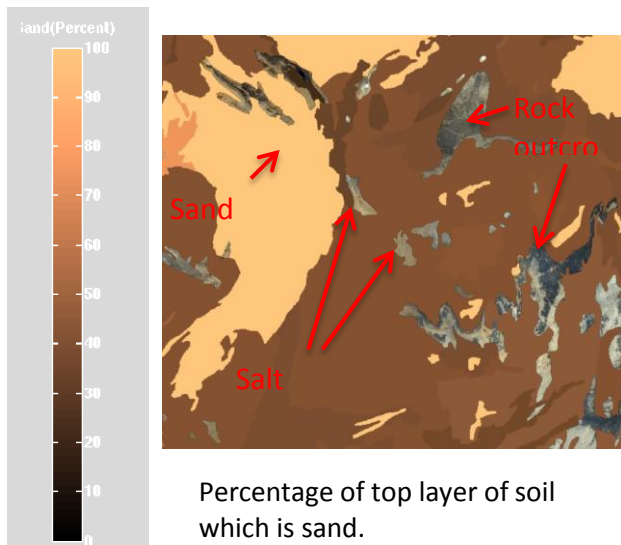




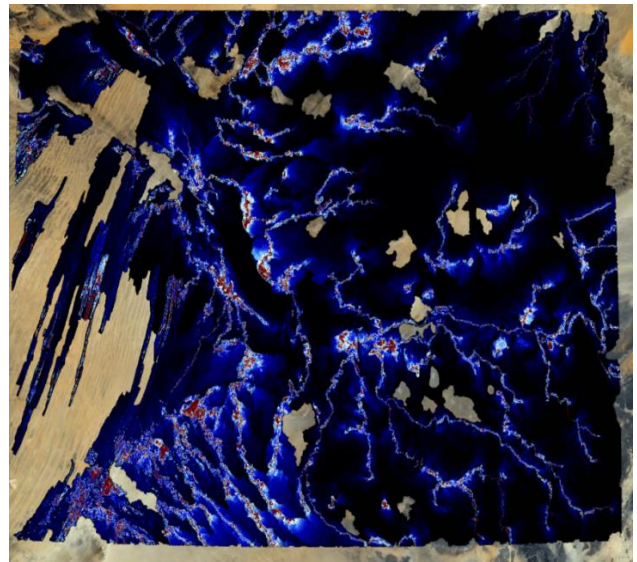
**Figure 101:** R11 time series in 2003 from Central Algeria. Both grid boxes have sharp R11 transitions in January- February and October associated with rainfall and flooding.



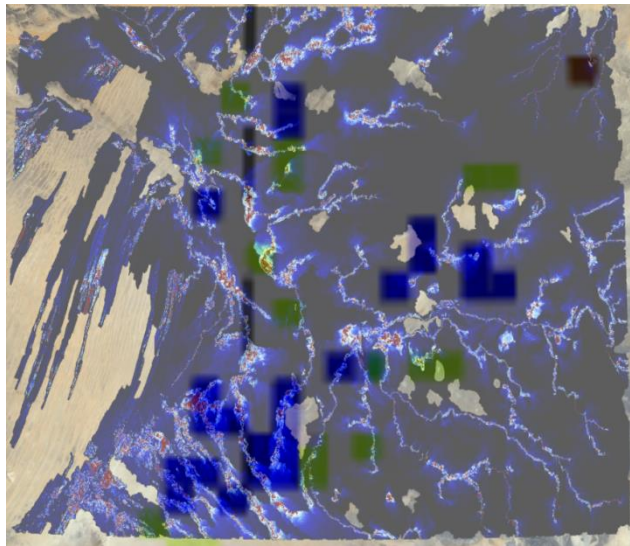
**Figure 102:** Google earth map of analyzed area (blue box) in central Algeria, the map on the right shows details of Box 8 in the left map



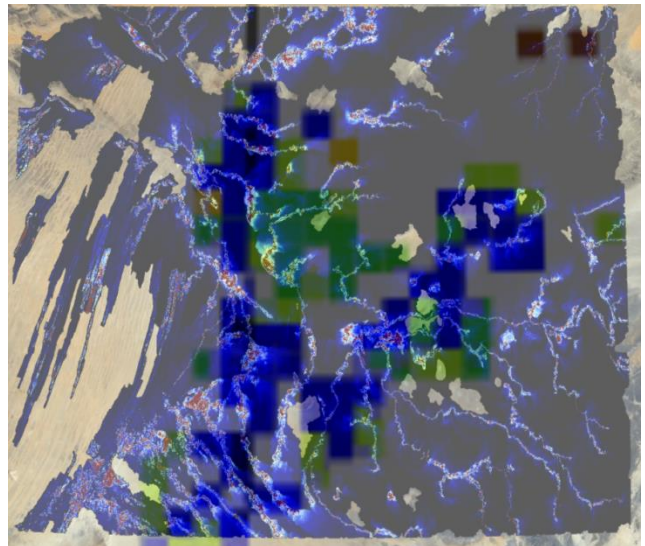
(a) Soil type of Box 8 in Figure 102



(b) Flood potential map of Box 8 in Figure 102



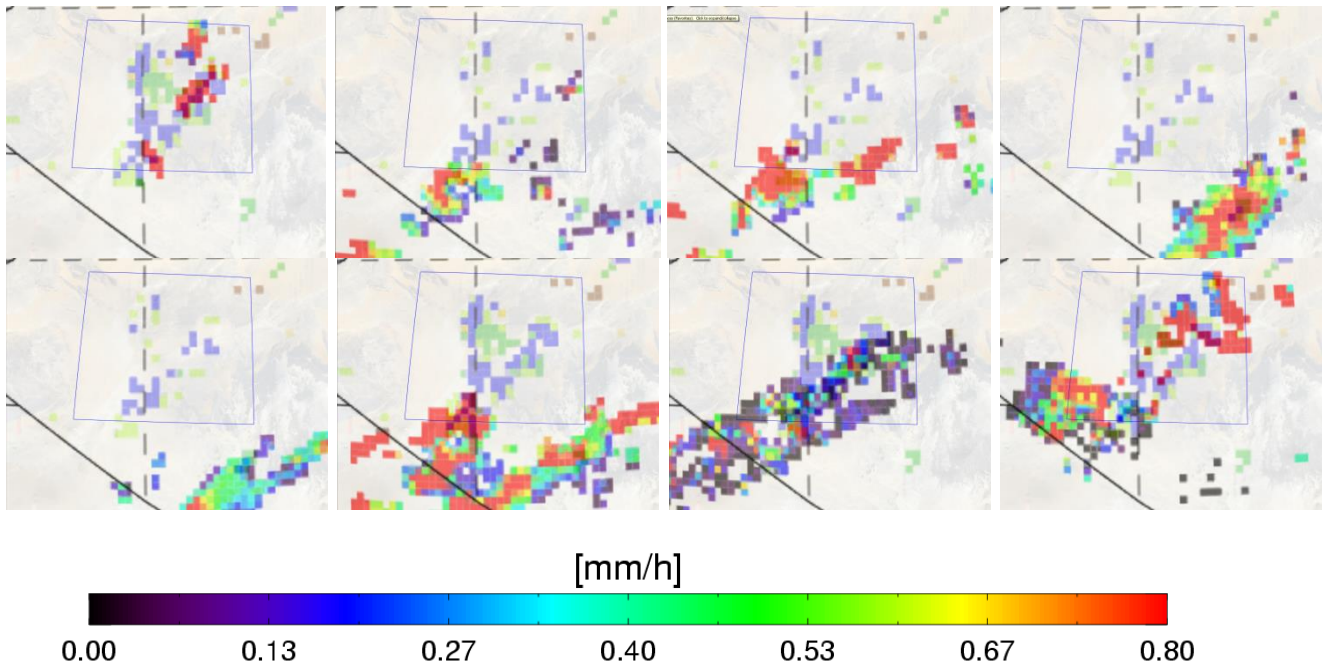
(c) Overlaid by Week 2 R11 transition map



(d) Overlaid by week 3 R11 transitions map

**Figure 103:** Flood potential map of Box 8. Areas of high flood probability are shown by red. Maps (a) and (b) show the soil type and original flood potential map; (c) and (d) show the R11 transition maps of week 2 and 3 in 2003 overlaid on the flood potential map.





**Figure 104:** TRMM (revisit time  $\sim 0.5$  day) precipitation rates from 1/13 to 1/16 2003 overlaid to R11 transition map over the area around Box 8 (in blue).

## 8 CONCLUSIONS

In this report, we provided a description of the Version1 AMSR-E land surface emissivity retrieval algorithm and products. The data base produced as part of this NASA funded effort represents a significant step forward compared to previous attempts to build similar microwave emissivity atlases. In particular, the improved treatment of penetration effects combined with the high degree of consistency between the MODIS LST product and AMSR-E observations results in a better separation of surface emissivities and thermal effects over arid areas. As a result our AMSR-E estimated emissivities are more stable in time than those produced in the past using SSM/I and ISCCP while neglecting penetration effects (Prigent 1997). Our advanced quality control procedure also results in better quality surface emissivity estimates over vegetated areas. Detailed description of this work can be found in Moncet *et al.* 2011a, Galantowicz *et al.*, 2011 and Moncet *et al.* 2011b.

As part of this phase we also proposed a number of key enhancements aimed at improving the quality of the emissivity product and yield in preparation of Version 2 of the emissivity atlas. These enhancements include:

- Mapping the AMSR-E Tb product on a fixed Earth grid which should minimize impact of surface inhomogeneity and provide temporally consistent infrared and microwave surface temperature estimates.
- Development of a R11-based time segmentation (RTS) which results in higher yield and higher quality surface emissivity estimates in stable areas
- Addition of the Terra MODIS  $T_{skin}$  products for improved characterization of surface forcing in 1b algorithm reduces noise in estimated thermal model parameters over deserts

The new procedures/enhancements have all been implemented but some aspects require further investigation (and refinements) before they can be definitively integrated into our process. For instance, the mapping of the AMSR-E Tb's on a fixed Earth grid results in a loss of flexibility with respect to being able to select the least cloud contaminated AMSR-E FOVs around a given grid box, which may impact both the yield and quality of the retained emissivity estimates for that grid box. The impact should be the largest over homogeneous, temporally stable areas and this impact remains to be quantified. It is expected that the implementation of the RTS time segmentation approach should compensate for this potential loss. We have shown that the new time segmentation technique significantly improves the quality of our surface emissivities, especially in areas of frequent cloudiness. The LST validation work presented in this report demonstrates that the agreement between retrieved AMSR-E LST and surface air temperature measurements both in clear and cloudy conditions over vegetated areas is excellent. In particular the use of microwave retrieved LSTs eliminates the clear bias in the monthly mean temperature estimates derived from infrared observations.

While evaluating the impact of the new RTS approach on the quality of our emissivity estimates, we observed that a number of areas which were previously classified as non-penetrating for some months during the year 2003 are now classified as (weakly) penetrating. This may occur in particular over vegetated areas experiencing some seasonal cycle that is only weakly impacting the R11 polarization ratio, which raises the broader question of the soundness of basing our time segmentation on R11 alone. Dew deposition, as we have shown here, is one example of phenomenon that might affect the surface emissivity and not necessarily induce measurable changes in the polarization ratio but in this case, emissivity changes occur on short time scales and at night only. It is also plausible that leaf wetness in general (caused by e.g. frequent precipitation over areas where evaporation is weak due to persistent cloudiness, conditions typically encountered over tropical forests) would produce fluctuations in emissivities on seasonal time scales. In addition, R11 over densely vegetated areas (with  $R11 > 0.98$ ) may not be responsive to small changes in vegetation greenness. As a result, with the new RTS scheme, emissivities derived from observations made during the clearest, hence driest, months may be applied to other months as well. This issue is easily overcome during the drier season since in this case one can use observed trends in instantaneous emissivity values as a basis for breaking down the analysis window into smaller time segments. One could also include auxiliary information (e.g. NDVI) at the initial segmentation stage. Once we enter the wet season it becomes more difficult to detect small emissivity changes because the noise caused by the interference of clouds in our emissivity retrieval process may be larger than the signal itself. Our preliminary evaluation shows that emissivities produced with the R11 segmentation, even if too stable, are of better (and in the worst cases, no worse) quality than the Version 1 emissivities. Also, we have not found cases so far where potential biases in our RTS emissivities would translate into mean retrieved LST errors exceeding our tolerance. Note that over the most questionable areas, we still have the option of attempting to simultaneously retrieve surface and atmospheric information from the microwave radiometric measurements alone, but whether this approach, when considering the combined impact of modeling and null space errors, would improve over the current scheme remains to be demonstrated.

Our plan for finalizing Version 2 of our emissivity retrieval algorithm includes first completing the evaluation of the impact of the Fixed Earth grid mapping of the AMSR-E Tbs on yield and combining this change with the R11-based time segmentation. In parallel, we will complete the evaluation of the new 1b algorithm, which makes use of Terra and Aqua MODIS surface temperature information, and the comparison with the Version 1 PTS algorithm. In the last step we will interface the new PTS algorithm with the new time segmentation and assess the impact on the quality of our emissivity estimates over desert. Note that we are likely to keep both Version 1 and Version 2 1b algorithm configurations running to accommodate the potential loss of Terra data.

As part of this work we also performed a preliminary evaluation of newly recalibrated AMSR-E Tbs produced by RSS. Some of the artifacts that had been previously identified in the V09 NSIDC distributed AE\_L2A product have been corrected in the new dataset. Some anomalies are still present in our retrieved surface emissivities and more work is required in order to determine whether they originate from instrumental problems or biases in our input data. This study will be conducted in conjunction with the assessment of the impact of replacing NCEP atmospheric correction with AIRS. These aspects are important for surface emissivity characterization but even more so for the planned subsequent application of the database to PW retrieval. Some of this work may be accomplished as part of our newly awarded NASA contract for the production of an updated microwave land surface emissivity database or as part of our involvement in the GPM calibration work.

At this stage it is believed that the improvements proposed here will lead to the best quality microwave land surface emissivity atlas produced to this date. The LST validation strategy developed as part of this work proves to be a very valuable tool for assessing the accuracy of the emissivity product. There are areas where no emissivity estimates are currently produced at certain times of the year. These generally correspond to areas where surface conditions (emissivity, penetration, soil temperature) change significantly on time scales of the order of a week (or less) and where cloudiness prevents one from accumulating enough clear observations to be able to track emissivity changes. As a byproduct of our analysis, we output flags that identify isolated events (outliers) and characterize the type of changes occurring in any particular area at any given time. These flags are particularly useful for retrieving from our database the information needed to study certain situations and develop alternate approaches or refinements (e.g. temporal emissivity interpolation) for dealing with those situations and eventually bridging the gaps in our coverage.

Note that LST retrieval over highly penetrating areas remains a challenge even when the surface is temporally stable. Unlike vegetated surfaces, where it is often sufficient to rely on the 11 GHz channel for estimating surface temperature (under non-precipitating conditions), retrieval of soil temperature profile (and in particular near surface temperature) over deserts requires accurate atmospheric correction in areas where data from other sources (e.g. IR sounders) may not be available (due to the presence of clouds or airborne dust). To this end we have added the capability to simultaneously extract cloud liquid water and water vapor profile information from the microwave observations. The technique relies on availability of a surface emissivity constraint, derived in the clear-sky using our 1b algorithm, and for microwave imagers, temperature profile data from NWP model. However, because of the difficulties in distinguishing the direct impact of CLW on the higher frequency channels (which sense less deeply underneath the surface) from the impact of clouds on insolation or surface cooling at night (especially over weakly polarized desert surfaces), there may be significant cross-talk between retrieved CLW and near surface temperature. Short of assimilating the microwave observations into a coupled land-atmosphere model, developing mechanisms for detecting the presence of CLW from the microwave observations over deserts (in complement to the IR imaging information) and flagging the surface retrieval product is a more reachable objective in the short term. We also demonstrated the potential usefulness of the microwave surface temperature data for quality control of IR derived LST in dust events.

Finally, we also demonstrated the usefulness of our emissivity retrieval methodology, beyond its main application to production of a surface information database, for assessing and comparing the quality of external products used as input into our process. This type of approach had been used for comparing the quality of different LST datasets (e.g. Moncet *et al.*, 2011c). In this study, it has also been applied to atmospheric products (e.g. AIRS).



## 9 REFERENCES

- Brown, S. T. and C. S. Ruf (2005), Determination of an Amazon hot reference target for the on-orbit calibration of microwave radiometers, *J. Atmos. Ocean. Tech.*, 22, 1340-1352.
- Clough, S.A., M.J. Iacono and J.-L. Moncet, 1992, Line-by-line calculation of atmospheric fluxes and cooling rates: Application to water vapor, *J. Geophys. Res.*, 97, 15761-15785, 1992.
- Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara and P. D. Brown, 2005, Atmospheric radiative transfer modeling: a summary of the AER codes, Short Communication, *J. Quant. Spectrosc. Radiat. Transfer*, 91, 233-244, 2005.
- Coll, C., Z. Wan, and J. M. Galve, 2009, Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product, *J. Geophys. Res.*, 114, D20102, doi: 10.1029/2009JD012038.
- Fily, M., A. Royer, K. Goïta, and C. Prigent, 2003: A simple retrieval method for land surface temperature and fraction of water surface determination from satellite microwave brightness temperatures in sub-arctic areas, *Remote Sensing of Environment*, 85, pp. 328–338, doi:10.1016/S0034-4257(03)00011-7.
- Galantowicz, J. F., Moncet, J.-L., P. Liang, A. E. Lipton, G. Uymin, C. Prigent and C. Grassotti, 2011, Subsurface emission effects in AMSR-E measurements: Implications for land surface microwave emissivity retrieval, *J. Geophys. Res.*, 116, D17105, doi:10.1029/2010JD015431
- Galantowicz, J. F., (2004): Metrics for the quality of footprint-matched passive microwave measurements, *Geosciences and Remote Sensing Symposium, 2004. IGARSS '04 Proceedings*, vol.6, no., pp.3763-3766 vol.6, 20-24. doi:10.1109/IGARSS .2004.1369941
- Galantowicz, J., T. Nehr Korn, R. Hoffman, A. Lipton (2003), Algorithm Theoretical Basis Document (ATBD) for the Conical-Scanning Microwave Imager/Sounder (CMIS) Environmental Data Records (EDRs), Vol. 1: Overview, Pt. 2: Footprint Matching, Version 1.4. Available at <http://www.aer.com/scienceResearch/mwrs/pubs.html>.
- Goudie, A.S., and N.J. Middleton (2001), Saharan dust storms: nature and consequences, *Earth-Science Reviews* 56, 179–204
- Holmes, T. R. H., R. A. M. de Jeu, M. Owe and A. J. Dolman, 2009, Land surface temperature from Ka band (37GHz) passive microwave observations, *J. Geophys. Res.*, Vol. 114, doi: 10.1029/2008JD010257
- Hulley, G. C., and S. J. Hook, 2009, Intercomparison of versions 4, 4.1 and 5 of the MODIS Land Surface Temperature and Emissivity products and validation with laboratory measurements of sand samples for the Namib desert, Namibia, *Remote Sens. Environ.*, 113, 1313-1318.

- Hunt, K. P., J. J. Niemeier, L. K. da Cunha, and A. Kruger (2011), Using cellular network signal strength to monitor vegetation characteristics, *IEEE Geosci. Rem. Sens.*, 8, 346-349.
- Isaacs, R. G., Y.-Q. Jin, R. D. Worsham, G. Deblonde, and V. J. Falcone Jr. (1989), The RADTRAN microwave surface emission models, *IEEE Trans. Geosci. Remote Sens.*, 27, 433-440.
- Jimenez, C. C. Prigent, J. Catherinot, W. Rossow, P. Liang and J.-L. Moncet (2011): A comparison of ISCCP land surface temperature with other satellite and in situ observations, submitted to *J. Geophys. Res.*
- Jones et al. (2009) "Daily land surface air temperature retrieval from AMSR-E: Comparison with AIRS/AMSU" *IEEE JSTARS*, special issue on Microwave Land Hydrology and Remote Sensing
- Jones, L. A., C. R. Ferguson, J. S. Kimball, K. Zhang, S. K. Chan, K. C. McDonald, E. G. Njoku, E. F. Wood, 2010 (in press). Daily Land Surface Air Temperature Retrieval from AMSR-E: Comparison with AIRS/AMSU, *IEEE J. Appl. Earth Obs. Rem. Sens.* (J-STARS).
- Jones, L. A., J. S. Kimball, E. Podest, K. C. McDonald, S. K. Chan, E. G. Njoku. 2009. A Method for Deriving Land Surface Moisture, Vegetation, and Open Water Fraction from AMSRE. *IEEE Int. Geosci. Rem. Sens. Symp. IGARSS '09*, July 13-17, Cape Town, South Africa.
- Kabela, E. D., B. K. Hornbuckle, M. H. Cosh, M. C. Anderson, and M. L. Gleason, 2009, Dew frequency, duration, amount, and distribution in corn and soybean during SMEX05, *Agric. and Forest Met.*, 149, 11-24.
- Karbou, F., C. Prigent, L. Eymard, and J. Pardo, 2005, Microwave land emissivity calculations using AMSU-A and AMSU-B measurements, *IEEE Trans. on Geosci. Remote Sens.*, vol. 43, no. 5, pp. 948-959.
- Liebe, H. J., P. W. Rosenkranz, and G. A. Hufford, 1992, Atmospheric 60 GHz oxygen spectrum: New laboratory measurements and line parameters, *J. Quant. Spectrosc. Radiat. Transfer*, 48, 629-643, doi: 10.1016/0022-4073(92)90127-P.
- Liebe, H. J., and T. A. Dillon, 1969, Accurate foreign-gas broadening parameters of the 22-GHz H<sub>2</sub>O line from refraction spectroscopy, *J. Chem. Phys.*, vol. 50, 727-732
- Meissner, T. and F. Wentz, 2010, Intercalibration of AMSR-E and WindSat brightness temperature measurements over land scenes, presented at Microrad 2010, Washington D.C., 1-4 March 2010.
- Moncet, J.-L., P. Liang, J. F. Galantowicz, A. E. Lipton, G. Uymin, C. Prigent and C. Grassotti, 2011, Land surface microwave emissivities derived from AMSR-E and MODIS measurements with advanced quality control, *J. Geophys. Res.*, 116, D16104, doi:10.1029/2010JD015429

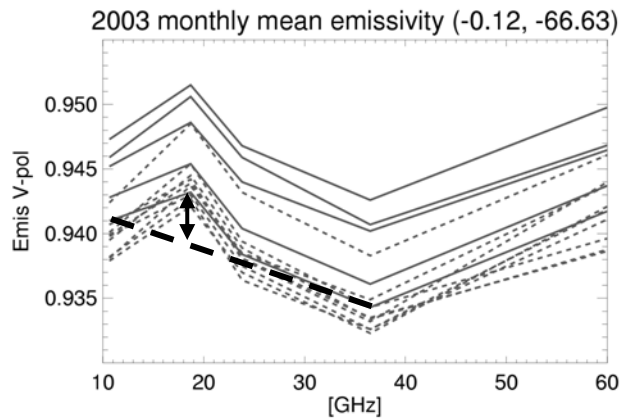
- Moncet J.-L., P. Liang, A. E. Lipton J. F. Galantowicz, and C. Prigent, 2011, Discrepancies between MODIS and ISCCP land surface temperature products analyzed with microwave measurements, *J. Geophys. Res.*, 116, D21105, doi: 10.1029/2010JD015432
- Njoku, E.G., P. Ashcroft, T. K. Chan and L. Li, 2005, Global Survey and Statistics of Radio-Frequency Interference in AMSR-E Land Observations, *IEEE Trans. GeoSci. Rem. Sens.* Vol. 43, No. 5, May 2005.
- Parinussa, R. M., R. A. M. de Jeu, T. R. H. Holmes and J. P. Walker, 2008, Comparison of microwave and infrared land surface temperature products NAFE'06 research sites, *IEEE Geos. Rem. Sens. Lett.*, Vol. 5, No. 4, pp.783.
- Payne, V. H., E. J. Mlawer, K. E. Cady-Pereira and J.-L. Moncet, 2011, Water vapor continuum absorption in the microwave, *IEEE Trans. Geosci. Remote Sens.*, vol 49 (6), 2194-2208
- Prigent C., W. B. Rossow, E. Matthews (1997), Microwave land surface emissivities estimated from SSM/I observations, *J. of Geophys. Res.*, 102, 21867-21890.
- Prigent, C., W. B. Rossow, E. Matthews, and B. Marticorena (1999), Microwave radiometer signatures of different surface types in deserts, *J. Geophys. Res.*, 104 (D10):12147-12158.
- Prigent C., F. Aires, and W. B. Rossow (2006), Land surface microwave emissivities over the globe for a decade, *Bull. Amer. Meteorol. Soc.*, 87, 1573–1584.
- Rosenkranz, P. W., 1998, “Water vapor microwave continuum absorption: A comparison of measurements and models”, *Radio Science*, vol. 33, 919-928
- Tedesco, M., R. E. J. Kelly, James L. Foster, and Alfred T. C. Chang (2004a): Updated daily. AMSR-E/Aqua Daily L3 Global Snow Water Equivalent EASE-Grids V002, [list the dates of the data used]. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.
- Tedesco, M., R. E. J. Kelly, James L. Foster, and Alfred T. C. Chang (2004b): Updated monthly. AMSR-E/Aqua Monthly L3 Global Snow Water Equivalent EASE-Grids V002, [list the dates of the data used]. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.
- Wan, Z., 2008, “New refinements and validation of the MODIS land-surface temperature/emissivity products,” *Remote Sens. Environ.*, 112, 59-74.
- Wan, Z., 2007, Collection-5 MODIS Land Surface Temperature Products Users' Guide, March 2007
- Wan, Z., 2006, Collection-4 MODIS Land Surface Temperature Products Users' Guide, March 2006
- Wan, Z., 1999, “MODIS Land-Surface Temperature Algorithm Theoretical Basis Document, Version: 3.3, available at [http://modis.gsfc.nasa.gov/data/atbd/atbd\\_mod11.pdf](http://modis.gsfc.nasa.gov/data/atbd/atbd_mod11.pdf)

Wan, Z. and Z.-L. Li, 2008, Radiance-based validation of the V5 MODIS land-surface temperature product, *Int. J. Remote Sens.*, 29, 5373-5395

## APPENDIX A

### SYSTEMATIC ANOMALIES IN AMSR-E DERIVED SPECTRAL EMISSIVITIES

Detailed inspection of our Version 1 retrieved AMSR-E emissivities had revealed the presence of systematic anomalies that may be related to instrument calibration errors. Figure A1 shows samples of AMSR-E 1a emissivities retrieved from NSIDC Version 09 AE\_L2A Tbs (Sec. 3) over tropical forest. The emissivities of forest canopies are expected to vary smoothly with frequency in the 11–37-GHz range (Brown and Ruf, 2005, Isaacs et al., 1989), but in this figure the 19-GHz emissivity does not appear to be in alignment with the 11, 24 and 37-GHz emissivities. In addition, we had observed that the average differences between 89V and 37V emissivities (not shown) are higher for AMSR-E than SSM/I, indicating a potential high bias at 89 GHz for AMSR-E. Meissner and Wentz (2010) recently compared AMSR-E and Windsat measurements over the Amazon and African forests and found that the two sets of measurements agreed very closely at 6, 11, 24, and 37 GHz, but the 19 GHz AMSR-E TBs were 2-2.5 K warmer than the corresponding Windsat measurements, which is consistent with the 19-GHz emissivity anomaly in Figure A1.



**Figure A1:** Typical tropical forest emissivity spectra retrieved from AMSR-E. The thick dashed line indicates the degree of alignment of the 11, 24 and 37 GHz V emissivities.

Remote Sensing System (RSS) has been working on recalibrating the SSM/I and AMSR-E instruments. A sample of the newest version of AMSR-E calibrated footprint matched Tbs (RSS Version 7) has recently been acquired from RSS (for the July 2003 period) for internal testing and evaluation purposes. The Tbs in this set were only provided up to 37 GHz. Given the timing of the delivery we only performed a cursory analysis of the new data set.

A comparison of the new RSS AMSR-E Tbs with the NSIDC product reveals that the 11GHz, 24 GHz and 37 GHz channel Tbs are on average 1 to 1.5K warmer than in the old version. Changes in the 19GHz Tbs are somewhat smaller. The order of magnitude of the observed differences (adjustments of the order of 1-2K after recalibration) is within RSS's expectations (Meissner, *personal communication*). Some occasional large differences (>2K) are often found in the first and last scan lines of the day, i.e. mostly over Polar Regions (Fig. A2). The cause of these artifacts is unknown at this point.

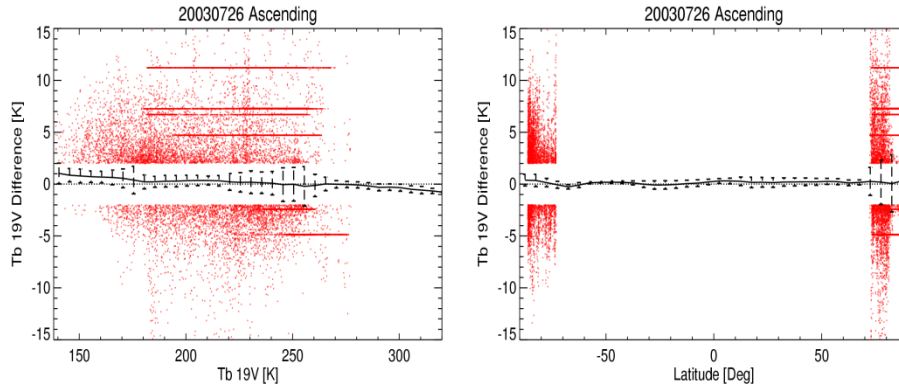


Figures A3 and A4 show typical emissivity spectra derived from the new RSS Tb dataset over vegetated areas in Brazil, Florida and over Northern hemisphere forests. It is apparent from these figures that the corrections introduced by RSS only have a small impact at 19 GHz. Although some features in the emissivity spectra - in particular over Florida - still appear unphysical, the increase in the 11 and 37 GHz emissivities acts in such a way as to reduce the differences between the actual 19 GHz emissivity values and the values obtained by linearly interpolating the 11 and 37 GHz emissivities in frequency, hence partially cancelling the 19GHz artifact shown in Fig. A1. Let  $D_{19}$  denote this difference, i.e.  $D_{19} =$

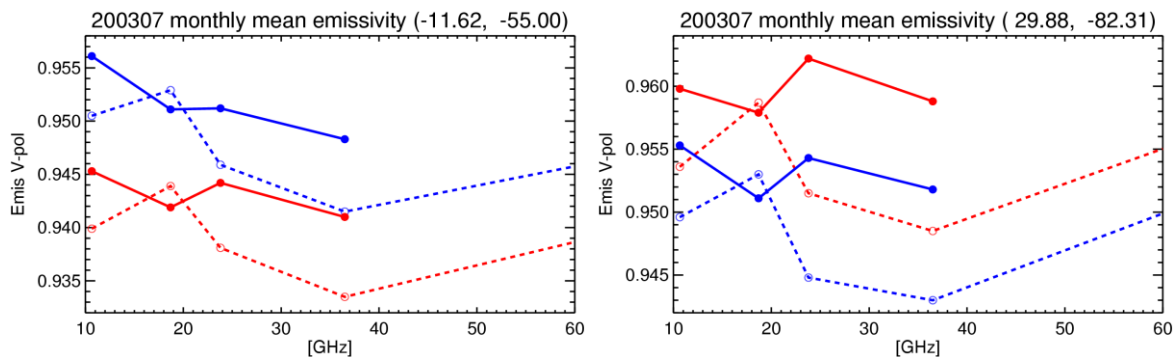
$$D_v = \varepsilon_v - [a_v \varepsilon_{11} - (1 - a_v) \varepsilon_{37}] \quad (\text{A1})$$

evaluated at  $\nu=19\text{GHz}$ . Plots of zonally averaged  $D_{19}$  computed from the mean July 2003 emissivity spectra over vegetated surfaces, with new and current calibration, are shown in Fig. A5. Note that  $D_v$  has little physical meaning. This shape index is introduced here mostly for visualization purposes.

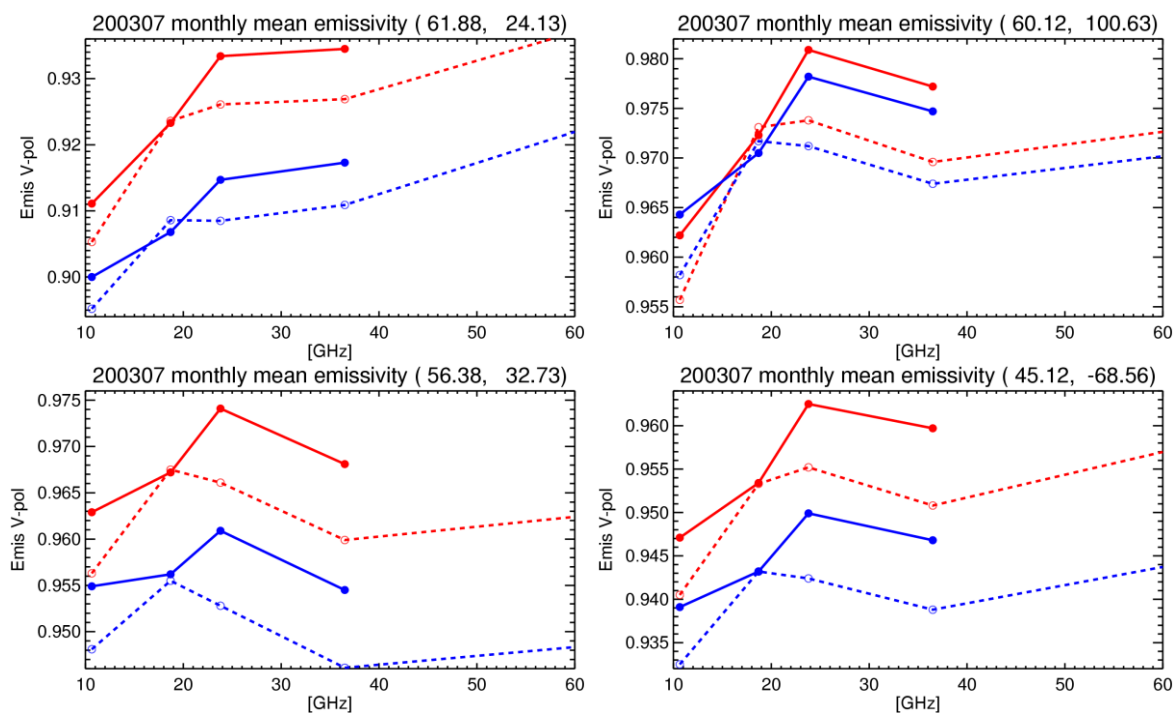
The first order effect of the new RSS calibration (Fig. A5) is to remove a 0.006 bias ( $\sim 2\text{K}$ ) in  $D_{19}$ . As one would expect, in both Fig. A5a and Fig. A5b, the nighttime  $D_{19}$  curves are similar to the daytime ones. The small systematic differences (of the order of 0.001 - or  $\sim 0.3\text{ K}$ ) between the ASC and DES values in the  $20^\circ\text{S} - 40^\circ\text{N}$  latitude range are within calibration uncertainties. However, these differences are seen to increase and reach over 0.002 in the higher Southern and Northern latitudes. The reason why  $D_{19}$  reaches a maximum around  $60^\circ\text{N}$  (over the Northern tundra and boreal forest) and then decreases again at higher latitudes (see Fig. A6) and the reason for the ASC - DES bias in this area are not fully understood. Note that this high northern latitudes feature was found to be most prominent in June and July and it tends to vanish in August.



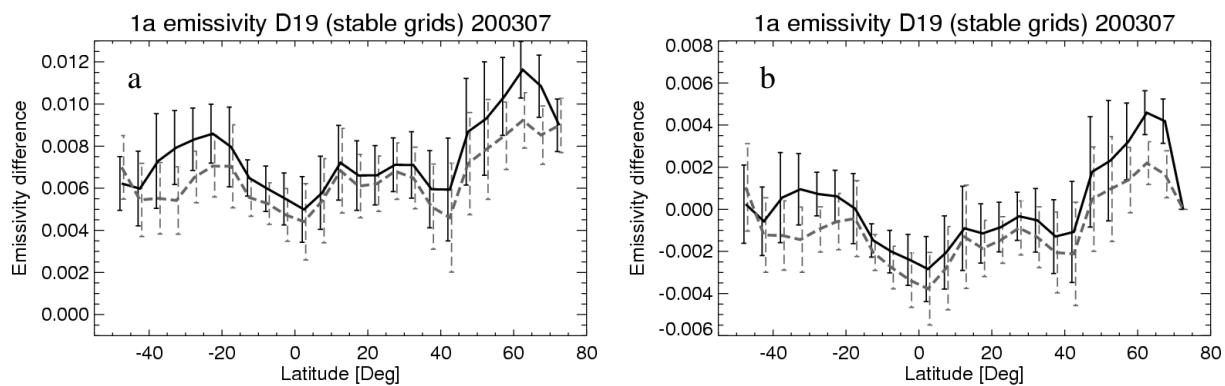
**Figure A2** Scatter plots of the difference between RSS V7 19V and NSIDC AE\_L2A V09 Tbs as a function of NSIDC Tbs (left) and latitude (right). The black lines are means and standard deviation of the Tb difference in each Tb and latitude bin computed from the global data.



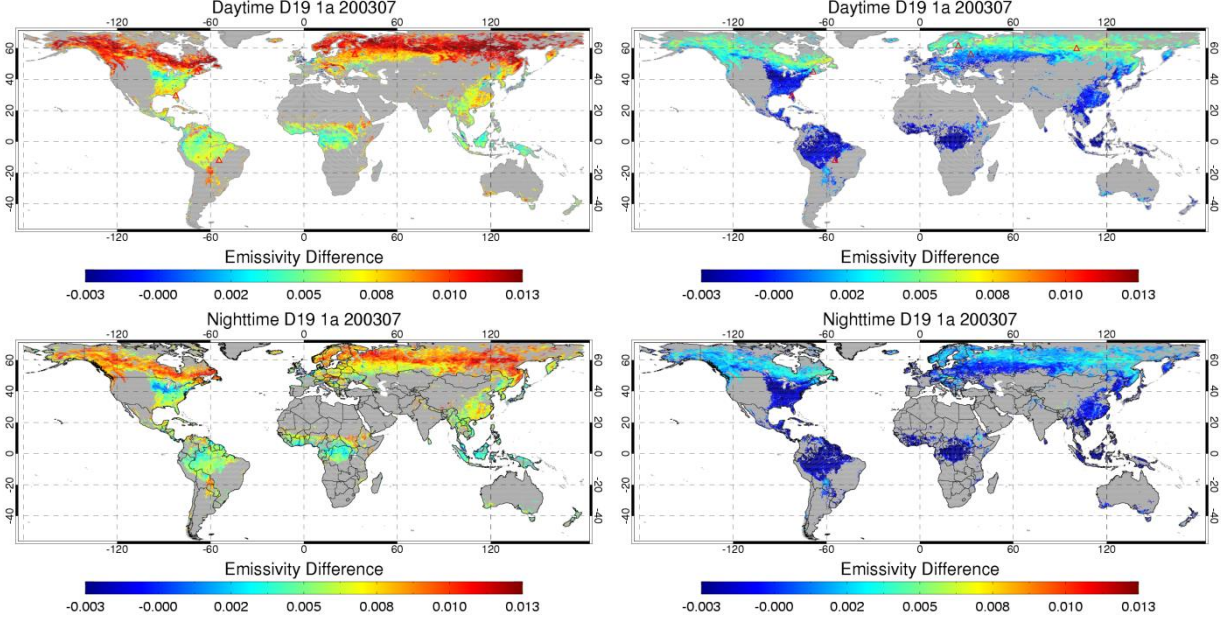
**Figure A3:** Selected V-polarized emissivity spectra over Brazil (left) and Florida (right) with new RSS calibration (solid) and V09 AE\_L2A product (dashed). Daytime and nighttime are identified in red and blue, respectively.



**Figure A4:** Same as Figure A3 for selected locations in Russia, Finland and Maine.

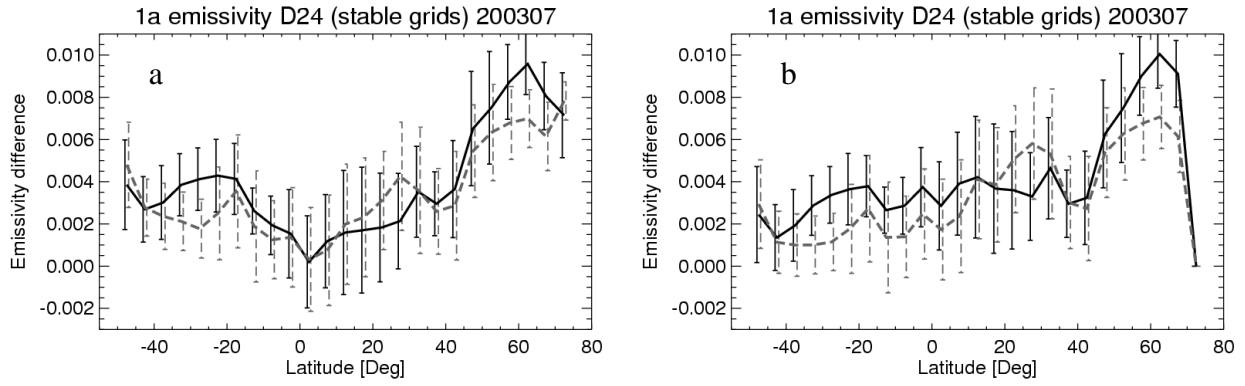


**Figure A5:** Zonally averaged monthly mean  $D_{19V}$  from V09 AE\_L2A product (left) and new RSS calibration (right) for ASC (black solid) and DES (gray dashed) Aqua overpasses (July 2003). The vertical bars represent the standard deviations in each latitude bin.

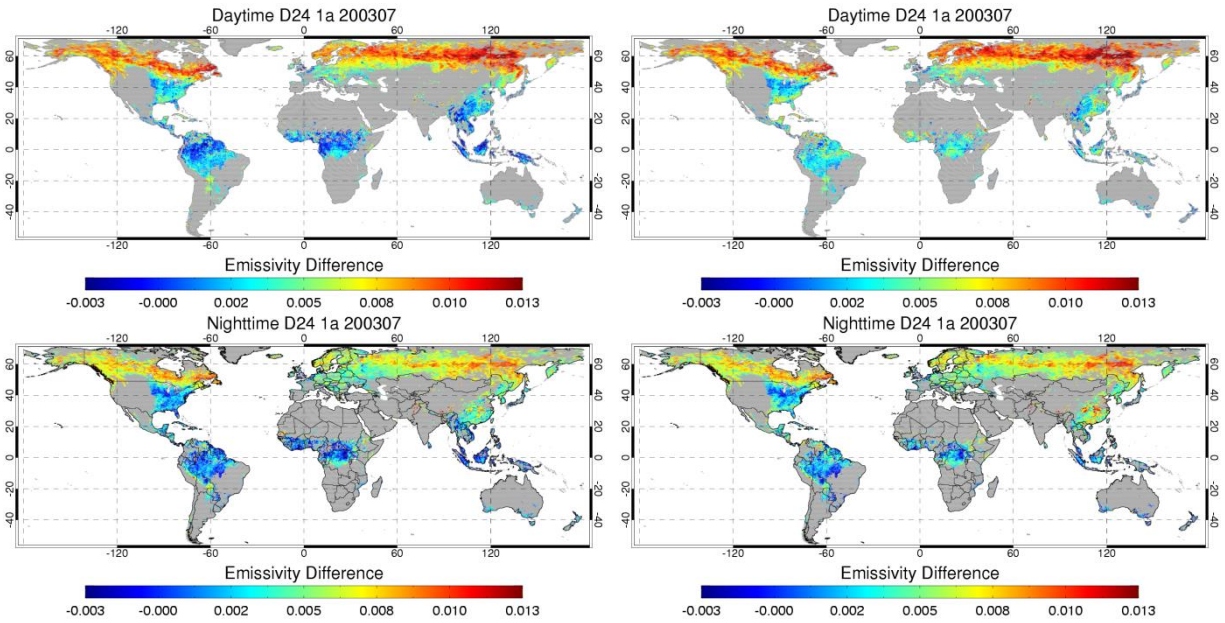


**Figure A6:** Global map of monthly mean  $D_{19}$  with V09 AE\_L2A product (left) and new RSS calibration (right) for ASC (top) and DES (bottom) passes.

Plots of zonally averaged  $D_{24} = D_{\nu}$  (Eq. A1) at  $\nu=24\text{GHz}$ , and corresponding geographical maps are shown in Figs. A7 and A8, respectively. It is apparent from Fig. A7 that the daytime (ASC pass) latitudinal variations of  $D_{24}$  in the  $50^{\circ}\text{S}$  to  $40^{\circ}\text{N}$  latitude range are significantly smaller with the new RSS Tb dataset than with the AE\_L2A Tbs. In particular, the areas of low  $D_{24}$  values over S. American and African tropical forest and over South East Asia in the daytime map of Fig. A8a are much reduced in size or have disappeared with the new calibration. It is also apparent that the  $D_{24}$  DES curves (Fig. A7) do not track the ASC curves as well as in the case of  $D_{19}$  (Fig. A5). The DES  $D_{24}$  goes from being smaller than the ASC  $D_{24}$  between  $\sim 20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  to being larger between  $\sim 20^{\circ}\text{N}$  and  $40^{\circ}\text{N}$ . The areas of significant day/night  $D_{24}$  differences are easily identified by inspecting the maps of Fig. A8. Some of these differences could be attributed in part to biases in our water vapor correction.



**Figure A7:** Same as Fig. A5 for  $D_{24}$ .

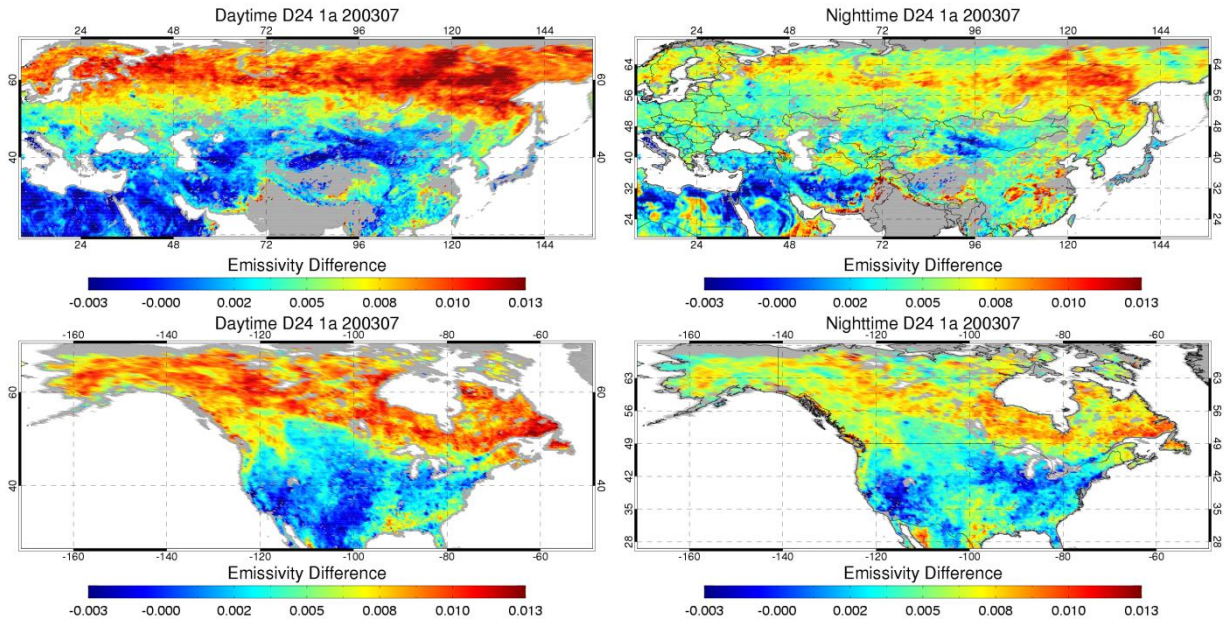


**Figure A8:** Same as Fig. A8 for  $D_{24}$ .

The new RSS calibration seems to result in a better alignment of our low frequency emissivities over vegetated surfaces over most latitudes. However, the 24 GHz emissivity now appears too high in some cases and the 19 GHz too low in other cases, leading to some unrealistic (or at least suspect) emissivity spectral shapes (e.g. Fig. A3). Other features such as the peak in the magnitude of the  $D_{19}$  and  $D_{24}$  terms and, more particularly, the systematic day/night differences in  $D_{19}$  and  $D_{24}$  in the northern high latitudes are still not understood (see also Fig. A9). As mentioned in Sec 6.2, the magnitude of this effect appears to be reduced when using AIRS atmospheric correction with the NSIDC AE\_L2A Tbs. However, the fact that it is still apparent when using an independent water vapor correction would tend to indicate that it may not be entirely caused by biases in the water vapor fields. We are aware of some artifact in the MODIS LST data at 60°N (this issue is documented in [http://landweb.nascom.nasa.gov/cgi-bin/QA\\_WWW/displayCase.cgi?esdt=MOD11&caseNum=SD\\_MOD11\\_10019&caseLocation=cases\\_data](http://landweb.nascom.nasa.gov/cgi-bin/QA_WWW/displayCase.cgi?esdt=MOD11&caseNum=SD_MOD11_10019&caseLocation=cases_data)), but the fact that the latitudinal trends are more pronounced at 24 than at 19 GHz is inconsistent with an attribution to surface temperature bias. Note that the observed latitudinal/seasonal trends and



day night differences may also be indicative of Tb biases that correlate with the orbital position of the spacecraft (such as residual impact of solar intrusion). In order to get more insight into these various aspects it would be useful to inspect emissivities retrieved from other sensors (e.g. WindSat or SSM/I). In the near future, we are also planning on examining the impact on these results of replacing the NCEP water vapor fields used here for our atmospheric correction with the AIRS atmospheric product. Note also that the RT model used to retrieve the spectral emissivity differs from the one used by RSS in their calibration by the magnitude of the continuum absorption (Appendix B). Although we would not expect much impact over land (high surface emissivity) it is possible that these might contribute to some degree to the observed features. This source of difference could be eliminated by upgrading to the latest version of the AER MonoRTM model (Appendix B) and by modifying the width of the 22GHz line to make it consistent with the RSS model or by adopting the same forward model as RSS in our study.



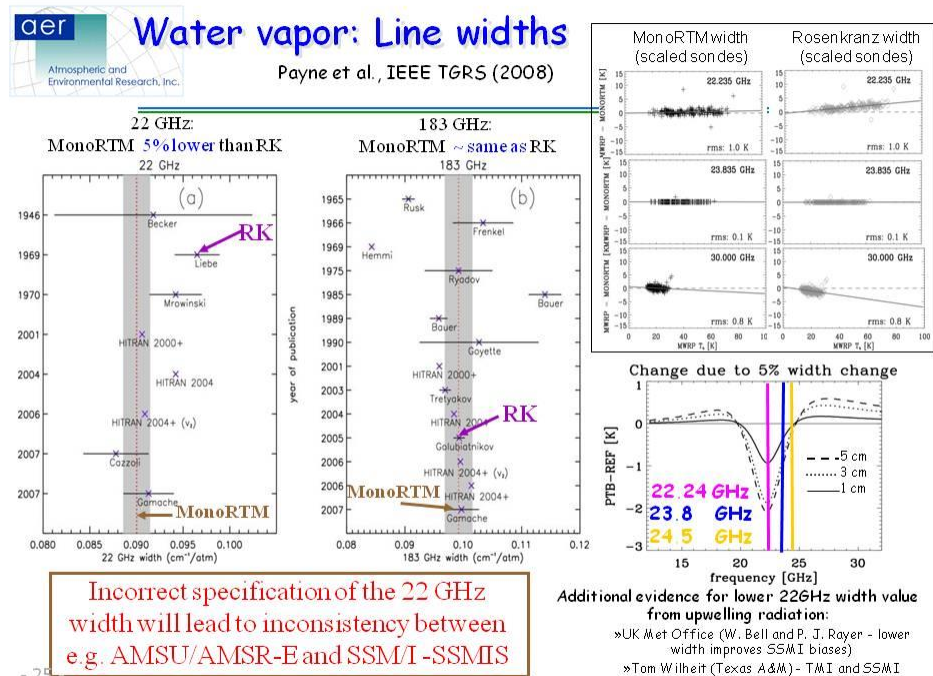
**Figure A9:** Regional daytime (left) and nighttime (right) maps of monthly mean  $D_{24}$  with V09 AE\_L2A product (left) over Asia (top) and N. America (bottom).



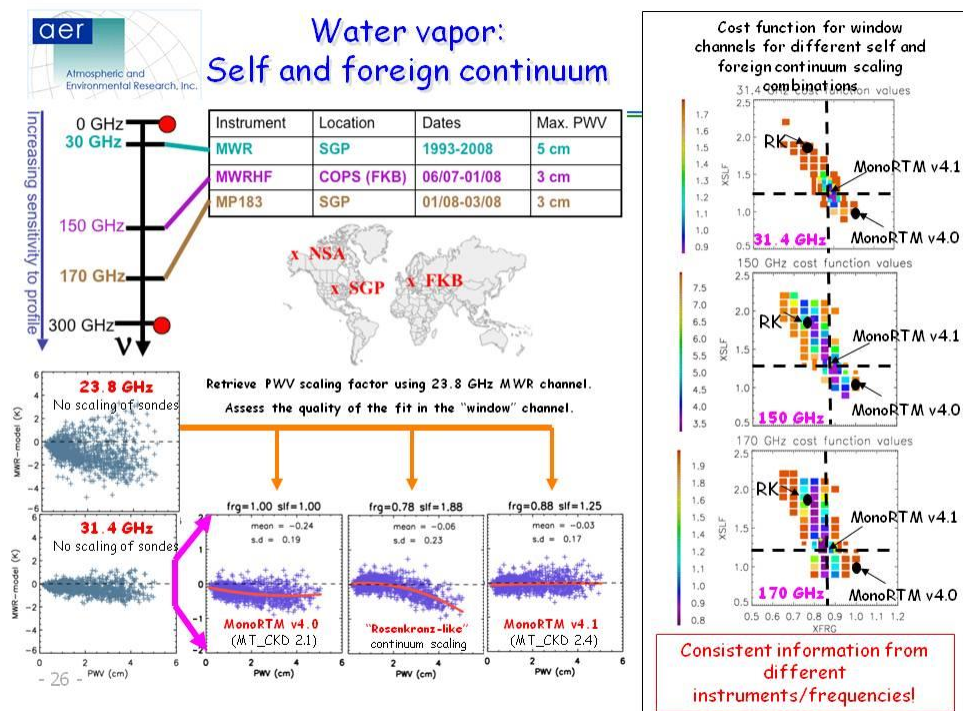
## APPENDIX B

### RELATED SPECTROSCOPIC RESEARCH

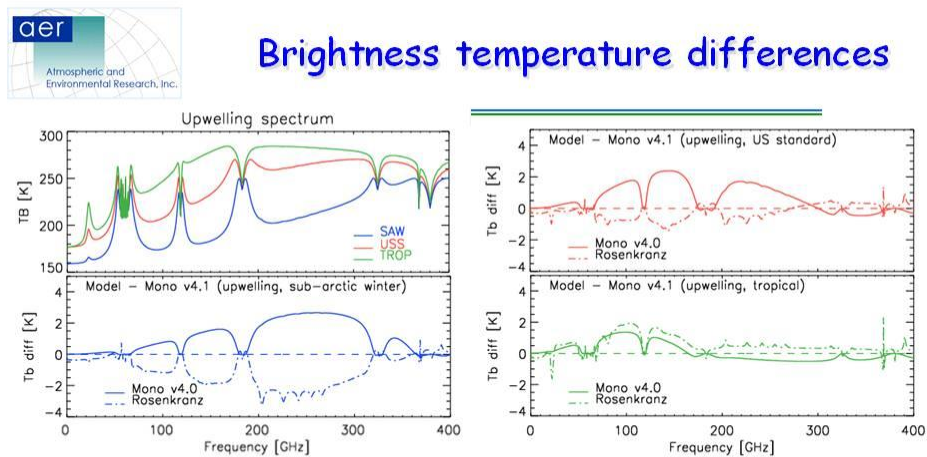
Spectroscopic work being conducted at AER is highly relevant to the calibration issues discussed above in Appendix A. The current status of AER's MonoRTM and LBLRTM line-by-line models (Clough *et al.*, 2005, Clough *et al.* 1992) is summarized in the slides of Fig. B1. This figure also compares the AER model to the Rosenkranz (RK) model (Rosenkranz, 1998) used in the development of our Version 1 emissivity database. The water vapor continuum in AER's models has recently been updated based on model/measurement comparisons using ground-based radiometer (up to 170 GHz) and radiosonde measurements at two different ARM sites (Payne *et al.*, 2011) (Fig. B1b). As shown in Fig. B1b, there are significant differences between the RK and the newer AER self and foreign water vapor continua. The impact of these differences at the AMSR-E frequencies is the largest at 89 GHz (Fig B1c). Independent experiments conducted by RSS have led to modifications to their water vapor continuum parameterization which go in the direction of bringing the AER and RSS models in close agreement in the continuum dominated parts of the spectrum (Meissner, Wentz, *personal communication*). It is our current understanding that the updated AMSR L2 product from RSS will reflect those changes. The only remaining significant difference between the RSS and AER models is in the specification of the 22 GHz water vapor line width, which differs by 5% between the two models. The 22 GHz line parameters used by RSS are inherited from the Liebe width (Liebe and Dillon, 1969) that is used in the RK model (Fig. B1a). This difference has the largest impact on SSM/I derived precipitable water estimates.



(a)



(b)



Note that RSS has recently readjusted their water vapor continuum to remove bias in CLW retrieval (Meissner and Wentz, *personal communication*). RSS and AER continua are now close together at 37 and 89 GHz (RSS model not valid above 89 GHz)

- 27 -

(c)

**Figure B1:** Differences between most widely used absorption models in the microwave region

## APPENDIX C

### AUTMOMATIC R11 TIME SERIES SEGMENTATION (RTS)

The objective of this operation is to produce a division of the time axis into contiguous segments such that the statistical properties (mean and standard deviation) of the R11 variable (denoted here by  $r_i$ , where  $i$  is the time index) within each segment are quasi-stationary. Conceptually, this is equivalent to fitting the R11 time series with a step function  $\gamma$  having the value

$$\gamma_k = \frac{\sum_{i=s_k}^{s_k+L_k-1} r_i}{L_k} \quad (C1)$$

over any one segment  $k$  of length  $L_k$  starting at position  $s_k$ .

The constraints applied in this process are: 1) no segment should be smaller than  $L_{min}$  days and 2) the total number of segments should be less than  $N_{max}$ .  $L_{min}$  and  $N_{max}$  are currently set to 7 and 16, respectively. These values are subject to change. The former constraint is meant to ensure that we have enough samples within each interval to build reasonable statistics. Another reason for imposing this constraint is that the shorter the length of the segment the less chances we have of being able to collect a sufficiently large number of clear observations for accurately estimating the segment's mean emissivity. Although our approach can operate on any specified time period we currently limit our analysis to one year of data.

This determination of  $\gamma$  is done in four steps:

- a) Fit the R11 time series with a smoothly varying string-like function  $f_i$ . This step has the advantage that it reduces noise in the data, thereby making the segmentation more robust.
- b) Apply the segmentation algorithm to  $f_i$
- c) Adjust the start and end points of each segments obtained in step (b) based on the actual R11 data
- d) Make final breakpoint adjustment to ensure that the *rms* difference between  $r_i$  and  $\gamma_i$ , in any given month, does not exceed the standard deviation of  $r_i$  for that month, which, when no transition occurs during the month, is equivalent to requiring that the mean offset between  $r$  and  $\gamma$  is within some prescribed tolerance. This last step ensures that the new segmentation does not miss subtle but potentially significant fluctuations in the monthly mean surface emissivities currently captured with our Versions 1 MTD approach.

The function  $f_i$  is derived in step (a) by minimizing the following generic cost function evaluated over the whole period of analysis:

$$\sum_i \left[ (f_i - r_i)^2 + a(f_{i+1} - f_i)^2 + b(f_{i+1} + 2f_i - f_{i-1})^2 \right] \quad (C2)$$

The first term in this expression is the deviation of the function  $f$  from the actual observations. The second term penalizes large slope values and the third term controls the local curvature of  $f$ . Coefficients  $a$  and  $b$  are tuning parameters which are currently set to 0 and 1, respectively.

The value of  $f$  minimizing the above cost function for any given day is:

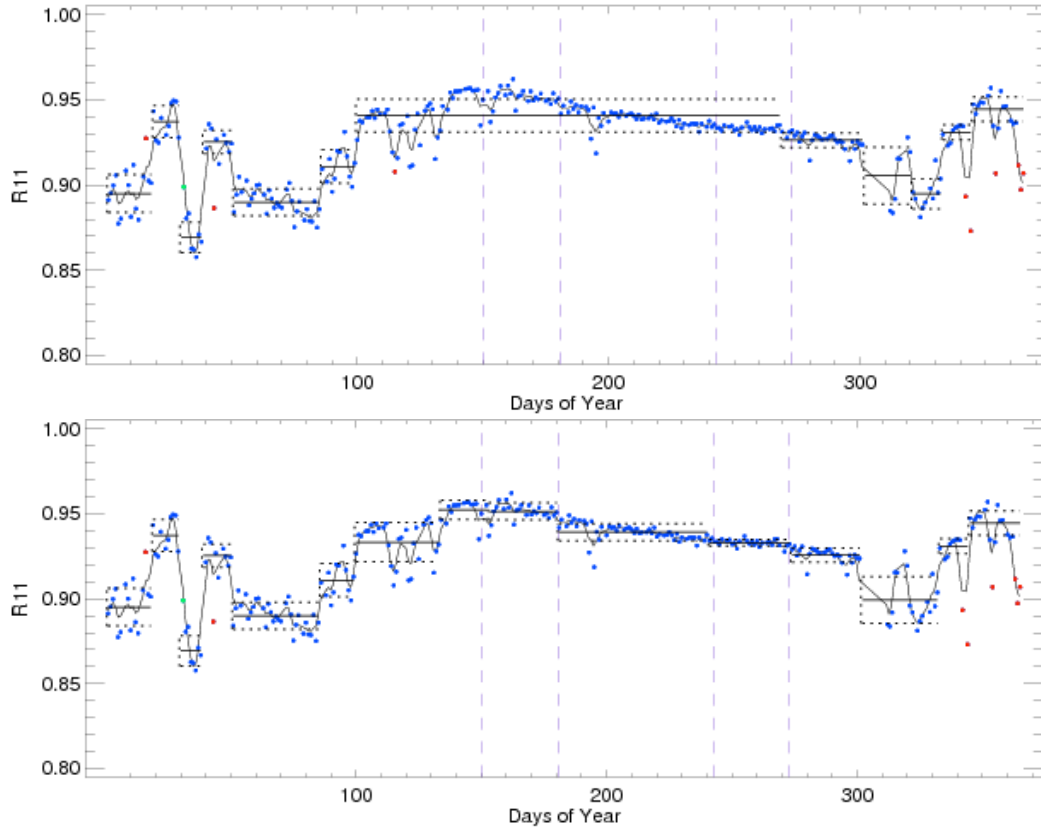
$$\hat{f}_i = [r_i + a(f_{i+1} + f_{i-1}) + b(-f_{i+2} + 4f_{i+1} + 4f_{i-1} - f_{i-2})] / (1 + 6b + 2a) \quad (C3)$$

Note Eq. (C3) takes a slightly different form at the beginning and end of the observation period. This special treatment of the edges can be avoided altogether by adding samples on both sides of the interval and, if desired, by ensuring that the resulting series is periodic.

In principle, one could solve for all values of  $f_i$  simultaneously but this approach becomes time consuming when the analysis periods exceeds  $\sim 30$  days. Instead, a solution is arrived at iteratively. Our fast procedure starts with setting  $f_i = f_{0,i} = r_i$  for all  $i$ 's. At each step  $m$ , a new function  $f_m$  is obtained by substituting the value of the function  $f_{m-1}$  at a single day  $i$ , chosen at random within the analysis period, with a new value derived by solving Eq. (C3). The final solution is obtained by repeating this operation a large number of times until convergence is reached. Note that in this process the same days are revisited several times. The maximum number of iterations is currently set to 20,000 but convergence is generally achieved well before reaching this limit.

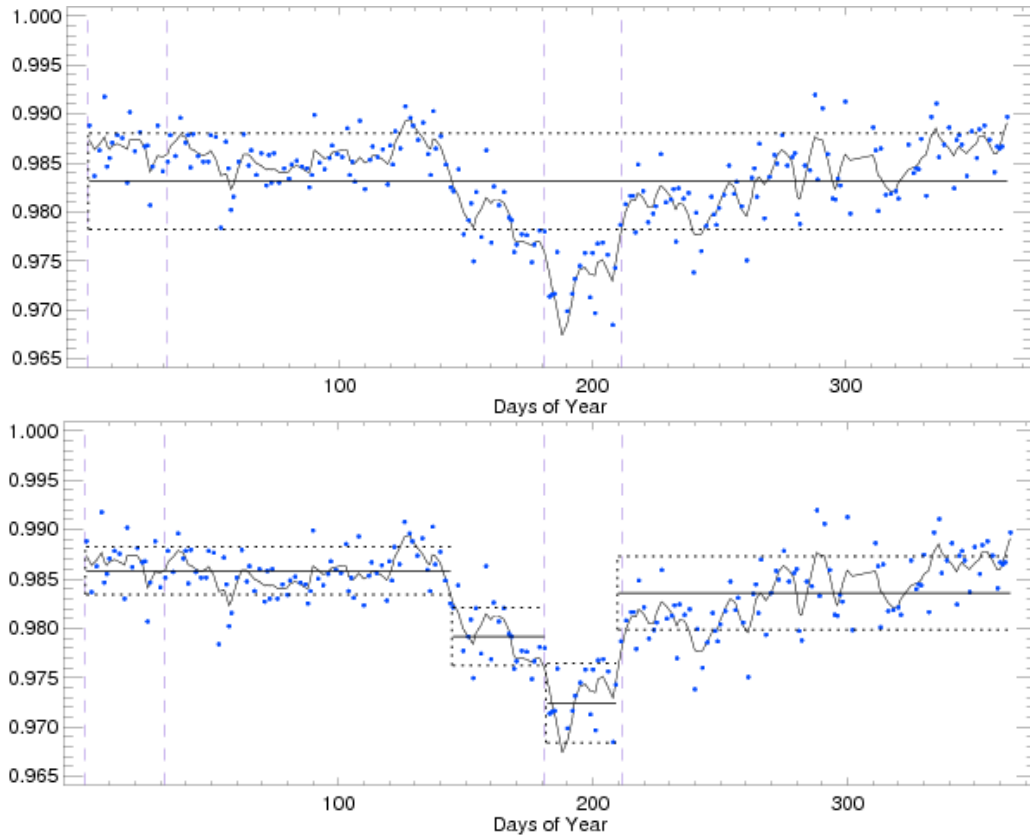
In step (b), a first guess solution for the total number of segments and position of start and end points of each segment is obtained by fitting  $f$  with a step function  $\gamma$ . This procedure starts with a single segment, i.e. by setting  $\gamma = \gamma_0$  where  $\gamma_{0,i} = m_0$ , the mean value of  $f$  over the year-long analysis period, and computing the sum of squares of the difference between  $f_i$  and  $\gamma_{0,i}$  (standard deviation of  $f_i$ ),  $\sigma_0^2$ . We then attempt to break this initial segment into two sub-segments by testing exhaustively all possible positions for the breakpoint between the sub-segments and retaining the solution that provides the maximum improvement in the fit. Once the position of the first breakpoint has been determined, the algorithm looks for a 2<sup>nd</sup> break point by considering each new sub-segment independently and applying the same procedure as above within the sub-segments. At each new step  $m$  in this procedure, the sum of squares of the difference between the new function  $\gamma_{m,i}$  and  $f_i$ ,  $\sigma_m^2$ , is evaluated and compared to its previous value,  $\sigma_{m-1}^2$ . If at any step the maximum improvement in the fit is smaller than some preset threshold  $\delta\sigma_{\min}^2$  the procedure stops. Otherwise the algorithm attempts to add a new breakpoint.  $\delta\sigma_{\min}^2$  is initially set to  $5 \times 10^{-6}$ . Note that this value corresponds to the delta-*rms* difference obtained with a mean offset between the R11 data and the fitting function  $\gamma$  in any given month approximately equal to 0.008. For our purpose, a value between 0.001 and 0.003 is more appropriate. However, when using a smaller  $\delta\sigma_{\min}^2$  the algorithm attempts to fit the fine structure in the  $f$  function which results in too many segments being produced. The last step of this procedure (step d) is a far more efficient way of meeting this objective. If the number of segments reaches the maximum limit  $N_{\max}$ ,  $\delta\sigma_{\min}^2$  is increased by equal increments of  $5 \times 10^{-6}$  and the last step is repeated until the total number of segment is  $\leq N_{\max}$ .

In Step (c), the actual R11 data is compared to the function  $\gamma$  and mean and standard deviation of the differences are computed for each segment. Outliers within a segment are tagged as such and removed from the statistics. There is also some uncertainty in the breakpoint position produced in step (b) that is inherent to the smoothing effect of the quadratic fit. A fine adjustment of the breakpoints is required in order to “absorb” the R11 samples located near the edges of two adjacent segments. If the value of a given sample in this transition zone is within two standard deviations of the mean R11 value for one of its two adjacent neighboring segments, it is absorb by that segment. Otherwise this sample is dropped and tagged as “orphan”. In this process, the length of a segment may become smaller than  $L_{min}$ . In this case, the segment is eliminated and the samples originally belonging to this segment are reassigned to the neighboring segments by extending the length of these two segments and setting a new breakpoint position so as to minimize the *rms* difference between  $r$  and  $\gamma$  over the two segments.



**Figure C1:** Results of R11 time series (blue dots) fit for a grid box in South Kazakhstan (43.625°N 69.59°E) after steps (a)-(c) (top panel) and step (d) (lower panel). In this figure, the outliers are shown as red dots and the orphans in green. The smoothing function  $f$  is represented by the black curve

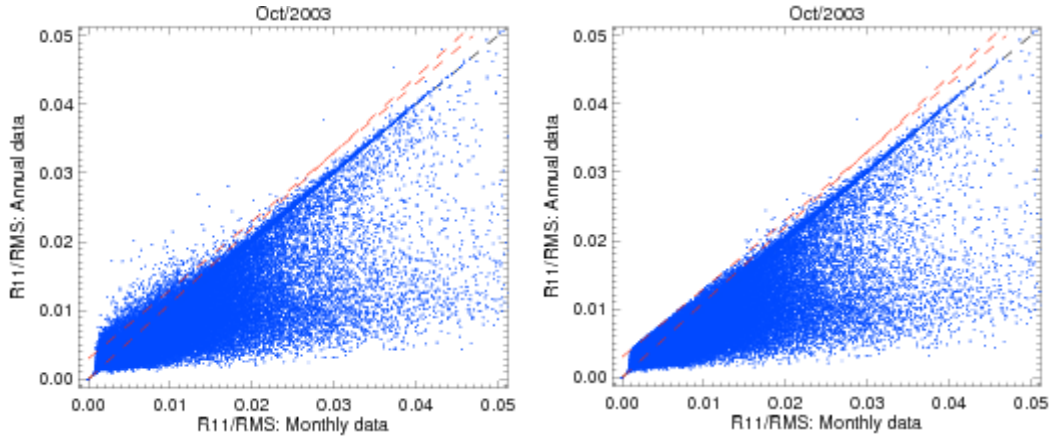




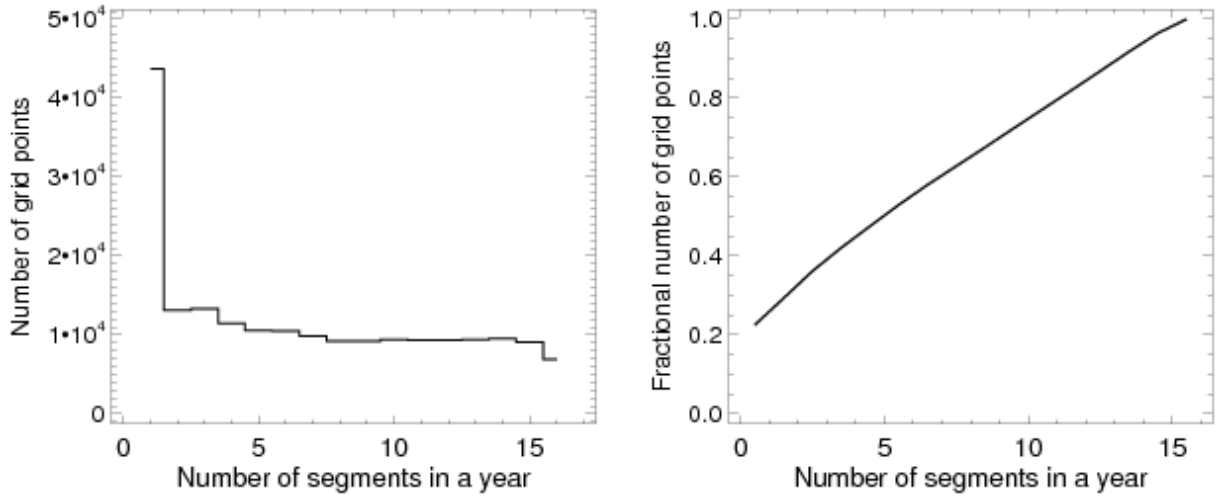
**Figure C2:** Same as Fig. C1 for another location in the Democratic Republic of the Congo (5.875 °S 17.22°E).

The final step (Step d) controls the performance of our intermediate solution produced in steps a-c to ensure that the monthly rms difference between  $r$  and  $\gamma$  does not exceed the standard deviation of R11 in any given month by more than some prescribed  $\varepsilon$ . If this condition is not met for a particular month (or months), we use beginning and ending of these months as our initial guess in step b and repeat steps b - d. Note that the final number of segments produced by this additional cycle is not significantly different from the initial number of segments obtained in the first pass. It is our experience that this procedure only has to be applied once.

Examples of time series fit before and after step (d) are shown in C1 and C2. Fig. C3 compares the monthly performance of new RTS segmentation with MTD for the month of October 2003. Statistics for the number of segments produced by the RTS approach for year 2003 are shown in Fig. C4. Note that the MTD approach uses 12 segments for all 196478 land grid points.



**Figure C3:** Comparison of monthly *rms* difference between R11 data and step function  $\gamma$  with monthly R11 standard deviation (MTD segmentation scheme) for the month of October for all 196478 land grid points. The left panel show the results obtained before step (d) and the right one after step (d) with tolerance  $\varepsilon=0.003$ .



**Figure C4:** Number of grid points versus number of segments with RTS approach (left) (total number of land grid points: 196478) and fraction of grid points described with fewer than  $N$  (in  $x$ -axis) segments.