# Microwave implementation of two-source energy balance approach for estimating evapotranspiration

Thomas R. Holmes[1], Christopher R. Hain[2], Wade T. Crow[3], Martha C. Anderson[3], William P. Kustas[3]

[1]Hydrologcal Sciences Lab, NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA
[2]Earth Science Office, NASA Marshall Space Flight Center, Huntsville, AL, USA
[3]Hydrology and Remote Sensing Laboratory, USDA-ARS, Beltsville, MD 20705, USA

*Correspondence to*: Thomas R. Holmes (thomas.r.holmes@nasa.gov)

**Abstract.** A newly developed microwave (MW) land surface temperature (LST) product is used to substitute thermal infrared (TIR) based LST in the Atmosphere Land Exchange Inverse (ALEXI) modelling framework for estimating ET from space. ALEXI implements a two-source energy balance (TSEB) land surface scheme in a time-differential approach, designed to minimize sensitivity to absolute biases in input records of LST through the analysis of the rate of temperature change in the morning. Thermal infrared (TIR) retrievals of the diurnal LST curve, traditionally from geostationary platforms, are hindered by cloud cover, reducing model coverage on any given day. This study tests the utility of diurnal temperature information retrieved from a constellation of satellites with microwave radiometers that together provide 6-8 observations of Ka-band brightness temperature per location per day. This represents the first ever attempt at a global implementation of ALEXI with MW-based LST and is intended as the first step towards providing all-weather capability to the ALEXI framework.

The analysis is based on 9-year long, global records of ALEXI ET generated using both MW and TIR based diurnal LST information as input. In this study, the MW-LST sampling is restricted to the same clear sky days as in the IR-based implementation to be able to analyse the impact of changing the LST dataset separately from the impact of sampling all-sky conditions. The results show that long-term bulk ET estimates from both LST sources agree well, with a spatial correlation of 92% for total ET in the Europe/Africa domain and agreement in seasonal (3-month) totals of 83-97 % depending on the time of year. Most importantly, the ALEXI-MW also matches ALEXI-IR very closely in terms of 3-month inter-annual anomalies, demonstrating its ability to capture the development and extent of drought conditions. Weekly ET output from the two parallel ALEXI implementations is further compared to a common ground measured reference provided by the FLUXNET consortium. Overall, the two model implementations generate similar performance metrics (correlation and RMSE) for all but the most challenging sites in terms of spatial heterogeneity and level of aridity. It is concluded that a constellation of MW satellites can effectively be used to provide LST for estimating ET through ALEXI, which is an important step towards all-sky satellite-based retrieval of ET using an energy balance framework.

## 1 Introduction

Estimating terrestrial evapotranspiration (ET) at continental to global scales is central to understanding the partitioning of energy and water at the earth surface and for evaluating modelled feedbacks operating between the atmosphere and biosphere. ET is an important flux that links the water, carbon, and energy cycles (Campbell and Norman, 1998). Approximately two-thirds of the precipitation over land is returned to the atmosphere by ET (Baumgartner & Reichel, 1975). Moreover, ET consumes 25-30% of the net radiation reaching the land surface (Trenberth et al., 2009). ET occurs as a result of atmospheric demand for water vapor and depends on the availability of water and energy. When plants are present, this balancing is controlled by leaf-level stomatal controls, and in agricultural areas the water availability may also be managed at the field scale through irrigation or drainage. The high spatial and temporal variability in the driving mechanisms in combination with possible field-scale management decisions poses a significant challenge to bottom-up modelling of ET at sub-monthly time scales, even at the spatial scales of numerical weather prediction (NWP) models (5-25 km). In order for NWP models to improve the characterization of the surface energy budget, there is

a need for timely diagnostic information on ET (Hain et al., 2015). This, in turn, could lead to a more timely and accurate identification of developing droughts (Anderson, 2011) which would aid farm-level management decisions as well as regional yield impact predictions.

ET is highly variable in space, so no amount of ground stations can provide an accurate estimate of the spatial average over larger domains, let alone the globe. Therefore, approaches have been developed to integrate satellite data with models to estimate ET from space. Surface energy balance approaches use surface temperature observations as the main diagnostic to estimate ET by partitioning the available energy into turbulent fluxes of sensible heating (H) and latent heating (LE). In the two source energy balance (TSEB) approach (Kustas and Norman, 1999; Norman et al., 1995) the partitioning is evaluated for the soil and the canopy separately. Anderson et al (1997) modified TSEB to leverage observations of the time evolution of surface temperature as a way to reduce the impact of biases in instantaneous temperature observations on the ET retrieval. This approach allowed for regional implementation of TSEB and came to be known as the Atmosphere Land Inverse Exchange (ALEXI) method (Anderson et al., 2007a; Mecikalski et al., 1999).

To date, ALEXI has always been implemented with land surface temperature (LST) retrievals from thermal infrared (TIR) imaging radiometers (Anderson et al., 2011). Most applications of ALEXI have utilized data products from geostationary satellites, for example the Geostationary Operational Environmental Satellite (GOES) with coverage over the Americas. More recently it has been applied to records from polar orbiting satellites to obtain consistent global coverage from a single sensor with short latency. This is based on day-night temperature differences from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua satellite from NASA's Earth Observing System (EOS) program (Hain and Anderson, 2017). Reliance on TIR effectively limits ET retrievals to clear skies (Rossow et al., 1989), and failure to completely mask cloud affected observations is shown to limit the precision in TIR-LST (Holmes et al., 2016). Continuous daily estimates of ET are generated from clear-sky ALEXI samples through temporal interpolation based on maintaining a normalized flux partitioning metric. In ALEXI this also accounts for daily evaporative losses (Anderson et al., 2007a). Recent work by Alfieri et al. (2017) analysed measurements from eddy-covariance towers and found the persistence for energy flux partitioning metrics to be short. In their analysis, they found that a return interval of no more than 5 days is necessary to keep the relative error in daily ET below 20 %.

In order to provide a more consistent and short return interval for daily ET retrievals at the global scale there is a need for accurate values during cloudy intervals. The approach we take here to address this challenge is to leverage passive microwave (MW) observations. The longer wavelengths (0.1-1 m) make MW observations of the land surface generally less susceptible to scattering and absorption by clouds than observations in the TIR spectral region (except for notable water and oxygen absorption windows; Ulaby et al. (1986)). One MW frequency band with a particularly high sensitivity to LST (Prigent et al., 2016) and high tolerance to clouds (Holmes et al., 2016) is Ka-band (36-37 GHz). MW radiometers with a Ka-band channel are available from several low Earth orbiting satellites that sample at different times of the day. Collectively they can be used to construct a diurnal cycle of brightness temperature for each location on Earth (Holmes et al., 2013b; Norouzi et al., 2012). This diurnal brightness temperature can then be scaled to match the diurnal temperature cycle as measured by TIR imagers (Holmes et al., 2015, 2016).

The methodology developed in Holmes et al (2015) was applied to create an 11-year record of MW-based LST (MW-LST) from various Ka-band sensors (see Section 2). Because this new dataset specifically includes diurnal information, it presents an opportunity to evaluate use of constellation-based MW-LST in a TSEB framework for estimating ET. For this purpose, we substituted MW-LST for MODIS LST in the global implementation of ALEXI as described in Hain and Anderson (2017) and generated a data record of weekly ET for the time-period 2003 to 2013 using each LST data source. No re-calibration of ALEXI was applied in this experiment to accommodate MW-LST. The only difference between the two resulting multi-year records of ET estimates are the spectral window (MW Ka-band vs. TIR) and spatial resolution of the LST inputs (0.25° for the MW implementation: ALEXI-MW, and 0.05° for the MODIS implementation: ALEXI-IR). In order to make the subsequent comparison

with ALEXI-IR as direct as possible, the MODIS cloud mask was also applied to MW-LST. This assures that potential issues related to the applicability of the ALEXI framework during cloudy conditions (particularly its assumptions regarding boundary layer development) are separated from the question of MW-LST performance within the two-source framework. The results are discussed by region and season, and in terms of bulk ET and its inter-annual variation. With this analysis, we hope to establish the degree to which ALEXI-MW resembles the ALEXI-IR under clear sky situations. The performance of the ALEXI model with all-sky LST observations will be the topic of subsequent investigations.

## 2 Methodology

### 2.1 ALEXI model

The ALEXI method is a comprehensive set of algorithms to diagnose the surface energy balance with the aim of retrieving ET (Anderson et al., 2007a; Mecikalski et al., 1999). ALEXI is based on the TSEB land-surface parameterization (Kustas and Norman, 1999; Norman et al., 1995) in which the partitioning of turbulent fluxes is evaluated for the soil (s) and the canopy (c) separately. This is accomplished by 1) partitioning the bulk net radiation (*Rnet*) between canopy and soil surface components and 2) attributing the observed composite surface radiometric temperature (*Trad*) to soil and canopy temperatures, $T_s$ and $T_c$ based on vegetation cover fraction. An initial guess for the canopy latent heat is based on the assumption that the green part of the canopy transpires at its potential rate ($LE_c = LE_{PT}$), where $LE_{PT}$ is estimated with a modified Priestley and Taylor approximation (1972). The sensible heat flux for the two source components ($H_s$ and $H_c$) is then calculated in a set of equations that accounts for their different resistance to heat transfer and that satisfy the observation-based $T_s$ and $T_c$ and air temperature $T_a$ (Norman et al., 1995). The final estimate of latent heat is determined in an iterative procedure in which $LE_c$ is reduced until a solution is found where the soil evaporation ($LE_s$) is non-negative. ET (in units of mass flux) is computed from the latent heat (units of energy flux) by dividing by the latent heat of vaporization.

ALEXI couples TSEB with an atmospheric boundary layer model to relate the morning rise in $T_{rad}$ to the growth of the overlying planetary boundary layer and simulate an internally consistent $T_a$. This removes the need for $T_a$ as an input dataset and limits the sensitivity of the method to biases in instantaneous satellite-based temperature estimates, while allowing for regional and global implementations of the model (Anderson et al., 1997). The ALEXI model is intended for coarse spatial grids (~5 km pixels) and provides the physical foundation to the multi-scale ALEXI/DisALEXI modelling system that has been applied to many satellite-based TIR data streams from 30-m to 10-km spatial resolutions (Anderson et al., 2011). The primary input to ALEXI is $T_{rad}$ at two times during the morning: 1.5 hours after sunrise (time 1) and 1.5 hours before solar noon (time 2). ALEXI computes the energy balance at both instantaneous points during the morning hours (post-dawn and pre-noon). The latent heat estimate at the second time is then upscaled to a daily flux, conserving a flux ratio metric. There are two pathways through which the $T_{rad}$ input affects ALEXI ET estimates: through the estimation of the morning rise in temperature between time 1 and time 2, $\Delta T_{rad}$, which affects the boundary layer growth and the strength of the sensible heat fluxes; and through the impact of $T_{rad}$ on the upwelling longwave component of $R_{net}$ at these times. Whereas the former is not sensitive to time-invariant biases in the diurnal temperature retrievals, the latter has a weak sensitivity to the absolute temperature at time 1 and time 2.

The experiment described in this paper is based on a recent global implementation of the ALEXI model (Hain and Anderson, 2017). This global ALEXI implementation differs from prior geostationary implementations in that its analysis is performed at weekly timescales. While a daily system is in preparation, at present, the global model is executed using 7-day averages of all inputs on "clear-sky" days to minimize computational load. In practice this means taking an average of all needed inputs (at time 1 and 2) on the "clear-sky" days in the 7-day period and running ALEXI. As in prior geostationary implementations the retrieved latent heat

estimate at time 2 is upscaled to a daily flux, conserving a flux ratio metric and using daily solar radiation retrievals. This accounts for changes in atmospheric demand while preserving the scaling flux ratio as determined on the clear-sky days. However, because the scaling flux ratio is held constant over the 7-day period the output is also reported as 7-day total ET (mm/week). The data sources for this version of ALEXI are listed in Table 1. This paper compares two sets of ALEXI ET estimates based on the exact same global model formulation but with alternative LST inputs to estimate the time integrated change in mid-morning $T_{rad}$. The baseline is a TIR version that makes use of MODIS-LST from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the polar orbiting satellites Aqua and Terra from NASA's Earth Observing System (EOS) program. This MODIS-based $T_{rad}$ estimates are used as the input in the current global ALEXI implementation (Hain and Anderson, 2017) described in Section 2.2. The alternative LST input from MW data is described in Section 2.3. The two separate implementations of ALEXI are identified by their temperature input source: ALEXI-IR (with MODIS-LST) and ALEXI-MW (with MW-LST). All other inputs needed to run ALEXI are identical for both implementations.

### 2.2 Temperature from MODIS

The MODIS instrument on the polar-orbiting Aqua satellite (July 2002 to present) with an equator overpass time of 1:30 a.m. / p.m. provides global TIR observations with spectral bands suitable for estimating LST. The specific LST product used for the ALEXI implementation is the MODIS Climate Modelling Grid (CMG) 0.05° daily LST product (MYD11C1 (Wan, 2008)), which is distributed by the Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov). Although the overpass times of this satellite do not correspond directly with ALEXI's time 1 and time 2, Hain et al. (2017) show that over the U.S., GOES-based $\Delta T_{rad}$ can be estimated with a 5-10 % relative error using a tree-based regression model based on independent variables including vegetation index, and landcover class. This regression model, trained over the GOES domain, is then applied globally to estimate $T_{rad}$ at time 1 and time 2 from MODIS LST.

### 2.3 Temperature from a constellation of MW satellites

The MW-LST product is based on vertical polarized Ka-band (36-37 GHz) brightness temperature ($T^{Ka}$), a spectral band commonly included in multi-frequency microwave radiometers in low-Earth orbit. The current MW-LST product integrates observations from six of these satellites. Most important are the Advanced Microwave Scanning Radiometer EOS (AMSR-E) on Aqua from mid-2002 to October 2011 and its successor AMSR2 on the Global Change Observation Mission 1st Water (GCOM-W) from July 2012 onward. Also included are the Special Sensor Microwave and Imager (SSM/I) on platforms F13, F14 and F15 of the Defense Meteorological Satellite Program; the Tropical Rainfall Measurement Mission (TRMM) Microwave Imager (TMI); and Coriolis-WindSat. Together this constellation of Ka-band radiometers allows for the estimation of the diurnal temperature cycle in a process that can be summarized in 4-steps, detailed below and diagrammed in Fig. 1.

#### 2.3.1 Inter-calibration of MW satellites

All available Ka-band observations are combined to create a global record with up to 8 observations per day for each 0.25° resolution grid box. The data are binned in 15-minute windows of local solar time (0:00-0:15 is first window of the day). The brightness temperatures are inter-calibrated using observations from the TRMM satellite (with an equatorial overpass) as a transfer reference. Individual 0.25° averages are masked if the spatial standard deviation of the oversampled Ka-band observations exceeds a prior determined threshold for each grid box. Both the inter-calibration and quality control procedures are described in detail in Holmes et al (2013a). The resulting global record of inter-calibrated Ka-band brightness temperatures spans the years 2003-2013.

### 2.3.2 Fitting of diurnal cycle model to sparse observations

For days with suitable MW observations (a minimum of 4, at least one of which is close to solar noon) and no $T^{Ka} < 250\ K$ (an indication of frozen soil), a continuous diurnal temperature cycle (DTC) is fitted. The DTC model combines a cosine and an exponential term to describe the effect of the sun and the decrease of surface temperature at night and is based on Göttsche and Olesen (2001) with slight adaptations to limit the number of parameters. This implementation (DTC3) is fully described in Holmes et al. (2015). DTC3 summarizes the DTC with four parameters: daily minimum ($T_0$) at start and end of day, diurnal amplitude $A$, and diurnal timing $\varphi$. The fitting procedure first determines $\varphi$ as a temporal constant (Holmes et al., 2013b) and subsequently $T_0$ and $A$ for each day individually. The success of the fit ($\varepsilon_d$) is expressed as the root mean square error (RMSE) between the modelled and observed $T^{Ka}$ for the $n$ observations (at times $t$) in any given day ($d$), calculated following Eq. (1):

$$\varepsilon_d = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(T_i - DTC3(\varphi, T_0, A, t_i)\right)^2} \quad (1)$$

This method was applied to the entire record of inter-calibrated Ka-band brightness temperatures (section 2.3.1) to create a database of annual maps of $\varphi^{Ka}$, and daily maps of $T_0^{Ka}$ and $A^{Ka}$.

### 2.3.3 Scaling of MW DTC parameters to match TIR-LST target

To relate the diurnal cycle in Ka-band brightness temperature to the composite radiative temperature of the land surface requires a set of DTC parameters that is equivalent to those derived from $T^{Ka}$ but derived from a TIR-LST product. In the present analysis, the TIR-LST that serves as a reference is produced at the Land Surface Analysis, Satellite Application Facility (LSA-SAF). LSA-SAF LST is based on TIR window channels of SEVIRI (Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) on board the Meteosat Second Generation (MSG) geostationary satellite. The same method for fitting a DTC model to sparse observations (section 2.3.2) was applied to the LSA-SAF LST to create a database of annual maps of $\varphi^{TIR}$ and daily maps of $T_0^{TIR}$ and $A^{TIR}$ (Holmes et al., 2015). This preparation step is diagrammed in Fig 1 as step '0'.

The Ka-band DTC parameters ($T_{0,d}^{Ka}$, $A_d^{Ka}$) are scaled so that the long-term mean matches that of the equivalent TIR-based parameters ($T_{0,d}^{TIR}$, $A_d^{TIR}$). Because $T_0^{Ka}$ is affected by the sensing depth, the scaling is performed by using daily mean temperature as an intermediate, which is defined as ($\bar{T}^{Ka} = T_0^{Ka} + A^{Ka}/2$) for this purpose.

$$A_d^{MW} = A_d^{Ka}/\delta \quad (2)$$

$$\bar{T}_d^{MW} = \beta_0 + \beta_1 \bar{T}_d^{Ka} \quad (3)$$

The scaled parameters are indicated with the superscript '*MW*'. The parameter $\delta$ represents the slope of the zero-order least squares regression line for estimating the amplitude of $A_d^{Ka}$ from TIR-LST ($A_d^{TIR}$). The intercept ($\beta_0$) and slope ($\beta_1$) to correct the mean daily temperature ($\bar{T}_d^{Ka}$) for systematic differences with TIR-LST ($\bar{T}_d^{TIR}$) are determined with a constrained numerical solver, as in Holmes et al. (2015). The constraint is based on radiative transfer considerations and assures that the scaling of the mean is in agreement with the prior scaling of the amplitude (Eq. 2).

The set of time-constant scaling parameters ($\delta$, $\beta_0$ and $\beta_1$) were determined for each 0.25° grid box based on all days in the period 2007-2012 where both MW and TIR-based DTC parameters were available (generally clear sky and above freezing). Because all three parameters are constant with time, Eqs 2-3 preserve their temporal independence of the TIR LST product. The consequence of using LSA-SAF LST as the reference product is that observation-based scaling parameters are limited to the domain covered by Meteosat (Africa, Europe, Middle-East). Outside this domain, the parameters must be extrapolated. The procedure for the extrapolation is still in development, and currently entails fitting linear regressions with vegetation characteristics. Because of the

limited confidence in the scaling parameters outside the MSG-domain, the analysis in this paper is focused on the Africa and Europe domain. Some results of the global set will be presented in the comparison with flux tower observations (Section 3.4).

#### 2.3.4 Constructing MW-LST

Global maps of the time-constant parameters ($\delta$, $\beta_0$ and $\beta_1$, section 2.3.3) are used to calculate the daily DTC parameters ($T_{0,d}^{MW}, A_d^{MW}$) in the scaled climatology of the TIR-LST product. This scaling (Eqs. 2 and 3) is applied to every day for which estimates of $\bar{T}^{Ka}$ and $A^{Ka}$ are available (see section 2.3.2). The methodology to scale the DTC parameters from this record of Ka-band observations to a physical temperature range is described in more detail in Holmes et al. (2015). The scaled parameters together with $\varphi^{TIR}$ are then used to construct the MW-LST based on the same DTC3 model as used in step 2:

$$\text{MW-LST}_i = DTC3(\varphi^{TIR}, T_{0,d}^{MW}, A_d^{MW}) \quad (4)$$

The use of the DTC model allows MW-LST to be diurnally complete for days when both $T_{0,d}^{MW}$ and $A_d^{MW}$ are available. MW-LST can therefore be generated at any time increment (i). The MW-LST database used for this paper was generated at 15-minute temporal interval. This allows $T_{rad1}$ and $T_{rad2}$ to be accurately interpolated from the database. $\varepsilon_d^{Ka}$ (Eq. 1) is used to flag days where the assumptions imposed by the shape of clear sky DTC3 are not valid or individual Ka-band observations have a large bias. In this experiment, MW-LST was only used if $\varepsilon_d^{Ka}$ is 2.5 K or lower.

#### 2.3.5 MW-LST in ALEXI

The continuous 7-day totals are achieved by temporal gap-filling of (clear sky) ET as a fraction of clear-sky latent heat flux to incoming solar radiation (Anderson et al., 2007a). To maximize similarity, the same MODIS cloud mask is applied to the ALEXI-MW implementation so that the mechanics of standard ALEXI can be evaluated under circumstances for which it has previously been developed and validated.

The fraction of days in a year where a clear sky MODIS-based $T_{rad1}$ and $T_{rad2}$ is available for ALEXI is below 0.3 for large parts of Europe and (sub)-tropical Africa (Fig 2a). In these areas the revisit time between observation days regularly exceeds 5 days, a threshold for temporal downscaling given the persistence of ET fraction (Alfieri et al., 2017). On average for the non-coast pixels, there is a MW-based estimate available for 69 % of those days where there is also a (clear sky) MODIS-based $T_{rad1}$ and $T_{rad2}$. The reason this percentage is not higher is mainly due to the requirement of a near-noon overpass for the fitting of the diurnal temperature curve (See Section 2.3.2), in combination with the 1 in 3 days without such overpass for a given location as determined by the orbit and coverage of Aqua and GCOM-W. The multi-year and global record of simultaneous MW-LST and MODIS LST during clear skies will support further calibration of MW-LST to MODIS LST in future investigations. This MW to MODIS calibration was not done in this study but is likely needed to maximize consistency between ALEXI implementations over the globe. In terms of potential for additional sampling through the use of MW-based LST, Fig 2c shows that MW-based estimates for the two ALEXI times are available for 54 % of the days where no MODIS-based estimate is available. Fig 2d depicts the fraction of days where either MODIS or MW-LST can be used to estimate the $T_{rad}$ inputs required to run ALEXI. This shows that the addition of MW-LST can bring the minimum average coverage in this domain to once every two days.

### 2.4 Flux tower observations

Tower measurements of latent heat flux obtained using the eddy-covariance (EC) technique are commonly used for ground truthing of remote sensing and model-based ET estimates (Baldocchi et al., 2001). Harmonized Fluxnet data are distributed in so-called synthesis datasets. They include the original observations at a half hour observation time, and aggregate values per day, week and month. For this work, we used the synthesis 2015 TIER 1 data as accessed in July 2016

(http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/) to serve as a common ground reference for the evaluation of the temporal characteristics of ALEXI-MW and ALEXI-IR. In particular, the part of the dataset of interest here are the daily aggregates of latent heat flux (variable name LE_F_MDS) which include quality control as described in Pastorello et al. (2014).

Based on these daily data, we computed the 7-day averages matching the window length of ALEXI. If not all days within a window have valid data, that window is disregarded. Overall, eddy-covariance observations of ET were available from 68 flux towers with at least one year of observations within the time period of this study.

### 2.5 Definition of regions

Although both MW and IR sets are available globally, the main analysis of this paper is focused on the domain encompassing Africa and Europe. This is because only in that region is the scaling of MW-LST to TIR-based LST currently supported by data (see Section 2.3.3). However, temporal comparisons (e.g., correlations) are much less affected by the mean absolute value of MW-LST product. Because of the limited availability of flux tower data, we include all available stations from across the globe which allows us to double the amount of stations available for the analysis compared to only the sites in Europe and Africa..

Within the main focus domain of this study we further highlight 11 climate-based domain subsets (see also Fig 3, bottom-right panel):

A. West-African Sahel, Arid
B. West-African Sahel, Semi-Arid
C. Guinean Coast, Dry sub-Humid
D. Central Africa, Humid
E. Horn of Africa, Arid
F. Southern Africa, semi-Arid (large bias in Fig 4)
G. Southern Africa, Arid (large bias in Fig 4)
H. Iberia, semi-Arid
I. Germany, continental Humid
J. European Russia, continental Humid, boreal forest (large bias in Fig 4)
K. France, Humid

These regions are selected to represent a wide variety of seasonal variation in precipitation and climate class, and are based on the work of Trambauer et al. (2014). Rather than attempting to cover the entire domain with these subsets, we selected smaller subsets in order to visualise the local deviations between MW and IR products that might otherwise be averaged out. We also added regions in Europe and several regions that showed a large bias in Fig. 4.

### 2.6 Metrics

Cumulative annual and seasonal fluxes are compared in terms of their relative deviation (RD (%)), calculated following Eq. 5:

$$RD = \frac{\bar{x}-\bar{y}}{(\bar{x}+\bar{y})/2} \times 100\% \quad (5)$$

where $\bar{x}$ represents the mean of the MW product and $\bar{y}$ the mean of the IR product, both sampled at the same times. This relative comparison is useful because neither product represents the truth and this formulation places the deviations in context of the size of the fluxes. Still, if the ET is very small (average ET below 14 mm/month) then the denominator becomes too small and the RD is not reported. The temporal agreement between the anomalies in the IR and MW-based ET products is analysed in terms of the Pearson's correlation (ρ), and the spatial agreement in terms of correlation coefficient ($R^2$).

The temporal agreement of the weekly ET estimates is further compared relative to the flux tower observations that serve as a common reference. For this assessment, MW- and IR-based ET estimates are again compared in terms of ρ but also in terms of root mean square error (RMSE) to quantify the absolute error. The RMSE is calculated following Eq. 7:

$$RMSE = \sqrt{\frac{1}{N}\sum(x-y)^2} \quad (7)$$

where x is the satellite estimate of ET and y is the tower-based measurement of ET. N is the number of data pairs.

## 3 Results

### 3.1 Comparing Multi-year means

The mean average $\Delta T_{rad}$ as calculated from MW-LST deviates from that calculated from MODIS LST by 0-20 %, which leads to a spatial $R^2$ of 0.90 (Fig 4. top row). These spatial variations in mean values arise from the different calibration targets. MW-LST is calibrated to match the LSA-SAF LST from MSG (Europe and Africa) with a precision of 2-3K (see Section 2.3.3), and MODIS $\Delta T_{rad}$ is trained on GOES (North-America) with an estimated precision of 5-10 % (see Section 2.2). These different calibration domains together with likely calibration differences between GOES and MSG LST products present sources of bias that can explain the regional variation we see in Fig 4. For example, the difference between $\Delta T_{rad}$ estimates in the North-East corner of this map may be an artefact of scaling with high incidence angles (θ) for the MSG geostationary satellite. In the farthest corner (θ > 60º), MSG observations were not used and the MW scaling is extrapolated based on land surface characteristics. The MW $\Delta T_{rad}$ also exceeds IR-based estimates by more than 10 % in Southern Africa, for which we do not currently have an explanation.

The general agreement in mean $\Delta T_{rad}$ translates into a high agreement between IR and MW-based ALEXI in terms of mean annual ET for the period 2003 – 2011. The spatial correlation between MW and IR in terms of ET is 92 % (Fig 4. bottom row), similar to that for $\Delta T_{rad}$.. Boreal Russion shows the most notable differences in absolute terms, where MW is lower by ~ 20 %. This is related to view angle impacts on the $\Delta T_{rad}$ retrieval, as noted above. MW ET is also much lower than IR ET in the Alps, which likely reflects an interaction between view angle and topography (e.g.., differences in pixel proportion of sunlit and shaded slopes) In the Horn of Africa, MW is higher by 20-30%., although little difference in DTRAD is apparent in this region. ALEXI ET becomes more sensitive to small changes in DTRAD near the dry end, where the iterative stress reduction in transpiration starts to kick in.

### 3.2 Regional/Seasonal Bulk Flux comparison

Figure 5 provides a more detailed comparison between the MW and IR products for the domain subsets as described in Section 2.5. For each domain subset, it shows the mean monthly total ET and the associated monthly means of $\Delta T_{rad}$. For European Russia (region J) and to a lesser extend Germany (I) and France (K), this shows higher MW-based $\Delta T_{rad}$ resulting in lower summertime ET estimates than for ALEXI-IR. Conversely, in the wintertime the lower MW-based $\Delta T_{rad}$ results in higher ET estimates compared to the TIR products in October-December. For Iberia (H), the semi-Arid Sahel (B) and Spain (H) there appears to be a difference in timing with MW estimating a later time for peak ET. The higher MW ET estimates in the Horn of Africa are rather uniform over the year, except for December and January where the difference is small. The size of the bias in ET for the Horn of Africa is relatively large compared to the modest bias in $\Delta T_{rad}$. Another disconnect between $\Delta T_{rad}$ deviation and ET can be seen in Southern Africa (regions F and G). Despite a general overestimation of $\Delta T_{rad}$ by MW the ET estimates are very close to those of ALEXI-IR. The small difference in ET estimates are the opposite of what would be expected from the $\Delta T_{rad}$ deviation. Finally, the humid tropical climates of Guinean coast and central Africa (regions C and D) have very little differences in both $\Delta T_{rad}$ and ET.

To provide some additional spatial and temporal context for these observations, the three-month total MW and TIR ET (averaged over 2003-2011) are shown in Fig. 6 for December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON). This shows that the cold season overestimation of MW-based ET, seen in the European regions, is present not only in Europe but also in East and Southern Africa in SON. The underestimation of MW-based ET in

summer is not as pronounced in terms of its relative difference. The apparent difference in timing, seen in the Sahel and Iberian regions, shows up across the southern border of the Sahara – MW-ET is higher in MAM, and TIR-ET is higher in JJA. The spatial correlation between MW and IR is higher in SON (96 %) and DJF (97 %) compared to the periods MAM (83 %) and JJA (84 %). Despite these localized differences, the transect averages are remarkably similar showing the general success of scaling MW-LST to TIR-LST (Section 2.3.3).

### 3.3 Inter-annual Variation

Because the long term mean of MW-LST is calibrated to match a TIR reference (see Section 2.3.3), a comparison in terms of anomalies is the real test of its performance in the ALEXI framework, especially in areas that are water limited (see Fig. 7). Of the subsets in water-limited regions, the Horn of Africa ($\rho$=0.78) and Spain ($\rho$=0.85) subsets show a high degree of correlation between MW and TIR-based ET anomalies. Semi-Arid Southern Africa (F) and the Sahel (B) show relatively poor correlation with $\rho$=0.48 and $\rho$=0.63 respectively. The size of the anomaly is much larger for ALEXI-MW in Southern Africa in January and February, reflecting a much larger inter-annual variation.

In energy-limited areas when ET is fully determined based on the meteorological forcing data, the effect of LST inputs is minimal. This is apparent in the Tropical region, where MW and ALEXI-IR have a correlation of 0.99 in Central Africa (region D). Figure 8 shows a map of the correlation between 3-month anomalies of MW and IR-based ALEXI ET.

Seasonal anomalies are calculated by taking the seasonal total ET for a given year and subtracting its corresponding long-term mean seasonal total (2003-2011 period, as shown in Fig 6). Examples of this are shown for a dry year (2008) and a wet year (2011), see Fig 9. Overall the two sets of anomalies agree very well – the MW ALEXI appears to identify roughly the same areas with anomalous high or low ET. The agreement is better in the wet year than in the dry year.

### 3.4 Comparison with flux tower observations

The availability of eddy-covariance observations of ET from 68 flux towers allows for a more detailed grid-level analysis of temporal agreement. Even at the 0.05-degree (~5 km) resolution of ALEXI-IR there is a large scale miss-match between remote sensing estimate and tower footprint. The impact of this scale difference will depend on the degree of spatial heterogeneity within the larger footprint. We therefore cannot use these flux tower observations to quantify absolute accuracy in either product, but instead focus on its use as a reference target to compare relative performance between two satellite products. To start, we compare the effect of the resolution degradation from 0.05 degree to 0.25 degree.

When 0.05° ALEXI-IR is averaged over its surrounding 0.25° grid (the average of the 5x5 0.05° grid cells) there is an overall improvement in $\rho$ (but not in RMSE), see Fig 10. Only at three sites does this spatial degradation lower the $\rho$ between the site and the 0.05° grid average higher than with the 0.25° grid box. The landscape heterogeneity is large at these sites (US-Ton, US-Var, and ES-Lgs). For most stations, the spatial degradation actually improves the $\rho$ with the site. In fact, 40 % of the difference in $\rho$ between MW and IR ALEXI is explained by the change in $\rho$ from ALEXI-IR (at 0.05°) to ALEXI-IR (at 0.25°). This indicates the presence of noise in the 0.05° MODIS LST input that is uncorrelated with the surrounding 0.25° grid average and negates any positive effect of its resolution advantage compared to a 0.25° grid average for most sites.

The following analysis compares MW and IR both at 0.25° grid resolution. The metrics we focus on are $\rho$ and RMSE which are computed for each flux tower site and listed in Table 2. For ALEXI-IR, $\rho$ is between 0.6 and 0.92 and RMSE is 12-33 mm/week for the majority of the sites. The impact of LST input varies from site to site (see also Fig 11), with some stations showing higher $\rho$ for ALEXI-MW, but most showing an advantage for ALEXI-IR, as expected. Overall, the mean $\rho$ is higher for ALEXI-IR ($\rho$=0.78 Vs $\rho$=0.74), even though the average RMSE comes out the same (24 mm/week).

It is interesting to investigate what drives the difference in temporal correlation at individual sites. The second row in Fig. 11 shows how the same data as presented in Fig. 11, but broken out based on geographic domain, climate or spatial agreement. The first panel splits the sites by geographic region. Europe and Africa (blue) is where MW-LST was calibrated with MSG SEVIRI and the North-American sites (green) is where MODIS ALEXI-IR has been calibrated with GOES data (see Section 2). Between these two groups of stations the relative improvement in ρ is higher in the North-American sites,. This is despite the MODIS ALEXI-IR being calibrated with GOES data. Panel 2 separates the sites based on climate, particularly in terms of the potential ET (PET) relative to the annual precipitation (P). The PET used for this classification is calculated following Priestley-Taylor (1972) with an alpha parameter of 1.26 and zero ground heat flux. Sites with humid climates (energy limited: PET<=P) have generally a higher ρ between station and satellite data and show only a modest impact of the change in LST input on ALEXI. In arid climates (water limited: PET>P) there is more variation in performance and correlations between satellite estimate and tower observation are generally lower. Partly, this reflects a lower signal to noise in areas with low overall ET, but it also reflects a more challenging environment for ET retrievals. The advantage of ALEXI-IR over ALEXI-MW is larger in these arid climates. Further subdividing the arid locations based on information on spatial heterogeneity reveals a still larger separation of performance (Fig 11, Panel 3 on bottom row). Taking the absolute bias (|b|) between ALEXI-IR at the 0.05° grid cell encompassing the tower site and the mean of the 0.25° surrounding grid box as proxy for spatial heterogeneity, we can see that for the sites that are both in a water limited region and have a high spatial bias, 11 in total, the average ρ for ALEXI-MW (ρ=0.55) is markedly lower than that for ALEXI-IR (ρ=0.65).

Six of the 68 sites have a markedly higher ρ with ALEXI-IR than with ALEXI-MW. All but one of these sites have an arid climate (See Table 2), and four of those stations also have a high spatial bias between the 0.05° grid box and 0.25° grid box mean (|b|>2 mm/week):

- US-Ton and US-Var (PET/P=2.5/1.7, b=-6.4), Woody Savannas, same 0.05 box. ALEXI-IR has a ρ=0.80/0.5, while ALEXI-MW has a lower value of 0.56/0.09. At US-Var, the site has an abrupt collapse of ET at end of summer. The satellite data miss this, especially the ALEXI-MW(0.25°) product.
- Zambia, Savannas. ZM-Mon. Water limited (PET/P=2.3), spatial heterogeneous (b=-2.9)
- ES-LgS, Woody Savannas (b=11, PET/P=2.8). The MODIS has a high ρ=0.84, while the MW has a poor ρ=0.62. The average comes in at ρ=0.82. This site is located on a mountain ridge. The smaller grid size of ALEXI-IR is able to capture the vegetation conditions at the mountain ridges whereas the 0.25 grid of ALEXI-MW has more bare soil which leads to lower ET values.

The station in Sudan (SD-Dem) is the only of these 6 stations that is in a water limited region (arid desert climate) and has low spatial bias. Despite the low bias, the station ET estimates are 2.5 times satellite estimates, so it could be that the near station land use is not representative of the wider area.

The final station that shows a large advantage in ρ for ALEXI-IR relative to ALEXI-MW is Fi-Hyy (No. 63 in Table 2) in a cold region climate. It is also one of only two stations with data availability at high latitude (above 60°N). This station has land cover dominated by evergreen needleleaf forest. The bias between the 0.05° and 0.25° grid box mean is also small (b=-0.6). The MW observations have relatively many weeks with very low ET estimates compared to the ALEXI-IR. The reason for this is not readily apparent but it could be that the MW product suffers from rainclouds that suppress temperature estimates during the morning hours around ALEXI time 1. This, in turn, leads to an overestimated morning temperature rise.

In contrast to these sites, there are two sites where the ALEXI-MW outperforms ALEXI-IR in terms of correlation with in situ sites despite being in a relatively arid climate with large spatial bias: US-SRG, US-NR1. For US-NR1, ρ is low because station records high values in winter time, and the site is located in an evergreen forest east of a mountain ridge, with high day to day variation, possibly due to varying wind direction or shading effects. Despite this, both satellite products pick up the seasonal cycle reasonably well, except that they both underestimate wintertime ET.

## 4 Discussion and Conclusion

This paper shows that a newly developed MW-LST product can be used to effectively substitute TIR-based LST in a two-source energy balance approach to estimate coarse-resolution ET (~25 km) from space. This particular TSEB approach, the ALEXI model framework, is an approach that minimizes sensitivity to absolute biases in input records of LST through the analysis of the rate of change in morning LST. It is therefore an important test of the ability to retrieve diurnal temperature information from a constellation of satellites that provide 6-8 observations of Ka-band brightness temperature per location per day. This represents the first ever attempt at a global implementation of ALEXI with MW-based LST and is intended as the first step towards providing all-weather capability to the ALEXI framework.

Because the long-term (7-year mean) diurnal features of MW-LST are calibrated to TIR-LST, it is perhaps not surprising that the long-term bulk ET estimates agree with a spatial correlation of 92 % for total ET in the Europe/Africa domain. A comparison with biases in the input datasets of $\Delta T_{rad}$ shows that a large part of the remaining differences can be mitigated by specifically calibrating MW-LST to MODIS LST. More convincing is the agreement in seasonal (3-month) averages of and 83-97 % because the calibration is based on time-constant parameters. Adding another layer of challenging complexity is the comparison in terms of 3-month anomalies. By this test, ALEXI-MW also matches ALEXI-IR very closely, demonstrating an ability to capture the development and extent of drought conditions.

The two parallel ALEXI implementations are further compared at the maximum temporal resolution of the current global ALEXI output (7 days) and relative to a common ground measured reference provided by the FLUXNET consortium. The 68 stations that were available for this analysis represent a wide range of land cover characteristics and climate conditions. Overall, they indicate a close match in both performance metrics (ρ and RMSE), especially considering the advantage of TIR-LST compared to MW-LST in these clear sky conditions. The most challenging conditions for MW-LST as input to ALEXI-ET according to these sites are locations with higher aridity levels and where the larger domain has a high spatial heterogeneity. Spatial heterogeneity places an obvious penalty on ALEXI-MW due to the coarser MW-LST input, even though in general ALEXI-IR improves in terms of its correlation with the tower data when it is spatially downgraded to 0.25° resolution. For future merging of IR- and MW-based ALEXI into a superior combined ET estimate this range in relative performance observed at these sites needs to be accounted for.

Based on the analyses presented in this paper, we outline the following roadmap for an all-sky implementation of ALEXI-MW. First of all, there is a need for global observation based calibration of MW-LST with MODIS-LST to reduce biases as identified at the high incidence angles of the MSG domain and avoid the need for extrapolation of scaling parameters. Second, the MW-LST could be used to improve the TIR cloud mask by attributing anomalous TIR-based $\Delta T_{rad}$ to the presence of clouds, with subsequent improvements in ALEXI-IR ET estimates. Finally, the all-sky implementation that is now within reach with ALEXI-MW will test the assumptions in new ways, which will require careful investigation. For example, the assumptions related to the boundary layer development may be tested as we move to include less stable conditions associated with cloudy skies. Similarly, evaporation of intercepted rain water will feature more prominently under cloudy skies and may require inclusion as a separate process within the current physical framework. With a combined MW+IR ALEXI estimates it appears entirely feasible to reduce the current window length for reporting MODIS ALEXI ET totals from 7-days to as low as 2. At a window length of 2 days the average satellite coverage would support each 2-day total with at least one ET retrieval (See Fig. 2). This would reduce the reliance on temporal downscaling and its associated assumptions and impact on estimation error. More independent estimates of ET would allow for more robust statistical analysis in the context of land-atmosphere exchange studies, even if the record length is not extended. Perhaps most importantly, a shorter reporting interval would also allow for earlier detection of agricultural drought as reflected in the ET-based drought indices (Anderson et al., 2011).

## Data Availability

The ALEXI-IR data is available from NASA SPoRT (MSFC). The ALEXI-MW is an intermediate research product available upon request. Time-series of ALEXI-MW and ALEXI-IR covering the site locations and time period of this paper are available upon request from the corresponding author. The Flux tower data is publicly available through the FLUXNET community as detailed in Section 2.3.


## Acknowledgements

We thank Carlos Jimenez and one anonymous reviewer for their insightful comments on this manuscript. This work was funded by NASA through the research grant "The Science of Terra and Aqua" (13-TERAQ13-0181).


## References


Alfieri, J. G., Anderson, M. C., Kustas, W. P. and Cammalleri, C.: Effect of the revisit interval and temporal upscaling methods on the accuracy of remotely sensed evapotranspiration estimates, Hydrol. Earth Syst. Sci., 21(1), 83–98, doi:10.5194/hess-21-83-2017, 2017.

Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P. and Mecikalski, J. R.: A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing, Remote Sens. Environ., 60(2), 195–216, 1997.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A. and Kustas, W. P.: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation, J. Geophys. Res., 112(D10), D10117, 2007a.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A. and Kustas, W. P.: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology, J. Geophys. Res. Atmospheres, 112(D11), doi:doi:10.1029/2006JD007507, 2007b.

Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., Gonzalez-Dugo, M. P., Cammalleri, C., d'Urso, G. and Pimstein, A.: Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery, Hydrol. Earth Syst. Sci., 15(1), 223, 2011.

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X. H., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K. and Wofsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities., Bull Am Meteorol Soc, 82, 2415–2434, 2001.

Campbell, G. S. and Norman, J. M.: An Introduction to environmental biophysics, Springer-Verlag, New York [etc., 1998.

Doelling, D.: CERES Level 3 SYN1DEG-DAYTerra+Aqua netCDF file - Edition 3A, , doi:10.5067/Terra+Aqua/CERES/SYN1degDAY_L3.003A, 2012.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A. and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sens. Environ., 114(1), 168–182, 2010.

Göttsche, F.-M. and Olesen, F. S.: Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data, Remote Sens. Environ., 76(3), 337–348, doi:10.1016/S0034-4257(00)00214-5, 2001.

Hain, C. R. and Anderson, M. C.: Estimating Morning Change in Land Surface Temperature from MODIS Day/Night Land Surface Temperature: Applications for Surface Energy Balance Modeling, Geophys. Res. Lett., 44(19), doi: 10.1002/2017GL074952..

Hain, C. R., Crow, W. T., Anderson, M. C. and Yilmaz, M. T.: Diagnosing Neglected Soil Moisture Source–Sink Processes via a Thermal Infrared–Based Two-Source Energy Balance Model, J. Hydrometeorol., 16(3), 1070–1086, doi:10.1175/JHM-D-14-0017.1, 2015.

Holmes, T. R. H., Crow, W. T., Yilmaz, M. T., Jackson, T. J. and Basara, J. B.: Enhancing model-based land surface temperature estimates using multiplatform microwave observations, J. Geophys. Res. Atmospheres, 118, 577–591, doi:10.1002/jgrd.50113, 2013a.

Holmes, T. R. H., Crow, W. T. and Hain, C.: Spatial patterns in timing of the diurnal temperature cycle, Hydrol. Earth Syst. Sci., 17(10), 3695–3706, doi:10.5194/hess-17-3695-2013, 2013b.

Holmes, T. R. H., Crow, W. T., Hain, C. R., Anderson, M. and Kustas, W. P.: Diurnal temperature cycle as observed by thermal infrared and microwave radiometers., Remote Sens. Environ., 158C, 110–125, doi:10.1016/j.rse.2014.10.031, 2015.

Holmes, T. R. H., Hain, C. R., Anderson, M. C. and Crow, W. T.: Cloud tolerance of remote-sensing technologies to measure land surface temperature, Hydrol. Earth Syst. Sci., 20(8), 3263–3275, doi:10.5194/hess-20-3263-2016, 2016.

Kustas, W. P. and Norman, J. M.: Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover, Agric. For. Meteorol., 94(1), 13–29, 1999.

Mecikalski, J. R., Diak, G. R., Anderson, M. C. and Norman, J. M.: Estimating fluxes on continental scales using remotely sensed data in an atmospheric-land exchange model, J. Appl. Meteorol., 38(9), 1352–1369, 1999.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R. and others: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83(1), 214–231, 2002.

Norman, J. M., Kustas, W. P. and Humes, K. S.: Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, Agric. For. Meteorol., 77(3–4), 263–293, doi:10.1016/0168-1923(95)02265-Y, 1995.

Norouzi, H., Rossow, W., Temimi, M., Prigent, C., Azarderakhsh, M., Boukabara, S. and Khanbilvardi, R.: Using microwave brightness temperature diurnal cycle to improve emissivity retrievals over land, Remote Sens. Environ., 123, 470–482, 2012.

Pastorello, G., Agarwal, D., Papale, D., Samak, T., Trotta, C., Ribeca, A., Poindexter, C., Faybishenko, B., Gunter, D., Hollowgrass, R. and others: Observational data patterns for time series data quality assessment, in 2014 IEEE 10th International Conference on e-Science, vol. 1, pp. 271–278, IEEE, Sao Paulo, Brazil., 2014.

Priestley, C. H. B. and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, Mon. Weather Rev., 100(2), 81–92, 1972.

Prigent, C., Jimenez, C. and Aires, F.: Towards “all weather”, long record, and real-time land surface temperature retrievals from microwave satellite observations, J. Geophys. Res. Atmospheres [online] Available from: http://onlinelibrary.wiley.com/doi/10.1002/2015JD024402/full (Accessed 17 May 2016), 2016.

Rossow, W. B., Garder, L. C. and Lacis, A. A.: Global, Seasonal Cloud Variations from Satellite Radiance Measurements. Part I: Sensitivity of Analysis, J. Clim., 2(5), 419–458, doi:http://dx.doi.org/10.1175/1520-0442(1989)002<0419:GSCVFS>2.0.CO;2, 1989.

Saha, S. and al: NCEP Climate Forecast System Version 2 (CFSv2) 6-hourly Products, [online] Available from: http://dx.doi.org/10.5065/D61C1TXF, 2011.

Saha, S. and et al: NCEP Climate Forecast System Reanalysis (CFSR) 6-hourly Products, January 1979 to December 2010., [online] Available from: http://dx.doi.org/10.5065/D69K487J, 2010.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d’Entremont, R. P., Hu, B., Liang, S., Privette, J. L. and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, Remote Sens. Environ., 83(1–2), 135–148, doi:10.1016/S0034-4257(02)00091-3, 2002.

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., Van Beek, L. P. H. and Uhlenbrook, S.: Comparison of different evaporation estimates over the African continent, Hydrol. Earth Syst. Sci., 18(1), 193, doi:doi:10.5194/hess-18-193-2014, 2014.

Trenberth, K. E., Fasullo, J. T. and Kiehl, J.: Earth’s global energy budget, Bull. Am. Meteorol. Soc., 90(3), 311, 2009.

Ulaby, F. T., Moore, R. K. and Fung, A. K.: Microwave Remote Sensing: Active and Passive. Vol. III. From theory to applications., Artech House, Norwood, MA., 1986.

Wan, Z.: New refinements and validation of the MODIS land-surface temperature/emissivity products, Remote Sens. Environ., 112(1), 59–74, doi:10.1016/j.rse.2006.06.026, 2008.

**Table 1. Primary inputs for current global implementation of ALEXI**

| Data | Purpose | Source | Spatial Resolution |
|---|---|---|---|
| LST | $T_{rad}$, Net Radiation | MODIS (MYD11C1) (Section 2.2) | 0.05° |
| | | MW-LST (Section 2.1) | 0.25° |
| Surface Longwave Radiation | Net Radiation | CFS-R[2], CFSRv2[3] | 0.5° |
| Surface Shortwave Fluxes | Net Radiation | CERES SYN1deg[4] | 1° |
| Albedo | Net Radiation | MODIS (MCD43B3)[5] | 0.05° |
| LAI | Trad partitioning | MODIS (MCD15A3)[6] | 0.01° |
| Landcover type | Canopy characteristics | MODIS (MCD12C1)[7] | 0.01° |
| Wind speed | Aerodynamic resistance | CFS-R, CFSRv2 | 0.5° |
| Lapse rate profile | Atmospheric Boundary layer growth model | CFS-R, CFSRv2 | 0.5° |

[2]NCEP Climate Forecast System Reanalysis (Saha and et al, 2010), [3] (Saha and al, 2011), [4] (Doelling, 2012), [5](Schaaf et al., 2002), [6](Myneni et al., 2002), [7](Friedl et al., 2010)

**Table 2. Time-series correlation of flux tower ET observations with three alternative satellite-based ET estimates (ALEXI_IR, ALEXI-MW, ALEXI- IR+MW). Comparison is based on weekly averages in the period of 2003 to 2011 (number of data pairs is noted in the table). Bias is mean difference between 0.05 grid and 0.25 surrounding grid average as measured by ALEXI-IR.**

| | | | | | | | Data | ALEXI-IR (0.05) | | ALEXI-MW (0.25) | | ALEXI-IR+MW (Avg) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. | Site ID | Latitude | Longitude | IGBP Vegetation Type | Climate PET/P | bias | pairs (wk) | ρ | RMSE | ρ | RMSE | ρ | RMSE |
| 1 | AT-Neu | 47.11667 | 11.3175 | Mix. Forests | 0.67 | -1.26 | 380 | 0.78 | 22.8 | 0.90 | 17.7 | 0.83 | 23.7 |
| 2 | AU-Ade | -13.0769 | 131.1178 | Savannas | N/A | -1.12 | 77 | 0.87 | 51.0 | 0.87 | 49.9 | 0.84 | 43.3 |
| 3 | AU-DaP | -14.0633 | 131.3181 | Savannas | 0.94 | -0.85 | 190 | 0.82 | 37.7 | 0.83 | 38.2 | 0.70 | 37.0 |
| 4 | AU-DaS | -14.1593 | 131.388 | Savannas | 0.85 | 3.09 | 178 | 0.77 | 32.3 | 0.79 | 34.5 | 0.70 | 25.6 |
| 5 | AU-Dry | -15.2588 | 132.3706 | Savannas | 6.63 | 4.41 | 107 | 0.56 | 32.2 | 0.58 | 35.3 | 0.48 | 26.6 |
| 6 | AU-Stp | -17.1508 | 133.3503 | Grasslands | 1.54 | -1.02 | 142 | 0.59 | 40.1 | 0.59 | 39.6 | 0.72 | 32.0 |
| 7 | AU-Tum | -35.6566 | 148.1516 | ENF | 0.96 | 2.52 | 289 | 0.85 | 22.6 | 0.83 | 25.3 | 0.79 | 24.6 |
| 8 | AU-Wac | -37.429 | 145.1873 | EBF | 0.13 | 8.53 | 132 | 0.80 | 20.1 | 0.80 | 24.2 | 0.80 | 20.6 |
| 9 | AU-Wom | -37.4222 | 144.0944 | EBF | 0.78 | 0.65 | 77 | 0.84 | 29.1 | 0.86 | 29.2 | 0.82 | 25.2 |
| 10 | BE-Bra | 51.30917 | 4.520556 | Mix. Forests | 0.57 | -0.74 | 308 | 0.81 | 14.0 | 0.83 | 12.5 | 0.73 | 14.1 |
| 11 | BE-Lon | 50.55219 | 4.744772 | Croplands | 0.61 | 0.09 | 309 | 0.84 | 15.4 | 0.85 | 15.2 | 0.86 | 15.3 |
| 12 | BE-Vie | 50.30507 | 5.998052 | Mix. Forests | 0.61 | 0.47 | 274 | 0.85 | 13.2 | 0.84 | 13.6 | 0.74 | 16.2 |
| 13 | CA-Qfo | 49.69247 | -74.342 | ENF | 0.56 | -0.32 | 308 | 0.69 | 14.8 | 0.76 | 13.0 | 0.82 | 11.6 |
| 14 | CA-SF1 | 54.48495 | -105.817 | ENF | 0.93 | 0.52 | 130 | 0.78 | 24.6 | 0.80 | 23.9 | 0.70 | 32.8 |
| 15 | CA-SF2 | 54.25392 | -105.878 | Mix. Forests | 1.82 | 0.77 | 114 | 0.70 | 25.5 | 0.75 | 24.2 | 0.59 | 32.0 |
| 16 | CA-SF3 | 54.09156 | -106.005 | ENF | 0.91 | 2.21 | 155 | 0.81 | 16.0 | 0.82 | 15.4 | 0.80 | 17.5 |
| 17 | CH-Cha | 47.21022 | 8.410444 | Mosaic | 0.46 | -5.44 | 257 | 0.87 | 38.8 | 0.93 | 32.2 | 0.89 | 29.3 |
| 18 | CH-Dav | 46.81533 | 9.855917 | ENF | 0.75 | -0.14 | 413 | 0.54 | 26.5 | 0.61 | 24.3 | 0.59 | 29.9 |
| 19 | CH-Fru | 47.11583 | 8.537778 | Mosaic | 0.39 | 3.77 | 241 | 0.90 | 21.6 | 0.89 | 25.3 | 0.80 | 28.6 |
| 20 | CN-Du2 | 42.04667 | 116.2836 | Grasslands | 1.99 | 1.10 | 101 | 0.62 | 24.8 | 0.65 | 25.1 | 0.68 | 24.8 |
| 21 | CZ-wet | 49.02465 | 14.77035 | Croplands | 0.94 | 1.65 | 101 | 0.92 | 13.0 | 0.92 | 13.8 | 0.88 | 15.8 |
| 22 | DE-Geb | 51.1001 | 10.9143 | Croplands | 0.81 | -1.18 | 411 | 0.85 | 15.5 | 0.86 | 14.8 | 0.77 | 19.4 |
| 23 | DE-Gri | 50.94947 | 13.51253 | Mixed Forests | 0.52 | -0.38 | 329 | 0.88 | 12.2 | 0.89 | 11.6 | 0.88 | 12.3 |
| 24 | DE-Hai | 51.07917 | 10.453 | Mix. Forests | 0.68 | 3.15 | 322 | 0.91 | 13.0 | 0.92 | 13.7 | 0.87 | 19.8 |
| 25 | DE-Kli | 50.89288 | 13.52251 | Croplands | 0.46 | -1.26 | 284 | 0.84 | 14.3 | 0.87 | 14.0 | 0.84 | 15.4 |
| 26 | DE-Lkb | 49.09962 | 13.30467 | ENF | 0.45 | 1.38 | 92 | 0.84 | 14.6 | 0.84 | 14.8 | 0.79 | 15.0 |
| 27 | DE-Obe | 50.78362 | 13.71963 | ENF | 0.53 | 3.09 | 167 | 0.87 | 13.4 | 0.86 | 13.3 | 0.83 | 14.0 |
| 28 | DE-Seh | 50.87062 | 6.449653 | Croplands | 0.63 | 0.48 | 132 | 0.84 | 24.1 | 0.85 | 24.0 | 0.80 | 29.3 |
| 29 | DE-Tha | 50.96361 | 13.56694 | ENF | 0.65 | 1.63 | 413 | 0.86 | 14.0 | 0.88 | 12.1 | 0.88 | 11.8 |
| 30 | ES-LgS | 37.09794 | -2.96583 | Woody Savannas | 2.01 | 11.28 | 92 | 0.84 | 11.3 | 0.77 | 17.1 | 0.62 | 18.3 |
| 31 | FI-Hyy | 61.8475 | 24.295 | ENF | 0.84 | -0.61 | 312 | 0.84 | 13.4 | 0.85 | 12.8 | 0.65 | 22.5 |
| 32 | FI-Sod | 67.36186 | 26.63783 | ENF | 0.49 | -0.08 | 185 | 0.47 | 22.5 | 0.46 | 22.2 | 0.43 | 25.9 |
| 33 | FR-Gri | 48.84422 | 1.95191 | Croplands | 0.75 | -1.18 | 232 | 0.82 | 21.6 | 0.83 | 20.5 | 0.81 | 19.8 |
| 34 | IT-Col | 41.84936 | 13.58814 | DBF | 0.88 | -0.67 | 180 | 0.72 | 20.4 | 0.74 | 18.6 | 0.79 | 18.1 |
| 35 | IT-Lav | 45.9562 | 11.28132 | ENF | 0.65 | 0.99 | 395 | 0.81 | 18.8 | 0.81 | 20.0 | 0.78 | 22.6 |
| 36 | IT-MBo | 46.01468 | 11.04583 | Grasslands | 0.52 | 3.15 | 419 | 0.82 | 20.9 | 0.84 | 21.4 | 0.87 | 17.6 |
| 37 | IT-PT1 | 45.20087 | 9.061039 | Croplands | 1.09 | 1.63 | 78 | 0.94 | 16.8 | 0.92 | 19.3 | 0.92 | 19.3 |
| 38 | IT-Ren | 46.58686 | 11.43369 | ENF | 0.62 | 0.53 | 360 | 0.75 | 33.0 | 0.80 | 32.6 | 0.77 | 34.5 |
| 39 | IT-Tor | 45.84444 | 7.578055 | ENF | 0.72 | -2.27 | 122 | 0.62 | 31.8 | 0.72 | 29.1 | 0.61 | 40.2 |
| 40 | NL-Loo | 52.16658 | 5.743556 | ENF | 0.83 | -0.70 | 380 | 0.70 | 28.9 | 0.70 | 28.4 | 0.67 | 26.1 |
| 41 | RU-Fyo | 56.46153 | 32.92208 | Mixed Forests | 0.80 | 0.82 | 314 | 0.83 | 17.5 | 0.84 | 17.1 | 0.71 | 25.2 |
| 42 | SD-Dem | 13.2829 | 30.4783 | Grasslands | 3.13 | 1.19 | 112 | 0.76 | 41.9 | 0.80 | 43.2 | 0.47 | 43.0 |
| 43 | US-AR1 | 36.4267 | -99.42 | Grasslands | 1.42 | 3.06 | 131 | 0.75 | 31.0 | 0.76 | 33.4 | 0.78 | 24.9 |
| 44 | US-AR2 | 36.6358 | -99.5975 | Grasslands | 1.48 | 0.38 | 121 | 0.75 | 20.2 | 0.77 | 19.9 | 0.76 | 15.7 |
| 45 | US-ARM | 36.6058 | -97.4888 | Croplands | 1.21 | -3.34 | 409 | 0.72 | 21.7 | 0.72 | 20.1 | 0.69 | 19.3 |
| 46 | US-ARb | 35.54974 | -98.0402 | Croplands | 1.28 | 1.87 | 78 | 0.82 | 31.0 | 0.82 | 32.2 | 0.89 | 28.3 |
| 47 | US-ARc | 35.54649 | -98.0401 | Grasslands | 1.27 | 1.87 | 84 | 0.86 | 37.5 | 0.85 | 39.5 | 0.89 | 35.1 |
| 48 | US-Blo | 38.89525 | -120.633 | ENF | 0.82 | 0.81 | 206 | 0.82 | 27.9 | 0.84 | 27.3 | 0.89 | 21.5 |
| 49 | US-Cop | 38.09 | -109.39 | Grasslands | 4.33 | -2.16 | 87 | 0.22 | 11.8 | 0.31 | 9.7 | 0.33 | 9.8 |
| 50 | US-GLE | 41.3644 | -106.239 | ENF | 0.41 | 10.44 | 261 | 0.46 | 27.2 | 0.46 | 27.5 | 0.37 | 36.1 |
| 51 | US-Los | 46.08268 | -89.9792 | Mixed Forests | 0.71 | 0.06 | 271 | 0.75 | 19.3 | 0.85 | 20.2 | 0.86 | 23.7 |
| 52 | US-MMS | 39.32315 | -86.4131 | DBF | 0.68 | 3.91 | 439 | 0.91 | 15.9 | 0.90 | 17.8 | 0.88 | 18.6 |
| 53 | US-Me2 | 44.4523 | -121.557 | ENF | 2.07 | 0.35 | 374 | 0.73 | 30.9 | 0.75 | 31.2 | 0.72 | 28.4 |
| 54 | US-NR1 | 40.03288 | -105.546 | ENF | 1.11 | -4.20 | 449 | 0.64 | 21.0 | 0.70 | 18.2 | 0.73 | 25.0 |
| 55 | US-Ne1 | 41.16506 | -96.4766 | Croplands | 0.89 | -0.86 | 427 | 0.88 | 34.0 | 0.89 | 32.9 | 0.87 | 31.7 |
| 56 | US-Ne2 | 41.16487 | -96.4701 | Croplands | 0.87 | -0.86 | 434 | 0.87 | 32.1 | 0.87 | 31.2 | 0.85 | 30.4 |
| 57 | US-Ne3 | 41.17967 | -96.4396 | Croplands | 1.00 | -0.42 | 420 | 0.88 | 27.4 | 0.89 | 26.6 | 0.87 | 25.8 |
| 58 | US-SRG | 31.7894 | -110.828 | Grasslands | N/A | 6.65 | 178 | 0.67 | 28.2 | 0.66 | 33.7 | 0.74 | 25.3 |
| 59 | US-SRM | 31.82143 | -110.866 | Open Shrublands | 2.78 | -1.89 | 397 | 0.50 | 28.3 | 0.58 | 26.6 | 0.67 | 19.6 |
| 60 | US-Syv | 46.24202 | -89.3477 | Mix. Forests | 0.87 | 0.57 | 171 | 0.88 | 17.8 | 0.89 | 15.2 | 0.91 | 15.6 |
| 61 | US-Ton | 38.4316 | -120.966 | Woody Savannas | 1.86 | -6.37 | 353 | 0.80 | 16.1 | 0.75 | 15.4 | 0.56 | 22.8 |
| 62 | US-Twt | 38.1055 | -121.652 | Croplands | 1.68 | 1.89 | 108 | 0.78 | 73.1 | 0.81 | 72.2 | 0.72 | 74.1 |
| 63 | US-Var | 38.40667 | -120.951 | Woody Savannas | 1.28 | -6.37 | 428 | 0.50 | 21.1 | 0.35 | 24.2 | 0.09 | 34.2 |
| 64 | US-WCr | 45.80593 | -90.0799 | DBF | 0.82 | 0.04 | 194 | 0.78 | 20.5 | 0.84 | 18.3 | 0.82 | 19.3 |
| 65 | US-Whs | 31.74383 | -110.052 | Open Shrublands | 3.32 | 4.02 | 228 | 0.69 | 17.4 | 0.76 | 19.1 | 0.64 | 20.0 |
| 66 | US-Wkg | 31.73653 | -109.942 | Grasslands | 2.78 | 0.67 | 389 | 0.61 | 19.0 | 0.68 | 19.3 | 0.73 | 15.5 |
| 67 | ZA-Kru | -25.0197 | 31.4969 | Savannas | 2.85 | -2.33 | 180 | 0.62 | 30.3 | 0.64 | 29.6 | 0.57 | 31.1 |
| 68 | ZM-Mon | -15.4378 | 23.25278 | Savannas | 1.68 | -2.88 | 92 | 0.76 | 24.8 | 0.78 | 22.9 | 0.53 | 26.8 |

ENF: Evergreen Needleleaf Forest, DBF: Deciduous Broadleaf Forest, EBF: Evergreen Broadleaf Forest, Mosaic: Cropland/Natural Vegetation Mosaic

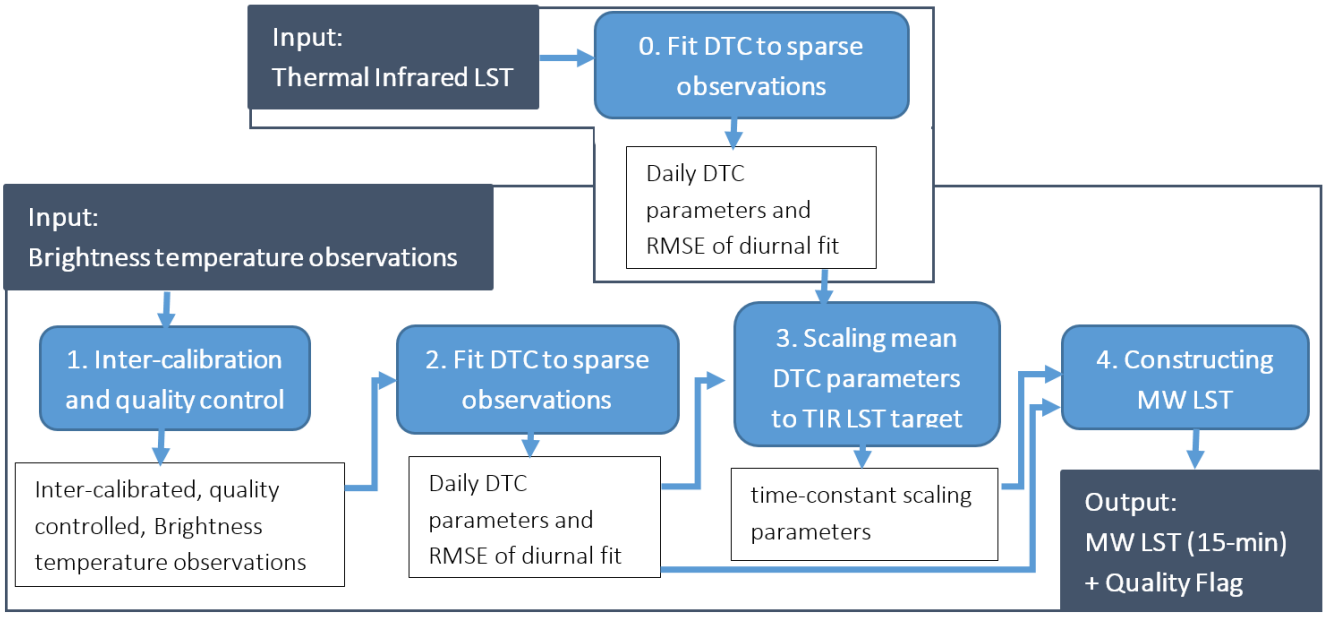


**Figure 1: MW-LST workflow**

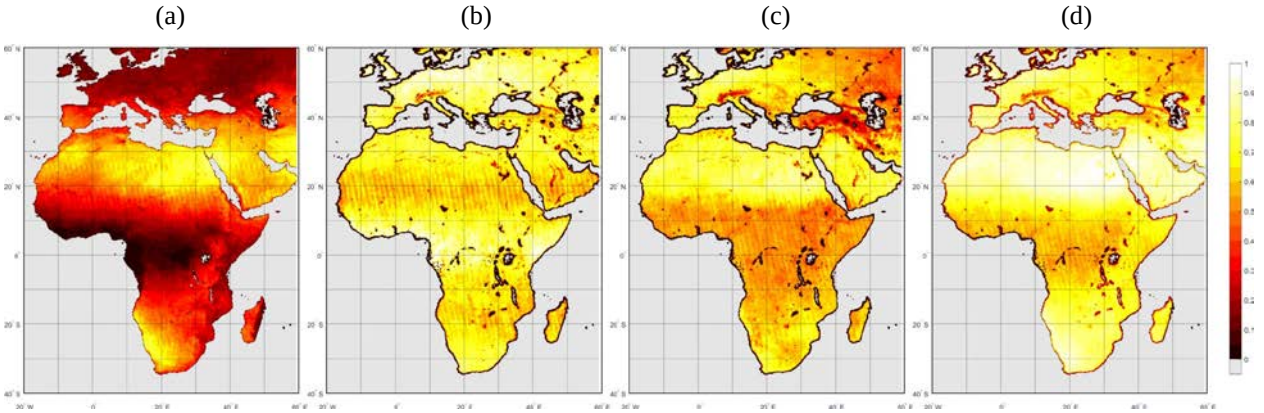


**Figure 2: Temporal coverage of MW and IR-based $\Delta T_{rad}$ in 2004. Panel a shows the fraction of total days where MODIS-based estimates of $\Delta T_{rad}$ are available. Panel b shows the fraction of this subset of days where there is also a MW-based $\Delta T_{rad}$ available. Panel c shows the fraction of days without a MODIS-based estimate but with availability of a MW-based estimate (potential for IR-gap coverage). Panel d shows the fraction of total days where either a MODIS- or a MW-based $\Delta T_{rad}$ is available.**

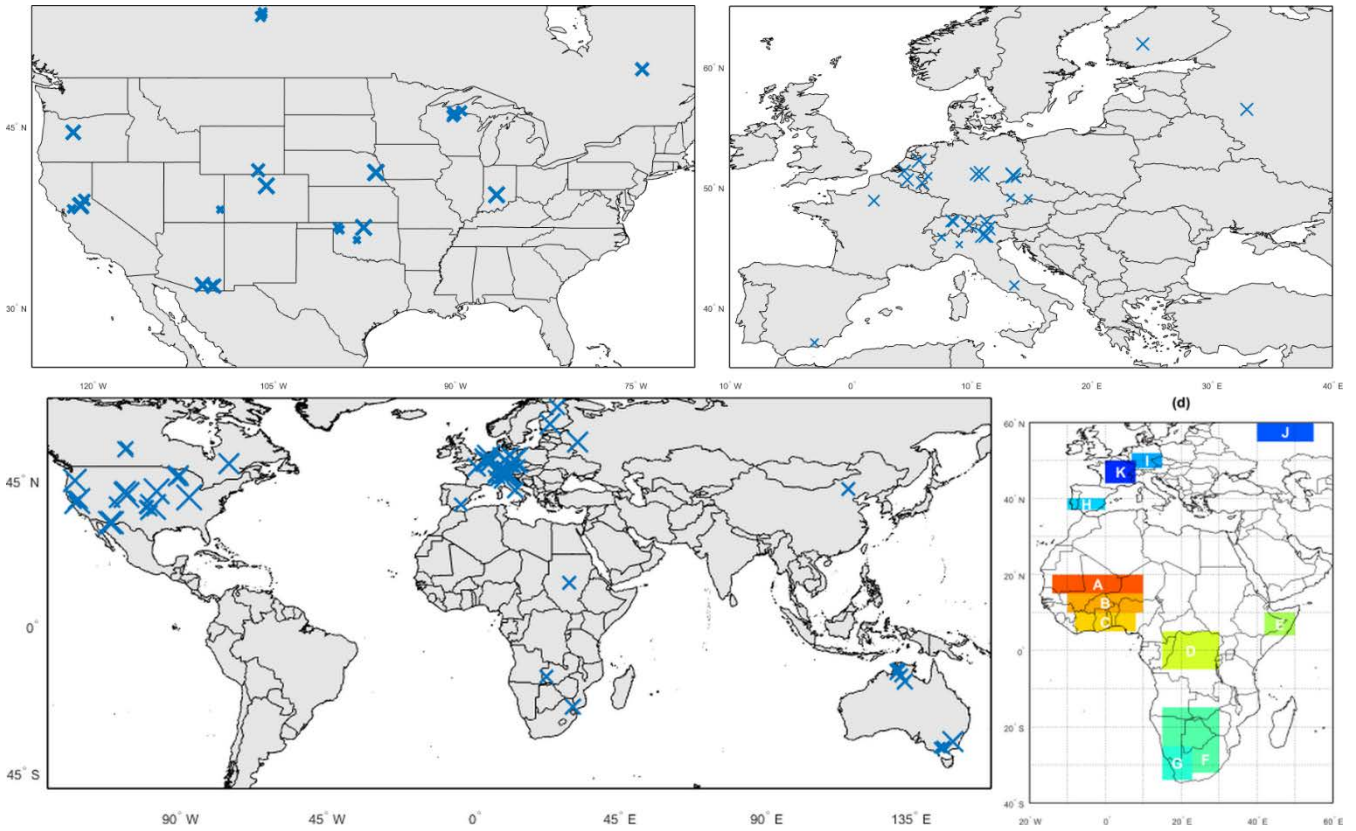


**Figure 3: Location of flux tower sites used in the analysis (see also Table 2): (a) North America, (b) Europe, (c) World. The size of the marker is in proportion to number of data days used in assessment. Panel (d) indicates the 11 regions selected based on annual precipitation cycle and geographic diversity.**

5

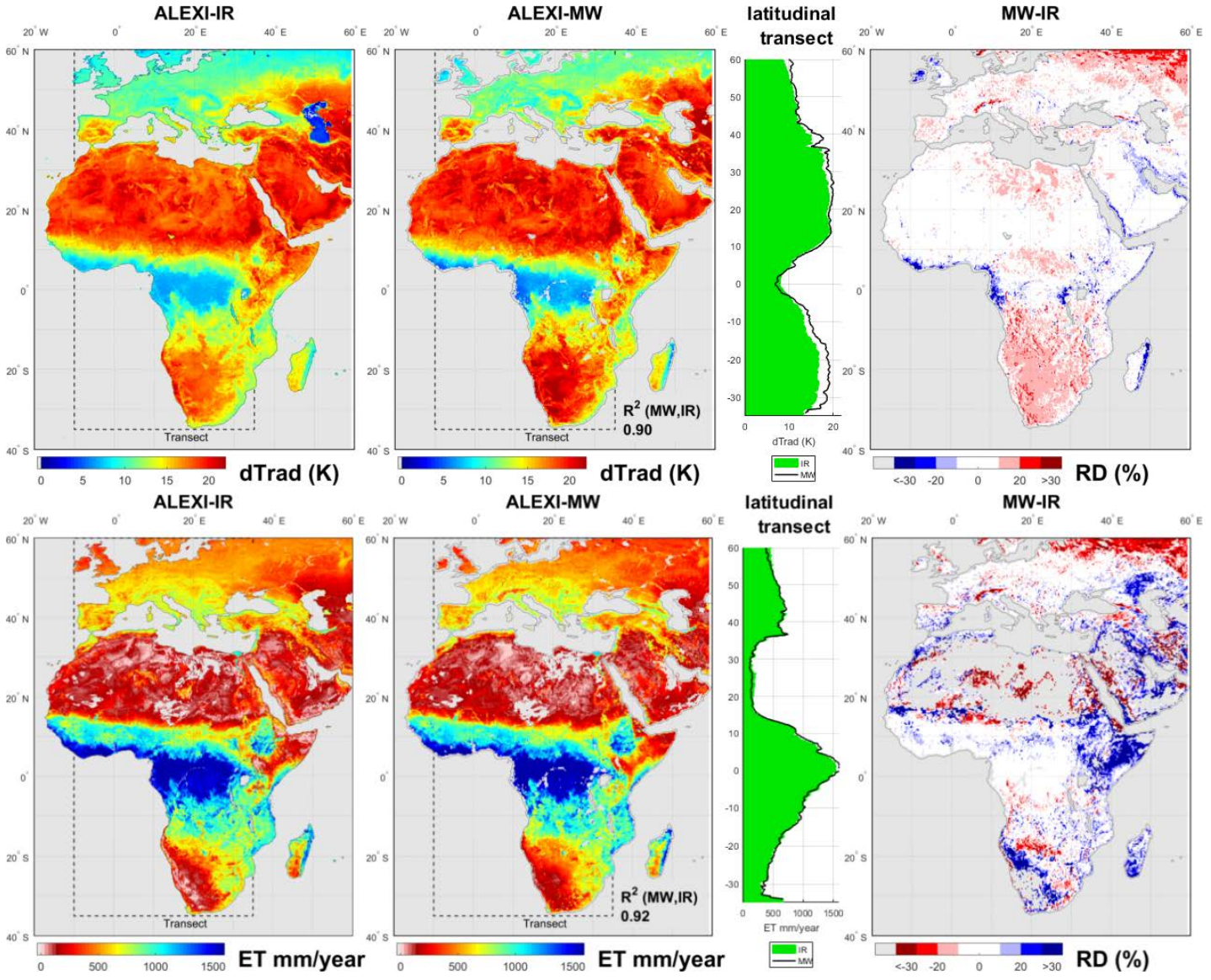


**Figure 4: Multi-year mean of clear-sky $\Delta T_{rad}$ (top row, 2004-2011) and mean annual clear-sky ET (bottom row, 2003-2011) for IR and MW. The transect shows the latitudinal average for longitude 10°W to 35°E. The right-hand panel shows the corresponding relative difference (RD) between the two estimates (MW - IR), with areas with ET below 14 mm/month greyed out. Note the reversed colour bars for $\Delta T_{rad}$ and ET to emphasize their negative correlation ($\Delta T_{rad}$ up, ET down). $\Delta T_{rad}$ is defined in the ALEXI framework as the temperature rise between 1.5 hr after sunrise to 1.5 hr before noon (see Section 2.1).**

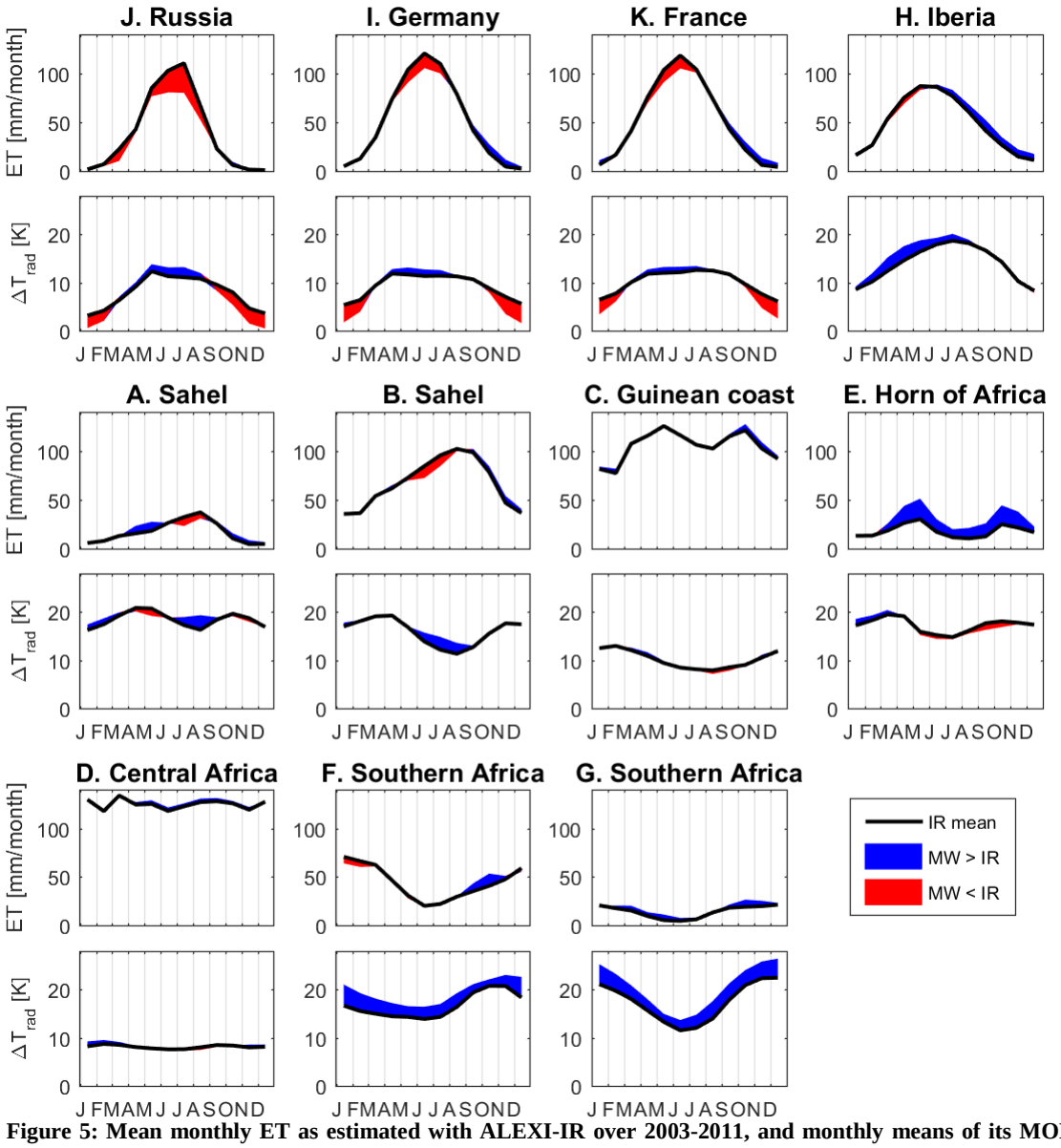


**Figure 5: Mean monthly ET as estimated with ALEXI-IR over 2003-2011, and monthly means of its MODIS-based $\Delta T_{rad}$ input (period 2004-2011), for selected regions. The deviation from these IR estimates when using the MW inputs is shown in blue for a positive deviation and red for a negative deviation.**

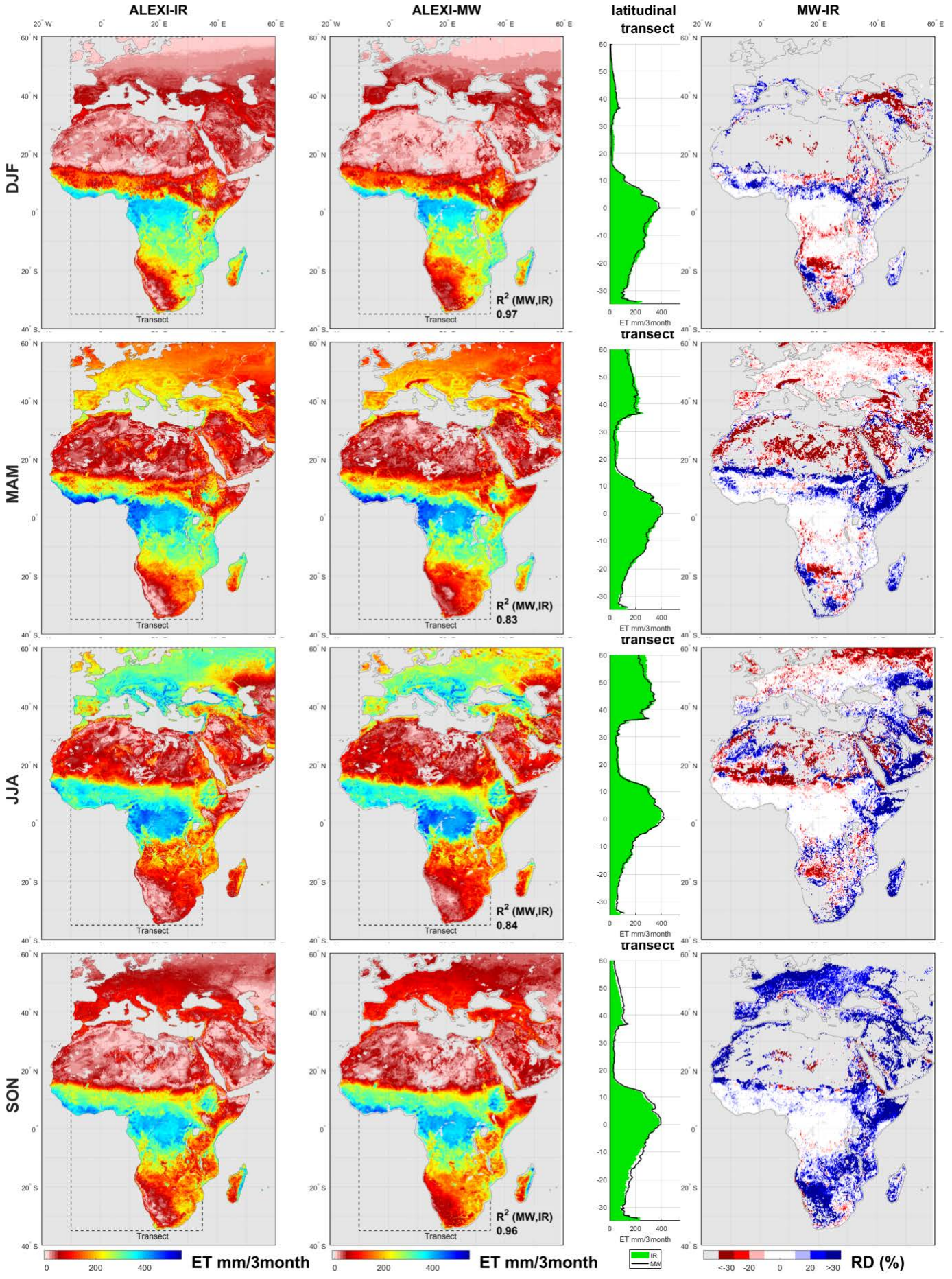


**Figure 6: As Fig 4, but now averaged by season.**

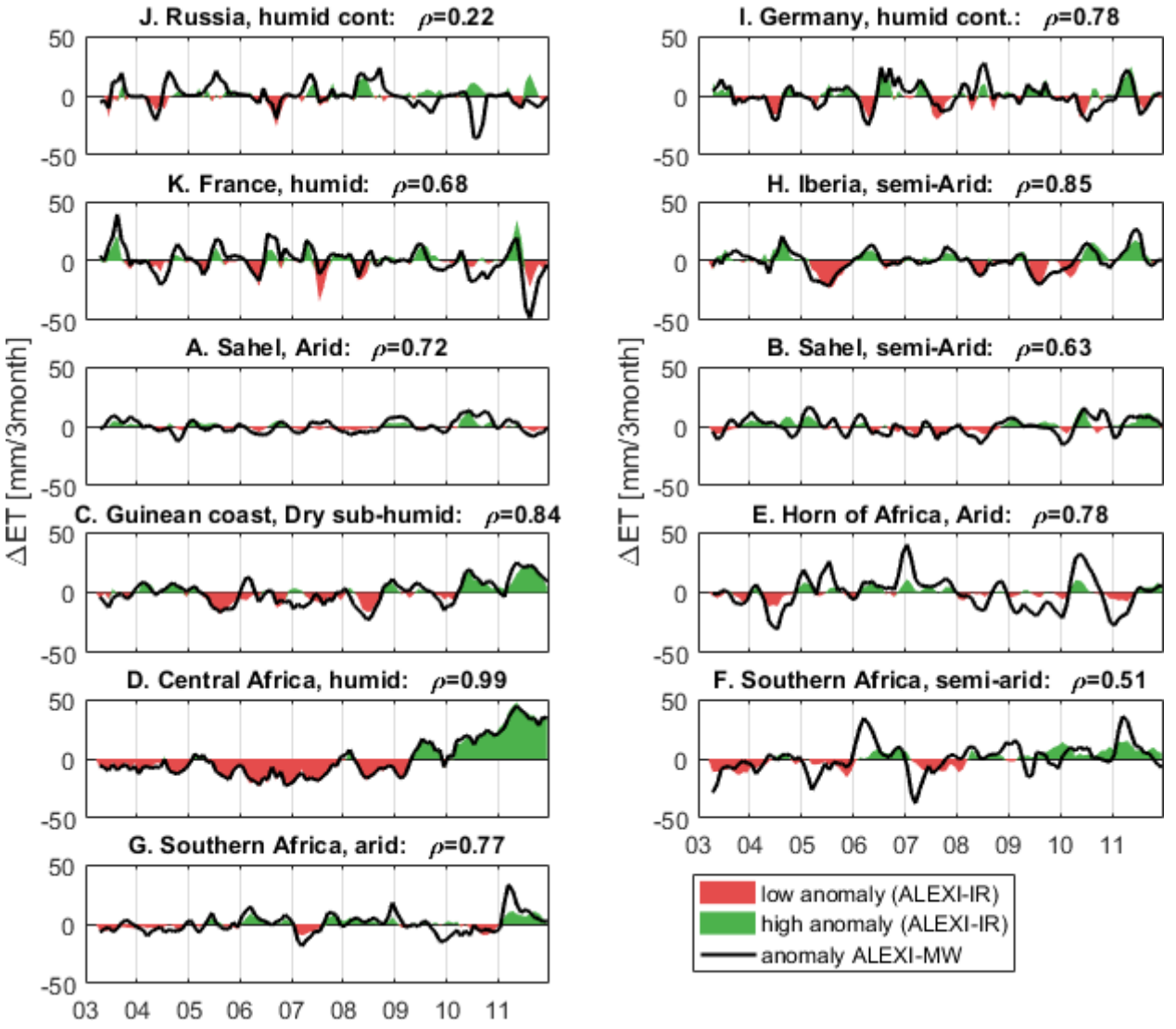


**Figure 7: Comparison of anomaly in 3-month ET totals as calculated from ALEXI-IR and ALEXI-MW for selected regions (see Fig. 3 for definition of regions).**

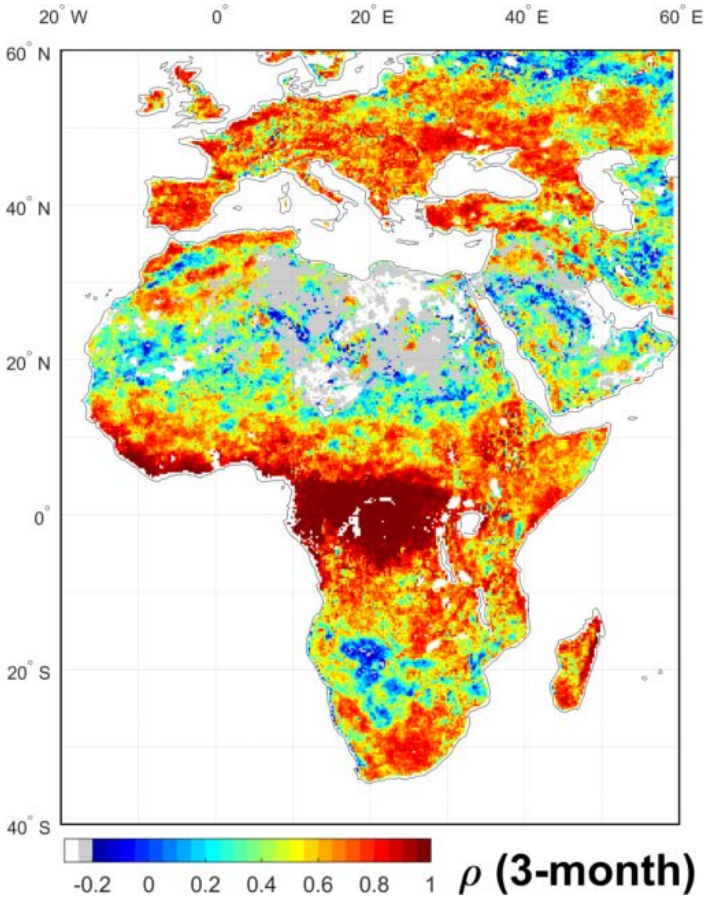


**Figure 8: Pearson's correlation between anomaly in 3-month ET totals as estimated by ALEXI-IR and ALEXI-MW, calculated at 0.25 degree resolution. White areas have no data, grey areas are masked because the standard deviation in 3-month anomaly was below 8 mm/3month in both estimates.**

5

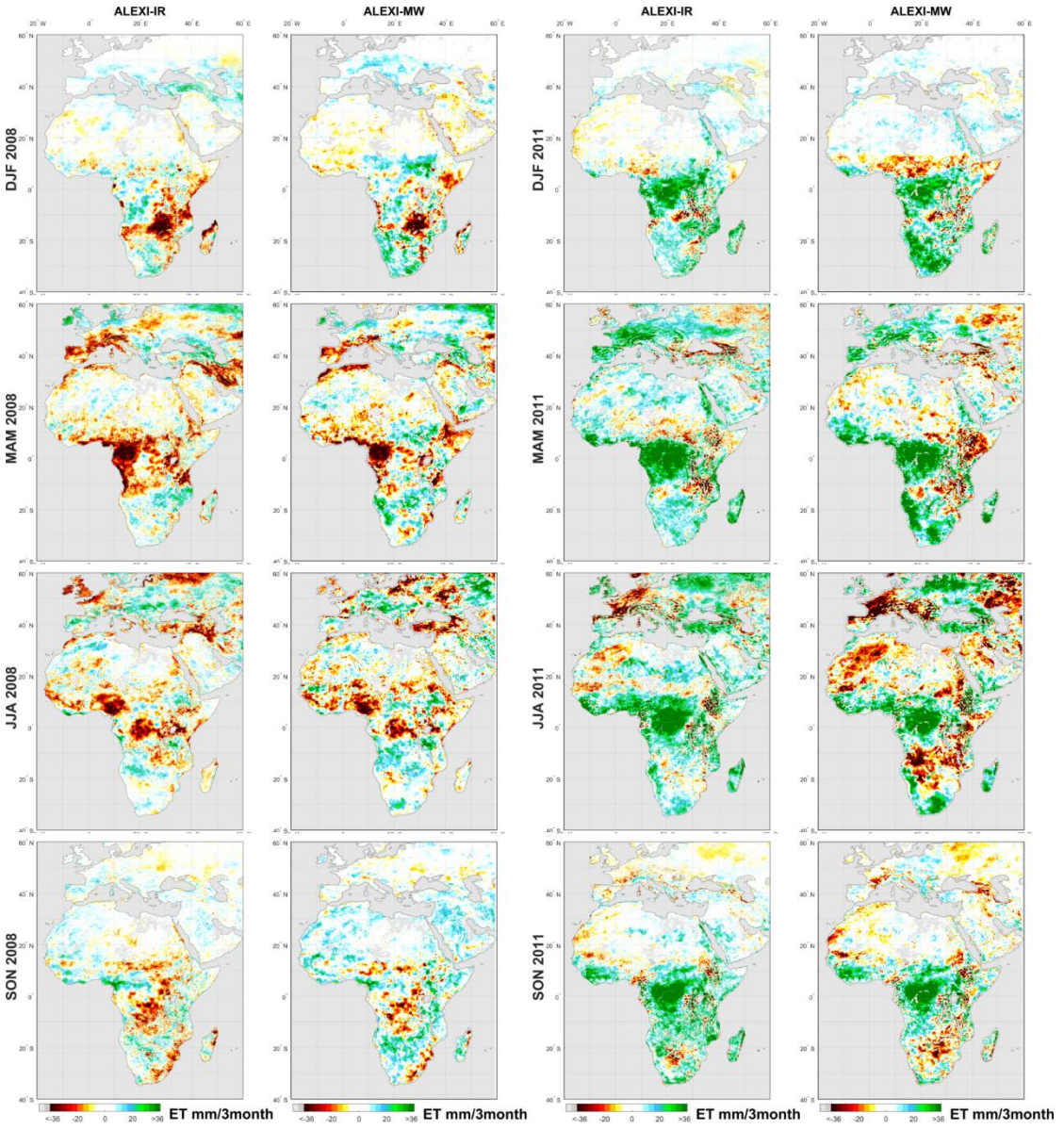


**Figure 9: Anomaly in seasonal ET compared to multi-year mean (2003-2011, see Fig 6) as retrieved by ALEXI-IR and ALEXI-MW. The first two columns show the anomalies for 2008 and the two right-hand columns show them for 2011.**

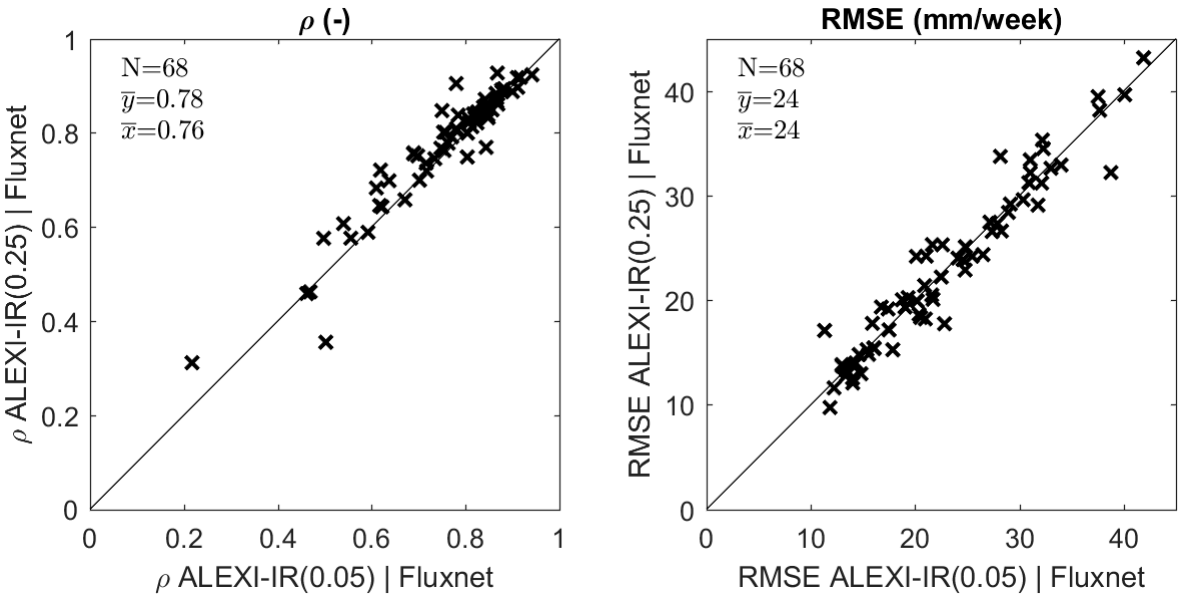


**Figure 10: The effect of spatial resolution in satellite product on Pearson correlation (ρ) and RMSE between weekly ET estimates from satellite data (ALEXI-IR) and flux-tower eddy-covariance measurements (Fluxnet). Each marker represents a single station and compares results at the original 0.05° resolution of ALEXI-IR (X-axis) with those calculated for 0.25° resolution ALEXI-IR (Y-axis).**

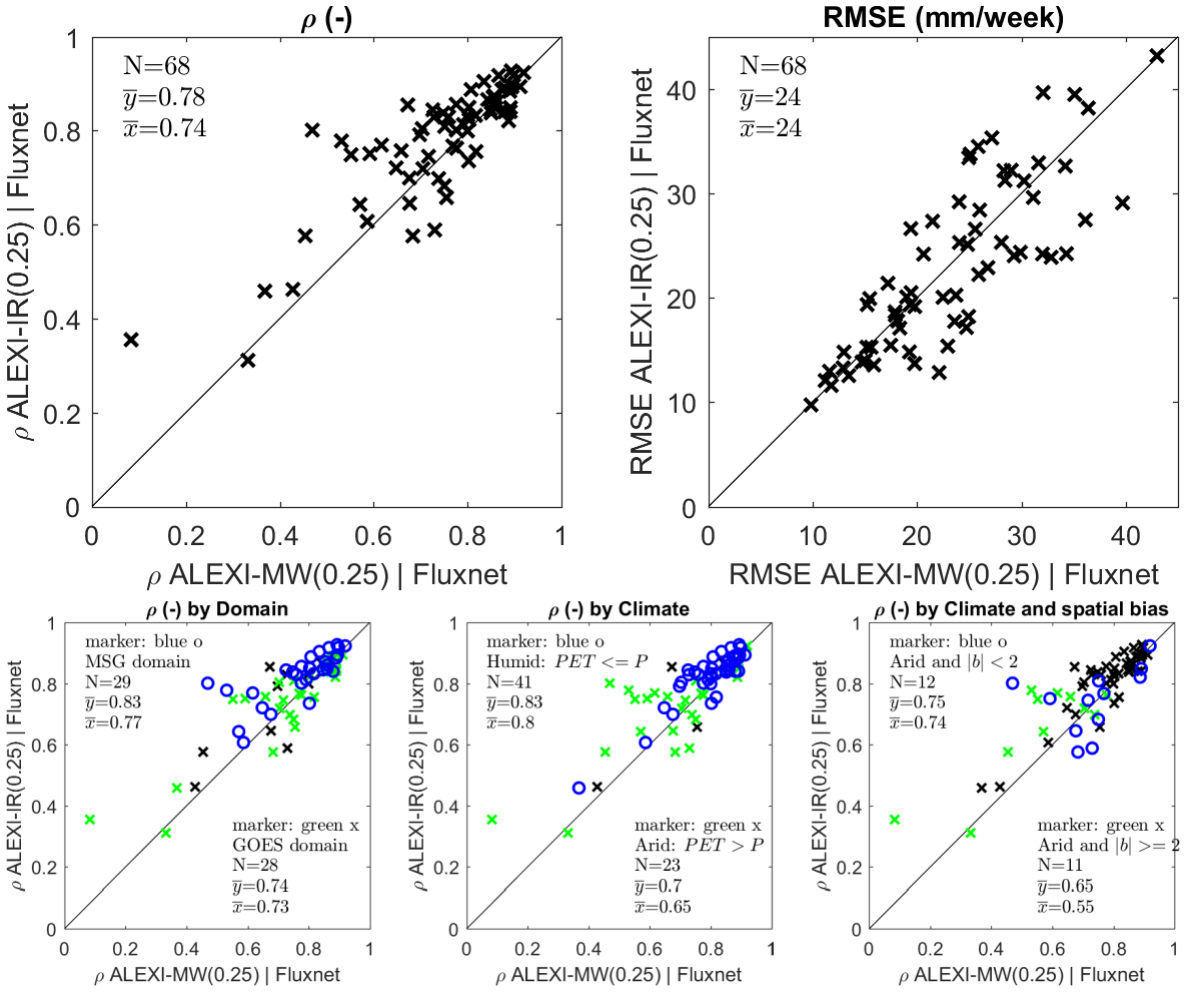


**Figure 11: The effect of switching from TIR to MW-LST as input to ALEXI on Pearson correlation (ρ) and RMSE between weekly ALEXI ET estimates and flux-tower eddy-covariance measurements (Fluxnet). Each marker represents a single station and compares results calculated for ALEXI-MW (X-axis) with those calculated for ALEXI-IR (Y-axis). Second row: same data as presented in the left-hand panel on the top row, but now distinct subsets of the tower sites are emphasized. The first panel splits the sites by geographic region, the second panel based on climate (Humid Vs Arid, see text for definition). Panel three splits the 'arid' sites further based on bias between the ALEXI-IR (0.05°) and the mean value for the encompassing 0.25° grid box with a threshold of |b|=2mm/week. The black x mark stations that are either below 60°N, or are not covered by the two contrasting selections.**