Remote Sensing of Environment 221 (2019) 157–172

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations

Radoslaw Guzinski[a,b,*], Héctor Nieto[c]

[a] European Space Agency, ESA Centre for Earth Observation (ESRIN), Largo Galileo Galilei 1, Casella Postale 64, Frascati 00044, Roma, Italy
[b] DHI GRAS, Agern Alle 5, Hørsholm 2970, Denmark
[c] IRTA Institute for Agriculture and Food Research and Technology, Edifici Fruitcentre, Parc de Gardeny, Lleida 25003, Spain

ARTICLE INFO

Keywords:
Sentinel-2
Sentinel-3
Thermal
Sharpening
Machine learning
Evapotranspiration
Heat flux

ABSTRACT

The Sentinel satellite missions are designed to provide remote-sensing observational capability to many diverse operational applications, including in the field of agriculture and food security. They do this by acquiring frequent observations from a combination of optical, thermal and microwave sensors at various spatial resolutions. However, one currently missing capability, that would enable monitoring of evapotranspiration, crop water stress and water use at field scale, is the lack of high-resolution (tens of meters) thermal sensor. In this study we evaluate a methodology for bridging this data gap by employing a machine learning algorithm to sharpen low-resolution thermal observations from the Sentinel-3 satellites using images acquired by high-resolution optical sensors on the Sentinel-2 satellites. The resulting dataset is then used as input to land-surface energy balance model to estimate evapotranspiration. The methodology is tested using Terra and Landsat satellite observations, due to lack of sufficiently long time-series of Sentinel observations, and benchmarked against fluxes derived with high-resolution thermal observations acquired by the Landsat satellites. We then apply the methodology to Sentinel-2 and Sentinel-3 images to confirm its applicability to this type of data. The results show that the fluxes derived with sharpened thermal data are of acceptable accuracy (relative error lower than 20%) and provide more information at flux-tower footprint scale than the corresponding low-resolution fluxes. They also replicate the spatial and temporal patterns of fluxes derived with high-resolution thermal observations. However, the increase in error of the modelled fluxes compared to using high-resolution thermal observations and the inherent limitations of the sharpening approach point to the need to add high-resolution thermal mission to the Sentinels' constellation.

1. Introduction

The Sentinel satellite missions are opening an operational era for the use of remote sensing observations in many domains, due to their range of spatial and spectral resolutions, frequent revisit times and long term guarantee of data continuity (Donlon et al., 2012; Drusch et al., 2012; Torres et al., 2012). Among those domains are agriculture and food security, for which Sentinel data has been used to operationally create dynamic cropland masks (Valero et al., 2016), perform crop type mapping (Inglada et al., 2015), monitor rice production (Torbick et al., 2017) or estimate plant parameters such as leaf area index (Campos-Taberner et al., 2017; Clevers et al., 2017), all at field-scale spatial resolution (10–20 m).

Estimates of evapotranspiration (ET) derived from satellite based observations are also of high interest in agriculture and food security domains as well as in other applications such as water resource management (Anderson et al., 2012). To be of most use, the ET maps should have a spatial resolution on the order of spatial scale of the predominant landscape feature (Guzinski et al., 2014; Kustas et al., 2004), which in terms of agriculture is a field that often ranges from hundreds (10 m by 10 m) to thousands (100 m by 100 m) of meters square (Lowder et al., 2016). At the same time, to reliably estimate the actual ET and crop water stress it is preferable to utilise satellite observations both in visible/near-infrared (VISNIR) and thermal-infrared (TIR) parts of the electromagnetic spectrum (Anderson et al., 2012; Hoffmann et al., 2016a). However, the Sentinel satellite constellation contains only one TIR sensor (Sea and Land Surface Temperature Radiometer - SLSTR) with a spatial resolution of 1000 m at nadir (on board of the Sentinel-3 satellite), while having multiple sensors with VISNIR observations capability ranging from a spatial resolution of 10 m to

* Corresponding author at: DHI GRAS, Agern Alle 5, Hørsholm 2970, Denmark.
E-mail addresses: rmgu@dhigroup.com (R. Guzinski), hector.nieto@irta.cat (H. Nieto).

https://doi.org/10.1016/j.rse.2018.11.019
Received 13 October 2017; Received in revised form 24 October 2018; Accepted 12 November 2018

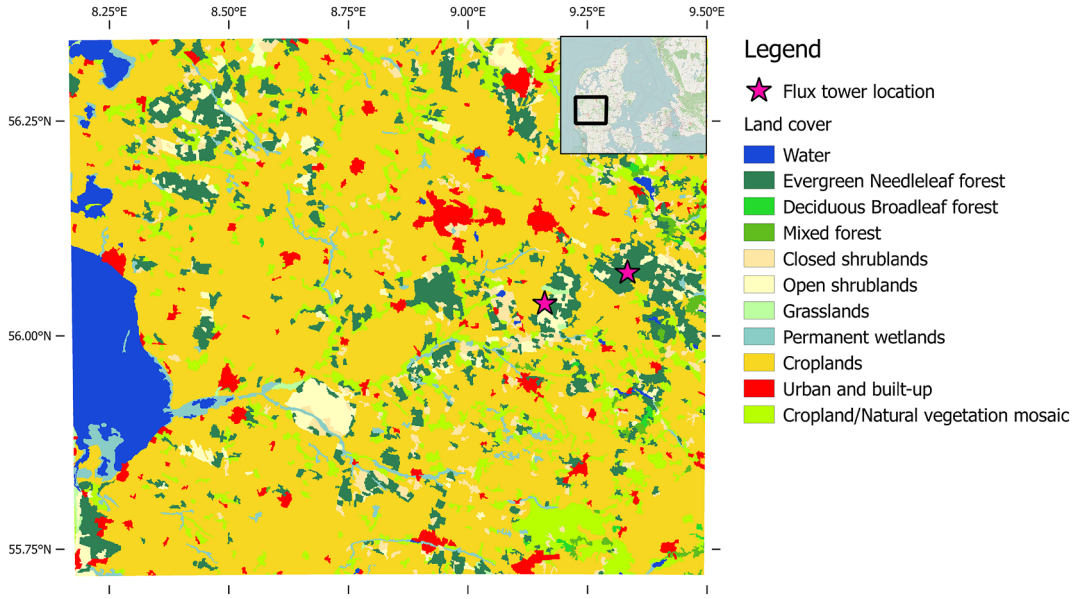

Fig. 1. Land cover map of the study area showing the extent of Landsat scene used in DMS.

1000 m (on board of Sentinel-2 and Sentinel-3 satellites). Even though recent studies have been highlighting the importance of high-resolution TIR observations for ET estimation (Fisher et al., 2017; Sobrino et al., 2016a) and a number of satellite mission concepts are (or have been) proposed (Crebassol et al., 2014; Lagouarde et al., 2013; Lee et al., 2015) or even launched (Hulley et al., 2017), a high-resolution TIR satellite mission which could act as a companion to the high-resolution VISNIR Sentinel-2 satellites is not expected to be launched within the closest number of years. Therefore, a method is required to bridge the spatial gap between the currently available Sentinel constellation's TIR and VISNIR observational capabilities in order to optimally exploit the synergies of both types of sensors for field-scale ET estimations.

A number of techniques have been developed for sharpening lower spatial resolution TIR observations with higher spatial resolution VISNIR observations (Chen et al., 2014; Bindhu et al., 2013; Bisquert et al., 2016; Gao et al., 2012). Most of those techniques consist of two general steps. In the first step, a relationship is derived between the TIR data and high-resolution VISNIR data aggregated to the resolution of the TIR data. In the second step, this relationship is applied to high-resolution VISNIR data to obtain TIR observations at the same high resolution. A post-processing step is sometimes also included which, for example, removes the bias between TIR datasets at the two resolutions. This sharpened TIR data can then be used in land-surface energy flux models to estimate high-resolution ET.

Similarly, a number of techniques have been developed for disaggregating low-resolution land surface energy fluxes to higher spatial resolutions (Guzinski et al., 2014; Norman et al., 2003). Those techniques rely on the higher spatial resolution TIR data (for example coming from the Landsat series of satellites), which are assumed to have lower accuracy than the coarser resolution TIR data. Furthermore, it should be also possible to sharpen the estimated low-resolution land surface fluxes using purely high-resolution VISNIR data following the same approach as for sharpening low-resolution TIR data.

Both of those approaches (i.e. sharpening low-resolution TIR data using high-resolution VISNIR data and using the sharpened output to estimate high-resolution ET; and directly sharpening estimated ET using high-resolution VISNIR data) are applicable to the Sentinel-2 (S2) –Sentinel-3(S3) satellite constellation. In this study we evaluate a number of combinations of the sharpening and disaggregation approaches for their ability to produce high-resolution ET maps in a river catchment in western Denmark. The evaluation includes the validation of the fluxes against two flux towers located in the catchment (one in an agricultural site and the other in a coniferous plantation forest) as well as comparison of spatial and temporal patterns of high-resolution ET produced by the different techniques. Due to the lack of sufficient number of Sentinel-3 images, we use Terra and Aqua (MODIS sensor) satellite data as a proxy for Sentinel-3 and Landsat data as a proxy for Sentinel-2. Even though the spectral bands of the sensors on board those pairs satellite are not exactly the same, the spatial resolutions are similar (≈1000 m for Terra/Aqua and S3 TIR data, and 30 m for Landsat VISNIR data compared to 10–20 m for S2 VISNIR data) and thus the study provides guidance on the optimal approach for deriving high-resolution ET with Sentinel-3 and Sentinel-2 observations.

2. Methods

2.1. Data sharpening scheme

The data sharpening method used in this study is based on Data Mining Sharpener (DMS) introduced by Gao et al. (2012). In the current study it was implemented as an open-source Python application, which is freely available online (https://github.com/radosuav/pyDMS, last accessed: 01/06-2017). The DMS is a machine-learning algorithm that derives a statistical relationship between the predictor high-resolution variables aggregated to low-resolution data pixels and the low-resolution variables that need to be sharpened. It then applies this relationship to the high-resolution data to derive high-resolution representation of the low-resolution data.

The current DMS implementation is based on an ensemble of decision-tree regressors, with each decision-tree trained with a random subset of the training samples drawn with replacement in a method known as Bagging (Breiman, 1996). The final value of the regression is an average of the values produced within the ensemble. The training sample is composed of the fine resolution data which passes a homogeneity threshold when aggregated to the coarser resolution (i.e. the coefficient of variation of high-resolution pixels falling within a low-resolution pixel must be below a certain threshold). All samples falling within a single regression tree leaf node are further fitted with a multivariate linear model. Therefore, once the DMS is trained an output value for a given input is determined by locating a regression tree leaf to which the input belongs and running the corresponding linear model, for an ensemble of regression trees.

The ensemble decision tree regression is performed both locally (in a moving-window fashion) and globally (i.e. to the whole study area

shown in Fig. 1). The results are then combined based on residuals between the regression outputs and the low-resolution training data (Gao et al., 2012). Also residual analysis and bias correction between regression outputs and low-resolution data is applied at the end of the sharpening process to ensure consistency (i.e. conservation of emitted longwave energy in case of thermal data) between the sharpened high-resolution pixels and their corresponding low-resolution pixel.

2.2. Land surface fluxes' modelling scheme

The land surface fluxes' modelling scheme is based on the Two-Source Energy Balance (TSEB) modelling scheme proposed by Norman et al. (1995) and further evolved in Kustas and Norman (1999). Briefly, the model splits the directional radiometric land surface temperature into its two main components: the temperatures of vegetation and soil. Based on this, net radiation and latent and sensible heat fluxes of vegetation and soil are estimated separately, before being combined to obtain the bulk surface fluxes. The transfer of heat between the two components and the atmosphere is modulated by "resistances" (in analogy to electrical systems) which depend on aerodynamic and meteorological conditions. Those resistances can be in parallel or series configuration, and in this study the latter one was used since it better accounts for the interaction between turbulent fluxes originating from the two components (Li et al., 2005; Morillas et al., 2013). In addition, the ground heat flux of the soil is estimated following the method of Santanello and Friedl (2003). At both the component and bulk levels the model enforces the principle of conservation of energy.

For more implementation details the reader is referred to the full source code of the Python implementation of the model used in this study which available online (https://github.com/hectornieto/pyTSEB/releases/tag/v1.4, last accessed 21.10.2018) and the referenced publications.

2.3. Flux disaggregation scheme

The flux disaggregation scheme is described in Guzinski et al. (2014) and is based on the assumption of evaporative fraction (EF - fraction of total available energy used for latent heat flux) remaining stable during the late-morning and early-afternoon hours under cloud-free conditions. The EF is estimated independently from TSEB fluxes derived using the low-resolution and high-resolution data. The high-resolution EF is then aggregated to the lower resolution and compared against the low-resolution EF. If the fine-resolution EF is higher (lower) than coarse-resolution EF in a given coarse-resolution pixel then the air temperature at blending height of that pixel is decreased (increased), in order to decrease (increase) the latent heat flux and the fine-resolution fluxes are recalculated. This is done under the assumption that the coarse-resolution temperature data produces more accurate flux estimates and therefore the high-resolution fluxes have to be adjusted to match the low-resolution ones (a valid expectation in case of Landsat and MODIS thermal sensors as well as in the case of sharpened thermal images).

2.4. Evaluated model configurations

In this study we compare the high-resolution (30 m) flux estimates derived in three different ways: using Landsat LST in TSEB together with the disaggregation scheme and low-resolution fluxes derived with MODIS LST (disTSEB$_L$); using DMS sharpened MODIS LST directly in TSEB (TSEB$_{M,DMS}$); and finally using DMS sharpened MODIS LST in TSEB together with the disaggregation scheme (disTSEB$_{M,DMS}$). The first model configuration is used as benchmarks for the latter two. In addition TSEB was also run with low-resolution MODIS data (TSEB$_M$) to allow evaluation of the benefit (or lack thereof) of deriving the fluxes at higher resolution. All those model runs were repeated twice: once with tower-based meteorological inputs (optimal case) and once with modelled meteorological inputs (typical case for regional applications). The two meteorological datasets are described in Section 3. Finally, the application of DMS to low-resolution fluxes (instead of low-resolution LST) was also tested (TSEB-DMS$_M$) but only when modelled meteorological data was used, since regional scale flux estimates are needed for training the DMS.

3. Data

3.1. Study sites and tower-based measurements

The study area is located in the Skjern river catchment on the western side of Denmark's Jutland peninsula (Fig. 1). It is a heavily agricultural area consisting mostly of fields of winter and summer barley, potatoes and plantation forest. Even though it experiences a mild maritime climate with an average temperature of 8 °C and average annual rainfall of 850 mm, irrigation of the crops is still required due to the predominantly sandy nature of the soil (Ringgaard et al., 2011).

The Skjern catchment is the location of the Danish Hydrological ObsErvatory (HOBE) which maintains three eddy-covariance (EC) flux towers within the area (Jensen and Illangasekare, 2011). The measurements from two of those towers, one located in a barley field near the town of Voulund (referred to as VOU tower) and the other in a conifer plantation forest with an average tree height of 20 m close to the town of Gludsted (referred to as GLU tower), are used in the current study. The area around the VOU tower is heterogeneous at 1 km scale, with multiple fields as well as shrubs and forests falling within a 1 km radius. The GLU tower, on the other hand, is surrounded by the plantation forest for at least 1 km in each direction but becomes heterogeneous upon closer inspection with different tree stands, clearings and forest roads visible at higher resolutions.

The VOU tower has the EC system located at 6 m height above ground level (agl) and an air temperature and humidity sensor located at a height of 4 m agl. At the GLU tower the EC system is at a height of 38 m agl and air temperature and humidity were measured at a height of 30 m agl. Both towers also have sensors for measuring four components of net radiation (incoming/outgoing and shortwave/longwave) as well as heat plates in the vicinity of the towers for the measurement of ground heat flux. More detailed description of the towers' location and set-up can be found in Ringgaard et al. (2011), while the pre-processing applied to tower measurements is described in Guzinski et al. (2014).

Meteorological measurements from the flux towers (air temperature and humidity, wind speed, and incoming solar radiation) were used as inputs to the TSEB model for the local-scale model runs. The fluxes measured by the towers (sensible heat, latent heat, ground heat and net radiation) were used to validate the TSEB model outputs. The energy balance closure at both towers was enforced by assigning the residual energy to the latent heat flux, as was done in previous studies utilising the same data (Guzinski et al., 2013, 2014; Hoffmann et al., 2016b). Flux tower footprint were estimated as described in Guzinski et al. (2014).

3.2. Satellite data

The low-resolution ($\approx$1000 m) data originates from the MODIS sensor on board the Terra and Aqua satellites. Four products from Collection 6 were used: MOD11A1 - instantaneous land surface temperature and emissivity (LST), at daily timestep (Wan, 2014); MCD15AH2 - 8 day leaf area index (LAI) composite (Yan et al., 2016); MCD43A3 - 16 day albedo composite; MOD13A2 - 16 day vegetation indices composite. The products that are provided with 500 m resolution (MCD15AH2 and MCD43A3) were resampled to 1000 m using pixel averaging. The albedo used in the models was derived from MCD43A3 white and black sky albedos with an assumption that in clear-sky atmosphere 80% of the light is direct and 20% diffuse. This was then further split into canopy and soil components using fraction of

Table 1
TSEB model parameters assigned by land cover class.

	Evergreen Needleleaf forest	Deciduous Broadleaf forest	Mixed Forest	Closed shrublands	Open shrublands	Grasslands	Wetlands	Croplands	Cropland/natural vegetation mosaic
Vegetation height (m)	20.0	15.0	20.0	1.5	1.5	0.5	1.0	1.2	1.2
Ratio of canopy height to width	2.0	1.0	1.5	1.5	1.2	1.0	1.0	1.0	1.0
Leaf size (m)	0.05	0.10	0.07	0.10	0.05	0.02	0.02	0.20	0.20
Leaf angle distribution	1.0	1.0	1.0	1.0	0.2	0.5	0.5	0.5	0.5
Leaf transmittance - VIS	0.05	0.15	0.10	0.15	0.15	0.15	0.15	0.15	0.15
Leaf transmittance - NIR	0.30	0.40	0.35	0.40	0.40	0.40	0.40	0.40	0.40

vegetation cover (see below) and the assumption that in the visible part of the spectrum albedo of soil is twice that of vegetation and in the near-infrared albedo of soil is 70% of that of vegetation. This assumption is based on standard spectral profiles of soil and vegetation (Baldridge et al., 2009) and while it might have an effect on the component fluxes originating from soil and vegetation, its impact on the bulk fluxes should be minimal, especially for moderate to dense canopies (e.g. LAI > 1). Emissivity was similarly split into soil and vegetation components by assigning vegetation the value of 0.995 and estimating the soil emissivity using fraction of vegetation cover and emissivity from MOD11A1 product. While MCD15AH2 and MCD43A3 products are derived from observations taken by both Terra and Aqua satellites, it was decided to use only Terra observations for the instantaneous land surface temperature model inputs since the overpass time of Terra is closer to that of Sentinel-3.

The high-resolution (30 m) data came from the sensors on board Landsat 5, Landsat 7 and Landsat 8 satellites. The atmospheric correction of both thermal and shortwave images was performed as described in Guzinski et al. (2014). Vegetation indices were calculated directly from the top-of-canopy (TOC) reflectances while Landsat LST was derived as described in Guzinski et al. (2014). It should be noted that while thermal data is acquired by the Landsat satellites at a resolution ranging from 60 m to 120 m (depending on the satellite), it is resampled to 30 m using cubic convolution before distribution to the users and therefore in this study was assumed to have a resolution of 30 m.

High resolution LAI and albedo were derived by running the DMS (see Section 2.1) with TOC VISNIR Landsat reflectances and MCD15AH2 and MCD43A3 MODIS products respectively. The DMS and TOC VISNIR Landsat reflectances were also used to sharpen the MOD11A1 MODIS LST products to be used as a replacement of Landsat LST, as well as to sharpen flux outputs of the low-resolution TSEB model runs. In all cases the moving-window size for local regression was set to 30 by 30 MODIS pixels and the homogeneity threshold for both local and global regressions was set automatically to select the 80% most homogeneous pixels from the aggregated fine-resolution image. The linear regression extrapolation was limited to 25% of the range of the training samples falling within the corresponding regression tree leaf node (see Section 2.1). The DMS was applied on a scene-by-scene basis, meaning that the Landsat scenes were not combined into a larger multi-temporal training data set but each Landsat scene was used individually to sharpen MODIS product or TSEB output corresponding to the scene acquisition date.

The Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) were used to derive fraction of vegetation that is green and actively transpiring (f_g). In most land cover classes (see Section 3.3 below) f_g was estimated as $f_g = 1.28\frac{EVI}{NDVI}(0 < f_g < 1)$ following the approach of Guzinski et al. (2013). The only exception was in croplands where $f_g = EVI$ during crop senescence (set as between day-of-year 160 and 230) and $f_g = 1$ during other periods, in order to avoid issues present in this land cover class detailed in Guzinski et al. (2013). MODIS LAI (both original resolution and sharpened) was assumed to represent green LAI and hence was divided by f_g to obtain Plant Area Index (PAI) which was then used as input into for modelling intercepted radiation and wind attenuation processes. NDVI was also used to estimate fractional vegetation cover (f_C) as $f_C = \frac{NDVI - NDVI_S}{NDVI_V - NDVI_S}$ where $NDVI_S$ is the vegetation of bare soil and assigned a value of 0.15 and $NDVI_V$ is NDVI of full vegetation cover and assigned a value of 0.9 (Gutman and Ignatov, 1998).

3.3. Ancillary data

Apart from the data coming from satellite observations, the TSEB model requires two other spatially distributed datasets. The first one is a land-cover map which, together with a look-up table, is used to assign input surface parameters that cannot be directly retrieved from satellite observations, e.g. vegetation height or average leaf size. In this study the land-cover map was based on Corine Land Cover (CLC) 2012 version 18.5, downloaded from the Copernicus Land Monitoring Service (http://land.copernicus.eu, last accessed 21.08.2017). The map was reclassified from Corine classes to those shown in Fig. 1 and resampled from the original 100 m resolution to either 30 m resolution (using nearest neighbour approach) for the high-resolution model runs or to 930 m resolution (using mode approach) for the low-resolution runs. Afterwards both maps were used to assign input parameters based on Table 1. Vegetation height (h_C) has the largest influence on the modelled fluxes from all the parameters listed in Table 1. Therefore in land cover classes where it changes throughout the growing season (grasslands and croplands) it was scaled with PAI using power law, with maximum value indicated in Table 1 reached when PAI was 5 in croplands and 4 in grasslands and a minimum value set to 10% of the maximum value.

The second ancillary dataset (required for regional-scale runs) is the input meteorological data, which in this study are obtained from the ERA-Interim reanalysis data set (Dee et al., 2011) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). Three ERA-Interim products were used: the 2 m air temperature (2T), 2 m dew point temperature (2D) used to calculate the vapour pressure, and the 10 m horizontal wind speed (10U and 10V). Although the 2T product nominally represents the air temperature at 2 m height, in this study it was assumed to represent regional-scale blending-height air temperature (due to low spatial resolution and smoothness of the product) and therefore was placed at 100 m agl. To compensate for the height difference the 2T air temperature was reduced by 1 K, in accordance with dry adiabatic lapse rate, before being used as input to the models. The wind field was assumed to be at 10 m above the canopy. The meteorological data are provided at a 0.75° spatial resolution and were subset and resampled for the high and low-resolution model runs. In the temporal domain the data were linearly interpolated between the 3-hourly ERA-Interim time steps. The only meteorological dataset which came from field-based observations instead of the reanalysis dataset was the incoming shortwave radiation, based on the assumption that clear-sky shortwave radiation is uniform at regional scale and on the findings of Guzinski et al. (2014).

Table 2
Accuracy statistics of fluxes modelled with TSEB driven with tower-based meteorological inputs at VOU site. Model runs are explained the in Section 2.4. Root Mean Square Error (RMSE) and Bias (measured minus modelled values) are in W/m^2, coefficient of variation (CV - RSME divided by the mean of the measured values) and correlation (r) are unitless.

	H				*LE*				*Rn*				*G*			
	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r
$TSEB_M$	70	23	0.41	0.71	51	7	0.18	0.75	27	1	0.05	0.98	31	−15	0.55	0.54
$disTSEB_L$	38	−19	0.23	0.92	40	−6	0.14	0.92	28	−17	0.06	0.98	33	11	0.65	0.40
$TSEB_{M,DMS}$	64	5	0.38	0.70	44	−9	0.15	0.85	31	−16	0.06	0.97	31	−2	0.56	0.40
$disTSEB_{M,DMS}$	56	−5	0.33	0.76	42	0	0.15	0.87	31	−18	0.06	0.97	31	−1	0.56	0.40

4. Results

In this section we validate the high-resolution modelled fluxes against flux tower measurements at the VOU and GLU sties. In addition we evaluate the differences between the fluxes modelled with Landsat LST and sharpened LST in the spatial and temporal domains. This analysis focuses on the agricultural site since this is where high-resolution observations provide the most utility. Finally, we apply the proposed methodology to Sentinel-2 and Sentinel-3 data and compare with fluxes obtained using Landsat-8 LST at regional scale.

4.1. Voulund agricultural area

Table 2 summarizes the performance of the TSEB algorithm when run at Landsat resolution using tower-based meteorological inputs, with the scatterplot of modelled versus measured fluxes shown in Fig. 2. Similarly, Table 3 and in Fig. 3 show the results when tower-based meteorological measurements are replaced by ERA-Interim fields. Concurrent cloud-free observations by Landsat and Terra satellites over the VOU site were present on 19 dates during the study period. However, due to issues with tower-based flux measurements, the modelled *H*, *LE*, R_n and *G* could be compared to measurements on 19, 15, 16 and 18 dates respectively.

The results at the agricultural site indicate that although the fluxes derived with sharpened MODIS LST do not provide the same accuracy as fluxes derived with Landsat LST (especially in case of sensible heat flux), they still provide an improvement compared to MODIS coarse scale fluxes, in particular if the disaggregation approach is being used. The RMSE of *H* decreases by 12 to 14 W/m^2(around 20%) and correlation increases by 0.03 to 0.05 between the $TSEB_M$ and $disTSEB_{M,DMS}$ runs, depending on which meteorological inputs were used. A smaller decrease in RMSE (6 to 9 /m^2) and larger increase in correlation (0.13) is also present between those two runs in case of *LE*. This implies that by sharpening the coarse LST with finer resolution optical data it is possible, to a certain degree, to separate the fluxes originating from within the flux tower footprint from the fluxes falling outside of it. Sharpened coarse-resolution modelled fluxes (configuration TSEB-DMS_M) are also closer to the tower measurements compared to the original resolution coarse fluxes but the improvements are not as large as when sharpened model inputs are used. The ground heat flux modelled at high spatial resolution has lower correlation compared to the low resolution *G*, regardless of the model configuration used. This might indicate inaccurate estimation of net radiation of the soil at higher spatial resolution but due to the small magnitude of *G* it should not significantly impact the other fluxes.

It is also worth noting that at the VOU site there is an increase in modelled *H* error when switching from Landsat LST to sharpened LST, while there is a slight decrease in error when switching from tower-based meteorological inputs to ERA-Interim dataset. For example, the decrease in RMSE of *H* for the $disTSEB_{M,DMS}$ runs driven by the two types of meteorological inputs is 3 W/m^2(correlation decreases from 0.76 to 0.73) while the increase in RMSE between $disTSEB_L$ and $disTSEB_{M,DMS}$ is 15 to 18 W/m^2(correlation decreases by 0.16). In case of latent heat flux the situation is somewhat different, with RMSE increasing by 16 W/m^2 and correlation decreasing by 0.14 when tower-based data is replaced by ERA-Interim for the $disTSEB_L$ run, while the RMSE of *LE* increases by 2 W/m^2 between the $disTSEB_L$ and $disTSEB_{M,DMS}$ runs when using tower-based dataset and decreases by 3 W/m^2 when using ERA-Interim datasets.

4.2. Gludset plantation forest

The statistical evaluation of modelled fluxes at the GLU site is shown in Table 4 and Fig. 4 for runs in which tower-based meteorological inputs were used and in Table 5 and Fig. 5 for runs in which ERA-Interim meteorological inputs were used.

At the GLU site there is a strong and consistent underestimation of *Rn* (bias ranging from −24 W/m^2 to −51 W/m^2) despite very strong correlation (r always above 0.97) for all model runs. Based on this high correlation, the underestimation can probably be attributed to consistent overestimation in the parameterization of reflectance (albedo) and/or transmittance of the short-wave radiation by canopy or soil. This bias seems to mostly transfer to the estimates of *H* in case of tower-based meteorological inputs and estimates of *LE* in case of ERA-Interim inputs. In addition the estimation of *G* has very low correlation, but due to low *G* values under the forest canopy that should not have a significant impact on the accuracy of other modelled fluxes.

At the forest site, the RMSE of *H* and *LE* from the $TSEB_M$ run is lower than that of $disTSEB_L$ and $disTSEB_{M,DMS}$ runs with both types of meteorological inputs, with the exception of high-resolution *H* estimated with ERA-Interim dataset which has lower error than the low-resolution *H*. This could be because the forest site is homogeneous at low spatial resolution while becoming heterogeneous at higher resolutions, which is directly opposite of the agricultural site. In addition, forest canopy is more complex than that of crops, with possibility of within-canopy heat storage or different wind attenuation compared to a vertically homogeneous canopy (Massman et al., 2017). Those factors could be more influential when modelled at higher spatial resolutions. This, together with large shifts in bias between *H* and *LE* when using the two types of meteorological inputs, make it hard to come to any firm conclusions with respect to the applicability of the sharpened LST in forested land covers. However, similarly to the situation at VOU, $disTSEB_L$ and $disTSEB_{M,DMS}$ provide the most accurate high-resolution flux estimates.

4.3. Spatial pattern comparison

In the above sections the fluxes modelled by different configurations of model inputs and approaches were evaluated against measurements from two flux towers. Even though tower footprints were taken into account, those evaluations can be essentially considered to be performed at a single point. However, it is also important to evaluate how similar or different are the fluxes produced by the different model runs in a spatial context. In particular, we focus on fluxes produced by the $disTSEB_L$ and $disTSEB_{M,DMS}$ model runs driven by tower-based meteorological observations (see Table 2 and Fig. 2) in the area surrounding the VOU flux tower. Since $disTSEB_L$ uses LST derived from

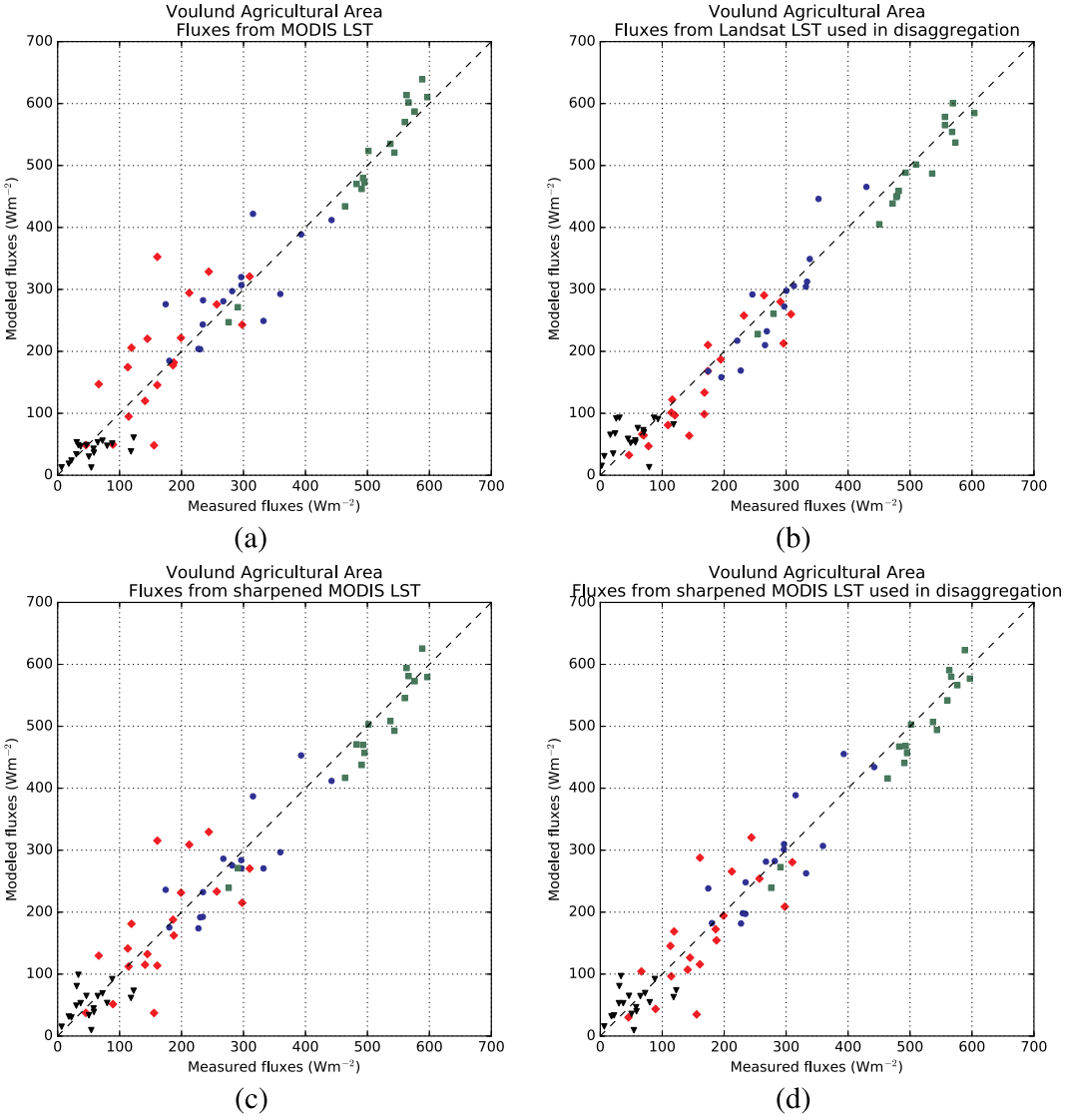

Fig. 2. Scatterplots of fluxes modelled with TSEB driven with tower-based meteorological inputs at VOU site. The fluxes are represented as follows: green squares - net radiation, blue circles - latent heat, red diamonds - sensible heat, black triangles - ground heat. Model runs are explained in Section 2.4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

actual thermal observation it is considered as the reference dataset. disTSEB$_L$ models fluxes at the time of Landsat overpass while disTSEB$_{M,DMS}$ at the time of MODIS overpass and therefore a direct comparison of the fluxes is not possible. Instead evaporative fraction is used, under the assumption that it remains constant during the morning hours in cloud free conditions.

Fig. 6 shows a visual comparison of EF on three selected dates, while Fig. 7 shows the scatterplot of disTSEB$_L$ versus disTSEB$_{M,DMS}$ EF on the same dates but only from pixels which were classified as belonging to cropland land cover class. The three dates were chosen such that on the first date (beginning of the growing season) the EF at VOU is high, in the second it is intermediate and in the third it is low (end of the growing season). Fig. 6 shows that in general the EF modelled with sharpened LST is able to capture the spatial patterns of the reference

Table 3
Accuracy statistics of fluxes modelled with TSEB driven with ERA-Interim meteorological inputs at VOU site. Model runs are explained in Section 2.4. Root Mean Square Error (RMSE) and Bias (measured minus modelled values) are in W/m^2, coefficient of variation (CV - RSME divided by the mean of the measured values) and correlation (r) are unitless.

	H				*LE*				*Rn*				*G*			
	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r
TSEB$_M$	65	32	0.39	0.70	59	−8	0.21	0.63	24	1	0.05	0.98	30	−15	0.55	0.55
disTSEB$_L$	38	−12	0.23	0.89	56	−13	0.20	0.78	25	−13	0.05	0.98	33	11	0.66	0.39
TSEB$_{M,DMS}$	60	11	0.35	0.66	58	−20	0.20	0.73	29	−16	0.06	0.98	31	−2	0.56	0.40
disTSEB$_{M,DMS}$	53	5	0.31	0.73	53	−8	0.19	0.76	27	−13	0.05	0.98	31	−1	0.56	0.41
TSEB-DMS$_M$	62	−2	0.36	0.69	53	−10	0.19	0.78	38	−24	0.07	0.97	28	−7	0.50	0.54

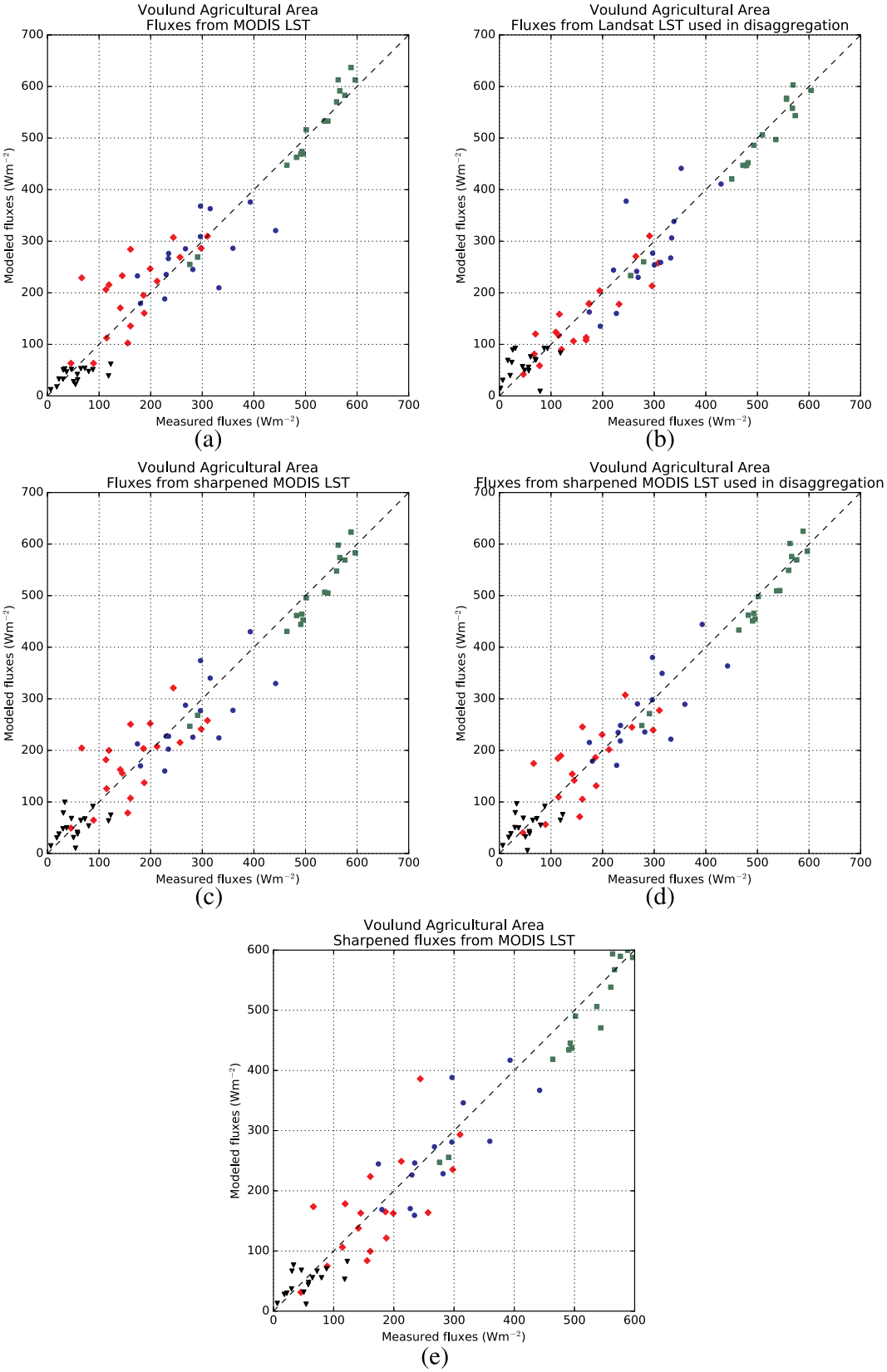

Fig. 3. Scatterplots of fluxes modelled with TSEB driven with ERA-Interim meteorological inputs at VOU site. The fluxes are represented as follows: green squares - net radiation, blue circles - latent heat, red diamonds - sensible heat, black triangles - ground heat. Model runs are explained in Section 2.4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Accuracy statistics of fluxes modelled with TSEB driven with tower-based meteorological inputs at GLU site. Model runs are explained in Section 2.4. Root Mean Square Error (RMSE) and Bias (measured minus modelled values) are in W/m^2, coefficient of variation (CV - RSME divided by the mean of the measured values) and correlation (r) are unitless.

	H				*LE*				*Rn*				*G*			
	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r
$TSEB_M$	69	−35	0.21	0.79	59	−8	0.20	0.50	35	−33	0.05	1.00	16	10	1.17	−0.04
$disTSEB_L$	83	−53	0.27	0.78	66	−14	0.22	0.53	32	−29	0.05	0.99	12	7	0.61	0.47
$TSEB_{M,DMS}$	112	−87	0.34	0.71	73	11	0.25	0.55	32	−30	0.05	1.00	16	11	1.20	0.10
$disTSEB_{M,DMS}$	88	−68	0.27	0.80	65	−8	0.22	0.51	33	−32	0.05	1.00	12	7	0.84	0.24

dataset, although with somewhat reduced range. This reduced range is evident at the beginning of the season (Figs. 6 top row and 7a) when the contrast of reflectance between the different fields is quite low due to limited amount of green crops present on the fields. In that case the higher values of LST (lower EF) are not well captured by the LST sharpening process. Similarly, in the middle of the season (Figs. 6 middle row and 7b), the contrast of reflectance between the different fields is also low (due to most fields being covered with green crops) and high EF values are not well represented. On the other hand, when the contrast in the reflectance between the fields is high (Figs. 6 bottom row and 7c) the variability of EF is very well captured by the sharpened LST. Therefore, one possible limitation of the sharpening method is that it might require contrasting surfaces at different vegetation densities.

Another contributing factor could be that the first two scenes were acquired by Thematic Mapper (TM) sensor on Landsat 5, while the last scene was acquired by Enhanced Thematic Mapper Plus (ETM+) on Landsat 7. This has two consequences: firstly ETM+ has increased radiometric sensitivity, accuracy and dynamic range over TM (Barsi

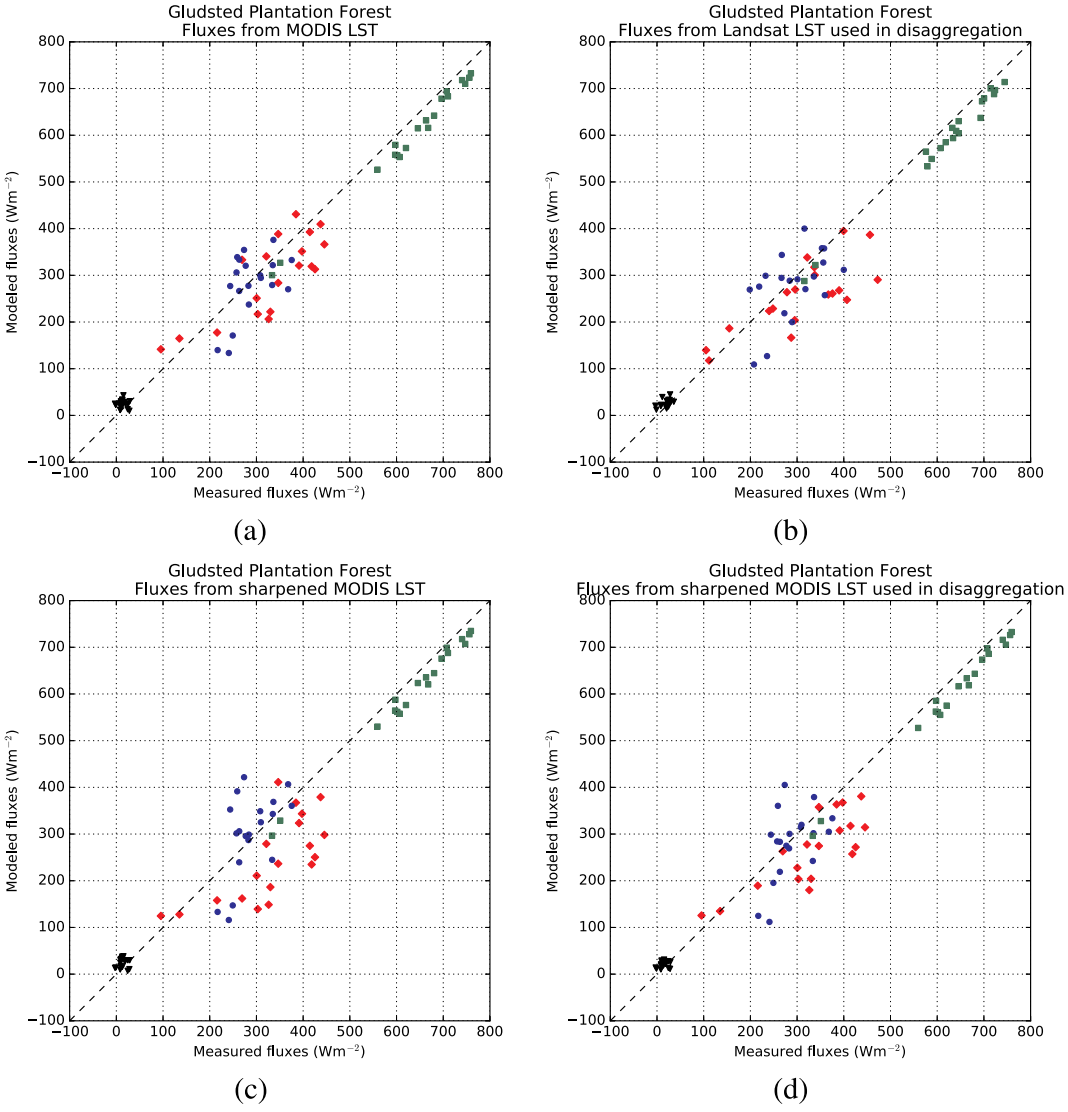

Fig. 4. Scatterplots of fluxes modelled with TSEB driven with tower-based meteorological inputs at GLU site. The fluxes are represented as follows: green squares - net radiation, blue circles - latent heat, red diamonds - sensible heat, black triangles - ground heat. Model runs are explained in Section 2.4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Accuracy statistics of fluxes modelled with TSEB driven with ERA-Interim meteorological inputs at GLU site. Model runs are explained in Section 2.4. Root Mean Square Error (RMSE) and Bias (measured minus modelled values) are in W/m^2, coefficient of variation (CV - RSME divided by the mean of the measured values) and correlation (r) are unitless.

	H				*LE*				*Rn*				*G*			
	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r	RMSE	Bias	CV	r
$TSEB_M$	68	27	0.21	0.81	86	−59	0.30	0.58	32	−30	0.05	1.00	11	2	0.77	−0.13
$disTSEB_L$	62	1	0.20	0.81	88	−58	0.30	0.59	26	−24	0.04	1.00	10	0	0.52	0.30
$TSEB_{M,DMS}$	74	−19	0.22	0.75	90	−48	0.31	0.62	29	−27	0.05	1.00	11	3	0.78	0.11
$disTSEB_{M,DMS}$	60	−4	0.18	0.81	94	−62	0.33	0.58	28	−26	0.04	1.00	9	1	0.67	0.21
$TSEB\text{-}DMS_M$	64	10	0.19	0.80	123	−97	0.43	0.53	60	−51	0.09	0.97	11	−3	0.82	−0.34

et al., 2007); and secondly due to orbit of Landsat 5 only two thirds of the catchment area shown in Fig. 1 are captured during an overpass which leads to less training points for the DMS. Both of those issues should be resolved when using Sentinel-2 optical observations to sharpen Sentinel-3 TIR data, hopefully leading to wider range of modelled ET: firstly Sentinel-2 optical sensor has 12 bit radiometric resolution (compared to 8 bit resolution for Landsat 5 and Landsat 7) and low radiometric noise; and secondly Sentinel-2 has wider swath width than Landsat satellites (Kääb et al., 2016).

One other difference observable between the EF derived by $disTSEB_L$ and $disTSEB_{M,DMS}$ is the sharpness of the latter (Fig. 6). This is due to the Landsat LST being resampled to 30 m from the acquisition resolution of between 60 m and 120 m (depending on the Landsat satellite), while the MODIS LST is sharpened using the Landsat reflectance observations which are acquired at 30 m resolution. When sharpening Sentinel-3 TIR data with Sentinel-2 optical data an even sharper (20 m resolution) image should be obtained.

4.4. Temporal pattern comparison

One of the main reasons for obtaining high-resolution estimates of land surface energy fluxes is to be able to follow the temporal evolution of agricultural water use and plant water stress at field scale. Therefore, it is important to evaluate how accurately the fluxes estimated with sharpened LST can follow that temporal evolution. Similarly to Section 4.3 we use EF for this evaluation and consider the fluxes derived with Landsat LST as the reference dataset. For both types of LST inputs the TSEB model was driven by tower-based meteorological observations and the disaggregation approach was used.

Fig. 8 shows the temporal evolution of EF for two agricultural fields throughout the study period. The relative location and size of the two fields (labelled (a) and (b)) are shown in the left panel of Fig. 9. It can be seen that both fields are much smaller than the spatial resolution of MODIS and SLSTR thermal observations (around 1 km) and that the distance between them is on the order of one low-resolution LST pixel. It can also be seen in Fig. 9 that the two fields do not always follow the same cultivation calendar. Despite this, EF modelled with sharpened LST is able to follow quite closely the temporal pattern of the reference EF, with r of 0.87 and Root Mean Square Difference (RMSD) of 0.08 for field (a) and r of 0.83 and RMSD of 0.09 for field (b).

However, upon closer inspection some deficiencies of the sharpened LST approach become evident. Fig. 9 shows that while in the 2013 growing season field (a) remained essentially bare until June (with crop cover visible only in the image from 21.07.2013), field (b) was sown with winter crop the previous season and so was covered with green vegetation from May through June and then being already harvested for the image on July. The EF shown on the right end of Fig. 8 reflects this temporal sequence in case of model runs using Landsat LST ($disTSEB_L$ - red crosses) but not necessarily when using sharpened LST ($disTSEB_{M,DMS}$ - green x'es). In field (a) $disTSEB_L$ EF starts around 0.6 at the beginning of May 2013 then drops to 0.4 and gradually increases to 0.5 as the vegetation grows, while in field (b) it remains around 0.6 throughout May and June 2013 as the green vegetation matures and drops to around 0.4 at the end of July when the field does not contain photosynthetic vegetation anymore. $disTSEB_{M,DMS}$ EF follows the reference EF quite closely in field (a), with the exception of the estimate at end of May 2013 when it is overestimated, but in field (b) drops to 0.4 at the end of May and remains constant during the following two modelled dates. Especially in field (b) the behaviour of $disTSEB_L$ EF appears more realistic, which might indicate that the sharpened LST is not able to fully capture the dynamics of this field.

4.5. Comparison of fluxes derived with sharpened Sentinel-3 LST and Landsat LST at regional scale

The suitability of applying the proposed methodology for deriving high-resolution fluxes with Sentinel-3 and Sentinel-2 images was also evaluated. The scenes were acquired on the morning of 17th of May 2017 and cover an agricultural area in the Po Valley in northern Italy. A Landsat-8 (L8) image was acquired over the area on the same morning, allowing for comparison of high-resolution fluxes estimated using S3 LST sharpened with S2 optical data against fluxes estimated using L8 LST in a region of interest (ROI) of 83 km by 40 km (Fig. 10).

Land surface energy fluxes were first derived at the Sentinel-3 resolution using the TSEB model and then at the Sentinel-2 resolution using the TSEB model and the disaggregation approach. The S3 LST was obtained following the split-window approach proposed by Sobrino et al. (2016b). For high-resolution application the LST was sharpened to 20 m using the DMS approach and all 12 S2 bands, atmospherically corrected using Sen2Cor tool (Louis et al., 2016) and (in case of bands with different spatial resolution) resampled to 20 m. LAI was retrieved from the atmospherically corrected Sentinel-2 reflectances using the biophysical processor available in the Sentinel Application Platform (SNAP) software and resampled to S3 resolution for the low-resolution run. Due to the lack of established S2 or S3 albedo product or algorithm, the albedo data was taken from a MODIS MCD43A3 product and resampled for the S3 resolution run or sharpened with DMS for the S2 resolution run. The meteorological inputs were taken from the ECMWF analysis data which is included in the Sentinel-3 SLSTR data file, with the exception of incoming shortwave radiation which was estimated using the approach of Weiss and Norman (1985). The landcover-based parameters were set according to reclassified Corine map, as described in Section 3.3. To derive the reference high-resolution fluxes the same inputs were used except DMS-sharpened S3 LST was replaced with L8 LST.

Maps of latent heat flux derived using Sentinel-3 LST, DMS-sharpened Sentinel-3 LST and Landsat-8 LST of three zoomed-in subsets of the ROI are shown in Fig. 11. It can be seen that at the original S3 resolution it is not possible to distinguish fluxes coming from individual fields or natural landscape features (e.g. riparian zones). However, at the S2 resolution those features become quite apparent both when using sharpened S3 LST and L8 LST. The spatial patterns shown in the two high-resolution latent heat flux maps are quite similar, although the fluxes estimated using L8 LST show larger dynamic range.

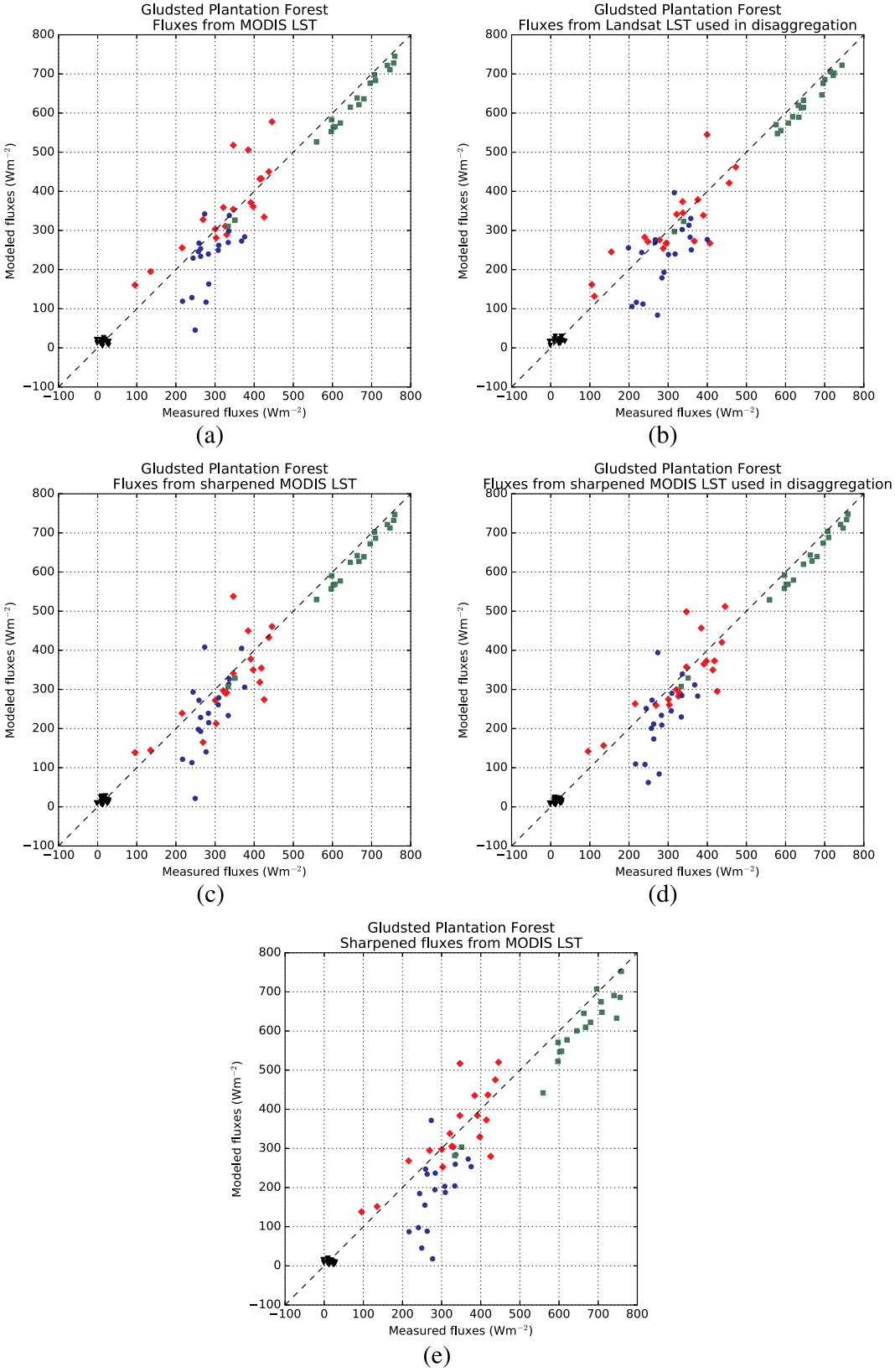

Fig. 5. Scatterplots of fluxes modelled with TSEB driven with ERA-Interim meteorological inputs at GLU site. The fluxes are represented as follows: green squares - net radiation, blue circles - latent heat, red diamonds - sensible heat, black triangles - ground heat. Model runs are explained in Section 2.4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

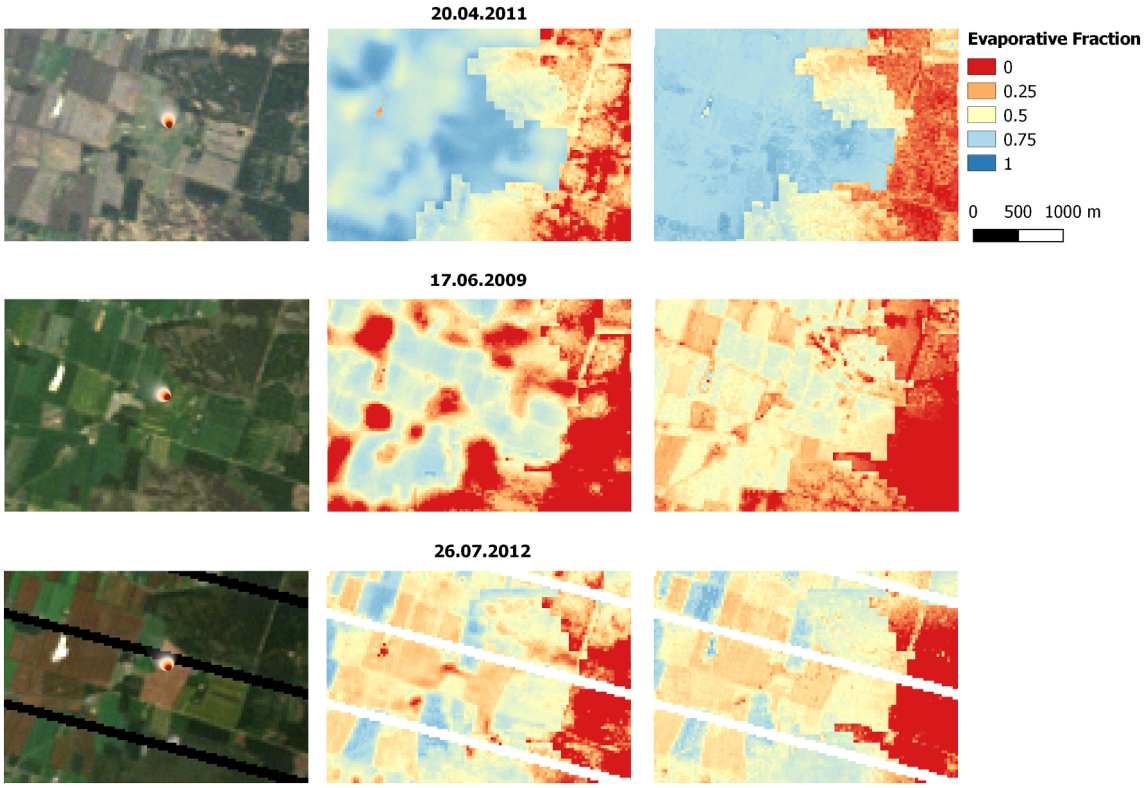

Fig. 6. Modelled evaporative fraction (EF) of the area surrounding VOU field on three dates: 20.04.2011 (beginning of growing season - high EF), 17.06.2009 (middle of growing season - medium EF) and 26.07.2012 (end of growing season - low EF). Left column shows Landsat true colour composite of the area with location of the flux tower and its footprint indicated by red-orange bubble. Middle column shows EF modelled using Landsat LST and TSEB with disaggregation. Right column shows EF modelled using sharpened MODIS LST and TSEB with disaggregation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Density scatterplots of latent and sensible heat fluxes and the evaporative fraction for all the agricultural pixels within the ROI (over 7 million points) are shown in Fig. 12. The plot and statistics for latent heat flux confirm the findings described above: the correlation between *LE* modelled with the two types of LST is high (0.89) and using DMS sharpened S3 LST leads to underestimation of high-value fluxes and overestimation of low-value fluxes. The correlation between the sensible heat fluxes is significantly lower (0.67) which leads to the correlation between EF estimates to lie between the two values (0.79). It is not immediately clear why the correlation of *H* is so much lower than that of the other fluxes (correlations of *G* and *Rn* are 0.99 and 0.90 respectively). However, the response of the TSEB model to the changes in input LST is non-linear and this contributes to the different correlations of the different fluxes.

5. Discussion

5.1. Potential for applicability in operational setting

High resolution estimates of evapotranspiration have many operational uses, in particular in the domains of agriculture and water resources management (Anderson et al., 2012). For example, they allow for field-scale water accounting or precision irrigation. In Section 4 we have shown that by using 1 km LST sharpened to 30 m it is possible to obtain instantaneous ET estimates at field scale with a relative RMSE of 15% (when using local meteorological observations) to 19% (when using reanalysis meteorological fields). The bias of the modelled ET is significantly smaller and those errors should reduce when instantaneous outputs are averaged to daily, weekly or monthly values. This level of accuracy is similar to what is achieved with field measurements and therefore the estimates can be considered as reliable for water accounting (Karimi et al., 2015). It was also demonstrated that fluxes derived using sharpened low-resolution (Terra or Sentinel-3) LST capture the spatial patterns of fluxes derived using high-resolution (Landsat) LST well but with a narrower dynamic range if the optimal conditions for sharpening are not met. The temporal patterns shown by the two types of modelled fluxes are also similar. This confirms the applicability of the proposed method to be used with Sentinel-2 and Sentinel-3 images to obtain high-resolution evapotranspiration estimates.

At both sites the fluxes derived from sharpening the low-resolution estimates (TSEB-DMS$_M$) were not as accurate as the fluxes derived with sharpened LST (disTSEB$_{M,DMS}$), with the exception of LE at the agricultural site where the RMSE of the two runs was the same. This indicates that it is preferable to sharpen the low-resolution data (in case of this study LAI, LST and albedo) and use them as inputs to the TSEB model, rather than sharpening the TSEB model outputs using DMS to achieve higher spatial resolution flux estimates. In case of applications utilising S2 data, LAI (and possibly albedo) can be derived directly at fine resolution (see Section 4.5) thus improving the high-resolution model parametrization. Furthermore, sharpening the fluxes directly may cause a lack of energy closure at fine scale that needs to be taken into account.

In the current study reanalysis meteorological dataset (ERA-Interim) was used. These data are usually released with up to 3 months delay, which is fine in case of end-of-season water accounting but not suitable for irrigation advice. In that case another meteorological dataset, e.g. ECMWF forecast products (Andersson, 2015) would have to be used. This could have a significant impact on the modelled flux accuracy, especially if the forecast air temperature or wind speed are not well captured and this should be evaluated in future studies. In case of regional or local applications, it is possible to use near-real-time

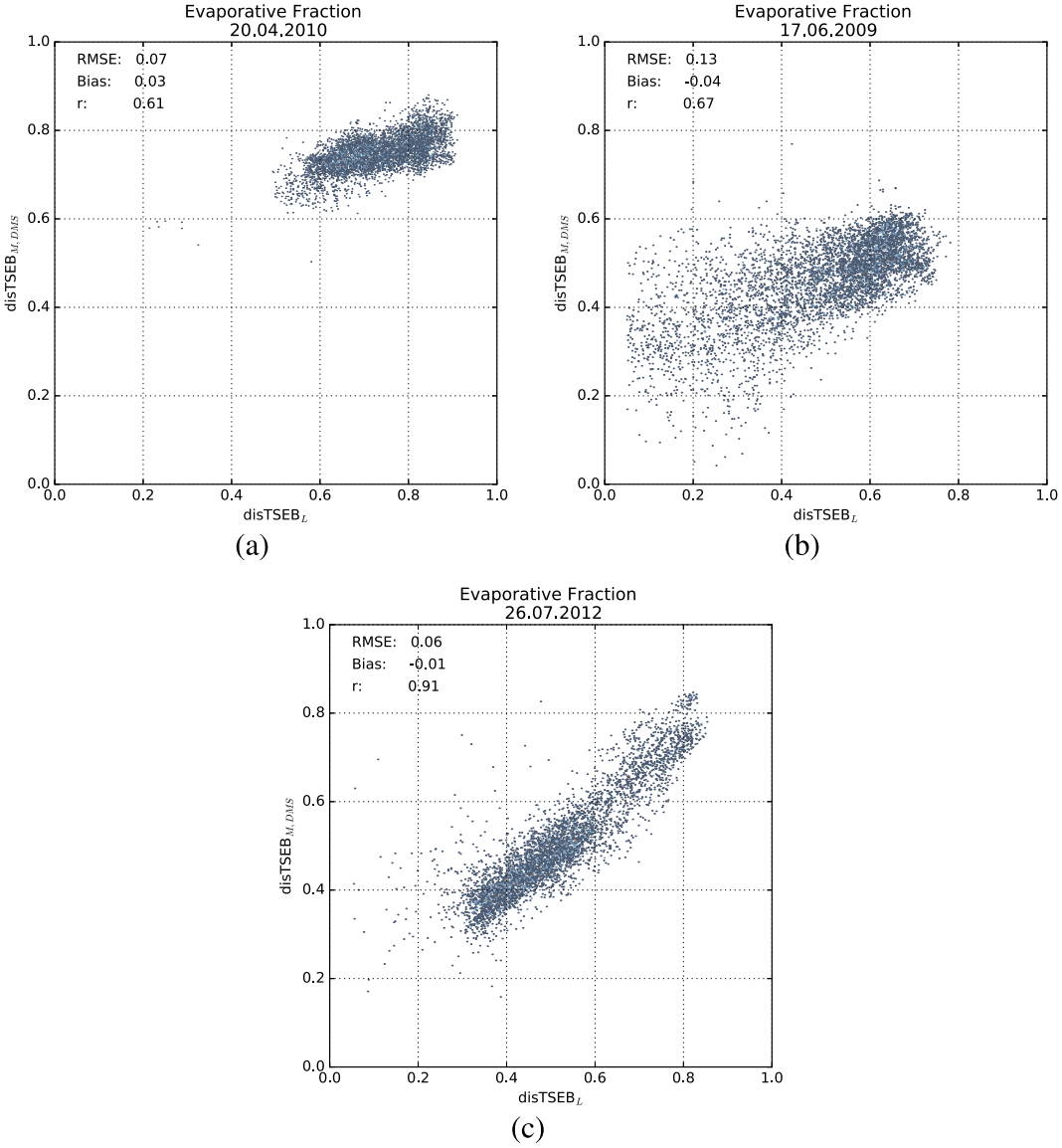

Fig. 7. Scatterplot of modelled evaporative fraction (EF) of the croplands surrounding VOU field on three dates: 20.04.2011, 17.06.2009 and 26.07.2012. EF modelled using Landsat LST and TSEB with disaggregation (x-axis, disTSEB$_L$) was plotted against EF modelled using sharpened MODIS LST and TSEB with disaggregation (y-axis, disTSEB$_{M,DMS}$). The plot includes all pixels from the area shown in Fig. 6 which fall into cropland landcover class.

observations from ground stations and thus avoid this issue.

5.2. Limitations of the proposed approach

The approach to sharpen low-resolution LST with high-resolution optical data is based on the assumption that there exists a statistical relation between the two dataset. The sharpening methodology used in this study (see Section 2.1) permits this relation to be non-linear and even very complex but it still must exist. While there is no general, globally transferable relation between TIR and optical data, it is possible to derive it when the two datasets are limited in space and time. This study shows that in such localised cases it is possible to use the sharpened LST to derive the land surface fluxes at high spatial resolution with acceptable accuracy. However, even then the range of high-resolution fluxes will not be as large as it would be if high-resolution LST observations were used, due to the aggregating effect present in the low-resolution observations.

Another inherent limitation of estimating ET using DMS sharpened LST is the underlying assumption that the TIR and optical datasets contain the same or related information about the modelled physical phenomena. In practice they might contain complementary information. For example, crop water stress can be observed in the TIR domain some weeks before it becomes visible in the optical domain but both types of observations are required to map the water stress in different stages of the crop phenological cycle (Hoffmann et al., 2016a). On the other hand, Sentinel-2 bands in the shortwave infrared (SWIR) part of the electromagnetic spectrum do contain information about vegetation water content (Ceccato et al., 2001) and (under certain assumptions) about soil water content (Sadeghi et al., 2017). Therefore, more research is required to determine whether the DMS sharpened LST is suitable for applications such as early detection of vegetation water stress and precision irrigation, in which crops are subject to certain stress without provoking a decrease in yield/biomass. In other applications, such as in regional field-scale water use mapping (e.g. Senay et al. (2016)) the use of sharpened LST as input to ET models should produce acceptable results.

The above mentioned limitations, together with results presented in Section 4, point to the need for direct high-resolution TIR observations to complement the existing capabilities of the Sentinel satellite constellation. While the methodology proposed in this study is valuable for

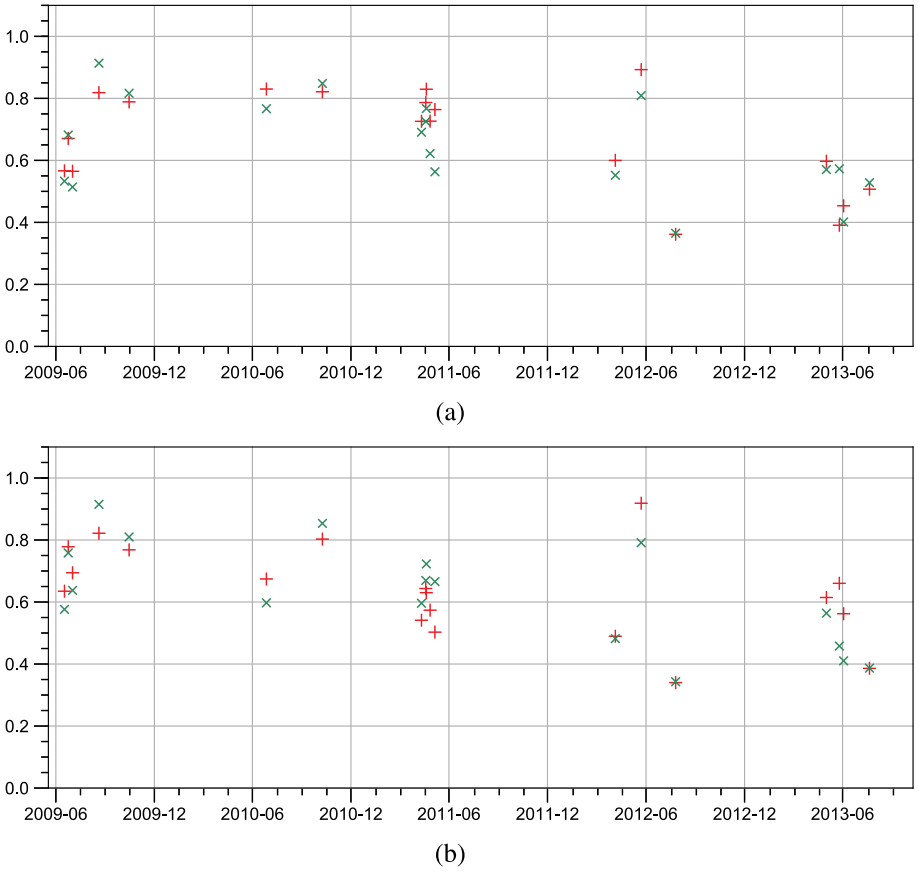

Fig. 8. Timeseries of modelled evaporative fraction (EF) of two crop fields in the vicinity of VOU tower: (a) the crop field containing the flux tower; (b) another field. The relative location of the fields is shown in Fig. 9. EF modelled using Landsat LST and TSEB with disaggregation is indicated with red crosses and EF modelled using sharpened MODIS LST and TSEB with disaggregation is indicated with green x'es. In both cases tower-based meteorological inputs were used. The figure includes the same 19 dates as in Fig. 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

filling in the currently present data gap, it does so with a reduced accuracy compared to what would be possible with a high-resolution TIR sensor and might not be suitable for all TIR-based applications. Further research into different types of sharpening procedures and ET models will probably improve on the results presented in this study but it is highly unlikely that the need for high-resolution evapotranspiration estimations will ever be fully met without a dedicated high-resolution TIR sensor.

6. Conclusion

The Sentinel constellation of satellites provides a suit of observational capabilities for operational use in agricultural applications. However, one currently missing feature is high-resolution TIR sensor which could be used to (among other applications) estimate evapotranspiration and water use at field scale. This study evaluates a method to fill this data gap by sharpening low-resolution TIR observations from Sentinel-3 satellites using high-resolution optical observations from Sentinel-2 satellites. The evaluation is performed at two study sites, both located in a temperate climatic zone: an agricultural site with field sizes smaller than low-resolution pixel size and plantation forest site larger than low-resolution pixel size. Due to lack of long enough timeseries of Sentinel-2 and Sentinel-3 observations and due to the need for high-resolution TIR data to use for benchmarking, the method was evaluated using MODIS and Landsat observations and later applied to Sentinel-3 and Sentinel-2 images to confirm its applicability to this type of data.

The results at the agricultural site indicate that although the fluxes derived using sharpened LST are not as accurate as those derived using high-resolution LST, they are still an improvement compared to low-resolution fluxes. For most accurate results the fluxes derived using sharpened LST should be refined with a disaggregation scheme which utilises low-resolution fluxes. Sharpening the low-resolution fluxes directly did not produce as good results. The fluxes derived with sharpened LST are able to satisfactorily capture the spatial patterns of different evapotranspiration rates of the different small-sized fields

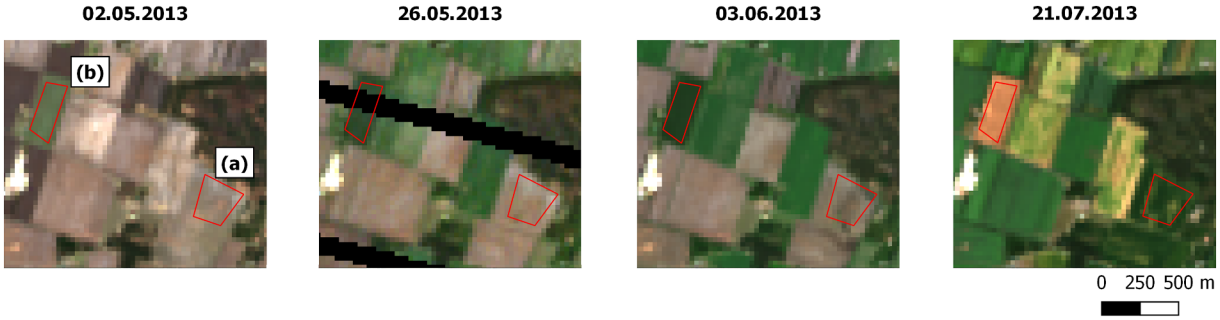

Fig. 9. True-colour composite Landsat images of the two fields used in Fig. 8 during the growing season of 2013. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

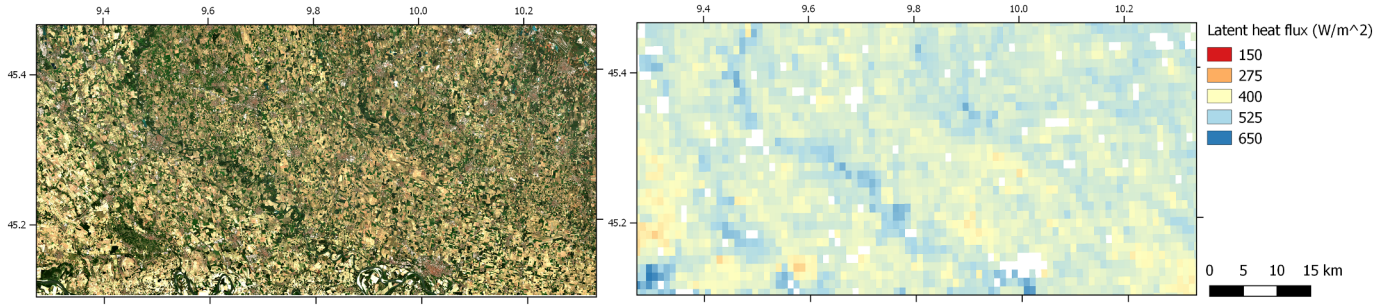

Fig. 10. True-colour composite Sentinel-2 image of the region of interest in the Po Valley (left panel) and instantaneous latent heat flux at Sentinel-3 pixels resolution of the same area derived using predominantly data from Sentinel-2 and Sentinel-3 satellites and acquired on the morning of the 17th of May 2017 (right panel). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(although with reduced range), especially when high contrast is present within the scene. The temporal evolution of evaporative fraction throughout multiple growing seasons is also generally well represented when using sharpened LST even though certain critical events are missed. At the forest site the utility of sharpened LST is harder to evaluate due to the nature of the site (homogeneous at coarse scale but heterogeneous at fine scale) and the difficulties in modelling ET in forest ecosystems.

Follow on studies are planned, which will utilise purely Sentinel-2 and Sentinel-3 observations and which will be conducted in irrigated croplands located in semi-arid and Mediterranean climates, where high LST contrast should be present at both low and high spatial resolutions and where water conservation issues are most critical. In addition, different LST sharpening techniques and ET modelling schemes will be tested.

This study demonstrates that although modern data fusion techniques can be employed to partially compensate for the current lack of high-resolution TIR observations from Sentinel satellites, they cannot fully cover this data gap. Therefore, Sentinel constellation should be complemented with TIR observations at high spatial (and temporal) resolution to optimally exploit the existing capabilities in agricultural applications.

Software access

The TSEB Python code, as used in this study, is accessible from https://github.com/hectornieto/pyTSEB/tree/v1.4. The latest code is accessible from https://github.com/hectornieto/pyTSEB (last accessed 09.10.2017).

The DMS Python code, as used in this study, is accessible from https://github.com/radosuav/pyDMS/tree/v1.0. The latest code is accessible from https://github.com/radosuav/pyDMS (last accessed 09.10.2017).

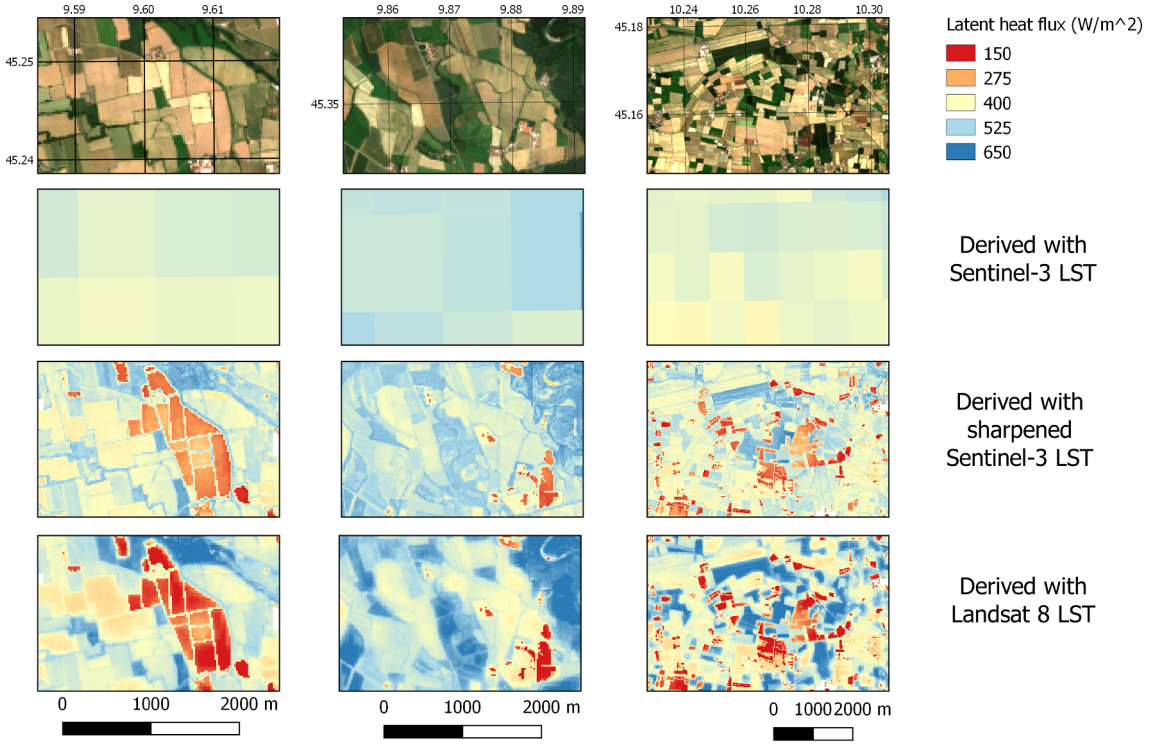

Fig. 11. Zooms from within the area shown in Fig. 10 showing Sentinel-2 true colour composites (top row) and instantaneous latent heat fluxes derived using Sentinel-3 LST, DMS sharpened Sentinel-3 LST and Landsat-8 LST using images acquired on the morning of the 17th of May 2017. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

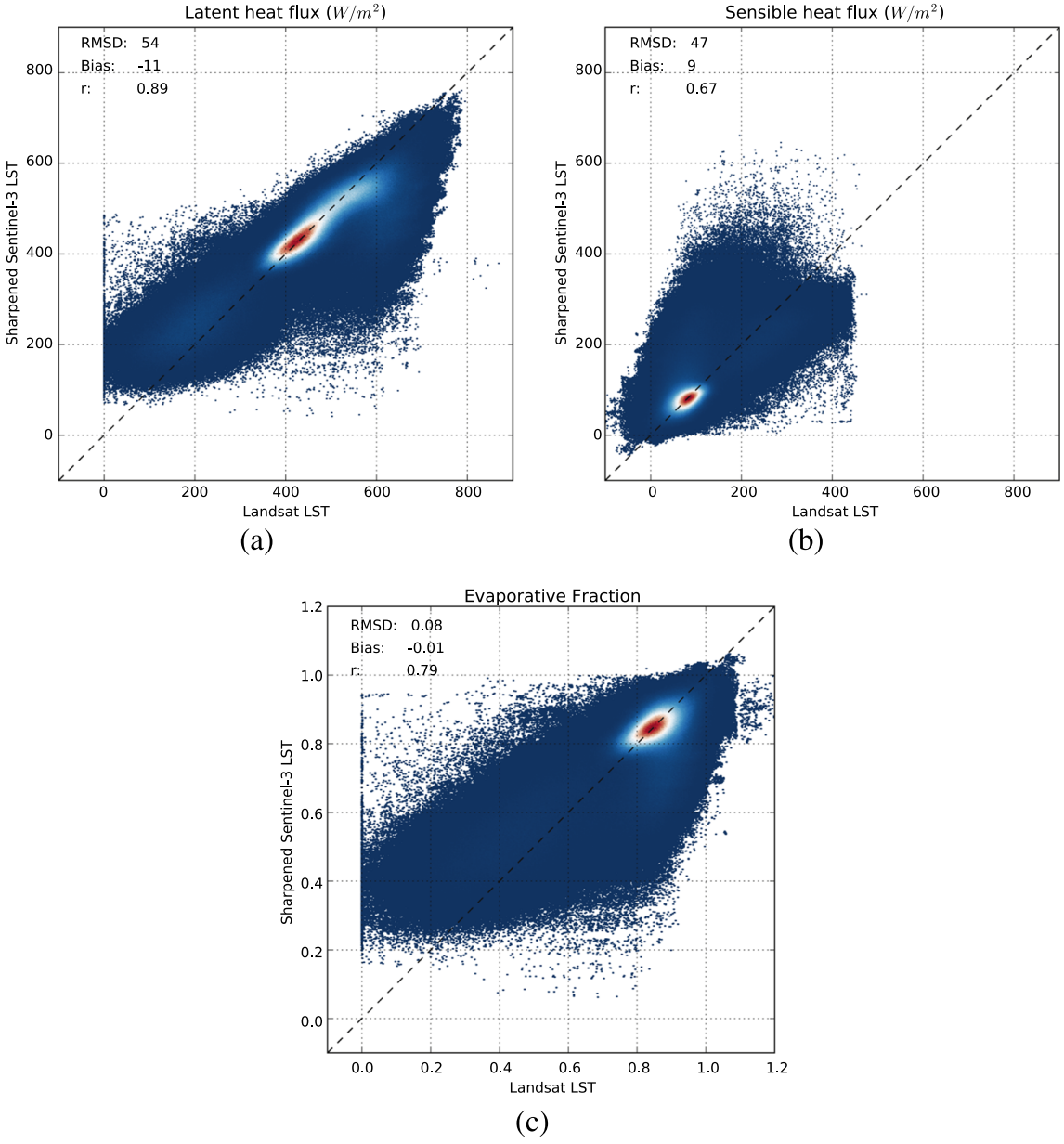

Fig. 12. Density scatterplots of modelled latent and sensible heat fluxes and evaporative fraction of the croplands in the Po Valley region of interest shown in Fig. 10. The fluxes were derived using Landsat-8 LST (x-axis) and DMS sharpened Sentinel-3 LST (y-axis) using images acquired on the morning of 17.05.2017. Red colour indicates higher density of points, blue colour lower density. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Acknowledgments

We would like to thank HOBE observatory, and in particular the personnel maintaining the flux towers and pre-processing the data, for providing the field measurements dataset.

pyTSEB was developed partly with funds under the Marie Curie COFUND Talentia Program N-267226 (TAPOST-203).

References

Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. Jul. Remote Sens. Environ. 122, 50–65. http://www.sciencedirect.com/science/article/pii/S0034425712000326.

Andersson, E., 2015 Nov. User guide to ECMWF forecast products, v1.2. https://www.ecmwf.int/files/user-guide-ecmwf-forecast-products.

Baldridge, A., Hook, S., Grove, C., Rivera, G., 2009. The aster spectral library version 2.0. Remote Sens. Environ. 113 (4), 711–715. http://www.sciencedirect.com/science/article/pii/S0034425708003441.

Barsi, J., Markham, B., Helder, D., Chander, G., 2007. Radiometric calibration status of Landsat 7 and Landsat 5. In: Proc. SPIE. 6744. pp. 67441F.

Bindhu, V.M., Narasimhan, B., Sudheer, K.P., 2013. Development and verification of a non-linear disaggregation method (NL-DisTrad) to downscale MODIS land surface temperature to the spatial scale of Landsat thermal data to estimate evapotranspiration. Aug. Remote Sens. Environ. 135, 118–129. http://www.sciencedirect.com/science/article/pii/S0034425713001028.

Bisquert, M., Sánchez, J.M., Caselles, V., 2016. Evaluation of disaggregation methods for downscaling MODIS land surface temperature to Landsat spatial resolution in Barrax test site. Apr. IEEE J. Selected Topics Appl. Earth Obs. Remote Sens. 9 (4), 1430–1438.

Breiman, L., 1996. Bagging predictors. Aug. Mach. Learn. 24 (2), 123–140. https://link.springer.com/article/10.1007/BF00058655.

Campos-Taberner, M., García-Haro, F.J., Camps-Valls, G., Grau-Muedra, G., Nutini, F., Busetto, L., Katsantonis, D., Stavrakoudis, D., Minakou, C., Gatti, L., Barbieri, M., Holecz, F., Stroppiana, D., Boschetti, M., 2017. Exploitation of SAR and optical sentinel data to detect rice crop and estimate seasonal dynamics of leaf area index. Remote Sensing 9 (3), 248. http://www.mdpi.com/2072-4292/9/3/248.

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., Grégoire, J.-M., 2001. Detecting vegetation leaf water content using reflectance in the optical domain. Jul. Remote

Sens. Environ. 77 (1), 22–33. http://www.sciencedirect.com/science/article/pii/S0034425701001912.
Chen, X., Li, W., Chen, J., Rao, Y., Yamaguchi, Y., 2014. A combination of TsHARP and thin plate spline interpolation for spatial sharpening of thermal imagery. Mar. Remote Sens. 6 (4), 2845–2863. http://www.mdpi.com/2072-4292/6/4/2845.
Clevers, J.G.P.W., Kooistra, L., van den Brande, M.M.M., 2017. Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. Apr. Remote Sens. 9 (5), 405. http://www.mdpi.com/2072-4292/9/5/405.
Crebassol, P., Lgouarde, J.-P., Hook, S., 2014, Jul, Jul. Thirsty thermal infrared spatial system. In: Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International, pp. 3021–3024.
Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K, Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Apr. Q. J. R. Meteorol. Soc. 137 (656), 553–597. http://onlinelibrary.wiley.com/doi/10.1002/qj.828/abstract.
Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. May. Remote Sens. Environ. 120, 37–57. http://www.sciencedirect.com/science/article/pii/S0034425712000685.
Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. May. Remote Sens. Environ. 120, 25–36. http://www.sciencedirect.com/science/article/pii/S0034425712000636.
Fisher, J.B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M.F., Hook, S., Baldocchi, D., Townsend, P.A., Kilic, A., Tu, K., Miralles, D.D., Perret, J., Lagouarde, J.-P., Waliser, D., Purdy, A.J., French, A., Schimel, D., Famiglietti, J.S., Stephens, G., Wood, E.F., 2017. The future of evapotranspiration: global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Apr. Water Resour. Res. 53 (4), 2618–2626. http://onlinelibrary.wiley.com/doi/10.1002/2016WR020175/abstract.
Gao, F., Kustas, W.P., Anderson, M.C., 2012. A data mining approach for sharpening thermal satellite imagery over land. Oct. Remote Sens. 4 (11), 3287–3319. http://www.mdpi.com/2072-4292/4/11/3287.
Gutman, G., Ignatov, A., 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. Jan. Int. J. Remote Sens. 19 (8), 1533–1543. http://www.tandfonline.com/doi/abs/10.1080/014311698215333.
Guzinski, R., Anderson, M.C., Kustas, W.P., Nieto, H., Sandholt, I., 2013. Using a thermal-based two source energy balance model with time-differencing to estimate surface energy fluxes with day-night MODIS observations. Jul. Hydrol. Earth Syst. Sci. 17 (7), 2809–2825. http://www.hydrol-earth-syst-sci.net/17/2809/2013/.
Guzinski, R., Nieto, H., Jensen, R., Mendiguren, G., 2014. Remotely sensed land-surface energy fluxes at sub-field scale in heterogeneous agricultural landscape and coniferous plantation. Sep. Biogeosciences 11 (18), 5021–5046. http://www.biogeosciences.net/11/5021/2014/.
Hoffmann, H., Jensen, R., Thomsen, A., Nieto, H., Rasmussen, J., Friborg, T., 2016a. Crop water stress maps for an entire growing season from visible and thermal UAV imagery. Dec. Biogeosciences 13 (24), 6545–6563. http://www.biogeosciences.net/13/6545/2016/.
Hoffmann, H., Nieto, H., Jensen, R., Guzinski, R., Zarco-Tejada, P., Friborg, T., 2016b. Estimating evaporation with thermal UAV data and two-source energy balance models. Feb. Hydrol. Earth Syst. Sci. 20 (2), 697–713. http://www.hydrol-earth-syst-sci.net/20/697/2016/.
Hulley, G., Hook, S., Fisher, J., Lee, C., 2017. Ecostress, a NASA earth-ventures instrument for studying links between the water cycle and plant health over the diurnal cycle. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5494–5496 July.
Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., Defourny, P., Koetz, B., 2015. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. Remote Sens. 7 (9), 12356–12379. http://www.mdpi.com/2072-4292/7/9/12356.
Jensen, K.H., Illangasekare, T.H., 2011. HOBE: a hydrological observatory. Feb. Vadose Zone J. 10 (1), 1–7. https://dl.sciencesocieties.org/publications/vzj/abstracts/10/1/1.
Kääb, A., Winsvold, S.H., Altena, B., Nuth, C., Nagler, T., Wuite, J., 2016. Glacier remote sensing using Sentinel-2. Part I: radiometric and geometric performance, and application to ice velocity. Jul. Remote Sens. 8 (7). http://www.mdpi.com/2072-4292/8/7/598.
Karimi, P., Bastiaanssen, W.G.M., Sood, A., Hoogeveen, J., Peiser, L., Bastidas-Obando, E., Dost, R.J., 2015. Spatial evapotranspiration, rainfall and land use data in water accounting - part 2: reliability of water acounting results for policy decisions in the Awash Basin. Hydrol. Earth Syst. Sci. 19 (1), 533–550. https://www.hydrol-earth-syst-sci.net/19/533/2015/.
Kustas, W.P., Li, F., Jackson, T.J., Prueger, J.H., MacPherson, J.I., Wolde, M., 2004. Effects of remote sensing pixel resolution on modeled energy flux variability of croplands in Iowa. Sep. Remote Sens. Environ. 92 (4), 535–547. http://www.sciencedirect.com/science/article/pii/S003442570400183X.
Kustas, W.P., Norman, J.M., 1999. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. Apr. Agric. For. Meteorol. 94 (1), 13–29. URL http://www.sciencedirect.com/science/article/pii/S0168192399000052.
Lagouarde, J.-P., Bach, M., Sobrino, J.A., Boulet, G., Briottet, X., Cherchali, S., Coudert, B., Dadou, I., Dedieu, G., Gamet, P., Hagolle, O., Jacob, F., Nerry, F., Olioso, A., Ottlé, C., Roujean, J.-l., Fargant, G., 2013. The MISTIGRI thermal infrared project: scientific objectives and mission specifications. May. Int. J. Remote Sens. 34 (9-10), 3437–3466. https://doi.org/10.1080/01431161.2012.716921.
Lee, C.M., Cable, M.L., Hook, S.J., Green, R.O., Ustin, S.L., Mandl, D.J., Middleton, E.M., 2015. An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. Sep. Remote Sens. Environ. 167, 6–19. http://www.sciencedirect.com/science/article/pii/S0034425715300419.
Li, F., Kustas, W.P., Prueger, J.H., Neale, C.M.U., Jackson, T.J., 2005. Utility of remote sensingbased two-source energy balance model under low- and high-vegetation cover conditions. J. Hydrometeorol. 6 (6), 878–891. https://doi.org/10.1175/JHM464.1.
Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. Sentinel-2 sen2cor: L2a processor for users. In: Ouwehand, L. (Ed.), ESA Living Planet Symposium 2016. ESA Special Publications (on CD) SP-740. Spacebooks Online, pp. 1–8. http://elib.dlr.de/107381/ August.
Lowder, S.K., Skoet, J., Raney, T., 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. Nov. World Dev. 87, 16–29. http://www.sciencedirect.com/science/article/pii/S0305750X15002703.
Massman, W., Forthofer, J., Finney, M., 2017. An improved canopy wind model for predicting wind adjustment factors and wildland fire behavior. Jan. Can. J. For. Res. 594–603. http://www.nrcresearchpress.com/doi/abs/10.1139/cjfr-2016-0354.
Morillas, L., Garca, M., Nieto, H., Villagarcia, L., Sandholt, I., Gonzalez-Dugo, M., Zarco-Tejada, P., Domingo, F., 2013. Using radiometric surface temperature for surface energy flux estimation in mediterranean drylands from a two-source perspective. Remote Sens. Environ. 136, 234–246. http://www.sciencedirect.com/science/article/pii/S0034425713001636.
Norman, J.M., Anderson, M.C., Kustas, W.P., French, A.N., Mecikalski, J., Torn, R., Diak, G.R., Schmugge, T.J., Tanner, B.C.W., 2003. Remote sensing of surface energy fluxes at 101-m pixel resolutions. Aug. Water Resour. Res. 39 (8), 1221. http://onlinelibrary.wiley.com/doi/10.1029/2002WR001775/abstract.
Norman, J.M., Kustas, W.P., Humes, K.S., 1995. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. Dec. Agric. For. Meteorol. 77 (34), 263–293. http://www.sciencedirect.com/science/article/pii/016819239502265Y.
Ringgaard, R., Herbst, M., Friborg, T., Schelde, K., Thomsen, A.G., Soegaard, H., 2011. Energy fluxes above three disparate surfaces in a temperate mesoscale coastal catchment. Feb. Vadose Zone J. 10 (1), 54–66. https://dl.sciencesocieties.org/publications/vzj/abstracts/10/1/54.
Sadeghi, M., Babaeian, E., Tuller, M., Jones, S.B., 2017. The optical trapezoid model: a novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. Sep. Remote Sens. Environ. 198, 52–68. http://www.sciencedirect.com/science/article/pii/S0034425717302493.
Santanello, J.A., Friedl, M.A., 2003. Diurnal covariation in soil heat flux and net radiation. Jun. J. Appl. Meteorol. 42 (6), 851–862. http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(2003)042%3C0851:DCISHF%3E2.0.CO%3B2.
Senay, G.B., Friedrichs, M., Singh, R.K., Velpuri, N.M., 2016. Evaluating Landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. Nov. Remote Sens. Environ. 185, 171–185. http://www.sciencedirect.com/science/article/pii/S0034425715302650.
Sobrino, J.A., Del Frate, F., Drusch, M., Jimenez-Munoz, J.C., Manunta, P., Regan, A., 2016a. Review of thermal infrared applications and requirements for future high-resolution sensors. IEEE Trans. Geosci. Remote Sens. 1–10. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7378483.
Sobrino, J.A., Jiménez-Muñoz, J.C., Sória, G., Ruescas, A.B., Danne, O., Brockmann, C., Ghent, D., Remedios, J., North, P., Merchant, C., Berger, M., Mathieu, P.P., Göttsche, F.M., 2016, Jun, Jun. Synergistic use of MERIS and AATSR as a proxy for estimating land surface temperature from Sentinel-3 data. Remote Sens. Environ. 179, 149–161. http://www.sciencedirect.com/science/article/pii/S0034425716301158.
Torbick, N., Chowdhury, D., Salas, W., Qi, J., 2017. Monitoring rice agriculture across Myanmar using time series Sentinel-1 assisted by Landsat-8 and PALSAR-2. Remote Sens. 9 (2), 119. http://www.mdpi.com/2072-4292/9/2/119.
Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24. http://www.sciencedirect.com/science/article/pii/S0034425712000600.
Valero, S., Morin, D., Inglada, J., Sepulcre, G., Arias, M., Hagolle, O., Dedieu, G., Bontemps, S., Defourny, P., Koetz, B., 2016. Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. Remote Sens. 8 (1), 55. http://www.mdpi.com/2072-4292/8/1/55.
Wan, Z., 2014, Jan, Jan. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45. http://www.sciencedirect.com/science/article/pii/S003442571300285X.
Weiss, A., Norman, J.M., 1985. Partitioning solar radiation into direct and diffuse, visible and near-infrared components. Apr. Agric. For. Meteorol. 34 (2), 205–213. http://www.sciencedirect.com/science/article/pii/0168192385900206.
Yan, K., Park, T., Yan, G., Chen, C., Yang, B., Liu, Z., Nemani, R.R., Knyazikhin, Y., Myneni, R.B., 2016. Evaluation of MODIS LAI/FPAR product collection 6. Part 1: consistency and improvements. Apr. Remote Sens. 8 (5), 359.