Hydrology and
Earth System
Sciences

Inter-comparison of energy balance and hydrological models for land surface energy flux estimation over a whole river catchment

R. Guzinski[1,2], **H. Nieto**[1,3], **S. Stisen**[4], **and R. Fensholt**[1]

[1]Department of Geosciences and Natural Resource Management, University of Copenhagen, Øster Voldgade 10, 1350 Copenhagen, Denmark
[2]DHI GRAS, Agern Allé 5, 2970 Hørsholm, Denmark
[3]Institute for Sustainable Agriculture (IAS), Spanish Research Council (CSIC), Campus Alameda del Obispo, Av. Menéndez Pidal s/n, 14004 Córdoba, Spain
[4]Geological Survey of Denmark and Greenland, Øster Voldgade 10, 1350 Copenhagen, Denmark

Correspondence to: R. Guzinski (rag@ign.ku.dk, rmgu@dhi-gras.com)

Abstract. Evapotranspiration (ET) is the main link between the natural water cycle and the land surface energy budget. Therefore water-balance and energy-balance approaches are two of the main methodologies for modelling this process. The water-balance approach is usually implemented as a complex, distributed hydrological model, while the energy-balance approach is often used with remotely sensed observations of, for example, the land surface temperature (LST) and the state of the vegetation. In this study we compare the catchment-scale output of two remote sensing models based on the two-source energy-balance (TSEB) scheme, against a hydrological model, MIKE SHE, calibrated over the Skjern river catchment in western Denmark. The three models utilize different primary inputs to estimate ET (LST from different satellites in the case of remote sensing models and modelled soil moisture and heat flux in the case of the MIKE SHE ET module). However, all three of them use the same ancillary data (meteorological measurements, land cover type and leaf area index, etc.) and produce output at similar spatial resolution (1 km for the TSEB models, 500 m for MIKE SHE). The comparison is performed on the spatial patterns of the fluxes present within the catchment area as well as on temporal patterns on the whole catchment scale in 8-year long time series. The results show that the spatial patterns of latent heat flux produced by the remote sensing models are more similar to each other than to the fluxes produced by MIKE SHE. The temporal patterns produced by the remote sensing and hydrological models are quite highly correlated ($r \approx 0.8$). This indicates potential benefits to the hydrological modelling community of integrating spatial information derived through remote sensing methodology (contained in the ET maps derived with the energy-balance models, satellite based LST or another source) into the hydrological models. How this could be achieved and how to evaluate the improvements, or lack of thereof, is still an open research question.

1 Introduction

Evapotranspiration (ET) acts as a coupling between two of the most important natural processes affecting the land surface: the water (mass) exchange and the energy exchange (Campbell and Norman, 1998). Therefore it has a strong impact on, and is impacted by, plant biophysics, weather and climate, and is an important component when modelling those processes. At the same time the knowledge of both the magnitude of water loss from the ground through evapotranspiration and spatial distribution of this flux has many practical applications, such as in agri- and aqua-culture, water resource management or drought monitoring (Anderson et al., 2012). This has led to an active interest from the research community in the spatially distributed modelling of evapotranspiration and to the development of a number of different methodologies. Two of the most common approaches are (1) the modelling of land surface energy fluxes, mostly with the use of land surface temperature (LST) maps derived from remote sensing observations, and (2) distributed physically based hydrological models.

The two types of modelling approaches have been compared previously, for example recently by Conradt et al. (2013) who compared ET patterns produced by hydrological model, remote sensing based model and ground measurements in sub-basins of the Elbe River. That paper concluded with a recommendation of further comparison studies of the different modelling approaches, especially using more than two independent models, to better understand their relative strengths and weaknesses. This should lead to improved model performance, but also to increased understanding of the errors (and their magnitudes) present in the different models, which is particularly important if the approaches are to be combined through, for example, data assimilation. Without the knowledge of errors, assimilating remotely sensed ET into hydrological models might not provide its full benefit. Pan et al. (2008), for example, found that assimilating ET derived with remote sensing observations into their hydrological model did not have a large impact on the modelled water budget since it was assumed that the calibrated hydrological model provided much more accurate ET values than the satellite observation based model. Schuurmans et al. (2011) found that although spatial patterns produced by the hydrological model became more realistic after the assimilation of satellite based ET, a major weakness was the lack of information about the standard error present in the ET estimates and lack of independent spatially distributed ET data set that could be used for validation. Therefore, in this study we compare two remote sensing based ET models and a hydrological model, with the aim of improving the understanding of their limitations and providing information that could be used in potential future integration of the approaches through data assimilation.

The remote sensing models of evapotranspiration (Kalma et al., 2008) aspire to minimize the calibration of site-specific parameters and the usage of locally derived data. Instead they aim to be applicable in a wide number of climatic and vegetation conditions without any major modifications, and to rely mostly on data acquired through satellite observations (e.g. LST) or regional-scale modelling (e.g. air temperature). This necessitates a number of assumptions and simplifications, which might lead to reduced accuracy of the modelled fluxes. Another feature of the remote sensing models is the treatment of each pixel within the modelling domain as a stand-alone sub-domain without any connections or interactions with the surrounding pixels. In some approaches, such as the triangle approach (see below), some of the model parameters are derived through common analysis of all the pixels in the domain, but the fluxes in the individual pixels are still derived individually. Similarly, the remote sensing models consider each satellite image as a stand-alone snapshot of the land surface conditions, with no memory of the past.

There are a number of remote sensing modelling methodologies being actively used by the research community ranging from simpler, empirical ones to more complex, physically based ones. One of the simpler approaches consist of the so-called "triangle" models, named after the shape resulting from plotting the pixel values of an LST map against pixel values of a vegetation index map. The evaporation fraction can then be derived by interpolating between the edges of the triangle (Jiang and Islam, 2001; Stisen et al., 2008). More advanced schemes characterize the ground surface as one layer (soil and vegetation combined) in one-source models (e.g. Surface Energy Balance System model, Su, 2002), or as two layers (soil and vegetation separately) in two-source models, the majority of which follow the two-source energy-balance (TSEB) modelling scheme (Norman et al., 1995). Both the one-source and two-source models characterize the fluxes of heat and moisture between the surface and the atmosphere in terms of a set of resistance equations, formulated from physically based models of boundary layer behaviour under different atmospheric conditions and vegetation covers. The two-source models have the advantage of explicitly representing the separate contributions of the soil and the vegetation, thus avoiding the need for parametrization of an "excess" resistance term whose value differs significantly from one reference to another (Norman et al., 1995; Matsushima, 2005; Kustas and Anderson, 2009; Boulet et al., 2012).

The distributed physically based hydrological models, in contrast to the remote sensing models, are heavily parametrized and calibrated for each individual catchment or study area (Refsgaard, 1997). Besides evapotranspiration, and other land surface fluxes, they can model a host of other hydrological processes such as channel flow, unsaturated zone flow or ground water flow and the interactions between those processes (Graham and Butts, 2005). This means that the modelling is performed in four dimensions (latitude, longitude, elevation or depth, and time) and that there is an interaction between the model pixels in both time and space. Due to their complex nature the hydrological models require significant computational resources and a large number of inputs for calibration and operation. During the operational stage the models require gridded meteorological input (including rainfall, air temperature and humidity), gridded soil hydraulic parameters and a digital elevation model among others (Stisen et al., 2011a). During the calibration stage (prior to the operational stage) additional information is required, for example measured hydrological parameters such as hydraulic head or stream outflow.

Since the hydrological models are calibrated using detailed hydrological observations, it is our hypothesis that the catchment-wide evapotranspiration estimated by those models is more accurate than the one derived with remote sensing models. On the other hand, we expect the energy-balance models driven by remote sensing observations to better represent the spatial patterns of the fluxes present within the catchment. We evaluate this hypothesis by running a hydrological model, MIKE SHE, described in Sect. 4.1, and two TSEB based models, Dual-Temperature-Difference (DTD – Norman et al., 2000; Guzinski et al., 2013) and TSEB-2 An-

gle Radiative Transfer (TSEB-2ART – Nieto et al., 2013), described in Sect. 4.2, over the Skjern river catchment located in western Denmark (see Sect. 2). Apart from being the largest river in Denmark, Skjern is also the study area of the Danish hydrological observatory HOBE (Jensen and Illangasekare, 2011), thus providing ample data for calibrating and validating the models. The three models are run with the same meteorological inputs, interpolated from field based observations, and the same land cover and leaf area index (LAI) maps, in order to minimize the uncertainties inherent in using different data sets. The differences between the models, and their input data, are described in Sect. 4.

The output land surface fluxes, and in particular the latent heat flux, from the three models are then inter-compared. The comparison is performed on a pixel-by-pixel basis as well as on catchment scale, and both systematic and unsystematic differences are analysed (Ji and Gallo, 2006). In addition, the catchment-scale temporal evolution of the evapotranspiration estimated from the three models is evaluated. Through this, we assess strengths and weaknesses of the different modelling approaches and in particular try to answer the following question: is there any additional (spatial) information present in inputs or outputs of remote sensing based surface energy-balance models that is missing from the physically based distributed hydrological model?

2 Study area

The study area covers the Skjern River catchment (Fig. 1) which is located on the western part of Denmark's Jutland peninsula and it is the largest river catchment in Denmark in terms of water volume. It has a roughly rectangular shape with the east–west length of around 60 km and north–south length of around 40 km. The Skjern River outlet is on the western side of the catchment with the discharge entering Ringkøbing Fjord. The terrain is mostly flat with a maximum elevation of 125 m a.s.l. and a gentle east to west slope. The soils are predominantly sandy with the main land use being agriculture and coniferous plantations. The catchment experiences a temperate maritime climate, with mean annual precipitation of 990 mm and mean annual temperature of 8.2 °C. Since 2007 the catchment is hosting the Danish Hydrological Observatory, HOBE, with numerous experiments and measurements concerning precipitation, evapotranspiration, greenhouse gas exchange, groundwater–surface water interactions and other related topics, making it highly suitable for calibrating and evaluating the distributed physically based hydrological models. For more details refer to Jensen and Illangasekare (2011).

3 Common model inputs

In order to compare the performance of the three models and not the accuracy of their inputs, the models used the same
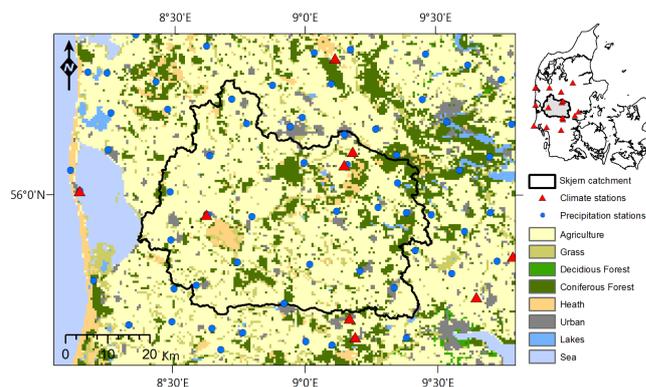
Figure 1. Land use map of the study area: the Skjern river catchment in the west of Denmark's Jutland peninsula. Model input meteorological data were interpolated from the measurements taken by the stations shown on the map.

auxiliary input data whenever possible. Those common inputs consisted of maps with meteorological forcings, LAI, albedo and land cover types. For the meteorological forcing data, kriged fields of wind and temperature corrected precipitation from 43 rain gauges were used (Stisen et al., 2011b) together with air temperature, relative humidity, incoming shortwave radiation, wind speed and pressure interpolated from 16 climate stations. The locations of the rain gauges and climate stations in relation to the study area are presented in Fig. 1. The vegetation-related inputs were derived using remote sensing data with LAI estimated from MODIS NDVI (MOD13A1 product) following the study of Boegh et al. (2009) and albedo estimated from narrow band MODIS reflectance following Liang (2001). It should be noted that albedo maps were only shared by MIKE SHE and DTD, with TSEB-2ART producing its own albedo maps as one of the outputs. A land cover map was taken from the Danish Areal Information System run by the Danish Ministry of Environment (http://www2.dmu.dk/1_Viden/2_Miljoe-tilstand/3_samfund/AIS/1a_Dynamisk_gis/Image_viewer/AAK_IMS_en.htm, last access: 29 January 2013), with the land cover dependent parameters listed in Table 1. The LAI correction factor, mentioned in the last row of Table 1, was derived during the calibration of the MIKE SHE model and is used as multiplicative factor for LAI estimated from MODIS NDVI for all land cover classes except for forests. Even though this is a MIKE SHE-specific parameter, it was also applied to LAI inputs to DTD and TSEB-2ART to ensure that comparable input data were used in all the models.

All common input data maps were delivered in UTM32-WGS84 projection. The LST observations used by the different models, as well as data used only by a single model, are described in the sections below.

Table 1. Land cover dependent parameters for the three models. The equations referred to in the table are (Eq. a) $0.15 \cdot \text{LAI}$ and (Eq. b) $0.12 \cdot \text{LAI} + 0.07$, where LAI is the leaf area index before multiplication by the LAI correction factor and has a minimum value of 0.5.

Parameter	Land cover class												Units
	Grass			Coniferous forest			Heath			Crop			
	MIKE SHE	DTD	TSEB-2ART	MIKE SHE	DTD	TSEB-2ART	MIKE SHE	DTD	TSEB-2ART	MIKE SHE	DTD	TSEB-2ART	
Vegetation heigh (h_C)	Eq. (a)	Eq. (a)	Eq. (a)	9.0	9.0	9.0	Eq. (a)	Eq. (a)	Eq. (a)	Eq. (a)	Eq. (a)	Eq. (a)	m
Clumping factor	NA	1.0	NA	NA	0.5	NA	NA	0.9	NA	NA	0.9	NA	unitless
Canopy height/canopy width	NA	1.0	1.0	NA	3.0	3.0	NA	1.0	1.0	NA	1.0	1.0	unitless
Leaf size	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	m
Temperature measurement height	10.0	10.0	10.0	18.0	18.0	18.0	10.0	10.0	10.0	10.0	10.0	10.0	m
Wind measurement height	10.0	10.0	10.0	18.0	18.0	18.0	10.0	10.0	10.0	10.0	10.0	10.0	m
Root depth	Eq. (b)	NA	NA	Eq. (b)	NA	NA	Eq. (b)	NA	NA	Eq. (b)	NA	NA	m
Minimum stomata resistance	90	NA	NA	150	NA	NA	120	NA	NA	90	NA	NA	s m^{-1}
Extinction coefficient	0.6	NA	NA	0.5	NA	NA	0.3	NA	NA	0.6	NA	NA	unitless
LAI correction factor	2.21	2.21	2.21	1.0	1.0	1.0	2.21	2.21	2.21	2.21	2.21	2.21	unitless

4 Models

4.1 MIKE SHE model

The implementation details of the hydrological model used in this study, MIKE SHE SW-ET, are presented in Stisen et al. (2011a) and Overgaard and Rosbjerg (2005). Briefly, the model couples ground-water and surface-water modules together with an ET module (Overgaard and Rosbjerg, 2005). The SW-ET module, based on the two-source model of Shuttleworth and Wallace (1985), uses hydrological modules' outputs of soil moisture, soil heat flux and fraction of soil and leaf covered by ponded water. Besides these parameters, meteorological observations of air temperature and humidity, wind speed and incoming shortwave radiation and maps of albedo, LAI and land cover are used to solve a set of 10 linear equations for the temperature and humidity of dry and wet soil, dry and wet leaf and inter-canopy air (see Appendices A and D in Overgaard and Rosbjerg, 2005, for more details). With those parameters it is possible to estimate the effective soil and leaf temperatures as well as the radiometric surface temperature (LST) and the latent and sensible heat fluxes. Since the model simulates LST, it is possible to calibrate the model against remotely sensed LST in addition to hydrological variables such as hydraulic head or stream outflow. The model used in this study was calibrated for the Skjern river catchment against the above mentioned hydrological variables, LST taken from the MYD11A1 Aqua-MODIS product, evapotranspiration measured at three flux tower sites placed within the catchment area and soil moisture measurements from a distributed sensor network. The calibration methodology will be a topic of a subsequent paper.

As input the model requires gridded meteorological forcing data, soil hydraulic parameters and a number of parameters related to vegetation. The meteorological forcing data, LAI, albedo and land use maps are described in Sect. 3. The soil hydraulic parameters came from a study of Greve et al. (2007). The derivation of other input parameters, such as soil surface roughness or irrigation water input, is described in Stisen et al. (2011a). The MIKE SHE SW-ET model, from now on referred to as MIKE SHE, was run at 500 m resolu-

tion, and as output provided the surface energy fluxes (sensible, latent and ground heat fluxes and net radiation) together with LST and soil and canopy temperatures, T_S and T_C respectively. The outputs were bilinearly interpolated from 500 m to 1 km to match the resolution of the outputs from both remote sensing models.

4.2 TSEB modelling scheme

The TSEB approach (Norman et al., 1995) splits the observed directional LST into its two main components, namely the temperature of soil and canopy:

$$T_R(\theta) \approx \left[f(\theta) T_C^4 + (1 - f(\theta)) T_S^4 \right]^{0.25}, \quad (1)$$

where $T_R(\theta)$ is the LST observed at the view zenith angle (VZA) of θ and $f(\theta)$ is the fraction of vegetation cover in the field of view of the sensor at VZA θ. This allows the model to estimate the latent and sensible heat fluxes from the soil and canopy separately, thus avoiding the need to parametrize the "excess" resistance term which is often present in single-source models but for which there does not yet exist an established methodology for estimating its value (e.g. Matsushima, 2005; Boulet et al., 2012).

In the single-angle TSEB models, the latent heat flux of the canopy, LE_C, is initially estimated using the assumption that the canopy is transpiring at the potential rate dictated by the divergence of net radiation in the canopy, $R_{n,C}$, and a modified Priestly–Taylor approach. This allows an initial estimation of the sensible heat flux of the canopy, H_C, and of T_C. If the model returns unrealistic results ($LE < 0$ meaning condensation during daytime) the transpiration of the canopy can be iteratively reduced until realistic results are obtained (Norman et al., 1995).

In the dual-angle TSEB models, T_S and T_C can be derived directly from the observation geometry, followed by H_S and H_C and finally LE_C as residual of the canopy energy balance. In both cases the total energy balance is ensured by estimating the latent heat flux from the soil, LE_S, as residual:

$$LE_S = R_{n,S} - H_S - G = (R_n - R_{n,C}) - (H - H_C) - G, \quad (2)$$

where $R_{n,S}$ is the net radiation of the soil, H_S is the sensible heat flux of the soil and G is the ground heat flux.

The two TSEB based models used in this study follow the principles described above but differ in other implementation details as described in the subsections below.

4.2.1 DTD model

The DTD model minimizes the influence of systematic error in the retrievals of LST and air temperature by replacing absolute temperature measurements with temperature change between two observations (Norman et al., 2000). In the original DTD model the first observations was early in the morning, when fluxes are minimal, and the second later in the morning or in the afternoon. Guzinski et al. (2013) demonstrated that replacing the early morning observations with night-time ones does not have a significant effect on the accuracy of the modelled fluxes, thereby facilitating the use of polar orbiting satellites with day and night overpasses, and introduced a simple scheme for accounting for vegetation phenology when estimating canopy transpiration. The model was further modified in Guzinski et al. (2014) where the resistance network to sensible heat flux was modified, to the so-called "series" configuration, to explicitly consider the in-canopy air temperature, thus improving the model performance during dry conditions. The DTD model formulation used in this study is as described in the Appendices of Guzinski et al. (2014), with the exception of the formulation of the resistance to heat transfer from the soil surface, R_S.

The R_S formulation used in the TSEB modelling scheme accounts for turbulent transport from free convection (Kustas and Norman, 1999):

$$R_S = \frac{1}{c(T_S - T_C)^{1/3} + bu_S}, \tag{3}$$

where c and b are constants given a value of $0.0025\,\mathrm{m\,s^{-1}\,K^{-1/3}}$ and $0.012\,\mathrm{m\,s^{-1}}$ respectively and u_S is the wind speed just above the soil surface. However, since DTD aspires to use just time differential temperature measurements, it was originally decided to remove the $(T_S - T_C)^{1/3}$ term from the resistance equation and instead to replace it with a LAI-dependent constant. For dense canopies $T_S - T_C$ was assumed to be 5 K, while for sparse canopies it was assumed to be 15 K (Norman et al., 2000). Those assumptions made sense for the data sets used to evaluate the model performance, taken in New Mexico over June and July 1997 (http://hydrolab.arsusda.gov/sgp97, last accessed 27 February 2014) and in Arizona from June to September 1990 (Kustas and Goodrich, 1994). However, in the current study area, dominated by croplands located in temperate maritime climate (see Sect. 2), sparse canopy conditions are usually present in early spring and autumn when the difference between T_S to T_C is significantly less than 15 K. Therefore the R_S formulation was further modified to make use of the difference in thermal inertia of

LST and air temperature:

$$R_S = \frac{1}{c\left[(T_{R,1} - T_{R,0}) - (T_{A,1} - T_{A,0})\right]^{1/3} + bu_S}, \tag{4}$$

where subscript 0 indicates temperatures estimate at night or early in the morning, and 1 indicates estimate at some other time during the day, and b and c have the same values as shown above. Thus, the use of both non-time differential temperature estimates and the assumptions about the magnitude of $T_S - T_C$ are avoided. This formulation implicitly takes into account the amount of vegetation cover, since vegetation has larger thermal inertia than soil and thus $(T_{R,1} - T_{R,0})$ is lower for dense canopies, while also reflecting the climatic conditions present in the study area.

The model uses MODIS LST estimates from the MYD11A1 product, together with land cover, LAI and albedo values derived as described in Sect. 3 and vegetation indices (normalized difference vegetation index and enhanced vegetation index) from the MOD13A2 product for estimating the fraction of vegetation that is green (Guzinski et al., 2013). The meteorological inputs are also as described in Sect. 3. The MODIS LST and vegetation indices products were provided by NASA in a georeferenced grid with 930 m resolution and Sinusoidal projection. This was bilinear resampled to 1 km resolution grid and reprojected to UTM32-WGS84 projection. The modelled fluxes are output at 1 km resolution.

4.2.2 TSEB-2ART model

When TSEB is applied with single-angle LST, some assumptions are needed based on the expectation that plants transpire at their potential rate. This assumption may lead to significant errors in cases when plants are stressed, or when the potential canopy transpiration is not well defined. For that reason, the green fraction of vegetation (f_g) is an important parameter within the model since it improves TSEB accuracy in forested ecosystems and during senescence by taking into account the phenological development of the vegetation (Guzinski et al., 2013; Chirouze et al., 2014). However, there does not yet exist an established method of estimating f_g using remote sensing data. To overcome this issue, dual-angle LST can be used for retrieving soil and canopy temperatures without employing any assumptions based on the canopy transpiration (Chehbouni et al., 2001; Kustas and Norman, 1997; Nieto et al., 2010a, b). Simple models for such retrieval have been proposed based on the proportion of vegetation and soil observed at two different viewing angles (Chehbouni et al., 2001; Kustas and Norman, 1997). However, since plant canopies are composed of finite leaves, multiple scattering of energy occurs within the canopy and therefore more physically complex methods for retrieving soil and canopy temperatures may be needed when using dual-angle LST measurements (François, 2002; Nieto et al., 2013).

The TSEB-2ART model (Nieto et al., 2013) couples a dual-angle version of TSEB introduced by Kustas and Norman (1997), with radiative transfer model (RTM) 4SAIL (Verhoef et al., 2007). Through this coupling it is possible to retrieve canopy and soil temperatures by analytically inverting the RTM with LST estimates of the same area but obtained through two different view zenith angles. 4SAIL takes into account the different emissivities of the end members (canopy and soil) and hence multiple scattering of the thermal radiation, as well as the downwelling longwave radiation reflected by the surface. Therefore, the coupling should result in more accurate temperature retrievals compared to just using the geometric configuration of the observations (Nieto et al., 2013). Similarly to other TSEB based models, the canopy and soil temperatures are then used by TSEB-2ART to estimate the sensible heat flux of the canopy and soil respectively. In addition the RTM is used to estimate the net radiation, and radiation divergence in the canopy, while taking into account multiple scattering of the shortwave/longwave radiation between the soil and the canopy and within the canopy. The inclusion of 4SAIL also allows for the use of different leaf inclination distribution functions, rather than the spherical leaf distribution of the original TSEB (Norman et al., 1995; Kustas and Norman, 1999). Ground heat flux is estimated as a fraction of net radiation reaching the soil based on Choudhury et al. (1987). Finally, transpiration and soil evaporation can be obtained as residual terms of the vegetation and soil energy budgets.

TSEB-2ART has been evaluated over three flux tower sites within the HOBE area, obtaining more accurate flux retrievals than both the original dual-angle (Kustas and Norman, 1997) and the single-angle TSEB (Norman et al., 1995) implementations when driven by LST estimates derived with the AATSR sensor on board the Envisat satellite (Nieto et al., 2013). Even though the Envisat satellite is no longer functional, the model can be applied to the dual-angle LST observations in the future Sentinel 3 mission (Donlon et al., 2012). Apart from the AATSR derived LST the model uses the same meteorological data as well as land cover and LAI maps as MIKE SHE and DTD models but produces its own albedo as part of the implementation of 4SAIL. The Envisat LST was derived from the ATS_TOA_1P, AATSR Gridded Brightness Temperature and Radiance, product, which is a full resolution data set resampled to a 1 km × 1 km grid for both the nadir and forward views by the European Space Agency (Scarpino and Cardaci, 2009). The split-window brightness temperatures (11 and 12 μm) for both forward and nadir were then reprojected to UTM32-WGS84 and resampled to the same 1 km resolution using a bilinear interpolation. LST at the two AATSR observation angles was then retrieved by the quadratic dual-channel split-window algorithm proposed by Coll et al. (2006) for AATSR. The modelled fluxes are output at 1 km resolution.

5 Comparison methodology

The spatial comparison was performed by selecting all the pixels in the Skjern catchment on all the days between 2003 and 2010 when at least 10 % of the catchment was cloud free during the night and day Aqua overpasses and which met the following conditions:

– the pixel is not classified as water or urban area (met by 96 % of the catchment area);

– all three models produce valid results, meaning $LE > 0\,W\,m^{-2}$ and $H \geq -100\,W\,m^{-2}$ (met by 85 % of modelled fluxes).

This resulted in over 95 000 sets to be compared. A median moving-window filter of 3 × 3 pixels was applied to the output maps to reduce noise caused by image misregistration while preserving the spatial patterns found in the maps.

The comparison was performed using the instantaneous modelled sensible heat flux, latent heat flux and available energy (AE) defined as the net radiation minus the ground heat flux. The magnitude of those fluxes is strongly influenced by the incoming solar radiation and so it has a cyclic annual component with generally larger fluxes during the summer months and lower during the winter months. This could potentially influence the correlation between the fluxes modelled with different models. To remove this time dependent component and instead to evaluate the influence of water availability on the partitioning of the available energy into latent and sensible heat fluxes by the different models, the evaporative fraction (EF), defined as the ratio of energy used for evapotranspiration to the total available energy, was also used during the comparison.

When comparing the fluxes estimated by the three different models the time at which the fluxes are estimated must be taken into account. The TSEB-2ART fluxes are estimated at the time of the Envisat overpass, which is around 11:30 local time (LT), while the DTD fluxes are estimated at the time of Aqua overpass, around 12:00–13:00 LT. The MIKE SHE fluxes are estimated at hourly intervals throughout the day. Therefore, when comparing the fluxes between MIKE SHE and one of the satellite based models a linear interpolation was performed between the two MIKE SHE estimates bracketing the satellite based estimate (e.g. if satellite overpass was at 11:48, MIKE SHE estimates from 11:00 and 12:00 would be used). When comparing the fluxes from two satellite based models there is an offset present due to this time difference, although it should be reduced when comparing EF (Peng et al., 2013). A decision was made to perform the comparison using instantaneous modelled fluxes, and not their daily estimates, since extrapolating to daily values would just involve multiplying EF by the daily available energy (or net radiation assuming negligible daily G). Therefore the multiplicative factor would be the same (if field-measured daily available energy were used) or very similar (if modelled daily available energy were used) for the three models and no additional

information or insight would be gained. On the contrary, the self-preservation of EF might not always hold over the whole study area due to frequently cloudy conditions, bringing additional complications and errors when extrapolating to daily values.

A number of statistical measures are used to explore the relation between the fluxes, and temperatures, estimated by the three models. The first one is the Pearson correlation coefficient, r, which measures the linear covariation of two data sets. To assess the differences between the data sets the root mean square difference (RMSD), which is the square root of the mean square difference (MSD), is used:

$$\text{RMSD} = \text{MSD}^{0.5} = \left[\frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)^2 \right]^{0.5}, \tag{5}$$

where X_i and Y_i are the ith points in the X and Y data sets. The MSD can be further split into systematic and unsystematic mean product differences (MPD), MPD_s and MPD_u respectively, where MPD_s measures the distance between the geometric mean regression line (Barker et al., 1988) of two data sets and the 1–1 line, while MPD_u measures the distance between the data sets' points and the geometric mean regression line (Ji and Gallo, 2006). The geometric mean regression is used instead of the linear regression since the former one assumes that both X and Y are subject to errors. Since $\text{MSD} = \text{MPD}_s + \text{MPD}_u$ it is also possible to calculate the relative contribution of the systematic and unsystematic difference to the total difference as $\text{MPD}_s / \text{MSD}$ and $\text{MPD}_u / \text{MSD}$ respectively (Ji and Gallo, 2006). The systematic component of the difference represents the variation between the data sets that can be corrected by simple linear transformation of one of the data sets, while the unsystematic difference can be thought of as noise caused by some unknown factors which is harder to correct for (Ji et al., 2008). For presentation purposes a square root is taken of MPD_s and MPD_u to obtain RMPD_s and RMPD_u respectively which are then shown in the results' tables. The last statistical measure used is the mean bias, calculated as the difference between the means of two data sets. With the exception of the sign of the bias, all the statistical measures are symmetric, meaning that no assumption is made about the correctness or otherwise of any of the data sets and that the same values are obtained if the order of the data sets is reversed when calculating the measures.

The temporal patterns of evapotranspiration were evaluated at catchment scale, meaning that all the valid non-urban and non-water pixels within the catchment were averaged to determine the catchment-scale fluxes. It should be noted that since MIKE SHE also simulates the fluxes over water and urban pixels, this average is not the whole catchment evapotranspiration as modelled by MIKE SHE. However, since the number of water and urban pixels is quite small (Fig. 1), the averaged value should be close to the whole catchment evapotranspiration. Only those days on which the remote sensing

models produced flux estimates in pixels representing at least 70 % of all non-urban and non-water catchment pixels were included in the analysis. In the case of DTD this condition was satisfied on 132 days over the 8-year period, while in the case of TSEB-2ART there were 68 valid days due to the less frequent revisit time of AATSR vs. MODIS. The catchment averages for each date were produced using the same set of pixels for the remote sensing models and MIKE SHE. The two data sets were compared using the r correlation coefficient, RMSD and bias and the ratio of RMSD and bias to the mean evapotranspiration estimated by MIKE SHE.

6   Results

6.1   Spatial patterns

The results of pixel-to-pixel comparisons of fluxes between the three model pairs are presented in Figs. 2 (MIKE SHE–DTD), 3 (MIKE SHE–TSEB-2ART), and 4 (TSEB-2ART–DTD) with statistics summarized in Table 2 and described for each model pair in the subsections below.

6.1.1   MIKE SHE vs. DTD

The bias between the turbulent fluxes modelled with MIKE SHE and DTD is significant with a value of $19 \, \text{W m}^{-2}$ and $-45 \, \text{W m}^{-2}$ for H and LE respectively. The RMSD is also quite large, at $78 \, \text{W m}^{-2}$ for H and $106 \, \text{W m}^{-2}$ for LE, and consequently the correlation coefficient between the modelled turbulent fluxes is relatively low, with a maximum value of 0.56. When the differences are split into systematic and unsystematic parts, 89 % of the error in H and 81 % of error in LE is unsystematic. The differences are propagated through to EF, leading to very low correlation although with a small bias. The differences in the turbulent fluxes cannot be caused mainly by differences in the parametrization of the available energy since in that case the correlation reaches 0.97. This was expected since the two models use the same incoming solar radiation forcing and the same albedo maps so the majority of the $35 \, \text{W m}^{-2}$ RMSD (55 % of MSD) is systematic and caused by the differences in the net longwave radiation estimation due to different LSTs, with DTD using MODIS LST and MIKE SHE the modelled LST from the SW-ET module, and by the ground heat flux calculations.

6.1.2   MIKE SHE vs. TSEB-2ART

The comparison of fluxes produced with MIKE SHE and TSEB-2ART follows a similar pattern as in the previous section, with relatively low correlation and significant RMSD but with much lower bias (maximum magnitude of $8 \, \text{W m}^{-2}$ in the case of H). The other statistics are similar to the ones from MIKE SHE–DTD comparison, with RMSD of $84 \, \text{W m}^{-2}$ and r of 0.38 for H and $85 \, \text{W m}^{-2}$ and 0.58 for LE. The correlation of EF is slightly higher than in the case
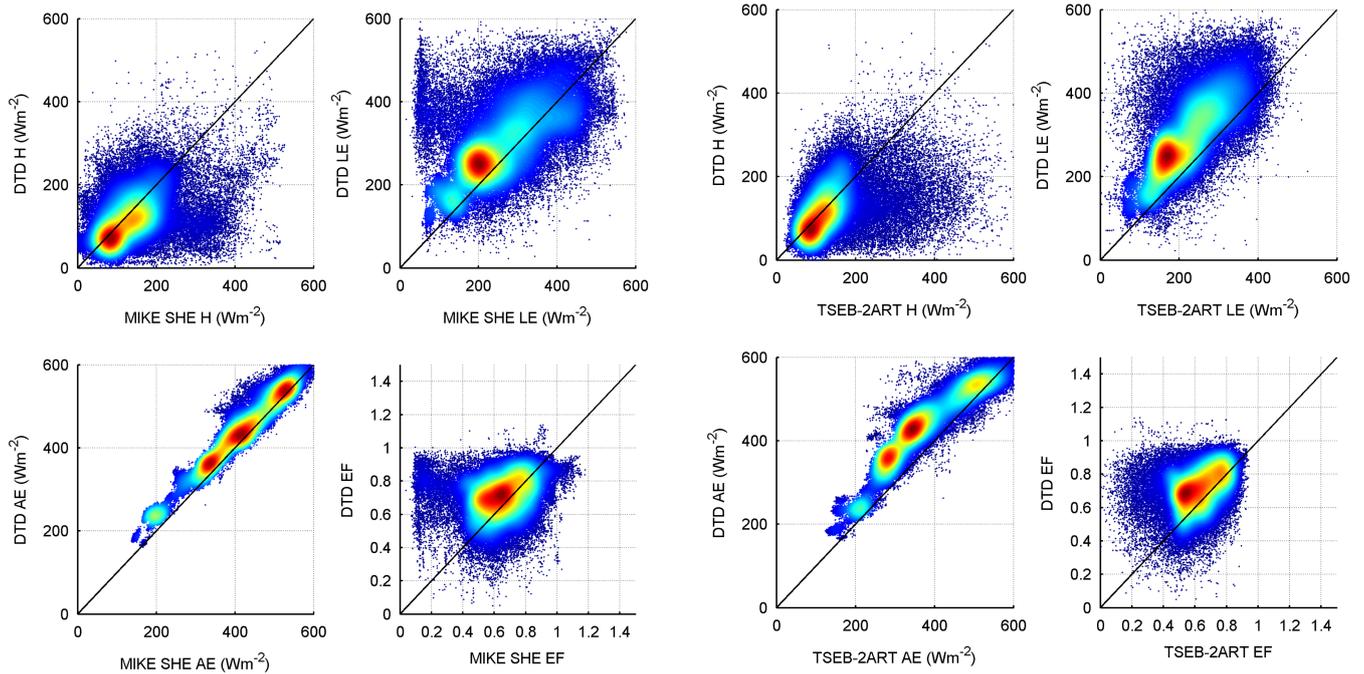
Figure 2. Density scatter plot of over 95 000 points comparing the sensible heat flux (top left), latent heat flux (top right), available energy (bottom left) and evaporative fraction (bottom right) modelled by MIKE SHE and DTD. Red colour indicates higher density of points, blue colour lower density.
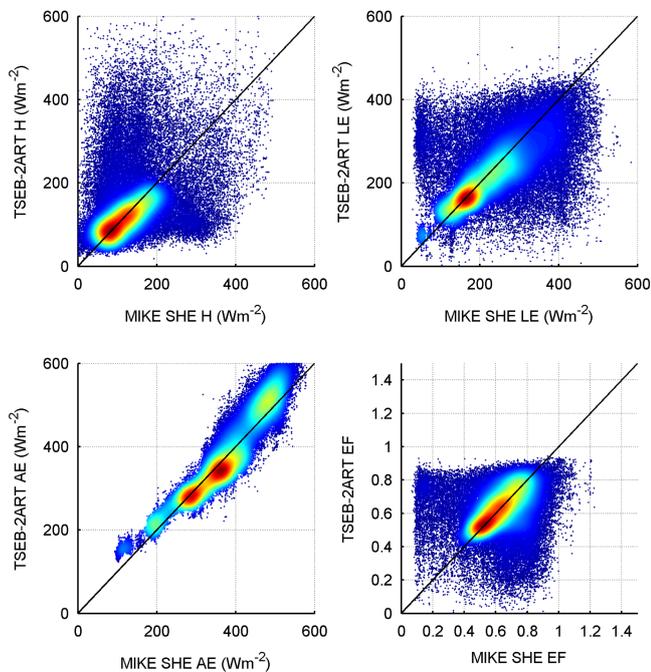
Figure 4. Density scatter plot of over 95 000 points comparing the sensible heat flux (top left), latent heat flux (top right), available energy (bottom left) and evaporative fraction (bottom right) modelled by TSEB-2ART and DTD. Red colour indicates higher density of points, blue colour lower density.

of DTD, with a value of 0.25, and the characterization of AE is consistent between the two models, with a correlation of 0.95 and a bias of $-7\,\mathrm{W\,m^{-2}}$, with only 8 % of MSD being attributed to systematic errors.

6.1.3 TSEB-2ART vs. DTD

The correlation between the turbulent fluxes modelled with TSEB-2ART and DTD is the highest of any model pairs, with correlation coefficient of 0.42 for H and 0.70 for LE, even though the fluxes were obtained at different times of the day. The time offset is evident in the bias of AE, with the value of AE during the later Aqua overpass time being on average $55\,\mathrm{W\,m^{-2}}$ higher than the value of AE during Envisat overpass time. The biases are also present in the other flux estimates, particularly of LE, with a value of $-77\,\mathrm{W\,m^{-2}}$. However, even though the biases are much higher than in any other pair, the RMSD between TSEB-2ART and DTD estimated turbulent fluxes is comparable to RMSD of those fluxes between the other pairs. As can be seen from the split of the difference into systematic and unsystematic components, a large component of the MSD between the fluxes is systematic with $\mathrm{RMPD_u}$ of LE reaching the lowest values of all the model pairs and $\mathrm{RMPD_u}$ of H being very close to the minimum. The correlation and RMSD of EF is also the best of all the model pairs, with values of 0.33 and 0.18 respectively.

Figure 3. Density scatter plot of over 95 000 points comparing the sensible heat flux (top left), latent heat flux (top right), available energy (bottom left) and evaporative fraction (bottom right) modelled by MIKE SHE and TSEB-2ART. Red colour indicates higher density of points, blue colour lower density.

Table 2. Statistical comparison between MIKE SHE, DTD and TSEB-2ART models for sensible and latent heat fluxes (H and LE), available energy (AE) and evaporative fraction (EF). Statistics used: correlation coefficient (r), root mean square difference (RMSD), systematic and unsystematic root mean product differences (RMPD$_s$ and RMPD$_u$ respectively), the percentage of mean square difference (MSD) attributed to systematic and unsystematic mean product differences (MPD) (MPD$_s$/ MSD and MPD$_u$/ MSD respectively) and bias. The statistics for H, LE and AE are in W m^{-2}, with the exception of MPD$_s$/ MSD and MPD$_u$/ MSD, which are percentages. The statistics for EF are unitless, with the exception of MPD$_s$/ MSD and MPD$_u$/ MSD, which are percentages.

		r	RMSD	RMPD$_s$	RMPD$_u$	MPD$_s$/ MSD	MPD$_u$/ MSD	Bias
MIKE SHE–DTD	H	0.42	78	26	74	11	89	19
	LE	0.56	106	46	95	19	81	−45
	AE	0.97	35	26	23	55	45	−25
	EF	0.21	0.20	0.08	0.18	18	82	−0.07
MIKE SHE–TSEB-2ART	H	0.38	84	12	83	2	98	−8
	LE	0.58	85	18	83	5	95	5
	AE	0.95	36	10	34	8	92	−7
	EF	0.25	0.19	0.03	0.18	2	98	0.01
TSEB-2ART–DTD	H	0.42	79	26	75	11	89	18
	LE	0.70	104	78	68	57	43	−77
	AE	0.94	67	56	37	69	31	−55
	EF	0.33	0.18	0.09	0.15	26	74	−0.09
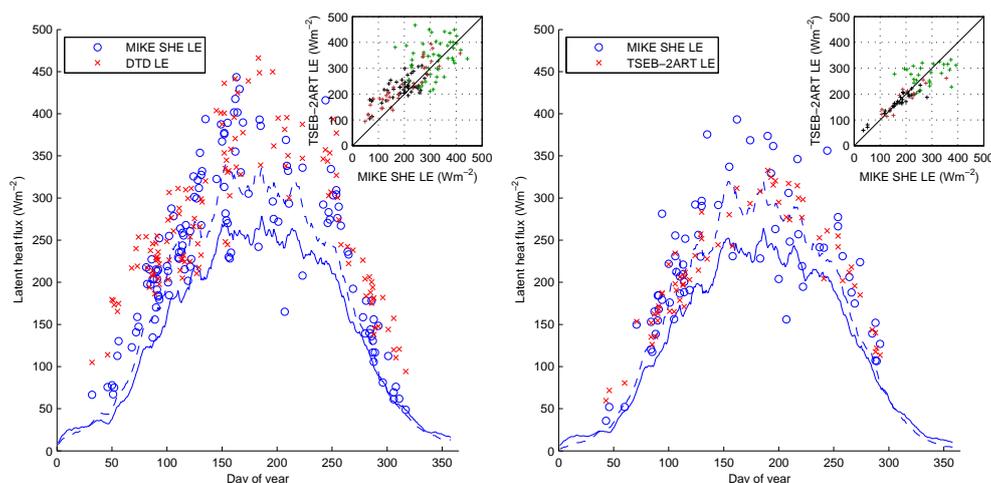
Figure 5. Average catchment-wide latent heat fluxes on the days when at least 70 % of non-water and non-urban pixels were modelled by either DTD (left) or TSEB-2ART (right). In the main graph the blue circles represent catchment fluxes modelled by MIKE SHE and the red crosses represent the catchment fluxes modelled by the remote sensing models on the same year and day of year (DOY) and at the same time of day. The figure contains dates from the 8 years under investigation and the blue solid line shows an 8-year averaged whole catchment ET for a particular DOY as modelled by MIKE SHE around the time of Aqua (left) or Envisat (right) overpass. The blue broken line shows potential ET for the same data set estimated using the Priestley–Taylor approach and MIKE SHE AE. The inset image contains a scatterplot of the MIKE SHE and remote sensing fluxes with black indicating fluxes from January to April, green from May to August and brown from September to December.

6.2 Temporal patterns

The results of comparing DTD and TSEB-2ART catchment-wide evapotranspiration estimates against MIKE SHE are presented in Fig. 5, with the statistics summarized in Table 3. The correlation between the latent heat fluxes modelled with DTD or TSEB-2ART and MIKE SHE are quite similar, with correlation coefficients having a value of 0.79 in the case of comparing MIKE SHE and DTD and 0.83 in the case of MIKE SHE and TSEB-2ART. The biases between the modelled fluxes are quite small, with the largest one present when looking at LE between MIKE SHE and DTD and having a value of −30 W m^{-2}, which represents just 13 % of the mean value of LE modelled by MIKE SHE. The RMSD values between DTD and MIKE SHE and between TSEB-2ART and MIKE SHE are 68 and 45 W m^{-2} respectively. This rep-

Table 3. Statistical comparison of catchment-wide latent heat fluxes estimated by the model pairs (MIKE SHE–DTD and MIKE SHE–TSEB-2ART) for predominantly cloud-free days over the period of 8 years. Statistics used: correlation coefficient (r), root mean square difference (RMSD), relative RMSD (%RMSD), bias and relative bias (%bias). RMSD and bias are in $W\,m^{-2}$ while %RMSD and %bias are calculated as the statistic divided by the mean of the MIKE SHE LE estimates and are percentages.

		r	RMSD	%RMSD	Bias	%Bias
MIKE SHE–DTD	LE	0.79	68	28	−30	−13
MIKE SHE–TSEB-2ART	LE	0.83	45	21	6	3
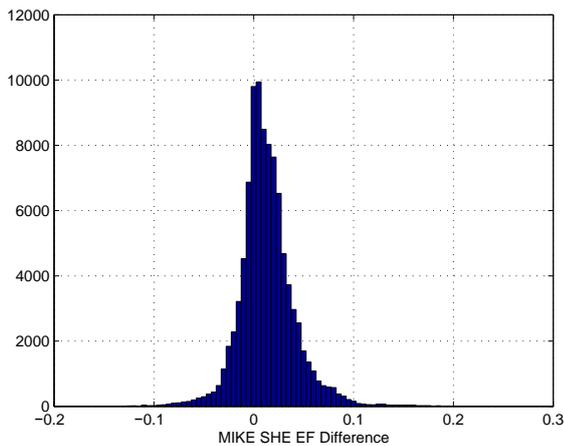
Figure 6. Histogram of the pixel-wise differences between evaporative fraction (EF) estimated by MIKE SHE at the time of Aqua overpass and Envisat overpass. The differences between the two sets were evaluated using the two-sample t test and are found to be statistically significant with a p value smaller than 0.001.

resents 28 % of the mean value of MIKE SHE LE in the case of DTD and 21 % in the case of TSEB-2ART.

7 Discussion

7.1 Spatial patterns

Even though DTD and TSEB-2ART estimate fluxes at different times during the day, the correlation between H and LE estimated by those two models is as strong (in the case of H) or stronger than between either of the models and MIKE SHE. In addition, the value of $RMPD_u$ between LE estimated with those two models is lower than for the other comparisons, even though the RMSD between LE modelled with TSEB-2ART and MIKE SHE is lower. This indicates that the spatial patterns produced by the remotely sensed models have a stronger agreement with each other than with the patterns produced by the hydrological model. It can be presumed that if the DTD and TSEB-2ART estimated the fluxes at the same time, the correlation would be even higher and the differences even smaller.

When the seasonal signal of the available energy is removed by replacing the turbulent fluxes by EF, the spatial

patterns produced by the remote sensing models are still more strongly correlated than when either of them is compared to the hydrological model. The correlation coefficient of TSEB-2ART and DTD EF is 0.33 compared to the second highest value of 0.25 between MIKE SHE and TSEB-2ART EF. However, it should once again be kept in mind that the remote sensing models estimate the fluxes at different times of the day. Usually it is assumed that during clear sky days the EF remains constant throughout the daytime and especially around noon (Peng et al., 2013). However, by comparing the differences of EF modelled by MIKE SHE at the Aqua and Envisat overpass times (Fig. 6), it can be seen that EF differs between the overpasses. This could be due to the fact that for the majority of the days used in this study there was some cloud cover over the Skjern river catchment (the threshold of inclusion in the study was 10 % of cloud free pixels during the night and day Aqua overpass) meaning that it is highly probable that clouds have passed over the study pixels between the two satellite overpasses, breaking the assumption of self-preservation of EF (Crago, 1996). Therefore it can be assumed that if the EF from TSEB-2ART and DTD were estimated at the same time the correlation would be higher still.

Figure 7 and Table 4 present the results of comparing just the vegetation transpiration (LE_C) produced by the three models. The correlation coefficient of the modelled transpiration is higher than for bulk LE (transpiration and soil evaporation combined) for all the model pairs, with the correlation between TSEB-2ART and DTD still remaining the highest. However, the RMSD of transpiration is higher than that of bulk LE when comparing TSEB-2ART with both DTD and MIKE SHE. This is also the case with bias, which additionally switches sign. The differences between transpiration produced by MIKE SHE and DTD are smaller than between the bulk ET. Those statistics could indicate that either the retrieval of canopy temperatures is more accurate than the retrieval of soil temperatures (which appears to be corroborated by the discussion in Sect. 7.1.2) or that the canopy component of two-source models is more physically sound. However, there are very few studies in which the two components of bulk LE are validated separately due to the difficulty of obtaining such measurements in situ, especially for comparison with models driven with satellite based observations. Therefore, the accuracy of those components has not yet been
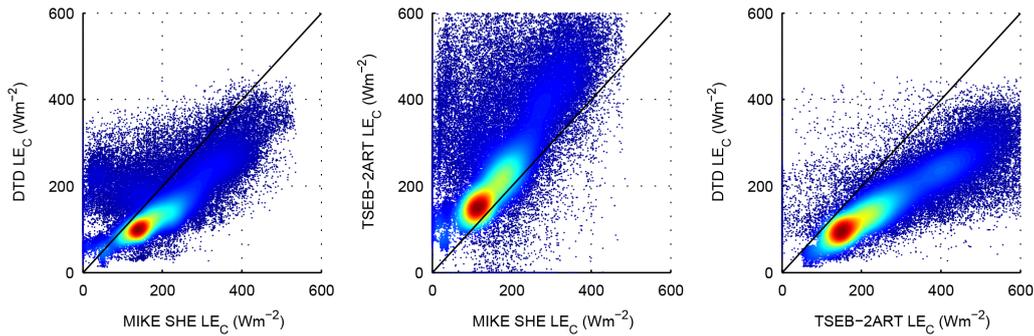
Figure 7. Density scatter plot comparing the vegetation latent heat flux modelled by MIKE SHE and DTD (left), MIKE SHE and TSEB-2ART (centre) and TSEB-2ART and DTD (right). Red colour indicates higher density of points, blue colour lower density.

Table 4. Statistical comparison between MIKE SHE, DTD and TSEB-2ART models for latent heat flux of the canopy (LE_C). Statistics used: correlation coefficient (r), root mean square difference (RMSD), systematic and unsystematic root mean product differences ($RMPD_s$ and $RMPD_u$ respectively), the percentage of mean square difference (MSD) attributed to systematic and unsystematic mean product differences (MPD) (MPD_s / MSD and MPD_u / MSD respectively) and bias. The statistics are in $W\,m^{-2}$, with the exception of MPD_s / MSD and MPD_u / MSD, which are percentages.

		r	RMSD	$RMPD_s$	$RMPD_u$	MPD_s / MSD	MPD_u / MSD	Bias
MIKE SHE–DTD	LE_C	0.67	85	41	75	23	77	29
MIKE SHE–TSEB-2ART	LE_C	0.63	145	104	101	51	49	−96
TSEB-2ART–DTD	LE_C	0.86	133	121	56	82	18	105

established and focused studies are required in order to come to more definitive conclusions.

When considering the causes of the remaining differences in the modelled fluxes, some factors can be directly removed. The three models used many of the same spatial data sets as input: LAI maps, land cover map and meteorological forcing data (air temperature, incoming solar radiation, humidity and wind speed). In addition, DTD and MIKE SHE used the same albedo maps and MIKE SHE was calibrated using the same Aqua MODIS LST observations as used by DTD. The mismatch caused by image misregistration was reduced by applying the median filter over the output maps, although on cloudy days there are many isolated pixels, making the filtering less efficient. The available energy is very highly correlated in all three comparisons, with small RMSD and bias in the case of the two comparisons for which fluxes are estimated at the same hour, so this is also not a major contributor to the differences between the turbulent fluxes.

The remaining major causes of the observed differences in the model outputs could be (1) parametrization used in different land cover classes; (2) the LST input maps estimated by different sensors, in the case of DTD (MODIS) and TSEB-2ART (AATSR), or modelled, in the case of MIKE SHE; and (3) the differences in the modelling approach between the three models even though all of them apply the two-source modelling scheme.

7.1.1 Differences due to parameterization of land cover classes

Figures 8–10 show box plots of the turbulent fluxes, AE and EF split according to the land cover class. The graphs indicate that the statistical distribution of fluxes in the different land cover classes is quite similar among the models, albeit with a large number of outlier points in sensible heat estimations of all models and latent heat estimations of DTD.

When looking at the median and 25th and 75th percentile values of evapotranspiration, the differences do not appear as significant as could be expected from the results shown in Table 2. Considering the DTD model, the ET estimated with the remote sensing model is generally larger than the MIKE SHE estimated ET across all land cover classes. The same can be observed in the case of AE but not in the case of H. This is largely due to the DTD soil heat flux (G) being affected by the LAI multiplicative factor (Table 1). The calculation of G in DTD is dependent on a fraction of the net radiation reaching the soil, which is fundamentally estimated based on the Beer–Lambert law. Therefore, an increase in the value of LAI input into the model leads to a decrease in the magnitude of G, which in turn means that AE is higher and that the magnitude of LE also increases.

In the case of TSEB-2ART, the range between the 25th and 75th percentile values of ET is smaller in croplands and grasslands when compared to MIKE SHE ET, while the median value of conifer forest ET is a bit larger. The range
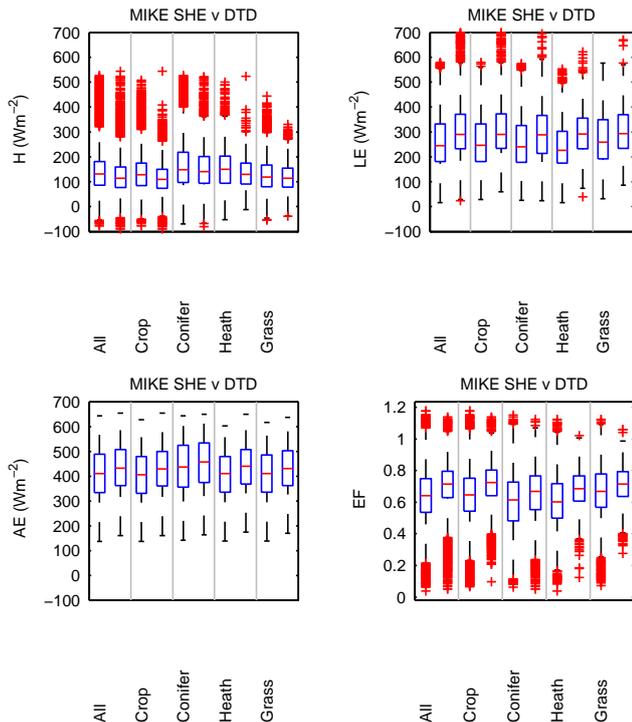
Figure 8. Box plots of sensible heat flux (top left), latent heat flux (top right), net radiation (bottom left) and evaporative fraction (bottom right) modelled by MIKE SHE (leftward box in each category) and DTD (rightward box in each category) and split by land cover class. The red horizontal line indicates the median value with the upper and lower box edges indicating the 75th and 25th percentiles respectively. The whiskers extend to the furthest point within 1.5 times the inter-box range above or bellow the box edges with points beyond that categorized as outliers and marked individually as a red crosses.

Figure 9. Box plots of sensible heat flux (top left), latent heat flux (top right), net radiation (bottom left) and evaporative fraction (bottom right) modelled by MIKE SHE (leftward box in each category) and TSEB-2ART (rightward box in each category) and split by land cover class. The red horizontal line indicates the median value with the upper and lower box edges indicating the 75th and 25th percentiles respectively. The whiskers extend to the furthest point within 1.5 times the inter-box range above or bellow the box edges with points beyond that categorized as outliers and marked individually as a red crosses.

of values between the 25th and 75th percentiles of H is also smaller in TSEB-2ART modelled fluxes, even though the range of AE is generally larger. This could be partially due to the different radiation scheme of TSEB-2ART. Firstly, TSEB-2ART estimates soil albedo as well as canopy albedo and transmissivity based on the spectral properties of the leaves and soil, whereas MIKE SHE uses the surface albedo derived from MODIS reflectances. Secondly, TSEB-2ART is also able to account for different leaf inclination distribution functions. Grass and cereal crops are characterized by a more erectophyll leaf distribution than the spherical distribution characteristic of other vegetation types, such as conifers and some broadleaved forests or shrubs, and implemented in MIKE SHE.

Finally, the time difference between Envisat and Aqua overpasses is clearly visible when comparing TSEB-2ART and DTD LE and AE, but it is not reflected in the values of H. This would indicate that the environment in the Skjern river catchment is not water limited, since the extra energy is generally used for evapotranspiration.

The large number of outliers present in the modelled H values can be partially attributed to the LAI multiplicative factor, especially in the case of TSEB-2ART. The model is quite sensitive to the increase in LAI, due to the physically based radiative transfer modelling, but also due to the large view zenith angle (55°) of the second AATSR LST observation. At this observation angle and with LAI above 4, the model assumes that almost all of the temperature signal comes from the vegetation cover. If the model simulations with LAI larger then 4 are removed (around 18 % of all model runs), the correlation between TSEB-2ART fluxes and the other modelled fluxes increases significantly and the errors decrease, while the comparison between DTD and MIKE SHE remains largely unaffected (Table 5).
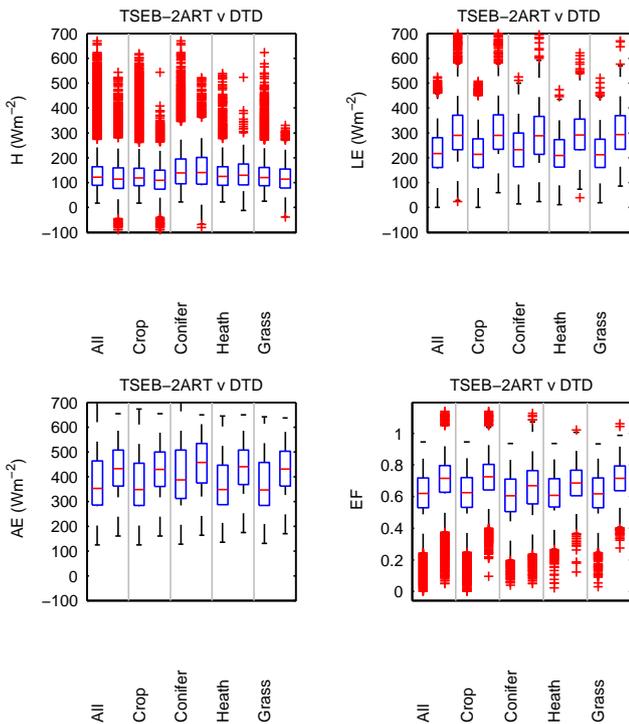
Figure 10. Box plots of sensible heat flux (top left), latent heat flux (top right), net radiation (bottom left) and evaporative fraction (bottom right) modelled by TSEB-2ART (leftward box in each category) and DTD (rightward box in each category) and split by land cover class. The red horizontal line indicates the median value with the upper and lower box edges indicating the 75th and 25th percentiles respectively. The whiskers extend to the furthest point within 1.5 times the inter-box range above or bellow the box edges with points beyond that categorized as outliers and marked individually as a red crosses.

Maps of correlation, RMSD and bias between LE modelled with different model pairs (Fig. 11) show clear spatial patterns, at least partly influenced by land cover (see Fig. 1). There is clear lack of statistically significant correlation between MIKE SHE and the remote sensing models over the forested areas, which is not present when the two remote sensing models are compared. Additionally RMSD is generally higher in forests for all model pairs, but particularly when comparing MIKE SHE and DTD. The bias between LE modelled with TSEB-2ART and DTD is negative throughout the catchment (due to the later overpass of Aqua as compared to Envisat), with the exception of most of the forest areas, where it is slightly positive. Apart from the land cover influenced patterns there is a larger-scale pattern when comparing the outputs of the hydrological and remote sensing models: the correlation is lower, RMSD higher and bias negative in the northern and eastern parts of the catchment (with the exception of the very north-western tip), while the opposite is true in the south-western part. This is due to the first area being classified as having predominantly clayey soil and sec-

ond as having predominantly coarse sandy soil (see Fig. 3 in Greve et al., 2007). This pattern is not visible when TSEB-2ART and DTD are compared which illustrates the sensitivity of the hydrological model to the proper characterization of soil hydraulic properties (which is difficult to do over large areas) and the advantage of the remote sensing models in not requiring this parameter.

7.1.2 Differences due to estimates of LST and its component temperatures

Table 6 shows the statistical comparison between the LST, which is used as input of the remote sensing models and is one of the outputs of the hydrological model, for all the pixels where flux comparison was also performed. The graphical representation is shown in the top left panels of Figs. 12–14. When comparing LST, it must be noted that it is dependent on the viewing geometry, such as VZA, which is quite different between the two satellites and also between the satellites and the hydrological model, for which the sensor is assumed to be directly at nadir. The correlation between the LST from the different model pairs is quite high, with r around 0.9 when comparing the remotely sensed LST from MODIS and AATSR with the MIKE SHE estimates, but reaching 0.97 when the two remotely sensed LSTs are compared. RMSD of LST is quite high, at 4.4 and 5.2 °C in the case of comparing MODIS and AATSR with MIKE SHE and around 3 °C when comparing MODIS with AATSR, although in this case the time difference between the observations should be kept in mind.

Although the high spatial correlation of LST would indicate that the different sources of LST are not a major component in the discrepancies between the modelled fluxes it must be noted that the fluxes are strongly dependent on the LST–T_a gradient and that this dependency is non-linear due to the turbulent transport of heat between the surface and the overlying air (Obukhov, 1971). Due to this non-linearity, the systematic differences in LST between models can potentially lead to larger unsystematic differences in flux estimations. An additional complication in the current study is the fact that the DTD uses the relative temperature difference between night and day observations (Guzinski et al., 2013) and that TSEB-2ART is based on the differences of temperature between the nadir and forward LST observations (Nieto et al., 2013) for flux estimation. It is also interesting to note that although MIKE SHE was calibrated with MODIS Aqua LST observations in the Skjern River catchment, the two satellite based LSTs have a better agreement with each other than with MIKE SHE, despite the overpass times of the two satellites being different. This indicates that the use of LST observations from a satellite sensor, either as a forcing input for MIKE SHE model or for data assimilation (in addition to it being used for calibration), could potentially improve the spatial performance of the hydrological model.

Table 5. Statistical comparison between MIKE SHE, DTD and TSEB-2ART models for sensible and latent heat fluxes (H and LE), available energy (AE) and evaporative fraction (EF) for flux estimates when LAI < 4. Statistics used: correlation coefficient (r), root mean square difference (RMSD), systematic and unsystematic root mean product differences (RMPD$_s$ and RMPD$_u$ respectively), the percentage of mean square difference (MSD) attributed to systematic and unsystematic mean product differences (MPD) (MPD$_s$/MSD and MPD$_u$/MSD respectively) and bias. The statistics for H, LE and AE are in W m^{-2}, with the exception of MPD$_s$/MSD and MPD$_u$/MSD, which are percentages. The statistics for EF are unitless, with the exception of MPD$_s$/MSD and MPD$_u$/MSD, which are percentages.

		r	RMSD	RMPD$_s$	RMPD$_u$	MPD$_s$/MSD	MPD$_u$/MSD	Bias
MIKE SHE–DTD	H	0.43	71	26	66	13	87	23
	LE	0.55	102	51	89	25	75	−51
	AE	0.97	37	28	24	58	42	−28
	EF	0.19	0.20	0.09	0.17	22	78	−0.09
MIKE SHE–TSEB-2ART	H	0.42	62	14	61	5	95	11
	LE	0.65	74	9	73	2	98	−8
	AE	0.95	32	1	32	0	100	0
	EF	0.35	0.16	0.04	0.16	6	94	−0.03
TSEB-2ART–DTD	H	0.45	57	4	57	1	99	2
	LE	0.77	92	69	60	57	43	−69
	AE	0.94	72	64	33	79	21	−64
	EF	0.41	0.15	0.07	0.13	21	79	−0.07
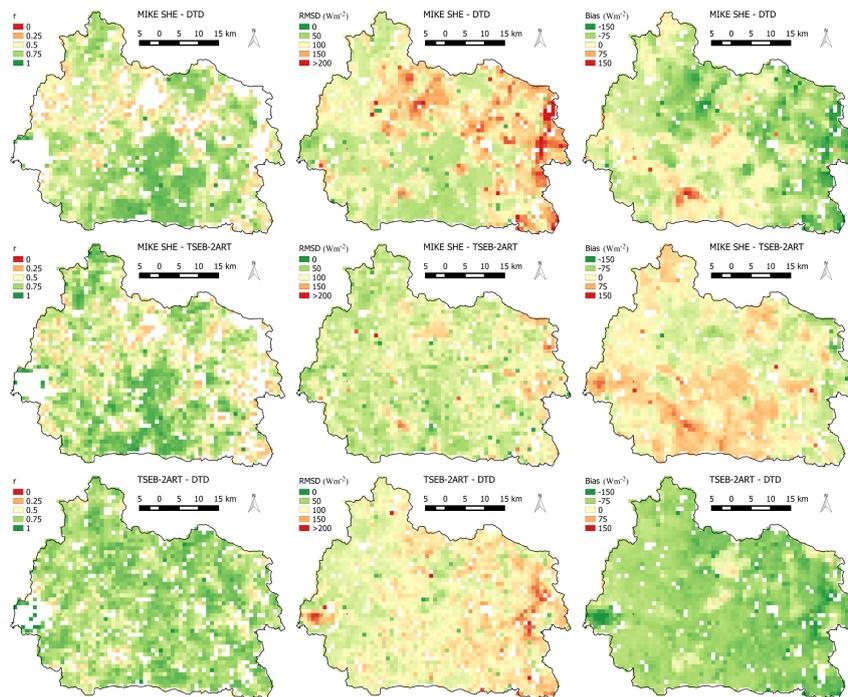
Figure 11. Maps of spatial patterns of correlation (r – first column), RMSD (second column) and bias (third column) calculated with LE output of MIKE SHE–DTD (first row), MIKE SHE–TSEB-2ART (second row) and TSEB-2ART–DTD (third row) in the whole Skjern river catchment. For each pixel in the RMSD and bias maps a mean of the statistics' values was taken from all the points satisfying the conditions stated in Sect. 5 (i.e. the same set of values was used as for producing statistics in Table 2 and density scatter plots in Figs. 2–4). Correlation maps used the same set of values but only the pixels where correlation was significant at 5 % level are shown.

In addition, the canopy, soil and in-canopy air temperatures (T_C, T_S and T_{AC} respectively) estimated by the different models are also compared in Table 6. The estimation of those temperatures could be considered to be an intermediate step during the estimation of the fluxes in the models (Norman et al., 1995; Overgaard and Rosbjerg, 2005), thereby allowing a deeper understanding of the internal model behaviour. The three models apply different methods for es-

Table 6. Statistical comparison between MIKE SHE, DTD and TSEB-2ART models for the land surface temperatures (LST), canopy temperatures (T_C), soil temperatures (T_S) and in-canopy air temperatures (T_{AC}). Statistics used: correlation coefficient (r), root mean square difference (RMSD), systematic and unsystematic root mean product differences (RMPD$_s$ and RMPD$_u$ respectively), the percentage of mean square difference (MSD) attributed to systematic and unsystematic mean product differences (MPD) (MPD$_s$/MSD and MPD$_u$/MSD respectively) and bias. LST comes from Aqua MODIS observations in the case of DTD and nadir view Envisat AATSR observations in the case of TSEB-2ART, and is modelled in the case of MIKE SHE. The other temperatures are estimated by all models. The statistics are in °C, with the exception of MPD$_s$/MSD and MPD$_u$/MSD, which are percentages.

		r	RMSD	RMPD$_s$	RMPD$_u$	MPD$_s$/MSD	MPD$_u$/MSD	Bias
MIKE SHE–DTD	LST	0.89	4.4	3.0	3.2	48	52	3.0
	T_C	0.88	3.1	0.5	3.1	2	98	−0.1
	T_S	0.63	7.2	2.8	6.6	15	85	−0.2
	T_{AC}	0.93	2.5	0.6	2.4	7	93	−0.6
MIKE SHE–TSEB-2ART	LST	0.91	5.2	4.3	2.9	69	31	4.2
	T_C	0.87	7.8	6.9	3.8	77	23	6.6
	T_S	0.72	9.8	7.2	6.6	54	46	−4.5
	T_{AC}	0.94	5.2	4.5	2.5	77	23	4.2
TSEB-2ART–DTD	LST	0.97	2.9	2.3	1.8	61	39	−2.1
	T_C	0.90	8.5	7.8	3.3	85	15	−7.5
	T_S	0.69	9.3	4.2	8.3	20	80	3.1
	T_{AC}	0.94	6.5	6.0	2.5	85	15	−5.8
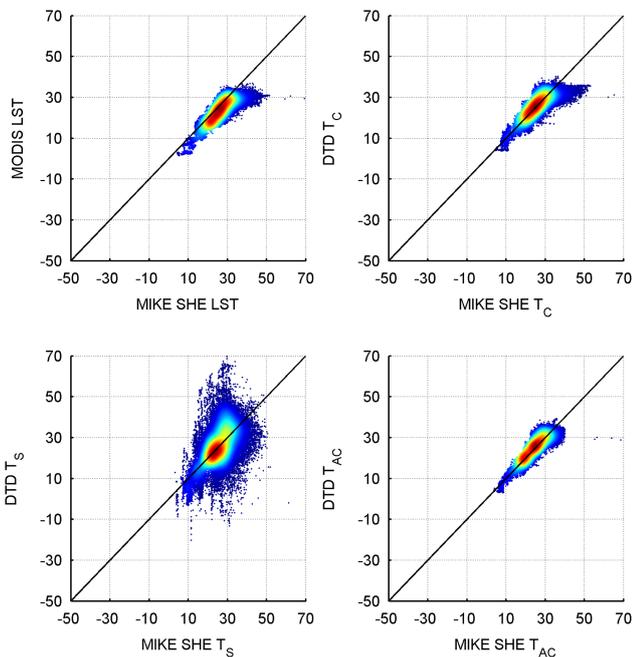
Figure 12. Density scatter plot of over 95 000 points comparing land surface temperature (top left), canopy temperature (top right), soil temperature (bottom left) and in-canopy air temperature (bottom right). Land surface temperature on the x axis was modelled by MIKE SHE and on the y axis came from daytime observations from the MYD11A1 MODIS product. The other temperatures were modelled by both MIKE SHE and DTD. Red colour indicates higher density of points, blue colour lower density.
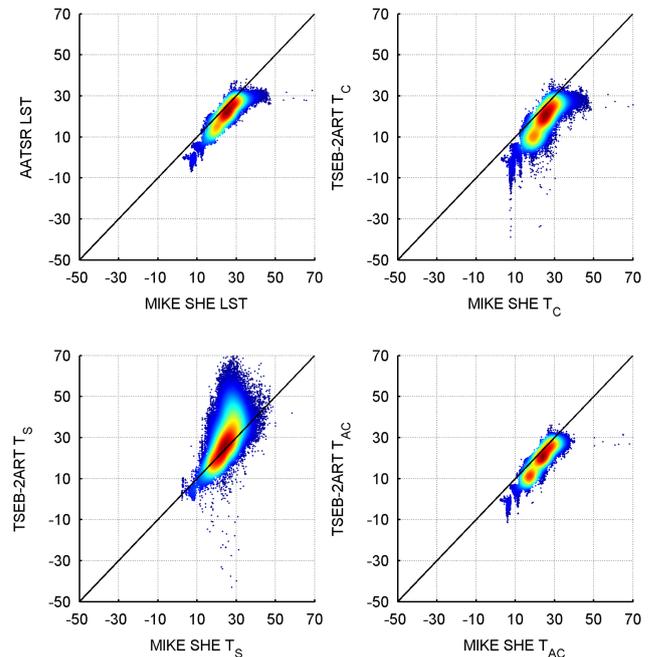
Figure 13. Density scatter plot of over 95 000 points comparing land surface temperature (top left), canopy temperature (top right), soil temperature (bottom left) and in-canopy air temperature (bottom right). Land surface temperature on the x axis was modelled by MIKE SHE and on the y axis came from nadir observations by AATSR sensor on the Envisat satellite. The other temperatures were modelled by both MIKE SHE and TSEB-2ART. Red colour indicates higher density of points, blue colour lower density.
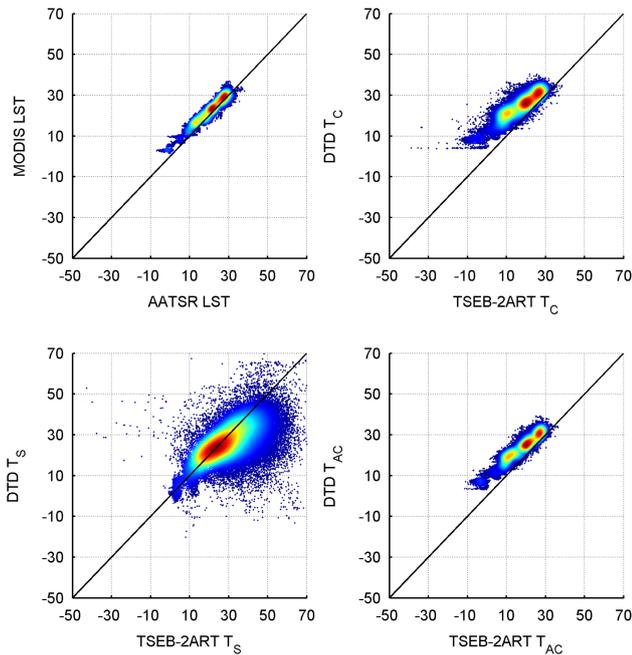
Figure 14. Density scatter plot of over 95 000 points comparing land surface temperature (top left), canopy temperature (top right), soil temperature (bottom left) and in-canopy air temperature (bottom right). Land surface temperature on the x axis came from nadir observations by AATSR sensor on the Envisat satellite and on the y axis came from daytime observations from the MYD11A1 MODIS product. The other temperatures were modelled by both TSEB-2ART and DTD. Red colour indicates higher density of points, blue colour lower density.

timating those temperatures. In the case of DTD, temperatures are not used directly during the flux estimation (the time differential temperature observations are used), but are derived as a final step when all the flux and resistance values are already established using rearranged Eqs. (A27), (A29) and (A33) from Guzinski et al. (2014). TSEB-2ART uses the viewing geometry of the two observation angles within a radiative transfer model framework to estimate T_C and T_S, which, together with the resistances to heat transport, are then used to calculate T_{AC} and the fluxes using the TSEB formulations (Nieto et al., 2013; Kustas and Norman, 1997; Norman et al., 1995). In MIKE SHE, the temperatures, together with humidity, are derived by solving a set of 10 linear equations involving the resistances and AE as parameters, after which the turbulent fluxes are derived (Overgaard and Rosbjerg, 2005).

Despite those three different methods the correlation between the temperatures is quite high (Table 6) which is surprising considering the much lower correlation between the modelled turbulent fluxes. Again, this is probably caused by the non-linearity between the gradient of temperatures and the heat flux due to turbulence. It could also be due to the heating effect that the interaction between soil temperature

and heat fluxes produces for the temperature of the air at the canopy interface, when the resistances are configured in series. The highest correlation, above 0.93 for all the pairs, is for T_{AC}, and the lowest, ranging from 0.63 to 0.72, is for T_S. Overall the two remote sensing models have most similar spatial patterns of T_C and T_{AC}, and MIKE SHE and TSEB-2ART have most similar spatial pattern of T_S.

Furthermore, since the TSEB-2ART model relies on the differences observed between the nadir and forward LST of AATSR in order to derive T_C and T_S, it is sensitive to errors in the estimation of LST at the two viewing angles. Those errors might be significant if, for example, atmospheric water vapour is not properly characterized and accounted for, due to the different optical path lengths of the forward (VZA $= 55°$) and nadir observations. Since the atmospheric path length of the forward view is longer, the a priori uncertainty in the estimation of forward LST is higher than in the case of nadir LST. In Figs. 13 and 14 it can be seen that this occurs in a number of cases, mostly leading to severe underestimation of T_C and T_{AC} and overestimation of T_S when compared to the other models. This large bias in TSEB-2ART estimated T_C is also present in the statistical comparison in Table 6. It also appears that MIKE SHE produces higher values at large magnitudes of T_C (Figs. 12 and 13), although those difference are not severe. This relative overestimation of T_C is reflected in MIKE SHE LST scatter plots, which indicates that it happens at high LAI values when vegetation cover fraction is close to 1.

7.1.3 Differences due to the modelling approach

Although there are differences in the estimated temperatures that could lead to larger unsystematic differences in the fluxes estimates, it is likely that there are also other factors contributing to the inconsistencies between fluxes. One of the factors could be the methodology employed by the different models for splitting of the available energy into the sensible and latent heat fluxes and in particular the way they estimate the resistances to heat and moisture transport. The two remote sensing models ensure the land surface energy balance by calculating latent heat flux as the residual of the other fluxes, i.e. LE $=$ AE $- H$ (Norman et al., 1995). The hydrological model, on the other hand, derives the latent and sensible heat fluxes concurrently (Overgaard and Rosbjerg, 2005). In addition the resistance network for LE in MIKE SHE has two extra resistance components compared to the resistance network for H: the resistance to soil evaporation and the stomata resistance to transpiration. Both of them depend on the soil moisture as modelled by the hydrological component of MIKE SHE. If the spatial patterns of the soil moisture estimated with MIKE SHE do not correspond closely to the spatial patterns seen by the satellites this could lead to the different spatial patterns of the estimated fluxes.

Another possible factor for the observed differences between the estimated fluxes could be the actual formulations

Table 7. Statistical comparison between MIKE SHE, DTD and TSEB-2ART models for sensible and latent heat fluxes (H and LE), available energy (AE) and evaporative fraction (EF) for model runs with resistance equations taken from Choudhury and Monteith (1988). Statistics used: correlation coefficient (r), root mean square difference (RMSD), systematic and unsystematic root mean product differences (RMPD$_s$ and RMPD$_u$ respectively), the percentage of mean square difference (MSD) attributed to systematic and unsystematic mean product differences (MPD) (MPD$_s$/MSD and MPD$_u$/MSD respectively) and bias. The statistics for H, LE and AE are in W m^{-2}, with the exception of MPD$_s$/MSD and MPD$_u$/MSD, which are percentages. The statistics for EF are unitless, with the exception of MPD$_s$/MSD and MPD$_u$/MSD, which are percentages.

		r	RMSD	RMPD$_s$	RMPD$_u$	MPD$_s$/MSD	MPD$_u$/MSD	Bias
MIKE SHE–DTD	H	0.38	84	30	78	13	87	30
	LE	0.52	115	58	100	26	74	−58
	AE	0.97	36	28	23	59	41	−27
	EF	0.17	0.23	0.11	0.20	22	78	−0.11
MIKE SHE–TSEB-2ART	H	0.30	80	6	80	1	99	−2
	LE	0.57	85	11	84	2	98	1
	AE	0.95	34	5	34	2	98	−3
	EF	0.23	0.18	0.02	0.18	1	99	0.00
TSEB-2ART–DTD	H	0.29	86	23	83	7	93	22
	LE	0.63	119	88	80	55	45	−87
	AE	0.94	71	61	36	75	25	−61
	EF	0.22	0.21	0.11	0.18	29	71	−0.11

used for resistances of heat transfer between the soil, vegetation, in-canopy air and above-canopy air. While the two remote sensing models use equations based on Norman et al. (1995), the hydrological model uses equations suggested by Choudhury and Monteith (1988). To evaluate whether those different formulations could be the reason for the fluxes estimated with the remote sensing models being more similar to each other than to the fluxes estimated with the hydrological model, it was decided to run the remote sensing models with the resistance equations taken from Choudhury and Monteith (1988).

The results are presented in Table 7. The correlation between the turbulent fluxes produced by all model pairs has decreased when compared to results in Table 2. This is surprising, as it could be expected that using the same resistance formulations would increase the correlation between the remote sensing models and the hydrological model. There was also a small increase in RMSD (with the exception of the MIKE SHE–TSEB-2ART H comparison). In addition, the number of valid pixels has been reduced from over 95 000 to over 83 000. This could indicate that the Norman et al. (1995) resistance formulations produce more realistic values than the Choudhury and Monteith (1988) formulations, which would point to the possibility of updating the equations used in the SW–ET module of MIKE SHE. Even when using the Choudhury and Monteith (1988) resistance formulations, the LE modelled with DTD and TSEB-2ART has the highest correlation. However, in the case of H, MIKE SHE and DTD produced the most correlated flux estimates, while the correlation of EF was very similar between the MIKE SHE–TSEB-2ART and DTD–TSEB-2ART model pairs.

Finally, Fig. 15 illustrates the effect of modifying the R_S formulation in the DTD, as proposed in Eq. (4). The R_S values estimated with DTD and TSEB-2ART are compared for all the pixels where the flux comparison was performed. In this case, the TSEB-2ART derived R_S can be thought of as the "true" value, since it uses the original R_S equation (Eq. 3), with T_S and T_C derived directly through the inversion of the RTM. The overestimation of R_S by DTD, visible as a "bubble" in the left panel for DTD R_S values between 150 and 200 s m^{-1}, is due to misparametrization of the differences between the canopy and soil temperatures in the original DTD formulation, and is mostly present in the coniferous forest. In the right panel this overestimation is less pronounced, indicating that the new R_S equation is better at parametrizing this temperature difference. The correlation parameter between the two resistances has increased from 0.62 in the case of the old formulation to 0.68 in the case of the new one, while the RMSD has decreased by around 15 %, from 36 to 31 s m^{-1}.

7.2 Temporal patterns

Both remote sensing models are reasonably accurate in matching MIKE SHE catchment-wide estimates of evapotranspiration, with the seasonal curve clearly visible for both models (Fig. 5) and reflecting the MIKE SHE seasonality well. The DTD model tends to produce larger latent heat fluxes before and after the growing season. This is probably due to the larger magnitude of AE estimated by this model (see Table 2), which is mostly assigned to LE, as it is calculated as a residual of the surface energy balance. On the other hand TSEB-2ART matches MIKE SHE fluxes quite
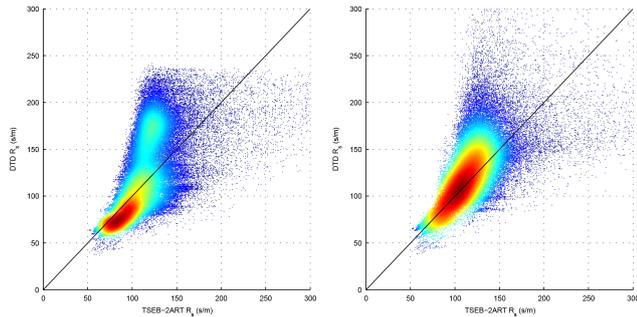
Figure 15. Density scatter plot of over 95 000 points comparing resistance of heat transfer from the soil surface (R_S) modelled by the TSEB-2ART and DTD. In the left panel the original DTD formulation is used (Norman et al., 2000) and in the right panel the new formulation is used (Eq. 4). Red colour indicates higher density of points, blue colour lower density.

well during that period. During the growing season there is the largest mismatch between LE modelled by the remote sensing models and the hydrological model. This is partly due to the fluxes having the largest magnitude during this time, but also due to LAI having a large value which blocks the temperature signal from the soil surface (see Sect. 7.1.1).

Figure 5 also highlights another weakness of the remote sensing models, namely that they only produce results on clear sky days. The great majority of latent heat fluxes estimated by the remote sensing models, and by the hydrological model on the same dates as the remote sensing models, lie above the line representing an average, all-weather ET for a particular DOY for all the years under study. This is also true when considering an 8-year averaged potential ET. The reason is because in the Skjern River environment the evapotranspiration is mainly driven by availability of energy (and not of water), and therefore on clear sky days the evapotranspiration will be higher than average. This has to be taken into account when extrapolating temporal patterns of evapotranspiration derived purely by the remote sensing input based models.

There are a couple of cases where the clear sky evapotranspiration modelled by MIKE SHE is much below the average line, even though the remote sensing models estimate much higher latent heat fluxes on those days. This most probably corresponds to days with soil drier than normal and could indicate: (1) a problem of the hydrological model in estimating the moisture of the upper layer of the soil or of the root zone during dry conditions, or (2) be related to uncertainties in the interpolated rainfall data due to omission by the rain gauges of local convective rainfall during the summer period.

8 Conclusions and outlook

Two remote sensing models and one hydrological model were run over an area covering a river catchment in west-

ern Denmark and the spatial and temporal patterns of the modelled evapotranspiration were compared. The spatial patterns of latent and sensible heat fluxes as well as EF produced by the remote sensing models were more strongly correlated with each other than the patterns produced by either of the remote sensing models compared to the hydrological model. This was the case even though the two remote sensing models use both different data (MODIS and AATSR LST) and different approaches to estimating the fluxes and, additionally, those estimates were produced at different time of the day, due to different overpass times of satellites. This indicates that the remote sensing models might contain some additional information that is not currently present in the hydrological model. At the same time, the temporal patterns of evapotranspiration produced by both of the remote sensing models and the hydrological model were strongly correlated, with relatively small RMSD and small bias. Those observations would appear to support the hypothesis that the remote sensing models would better represent the spatial patterns of evapotranspiration present throughout the catchment, while the hydrological model would better represent the catchment-wide evapotranspiration.

This points towards a possibility of using the remotely sensed evapotranspiration to improve the spatial accuracy of distributed, physically based hydrological models. This could be achieved either through using the estimated latent heat flux as one of the calibrating parameters or through data assimilation during the model run. Certain attempts at incorporating spatial distributed data derived through remote sensing into hydrological models, either through data assimilation or calibration, have already been made but they were mostly focused on soil moisture (e.g. Draper et al., 2011; Corato et al., 2013), LST (Stisen et al., 2011a; Ridler et al., 2012) or LAI (Boegh et al., 2004). Pipunic et al. (2008) have looked at assimilating simulated H and LE estimates into a land surface model, however this was done with a one-dimensional single column model, i.e. without considering spatial patterns. Others have assimilated ET maps into distributed hydrological models but the impact of that assimilation was inconclusive (Pan et al., 2008; Schuurmans et al., 2011). Therefore, further studies are needed to establish whether ET, and in particular its spatial distribution, would bring any additional information beyond what is provided by soil moisture or LST estimates alone. In the case of MIKE SHE it might also be useful to use the T_C, T_S and T_{AC} estimates from the remote sensing models to constraint the number of unknowns that need to be addressed in the model. Methodologies for validating the accuracy of spatial patterns at the catchment scale, while at the same time remaining independent of the model used, would also have to be investigated.

Edited by: H. Cloke

References

Anderson, M. C., Allen, R. G., Morse, A., and Kustas, W. P.: Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources, Remote Sens. Environ., 122, 50–65, 2012.

Barker, F., Soh, Y., and Evans, R.: Properties of the geometric mean functional relationship, Biometrics, 44, 279–281, 1988.

Boegh, E., Thorsen, M., Butts, M., Hansen, S., Christiansen, J., Abrahamsen, P., Hasager, C., Jensen, N., van der Keur, P., Refsgaard, J., Schelde, K., Soegaard, H., and Thomsen, A.: Incorporating remote sensing data in physically based distributed agro-hydrological modelling, J. Hydrol., 287, 279–299, doi:10.1016/j.jhydrol.2003.10.018, 2004.

Boegh, E., Poulsen, R., Butts, M., Abrahamsen, P., Dellwik, E., Hansen, S., Hasager, C. B., Ibrom, A., Loerup, J.-K., Pilegaard, K., and Soegaard, H.: Remote sensing based evapotranspiration and runoff modeling of agricultural, forest and urban flux sites in Denmark: from field to macro-scale, J. Hydrol., 377, 300–316, 2009.

Boulet, G., Olioso, A., Ceschia, E., Marloie, O., Coudert, B., Rivalland, V., Chirouze, J., and Chehbouni, G.: An empirical expression to relate aerodynamic and surface temperatures for use within single-source energy balance models, Agr. Forest Meteorol., 161, 148–155, 2012

Campbell, G. and Norman, J.: Introduction to Environmental Biophysics, Springer, New York, New York, USA, 1998.

Chehbouni, A., Nouvellon, Y., Lhomme, J.-P., Watts, C., Boulet, G., Kerr, Y., Moran, M., and Goodrich, D.: Estimation of surface sensible heat flux using dual angle observations of radiative surface temperature, Agr. Forest Meteorol., 108, 55–65, 2001.

Chirouze, J., Boulet, G., Jarlan, L., Fieuzal, R., Rodriguez, J. C., Ezzahar, J., Er-Raki, S., Bigeard, G., Merlin, O., Garatuza-Payan, J., Watts, C., and Chehbouni, G.: Intercomparison of four remote-sensing-based energy balance methods to retrieve surface evapotranspiration and water stress of irrigated fields in semi-arid climate, Hydrol. Earth Syst. Sci., 18, 1165–1188, doi:10.5194/hess-18-1165-2014, 2014.

Choudhury, B., Idso, S., and Reginato, R.: Analysis of an empirical model for soil heat flux under a growing wheat crop for estimating evaporation by an infrared-temperature based energy balance equation, Agr. Forest Meteorol., 39, 283–297, 1987.

Choudhury, B. J. and Monteith, J. L.: A four-layer model for the heat budget of homogeneous land surfaces, Q. J. Roy. Meteor. Soc., 114, 373–398, doi:10.1002/qj.49711448006, 1988.

Coll, C., Caselles, V., Galve, J. M., Valor, E., Niclòs, R., and Sánchez, J. M.: Evaluation of split-window and dual-angle correction methods for land surface temperature retrieval from Envisat/Advanced Along Track Scanning Radiometer (AATSR) data, J. Geophys. Res., 111, D12105, doi:10.1029/2005JD006830, 2006.

Conradt, T., Wechsung, F., and Bronstert, A.: Three perceptions of the evapotranspiration landscape: comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and sub-basin water balances, Hydrol. Earth Syst. Sci., 17, 2947–2966, doi:10.5194/hess-17-2947-2013, 2013.

Corato, G., Matgen, P., Giustarini, L., and Fenicia, F.: On the effects of hydrological model structure on soil moisture data assimilation, EGU General Assembly, 7–12 April 2013, Vienna, Austria, EGU2013-11602, 2013.

Crago, R. D.: Conservation and variability of the evaporative fraction during the daytime, J. Hydrol., 180, 173–194, doi:10.1016/0022-1694(95)02903-6, 1996.

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., and Sciarra, R.: The global monitoring for environment and security (GMES) sentinel-3 mission, Remote Sens. Environ., 120, 37–57, 2012.

Draper, C., Mahfouf, J.-F., Calvet, J.-C., Martin, E., and Wagner, W.: Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France, Hydrol. Earth Syst. Sci., 15, 3829–3841, doi:10.5194/hess-15-3829-2011, 2011.

François, C.: The potential of directional radiometric temperatures for monitoring soil and leaf temperature and soil moisture status, Remote Sens. Environ., 80, 122–133, 2002.

Graham, D. N. and Butts, M. B.: Flexible, integrated watershed modelling with MIKE SHE, in: Watershed Models, edited by: Singh, V. P. and Frevert, D. K., Taylor and Francis Group, Boca Raton, FL, USA, 245–272, 2005.

Greve, M. H., Greve, M. B., Bøcher, P. K., Balstrøm, T., Breuning-Madsen, H., and Krogh, L.: Generating a Danish raster-based topsoil property map combining choropleth maps and point information, Geogr. Tidsskr., 107, 1–12, 2007.

Guzinski, R., Anderson, M. C., Kustas, W. P., Nieto, H., and Sandholt, I.: Using a thermal-based two source energy balance model with time-differencing to estimate surface energy fluxes with day–night MODIS observations, Hydrol. Earth Syst. Sci., 17, 2809–2825, doi:10.5194/hess-17-2809-2013, 2013.

Guzinski, R., Nieto, H., Jensen, R., and Mendiguren, G.: Remotely sensed land-surface energy fluxes at sub-field scale in heterogeneous agricultural landscape and coniferous plantation, Biogeosciences, 11, 5021–5046, doi:10.5194/bg-11-5021-2014, 2014.

Jensen, K. H. and Illangasekare, T. H.: HOBE: a hydrological observatory, Vadose Zone J., 10, 1–7, 2011.

Ji, L. and Gallo, K.: An agreement coefficient for image comparison, Photogramm. Eng. Rem. S., 72, 823–833, 2006.

Ji, L., Gallo, K., Eidenshink, J., and Dwyer, J.: Agreement evaluation of AVHRR and MODIS 16-day composite NDVI data sets, Int. J. Remote Sens., 29, 4839–4861, 2008.

Jiang, L. and Islam, S.: Estimation of surface evaporation map over Southern Great Plains using remote sensing data, Water Resour. Res., 37, 329–340, doi:10.1029/2000WR900255, 2001.

Kalma, J. D., McVicar, T. R., and McCabe, M. F.: Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data, Surv. Geophys., 29, 421–469, 2008.

Kustas, W. and Anderson, M.: Advances in thermal infrared remote sensing for land surface modeling, Agr. Forest Meteorol., 149, 2071–2081, 2009.

Kustas, W. P. and Goodrich, D. C.: Preface [to special section on Monsoon '90 Multidisciplinary Experiment], Water Resour. Res., 30, 1211–1225, doi:10.1029/93WR03068, 1994.

Kustas, W. P. and Norman, J. M.: A two-source approach for estimating turbulent fluxes using multiple angle thermal infrared observations, Water Resour. Res., 33, 1495–1508, 1997.

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, Agr. Forest Meteorol., 94, 13–29, 1999.

Liang, S.: Narrowband to broadband conversions of land surface albedo I Algorithms, Remote Sens. Environ., 76, 213–238, 2001.

Matsushima, D.: Relations between Aerodynamic Parameters of Heat Transfer and Thermal-Infrared Thermometry in the Bulk Surface Formulation, J. Meteorol. Soc. Jpn., Ser. II, 83, 373–389, doi:10.2151/jmsj.83.373, 2005.

Nieto, H., Sandholt, I., Rasmussen, M., Herbst, M., Ringgaard, R., and Friborg, T.: Optical and thermal multiangular measurements for the estimation of evapotranspiration on an eddy covariance flux tower, in: 3rd International Symposium on Recent Advances in Quantitative Remote Sensing, 27 September–1 October 2010, Valencia, Spain, edited by: Sobrino, J. A., 2010a.

Nieto, H., Sandholt, I., Ringgaard, R., Herbst, M., Stisen, S., and Friborg, T.: Combining ENVISAT AATSR and MERIS imagery for the retrieval of evapotranspiration in a two-source resistance energy balance model, in: Remote Sensing and Hydrology 2010 Symposium, 27–30 September 2010, Jackson Hole, USA, 2010b.

Nieto, H., Guzinski, R., Jensen, R., Sandholt, I., and Jensen, K.: TSEBRTM: Coupling a canopy Radiative Transfer model with a Two Source Energy Balance Model for the Estimation of Surface Energy Fluxes with Dual-Angle Land Surface Temperature, TR32-HOBE International Symposium, 11–14 March 2013, Bonn, Germany, S4-128, available at: http://tr32meeting.uni-koeln.de/images/abstract/s4/TR32-HOBE-S4-Thursday-1140-1200_Nieto.pdf, last access: 27 May 2014, 2013.

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, Agr. Forest Meteorol., 77, 263–293, 1995.

Norman, J. M., Kustas, W., Prueger, J., and Diak, G.: Surface flux estimation using radiometric temperature: a dual-temperature-difference method to minimize measurement errors, Water Resour. Res., 36, 2263–2274, 2000.

Obukhov, A.: Turbulence in an atmosphere with a non-uniform temperature, Bound.-Lay. Meteorol., 2, 7–29, 1971.

Overgaard, J. and Rosbjerg, D.: Energy-Based Land-Surface Modelling: New Opportunities in Integrated Hydrological Modelling, Technical University of Denmark – Danmarks Tekniske Universitet, Department of Hydrodynamics and Water Resocurces – Strømningsmekanik og Vandressourcer, Copenhagen, Denmark, 2005.

Pan, M., Wood, E. F., Wójcik, R., and McCabe, M. F.: Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation, Remote Sens. Environ., 112, 1282–1294, doi:10.1016/j.rse.2007.02.039, 2008.

Peng, J., Borsche, M., Liu, Y., and Loew, A.: How representative are instantaneous evaporative fraction measurements of daytime fluxes?, Hydrol. Earth Syst. Sci., 17, 3913–3919, doi:10.5194/hess-17-3913-2013, 2013.

Pipunic, R., Walker, J., and Western, A.: Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: a comparative synthetic study, Remote Sens. Environ., 112, 1295–1305, doi:10.1016/j.rse.2007.02.038, 2008.

Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, J. Hydrol., 198, 69–97, 1997.

Ridler, M. E., Sandholt, I., Butts, M., Lerer, S., Mougin, E., Timouk, F., Kergoat, L., and Madsen, H.: Calibrating a soil–vegetation–atmosphere transfer model with remote sensing estimates of surface temperature and soil surface moisture in a semi arid environment, J. Hydrol., 436–437, 1–12, doi:10.1016/j.jhydrol.2012.01.047, 2012.

Scarpino, M. and Cardaci, M.: Envisat-1 product specifications, Vol. 7, AATSR products specifications, ESA Doc Ref, Tech. rep., PO-RS-MDA-GS-2009, 2009.

Schuurmans, J. M., van Geer, F. C., and Bierkens, M. F. P.: Remotely sensed latent heat fluxes for model error diagnosis: a case study, Hydrol. Earth Syst. Sci., 15, 759–769, doi:10.5194/hess-15-759-2011, 2011.

Shuttleworth, W. J. and Wallace, J.: Evaporation from sparse crops-an energy combination theory, Q. J. Roy. Meteor. Soc., 111, 839–855, 1985.

Stisen, S., Sandholt, I., Nørgaard, A., Fensholt, R., and Jensen, K. H.: Combining the triangle method with thermal inertia to estimate regional evapotranspiration, applied to MSG-SEVIRI data in the Senegal River basin, Remote Sens. Environ., 112, 1242–1255, 2008.

Stisen, S., McCabe, M. F., Refsgaard, J. C., Lerer, S., and Butts, M. B.: Model parameter analysis using remotely sensed pattern information in a multi-constraint framework, J. Hydrol., 409, 337–349, 2011a.

Stisen, S., Sonnenborg, T. O., Højberg, A. L., Troldborg, L., and Refsgaard, J. C.: Evaluation of climate input biases and water balance issues using a coupled surface–subsurface model, Vadose Zone J., 10, 37–53, 2011b.

Su, Z.: The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes, Hydrol. Earth Syst. Sci., 6, 85–100, doi:10.5194/hess-6-85-2002, 2002.

Verhoef, W., Jia, L., Xiao, Q., and Su, Z.: Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies, IEEE T. Geosci. Remote, 45, 1808–1822, 2007.