Intercomparison of four remote-sensing-based energy balance methods to retrieve surface evapotranspiration and water stress of irrigated fields in semi-arid climate

Jonas Chirouze, G. Boulet, L. Jarlan, R. Fieuzal, J. C. Rodriguez, J. Ezzahar,
S. Er Raki, G. Bigeard, O. Merlin, J. Garatuza-Payan, et al.

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hydrology and Earth System Sciences

Intercomparison of four remote-sensing-based energy balance methods to retrieve surface evapotranspiration and water stress of irrigated fields in semi-arid climate

J. Chirouze[1], G. Boulet[1], L. Jarlan[1], R. Fieuzal[1], J. C. Rodriguez[2], J. Ezzahar[3], S. Er-Raki[4], G. Bigeard[1], O. Merlin[1], J. Garatuza-Payan[5], C. Watts[2], and G. Chehbouni[1]

[1]Centre d'Etudes Spatiales de la Biosphère, UPS, CNRS, CNES, IRD – UMR5126, Toulouse, France
[2]Universidad de Sonora, Hermosillo, Sonora, México
[3]Centre National de l'Energie, des Sciences et des Techniques Nucléaires, Kénitra, Morocco
[4]LP2M2E, Département de physique, Faculté des Sciences et Techniques, Université Cadi Ayyad, Marrakech, Morocco
[5]Instituto Tecnológico de Sonora, Ciudad Obregón, Sonora, México

Correspondence to: G. Boulet (gilles.boulet@cesbio.cnes.fr)

Abstract. Instantaneous evapotranspiration rates and surface water stress levels can be deduced from remotely sensed surface temperature data through the surface energy budget. Two families of methods can be defined: the contextual methods, where stress levels are scaled on a given image between hot/dry and cool/wet pixels for a particular vegetation cover, and single-pixel methods, which evaluate latent heat as the residual of the surface energy balance for one pixel independently from the others. Four models, two contextual (S-SEBI and a modified triangle method, named VIT) and two single-pixel (TSEB, SEBS) are applied over one growing season (December–May) for a 4 km × 4 km irrigated agricultural area in the semi-arid northern Mexico. Their performance, both at local and spatial standpoints, are compared relatively to energy balance data acquired at seven locations within the area, as well as an uncalibrated soil–vegetation–atmosphere transfer (SVAT) model forced with local in situ data including observed irrigation and rainfall amounts. Stress levels are not always well retrieved by most models, but S-SEBI as well as TSEB, although slightly biased, show good performance. The drop in model performance is observed for all models when vegetation is senescent, mostly due to a poor partitioning both between turbulent fluxes and between the soil/plant components of the latent heat flux and the available energy. As expected, contextual methods perform well when contrasted soil moisture

and vegetation conditions are encountered in the same image (therefore, especially in spring and early summer) while they tend to exaggerate the spread in water status in more homogeneous conditions (especially in winter). Surface energy balance models run with available remotely sensed products prove to be nearly as accurate as the uncalibrated SVAT model forced with in situ data.

1 Context and objectives

Evapotranspiration is the largest water loss component of continental surfaces (Oki and Kanae, 2006; Shiklomanov, 1998). In semi-arid areas, more than 80 % of the annual available water is lost through evapotranspiration (Chehbouni et al., 2008a). In most countries, and especially the least developed ones, irrigation consumes the largest amount of water. It often represents more than 80 % of all uses (World Water Assessment Programme, 2012). In many cases, water use efficiency is lower than 50 %. For countries facing water shortage, or likely to suffer from more frequent drought spills under climate change scenarios, there is a great need to rationalize this use, and therefore to monitor more closely the different terms of the water budget (Oki and Kanae, 2006). Among them, evapotranspiration is of major importance.

Although the water budget can be fairly easily monitored by the farmer at plot scale, it is much more difficult for regional authorities or national planners to monitor water allocation and use at the relevant scales, i.e., the irrigated perimeter and the basin scales. To do so, remote sensing data is increasingly used, because it allows for the description of the surface at most scales ranging from plot to region, at a temporal scale no greater than a few weeks, which is particularly important to follow the growth of vegetation.

There are many methods that compute evapotranspiration with the help of remote sensing data (Courault et al., 2005). Some methods rely only on the atmospheric demand through different radiation and atmospheric variables derived from remote sensing (Venturini et al., 2008). Since evapotranspiration depends largely on the availability of water, which is often greater in the root zone than at the soil surface, surface losses depend mostly on the intensity of transpiration. Many methods, especially those designed for irrigated agriculture, which is usually not short of water, compute a potential or reference evapotranspiration rate and weight the latent heat flux by the amount of vegetation present for each pixel, through the use of a vegetation index such as the NDVI (normalized differential vegetation index; Cleugh et al., 2007). But these methods are of little help when vegetation suffers from water stress, which means that they have little applicability for operational management of irrigation water when the objective is to prevent stress.

Since evaporation is the most efficient way to dissipate extra energy at the surface, there is a tight coupling between water availability and surface temperature under water stress conditions. Therefore, the use of information in the thermal infrared (TIR) domain (3–15 µm) is an appropriate way to assess actual evaporation and soil moisture status at relevant space and timescales (Boulet et al., 2007; Hain et al., 2009). Methods to estimate evapotranspiration from satellite data in the TIR domain are reviewed in Kalma et al. (2008) and Kustas and Anderson (2009). Estimates of instantaneous evaporation rates at the time of the satellite overpass can be converted into daily values through the use of extrapolation algorithms based on the diurnal self-preservation of quantities like the evaporative fraction, i.e., the ratio between the latent heat flux and the available energy (Delogu et al., 2012).

TIR-based methods to compute instantaneous evapotranspiration can be broadly divided into two families: contextual and single-pixel methods. Contextual methods cover all approaches based on the simultaneous presence, at the time of acquisition, of hot/dry and cold/wet pixels within the satellite image, for a sufficiently large range of vegetation covers or surface states. The latter are usually inferred in other optical wavelengths so that for given vegetation type/extent or a given value of the scaling surface parameter one can securely associate contrasted temperature patterns with contrasted soil moisture conditions. Those methods use synchronous information of several pixels of a given image. Each intermediate temperature for a given vegetation class is then scaled

to these extremes to provide an intermediate water stress condition.

On the other hand, single-pixel methods mostly solve the surface energy budget for each pixel independently from the others. They are more sensitive to absolute errors in surface temperature estimates (Cammalleri et al., 2012; Norman et al., 2000). However, one usually expects that they are well adapted to uniform landscapes with fairly homogeneous vegetation and surface water conditions, such as natural landscapes or rain-fed monoculture, for which contextual methods are likely to fail. They are also more adapted to the use of low-resolution data. For the latter indeed, pixels are mixed and often cover many individual plots with contrasted levels of NDVI and soil moisture. The resulting conditions tend to blur into average effective moisture/vegetation characteristics at the pixel scale. They are therefore applied to produce global maps of evapotranspiration (Jiménez et al., 2011).

Contextual models, which use temperature differences instead of absolute values, are less sensitive to absolute errors in surface temperature estimates. But they rely on the assumption that all soil moisture conditions are present within one image for a large enough range of vegetation fraction cover. This hypothesis can be sometimes misleading (Choi et al., 2009; Gonzalez-Dugo et al., 2009). For instance, just after rainfall or after a long dry-down, this assumption can be challenged for natural landscapes or rain-fed agriculture. Moreover, wet bare soils or fully stressed vegetation are not always present on a single image, especially in irrigated agricultural areas with sufficient water supply.

Many studies have tested the performance of these models at various scales, from very high (Gomez et al., 2005; Jacob et al., 2002; Long and Singh, 2012; Minacapilli et al., 2009; Su et al., 2005; Timmermans et al., 2007) to low spatial resolution (Jia et al., 2003; McCabe and Wood, 2006; Su et al., 2007; Verstraeten et al., 2005; Yang and Wang, 2011). But in most cases, those studies lack the temporal representativeness of a surface evolution during a full growing season (Choi et al., 2009; French et al., 2005; Gonzalez-Dugo et al., 2009; van der Kwast et al., 2009; Li et al., 2005, 2008; Long and Singh, 2012; Ma et al., 2011; McCabe and Wood, 2006; Minacapilli et al., 2009; Su et al., 2005, 2007). A nonexhaustive list of validation and intercomparison studies of surface energy budget (SEB) models is shown in Table 1. Depending on the experiments, root mean square differences (RMSD) for instantaneous retrievals of turbulent fluxes range from $40\,\mathrm{W\,m^{-2}}$ (Gonzalez-Dugo et al., 2009; Jia et al., 2003; Li et al., 2008; Long and Singh, 2012; McCabe and Wood, 2006; Verstraeten et al., 2005) to more than $150\,\mathrm{W\,m^{-2}}$ (Choi et al., 2009; Li et al., 2005; Oku et al., 2007). Understanding of the reasons for such a wide range of performance is crucial in order to implement thermal data in automated operational algorithms.

Amongst well-known models, two algorithms emerge in the "single-pixel models" category: TSEB (two-source energy balance; Norman et al., 1995) and SEBS (surface

Table 1. Nonexhaustive review of validation exercises of instantaneous models. One site covers less than 30×30 low-resolution pixels, 1 station covers less than 10 high-resolution pixels.

Publication	Model(s)	Resolution	Precision on λE (W m^{-2})	Spatial and temporal coverage
Anderson et al. (2011)	TSEB	10 m to 10 km (airborne)	–	1 station on 1 site, 1 flight
Choi et al. (2009)	TSEB, METRIC, TRIM	60 m (LANDSAT 7)	RMSD 50–150	14 stations on 1 site, 2 dates
French et al. (2005)	TSEB, SEBAL	15–90 m (ASTER)	Bias 10–80	8 stations on 1 site, 1 date
Galleguillos et al. (2011)	S-SEBI, WDI	90 m (ASTER)	–	Comparison with SVAT (Hydrus 1-D) and 1 station on 1 site
Gomez et al. (2005)	S-SEBI	20 m (airborne)	RMSD ~ 90	7 stations on 1 site, 19 dates
Gonzalez-Dugo et al. (2009)	TSEB, METRIC	120–60 m (Landsat 5 & 7)	RMSD ~ 50	12 stations on 1 site, 3 dates
Jia et al. (2003)	SEBS, SEBI	1 km (ATSR-2)	RMSD 10–40 (BR)	11 stations on 1 site, 11 dates, 1 scintillometer
Li et al. (2005)	TSEB	120–60 m (Landsat 5 & 7, aircraft)	RMSD 40–120	5 stations on 1 site, 4 dates
Li et al. (2008)	TSEB	120 m (Landsat 5)	RMSD ~ 40	3 stations on 1 site, 3 dates
Long and Singh (2012)	TTME	90–60 m (ASTER & Landsat)	RMSD ~ 50	12 stations, 1 site, 4 images
Ma et al. (2011)	SEBS	90 m (ASTER)	Bias > 80	3 stations on 1 site, 4 dates
McCabe and Wood (2006)	SEBS	990–1020 m (aggregated ASTER & Landsat + MODIS)	RMSD 60 (HR) -80 (BR)	7 stations on 1 site, 1 date
Minacapilli et al. (2009)	TSEB, SEBAL	15 m (airborne)	–	1 date
Oku et al. (2007)	SEBS	5–7 km (GMS-5)	RMSD > 100	
Su et al. (2007)	SEBS	1 km (MODIS)	RMSD 40–60	2 dates, sites CEOP EOP-1
Su et al. (2005)	SEBS	30 m (LANDSAT 7) & 20 km (GOES, MODIS)	RMSD 30 (HR) -140 (BR)	1 date, 8 stations on 8 sites
Timmermans et al. (2007)	TSEB, SEBAL	6–12 m (airborne)	RMSD 60–70	2 stations on 1 site, 5 airborne flights
Van der Kwast et al. (2009)	SEBS	90 m (ASTER)	–	6 stations on 1 site, 1 date
Verstraeten et al. (2005)	S-SEBI	1.1 km (NOAA/AVHRR)	RMSD ~ 40	13 stations on 13 sites, AVHRR series
Yang and Wang (2011)	S-SEBI	1 km (MODIS)	Only evaporative fraction evaluated	12 stations, 12 sites, 16 dates
Jacob et al. (2002)	SEBAL	20 m (airborne)	RMSD ~ 85	7 stations, 1 site, ~ 15 airborne flights

energy balance system; Su, 2002); amongst the contextual approaches, one can cite the popular S-SEBI (simplified surface energy balance index; Roerink et al., 2000) and "triangle" or "trapezoidal" (Moran et al., 1994) approaches, for the most simple ones, or two complex but widely used methods, SEBAL (surface energy balance algorithm for land;

Bastiaanssen et al., 1998) and METRIC (mapping evapotranspiration at high resolution with internalized calibration; Allen et al., 2007).

Due to the limited availability of high-resolution images, these models have not yet been tested for a wide range of climates and landscapes, spanning various climatic and

vegetation conditions. Indeed, it is difficult to build a comprehensive yet exhaustive protocol to validate the SEB models with enough data in space and time. The main reason is that observations of turbulent fluxes are available for long-term applications, but for a few locations only, and except for intensive international campaigns, no more than one or two points within each image.

Soil–vegetation–atmosphere transfer (SVAT) models, however, are able to simulate the surface temperature from irrigation and rainfall inputs. In the case of local-scale applications where all water-related variables are well known (i.e., water inputs, initial soil moisture, etc.), they are expected to perform better than surface energy balance methods. However, since information about irrigation at regional scale is very difficult to assess, those models are difficult to implement in large irrigated perimeters. Because surface temperature is related to water stress (Hain et al., 2009), one can constrain model prediction through the assimilation of the observed surface temperature into SVAT models (Coudert and Ottlé, 2007; Olioso et al., 2005) and compensate for errors in water inputs estimates. Even for most hydrological models with daily time steps, which do not simulate the equilibrium surface temperature, a remotely sensed evapotranspiration product could be used in an assimilation scheme (McCabe et al., 2008; Schuurmans et al., 2003). Again, even if most SVAT models are able to assimilate directly the surface temperature, it is often hard to specify model errors, observation errors, and, particularly, spatial model error covariance when assimilating surface temperature images into distributed SVAT models. If contextual models show robust and reliable performance, the TIR-based evapotranspiration products could be assimilated either in SVAT or hydrological models by providing additional information about distributed constraints of studied areas.

The main objective of this paper is to test the relative performance of four TIR-based instantaneous evapotranspiration and water stress simulation models (two contextual and two single-pixel) to retrieve surface fluxes and water stress levels from remote sensing data over an intensive irrigated perimeter in semi-arid land throughout the main agricultural season. It is a necessary preliminary study before assimilating evapotranspiration products and their relative error in distributed SVAT models, which is our ultimate goal beyond the scope of this study. The performance of those four models will be assessed through data collected at seven flux stations and compared to outputs of an uncalibrated SVAT model forced with in situ vegetation, climate and irrigation input data measured at those seven stations.

2   Material and methods

2.1   The surface energy balance for estimating evapotranspiration

The first three models (SEBS, TSEB and S-SEBI) compute evapotranspiration as the residual of the energy balance, written as follows:

$$\lambda E = R_n - H - G, \tag{1}$$

where R_n is the net radiation, H the sensible heat flux, λE the latent heat flux and G the soil heat flux, all expressed in watts per square meter (W m^{-2}).

Models differ primarily in the partitioning of available energy $R_n - G$ into turbulent fluxes H and λE, and secondly on the way they compute the available energy.

2.1.1   Available energy

There are many different ways in the literature to estimate the net radiation R_n and soil heat flux G as well as their components from remote sensing data. However, in our model intercomparison, we focus more precisely on the way available energy is partitioned between latent heat flux λE and sensible heat flux H. That is why we made the choice to use the same formulation of R_n and G for the three models requiring available energy $R_n - G$ estimates in the computation of evapotranspiration. Starting from the same basis, it will be easier to analyze the different model behaviors.

The general formulation of net radiation is:

$$R_n = (1 - \alpha) \cdot R_{sw} + \varepsilon \cdot \left(R_{lw} - \sigma \cdot T_0^4 \right). \tag{2}$$

R_{sw} and R_{lw} are the shortwave and longwave incoming radiation (respectively), T_0 the surface temperature, α and ε are the albedo and the emissivity of the surface (respectively) and σ the Stefan–Boltzmann constant. For all SEB models, the emissivity value of the surface is fixed at 0.98 for the whole scene and for each date.

The amount of available formulation in the literature to estimate the soil heat flux is much larger than for the net radiation (for example, see Bastiaanssen, 2000; Norman et al., 1995; Roerink et al., 2000; Santanello and Friedl, 2003; and Su, 2002). It has been shown that for a fixed time around midday, G can be considered as proportional to the net radiation, the proportionality factor being determined by the surface vegetation and soil properties.

We tested each formulation of G proposed in the three SEB models (Norman et al., 1995; Roerink et al., 2000; Su, 2002), which require an estimate of the ground heat flux, as well as the Bastiaanssen (2000) formulation implemented in the METRIC model. The latter proved to be the most accurate, with a RMSD of 57 W m^{-2} when applied to the current data set (which will be detailed later in the paper) whereas the other formulations showed RMSD greater than 90 W m^{-2}.

Thus we chose to implement it for all methods. Its formulation is detailed in Eq. (3).

$$G = R_n T_0 \times (0.0038 + 0.0074\alpha) \times (1 - 0.98\,\mathrm{NDVI}^4) \quad (3)$$

2.1.2 Single-pixel models

Both single-pixel models compute turbulent exchanges between the source at the aerodynamic level and the reference level z where atmospheric forcing data is available, usually just a few meters above the crop height. Heat transfers are based on the Monin–Obukhov similarity theory in the atmospheric boundary layer (ABL). Bulk ABL similarity functions proposed by Brutsaert (1999) describe the wind and temperature profiles in the turbulent environment (Eqs. 4, 5; see Brutsaert, 1999; and Su, 2002).

$$T_{\mathrm{aero}} - T_a = \frac{H}{\rho\,C_p} \times \frac{1}{k\,u_*}\left[\ln\left(\frac{z - d_0}{z_{0h}}\right) - \Psi_h\left(\frac{z - d_0}{L}\right)\right.$$
$$\left. + \Psi_h\frac{z_{0h}}{L}\right] = \frac{H}{\rho\,C_p} \times r_a \quad (4)$$

$$u = \frac{u_*}{k}\left[\ln\left(\frac{z - d_0}{z_{0m}}\right) - \Psi_m\left(\frac{z - d_0}{L}\right) + \Psi_m\frac{z_{0m}}{L}\right] \quad (5)$$

r_a is the aerodynamic resistance to heat transfer at the surface–atmosphere interface, u is the wind velocity at level z, $k = 0.4$ is the von Karman constant, d_0 is the displacement height ($d_0 \approx h_c \times 2/3$, h_c being the crop height), and z_{0m} is the roughness height for momentum transfer ($d_0 \approx h_c \times 0.123$). T_a and T_{aero} are the temperature of the air at reference and aerodynamic levels (respectively), ρ is the density of the air, C_p is the heat capacity of air and z_{0h} is the roughness height for heat transfer. Ψ_m and Ψ_h are the stability correction functions for momentum and sensible heat transfer and L is the Monin–Obukhov length (Su, 2002).

If one agrees in general that Eq. (4) provides a fairly robust estimate of the aerodynamic resistance to heat transfer r_a, estimating the aerodynamic temperature remains a difficult issue (Kustas and Anderson, 2009). There is currently no way to measure this temperature directly, and models rely usually on the only available data, the surface temperature obtained by satellite imagery. The difference between the two single-pixel models will thus lie in the approximation made in order to relate T_{aero} to the radiometric surface temperature T_0.

The surface energy balance system (SEBS) model

SEBS (Su, 2002) computes the latent heat flux as the residual of the energy balance for a mixed pixel. The particularity of the SEBS model resides in two points. First, it proposes a new formulation of the difference between the roughness for heat transfer z_{0h} (Eqs. 6–7) and momentum z_{0m}, which allows one to replace T_{aero} by T_0 in Eq. (4):

$$z_{0h} = \frac{z_{0m}}{\exp\left(k\,B^{-1}\right)}, \quad (6)$$

where:

$$k\,B^{-1} = A_1 \times f_c^2 + A_2 \times f_c\,f_s + k\,B_s^{-1} \times f_s^2. \quad (7)$$

The first term A_1 describes the full canopy aerodynamic properties (Choudhury and Monteith, 1988), $k\,B_s^{-1}$ is representative of the bare-soil properties and the A_2 term takes into account the interactions between the vegetation and the bare soil. f_c is the canopy coverage fraction and $f_s = 1 - f_c$.

The other specificity of SEBS lies in the retrieval of the latent heat flux. Once the sensible heat flux is computed using Eq. (4) the model ensures that the retrieved latent heat flux is bounded by two theoretical, extreme conditions under the given climate forcing (null and potential evapotranspiration rates, Eqs. 8 and 9, respectively). This is achieved through the substitution of any outlier with the corresponding theoretical minimum (Eq. 8) or maximum (Eq. 9):

$$\lambda E_{\mathrm{dry}} = 0, \quad (8)$$

$$\lambda E_{\mathrm{wet}} = \left[R_n - G - \frac{\rho\,C_p\,(e_s - e)}{r_{\mathrm{ew}}\,\gamma}\right]\bigg/\left(1 + \frac{\Delta}{\gamma}\right). \quad (9)$$

λE_{dry} and λE_{wet} are the latent heat fluxes in those extreme dry and wet conditions. e and e_s are respectively the actual and saturation vapor pressure, γ is the psychrometric constant, Δ is the slope of saturation vapor pressure curve at temperature T_a and r_{ew} is the aerodynamic resistance for wet conditions.

The two-source energy balance (TSEB) model

TSEB computes two separate energy budgets for the soil and the vegetation, and estimates evaporation and transpiration (respectively) as residual terms of the energy balance (see Eqs. 10 and 11). Net radiation is computed according to Eq. (2). It is then partitioned based on fraction cover into the two main components, the net radiation of the canopy $R_{n,c}$, and the bare-soil $R_{n,s}$ (Norman et al., 1995) respectively. Both energy balance equations read

$$R_{n,c} = H_c + \lambda E_c, \quad (10)$$

$$R_{n,s} = H_s + \lambda E_s + G. \quad (11)$$

H_s and λE_s are the sensible and latent heat flux at the soil–atmosphere interface and H_c and λE_c the fluxes at the canopy–atmosphere interface (partitioning of the energy budget is summarized in Fig. 1).

Similarly to Eq. (1), both latent heat flux components λE_s and λE_c can be retrieved as residuals of both energy balance equations provided that the soil (T_s) and vegetation (T_c) component surface temperatures are known. The "trick" to get two unknowns out of one single piece of information (mixed-pixel temperature) is to assume that in many cases the vegetation is unstressed and transpires at a potential rate λE_c, which is obtained with a Priestley–Taylor formulation (see Eq. 12).

$$\lambda E_c = 1.3\,f_g\,\frac{\Delta}{\Delta + \gamma}\,R_{n,c}, \quad (12)$$
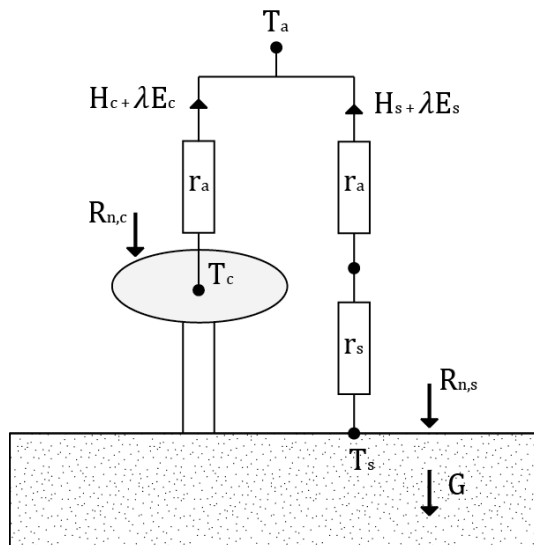
Fig. 1. Scheme of the resistance network of TSEB following Norman et al. (1995).

where f_g is the green fraction of leaf area index (LAI). Here, LAI is retrieved from the NDVI and is thus considered as green LAI. As a consequence, f_g is set to one in our application. Fraction cover is estimated from LAI with a Beer–Lambert law using a fixed extinction coefficient of 0.4. This first computation of λE_c gives us a first guess of H_c, as a residual of the energy budget of the canopy (Eq. 10). The canopy temperature T_c can be deduced from the canopy's sensible heat flux H_c through a formulation similar to Eq. (4). The total sensible heat flux H is

$$H = H_s + H_c = \rho C_p \frac{T_s - T_a}{r_a + r_s} + \rho C_p \frac{T_c - T_a}{r_a}, \qquad (13)$$

where the resistance r_s takes into consideration the resistance to heat transfer in the boundary layer immediately above the soil surface.

Equation (14) (Norman et al., 1995) links canopy and soil temperatures to the observed radiometric temperature T_0 and f_c and allows one to calculate T_s:

$$T_0 = \left[f_c T_c^4 + (1 - f_c) T_s^4 \right]^{1/4}. \qquad (14)$$

Combining Eqs. (11) and (13) gives H_s and λE_s. If λE_s is positive then a balance is reached. If $\lambda E_s < 0$, the assumption that the vegetation transpires at a potential rate is no longer valid, the soil is considered as dry and λE_s set to zero and the other parameters are computed from Eqs. (11), (13) and (14). λE_c is obtained as a residual of Eq. (10). If λE_c is negative then soil and vegetation are dry and λE_c is set to zero. Then $H_c = R_{n,c}$ and Eqs. (13) and (14) lead to a new H_s. G is computed as the residual of the energy balance for the bare soil (and thus differs in this specific case from the other models).

2.1.3 Contextual models

The simplified surface energy balance index (S-SEBI) model

The S-SEBI (Roerink et al., 2000) model is based on the observation that for homogeneous atmospheric conditions over a scene, the scatter plot between surface temperature T_0 and surface albedo α is bounded by two theoretical lines corresponding to extreme soil moisture conditions (see Fig. 2, upper panels). The method then considers for each pixel α the corresponding extreme conditions (with the same available energy) with the respective surface temperatures T_{hot} and T_{cold}. Those conditions correspond respectively to dry ($\lambda E_{min} = 0$, $H = H_{max}$) and wet ($\lambda E = \lambda E_{max}$, $H_{min} = 0$) areas where all the available energy is used to evaporate. The evaporative fraction is then computed using those extreme temperatures according to

$$\Lambda = \frac{\lambda E}{R_n - G} = \frac{R_n - G - H}{R_n - G} = \frac{H_{max} - H}{H_{max}}. \qquad (15)$$

From Eq. (4), we can deduce that, for a given range of albedo, the wet boundary conditions ($H = 0$) implies $T_{cold} \approx T_a$. So if we replace T_a by T_{cold}, we can deduce from Eqs. (4) and (15) the following formulation of the evaporative fraction:

$$\Lambda = \frac{T_{hot}(\alpha) - T_0}{T_{hot}(\alpha) - T_{cold}(\alpha)}. \qquad (16)$$

The turbulent fluxes H and λE are then deduced from the evaporative fraction (Eq. 16) and an estimate of the available energy (Eq. 2).

A vegetation index-temperature trapezoid method (VIT)

Similarly, Moran et al. (1994) proposed a method for retrieving λE on large areas by combining information on the temperature difference, $\Delta T = T_0 - T_a$, and vegetation extent. In the original method, the Penman–Monteith equation is applied to calculate equilibrium surface-to-air temperature differences ($T_0 - T_a$) in four extreme conditions of soil moisture and vegetation cover (well watered and fully stressed canopies, wet and dry bare soils). Four theoretical vertices of a trapezoid in the T_0/VI (VI for vegetation index) space are obtained (see Fig. 2, lower left panel). The other assumption is that $T_0 - T_a$ is linearly related to the vegetation cover, which itself is linearly related to the vegetation index used by Moran et al. (1994), the soil-adjusted vegetation index (SAVI). This allows straight lines to be drawn between the vertices 1 and 3 and the vertices 2 and 4. The third assumption that us allows to link this graphic representation to the water status of the surface is that, for uniform available energy, vapor pressure deficit of the air and aerodynamic resistance values, $T_s - T_a$ and $T_c - T_a$, are linearly dependent on evaporation and transpiration, respectively. The land surface
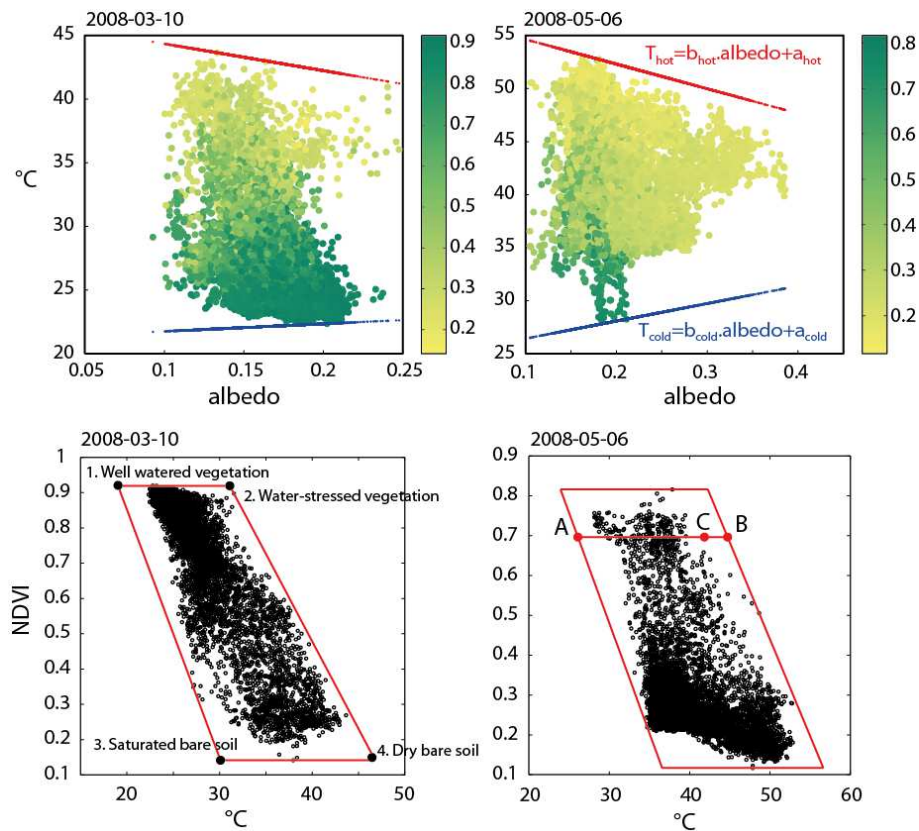
Fig. 2. Upper panels: ASTER surface temperature vs. FORMOSAT-2 albedo scatter plot and dry and hot regression law determined with the SPLIT method in S-SEBI. The points' colors correspond to the value of FORMOSAT-2 NDVI for each pixel. Lower panels: FORMOSAT-2 NDVI vs. ASTER surface temperature scatter plot and the trapezoid determined with the SPLIT method in VIT. For both methods, the plots are made for 10 March and 6 May.

temperature T_0 being directly linked to T_s and T_c, the linear relation between T_0 and evapotranspiration is established as follows:

$$\text{WDI} = \frac{\Delta T - \Delta T_{\min}}{\Delta T_{\max} - \Delta T_{\min}}, \qquad (17)$$

where WDI is the water deficit index; ΔT, $\Delta T_{\min}$ and $\Delta T_{\max}$ are respectively the surface-to-air temperature difference at points C (actual case), A (well watered case) and B (stressed case). Graphically, it corresponds to the ratio of the distances AC and AB. WDI can be assimilated to surface water stress. Other applications of the VIT method also use NDVI or fraction cover instead of SAVI.

In our case, a preliminary study had shown that the theoretical vertices calculated with the Moran et al. (1994) approach were not usable. Actually, the theoretical trapezoid did not include all the situations present in the images (not shown). Thus, pixels with unrealistic values of WDI were present, i.e., either with negative values or values larger than one. We then chose to make an assumption similar to the one made in the S-SEBI method, which is that the variety of relative moisture conditions present in the image is sufficient

to assess, if not the four extreme points, at least the bounding linear relationships (lines 1–3 and 2–4) corresponding to well watered cold points and stressed hot ones. We thus decided to derive the four vertices graphically and not theoretically by drawing the bounding trapezoid for each ΔT–NDVI scatter plot. Since we have a flat study zone of limited extent, we consider that a single air temperature measurement at 10 m is representative of the whole area. We can then work on the NDVI–T_0 scatter plots directly.

2.1.4 Method to retrieve water status extremes in contextual models

Various methods to retrieve linear relationships corresponding to extreme boundary conditions between surface temperature and albedo (S-SEBI) or NDVI (VIT) have been tested (not shown). They are either manual or automatic, with different levels of complexity. An entirely manual method would allow the user to take into account qualitative a priori information on the surface (type of cultivated crop, sewing, harvesting and irrigations dates, etc.) to decide whether extremes correspond to target situations (for instance, wet bare

soils or dense vegetation under water stress) or anecdotic outliers. However, manual methods are not appropriate for integration into operational automated retrieval algorithms. That is why we chose to use the automatic method presented in Verstraeten et al. (2005) and named SPLIT method. In this method, a classification of albedo values is performed. We chose a subdivision in ten classes of albedo values for each image. For each class, the median of the 5 % maximum and the median of the 5 % minimum surface temperature values are identified, then the least square method is applied to those median values to retrieve each bounding linear relation. In our application of the VIT and S-SEBI methods, we infer the bounding relationships on a subset of the original image, corresponding to our area of interest (4 km × 4 km). This is justified by the fact that our landscape is an irrigated agricultural area, which is by nature heterogeneous but holds a fairly uniform distribution of this heterogeneity (in terms of land cover and irrigation practices).

2.1.5 The soil–vegetation–atmosphere transfer (SVAT) model

In this study we use the outputs of the interactive canopy radiative exchange (ICARE) SVAT model as a reference or benchmarking tool to relate the SEB models' performance. ICARE is a classic dual-source SVAT model that solves both water and energy balance of the surface. It is forced with climatic and vegetation growth data. The main differences between the SVAT model and the SEB models reside in two points.

First, the water balance module simulates the evolution of soil moisture for two soil layers (shallow surface and root zone) using the force–restore method (Noilhan and Planton, 1989):

$$\begin{cases} \frac{\partial w_s}{\partial t} = \frac{C_1}{\rho_w d_1} \left(P + I - E_s \right) - \frac{C_2}{\tau} \left(w_s - w_{eq} \right) \\ \frac{\partial w_2}{\partial t} = \frac{P + I - E_s - E_c}{\rho_w d_2} \end{cases} \qquad (18)$$

where w_s and w_2 are respectively the surface layer and root zone volumetric soil moisture, P the precipitation rate, I the irrigation rate, E_s the soil evaporation rate, E_c the plant transpiration rate, d_1 and d_2 the thickness of surface layer and root zone respectively set at 0.05 and 1 m for all crops, τ the diurnal time period (1 day) and ρ_w the density of water. C_1 and C_2 are coefficients estimated following the ISBA model (Noilhan and Mahfouf, 1996) depending on soil texture (pedotransfer functions) and w_{eq} is an equilibrium soil surface moisture representing the case where capillarity and gravity forces compensate each other. E_s and E_c are computed with a coupled two-source energy balance equation (Shuttleworth and Wallace, 1985); and they depend respectively on a surface resistance to evaporation (a function of surface soil moisture, Passerat de Silans, 1986) and a stomatal resistance (a function of root zone soil moisture, Noilhan and Planton, 1989). Surface and root zone temperatures in

the soil are also computed using the force–restore method based on a classical heat diffusion law. As any dynamic model, initial conditions for the surface and the root zone temperature and moisture levels are required. In our application, initial conditions are taken from in situ measurements at each station.

Second, contrary to SEB models, the surface temperature is not an input but an output of the SVAT model. Direct information about soil moisture is absent from the SEB models, in opposition to the SVAT model, which is forced by rain (null over the season in our case) and irrigation time series. Next is the computation of the radiation budget. The model used by ICARE is not forced by a remotely sensed albedo value but computes its own broadband albedo with a multiple reflections model and fixed values of soil and vegetation broadband reflectances; and it is likely to introduce differences in the computation of the net radiation. Finally, the soil heat flux G is not calculated as a fraction of the net radiation but as a residual of the energy budget; and it provides an upper boundary condition to the heat diffusion law between the soil layers. Soil texture was determined from in situ measurements at each station and the main physical parameters of soil were estimated following ISBA's own pedotransfer functions (Noilhan and Mahfouf, 1996, Appendix A4). Estimation of the available energy differs significantly from the SEB models. A more detailed presentation of the model is available in Gentine et al. (2007).

As any complex physically based SVAT model, ICARE requires a large set of input parameters describing the different properties of the surface (soil and vegetation). Those parameters need to be calibrated in order to obtain consistent results. However, we chose to run the model in its most standardized version, with literature or measured values, when available. Vegetation parameters such as leaf aerodynamic properties were determined using standard values from the literature. In situ measured input data for vegetation includes vegetation height and green and dry LAIs, respectively. Total LAI is used to compute fraction cover and within canopy aerodynamic resistances, whereas green LAI weights the stomatal resistance. The only soil parameter that has been calibrated was the soil resistance to evaporation r_{ss}, because soil texture and composition are almost uniform over the whole area. This choice was also made because future implementation of data assimilation in ICARE would provide a way to calibrate the model. Therefore, we wished to compare the SEB models to a SVAT model running with the most "standard" set of parameters possible. The formulation of r_{ss} is given in Eq. (19) from Passerat de Silans (1986). w_s is the soil water content index and w_{sat} the soil saturation soil water content index. A_{rss} and B_{rss} are therefore the only two empirical constants that have been calibrated.

$$r_{ss} = \exp \left(A_{rss} - B_{rss} \frac{w_s}{w_{sat}} \right) \qquad (19)$$

The calibration of A_{rss} and B_{rss} has been done simultaneously on the sole northeastern wheat eddy covariance (EC) station (wheat(1), see Sect. 2.2.1) when the surface was almost bare using a multicriteria approach on two observed variables: the latent heat flux and the surface (5 cm) soil moisture. The minimum value of the sum of the relative absolute difference of these two variables, scaled by their mean value, was searched systematically for all parameters combinations within the 0–20 range. Results of this calibration lead to values of 12 and 19 for A_{rss} and B_{rss} (respectively). The minimum stomatal resistance, which is a very sensitive parameter for the estimation of latent heat flux, has been set to $100 \, \text{s m}^{-1}$ (Gentine et al., 2007; Martin et al., 1999).

The SVAT model has been run at each station at a half hour time step for the whole growing season, except for the chili pepper station. The latter was covered by a white cap during half of the season, which is not accounted for by ICARE. Since irrigation patterns are not known with enough precision, ICARE is not run spatially to produce maps, but locally at each flux station.

2.2 Site and data acquisition

2.2.1 Site description

This study was conducted in the Yaqui Valley (27.4° N, 109.9° W), in the state of Sonora, northwestern Mexico. With an area of 225 000 ha, bordered on the southwest by the Sea of Cortez and on the northeast by the Sierra Madre, it is the largest agricultural district of the state. The main cultivated crop is winter wheat. The climate is semi-arid with an average annual potential evapotranspiration of about 2233 mm (1971–2000 average; Servicio Meteorológico Nacional, México, http://smn.cna.gob.mx), far greater than the average annual precipitation, which is 290 mm (1981–2000 average; see http://smn.cna.gob.mx/observatorios/historica/obregon.pdf). Rainfall is brought essentially during the monsoon season (from June to September) with only 42.8 mm of precipitation from January to June. About 90 % of the water consumption in the valley comes from irrigation (Chehbouni et al., 2008b). The estimation of water losses by evapotranspiration is consequently a key factor in the management of water at the regional scale.

From December 2007 to May 2008, an international cooperative experiment was carried out over a square of 4 km × 4 km, located south of the city of Ciudad Obregón (center of the zone: 27.263° N, 109.892° W). Around 50 % of the cultivated crops are wheat. The remaining part is divided between broccoli, beans, chili pepper, potatoes, chickpea, safflower, orange and corn. Seven micro-meteorological stations equipped with an EC flux measurement system were installed in different fields. Their location is shown in Fig. 3.

Fig. 3. Satellite view of the studied zone with respective positions of the EC stations (Imagerie ©2012 Cnes/Spot Image, DigitalGlobe, Données cartographiques ©2012 Google, INEGI).

2.2.2 Automated data acquisition

Meteorological data were acquired at a height of 10 m from a weather station installed at the center of the zone. It provided measurements of wind speed and direction (R. M. Young anemometer), air temperature and moisture (Vaisala humidity and temperature sensor). Missing data were replaced with a combination of the meteorological data also available at each EC station. Data were acquired with half hourly time steps from 27 December 2007 to 13 May 2008.

At each of the seven sites, the net radiation was acquired using CNR1 (Kipp & Zonen, stations (1), (3) and (7)) and Q7.1 (REBS, stations (2), (4), (5) and (6)) radiometers. The soil heat flux was estimated with HUKSEFLUX HFP-01 plates buried at 0.05 m at the top and bottom of the furrow (when applicable). Surface temperature was measured at each site with Apogee infrared radiometers at nadir, at 2m height (which corresponds to a footprint radius of 0.66 m). Soil moisture was acquired at 0.05 and 0.3 m depths using a CS616 TDR (time domain reflectometer, Campbell Scientific Inc., UT, USA). Those data were acquired at a sampling interval of 10 s then averaged and recorded every 30 min.

Latent and sensible heat fluxes were measured with KH20 fast response hygrometers (Campbell) and Campbell CSAT3 or RM Young 81000 3D sonic anemometers at a frequency of 10 Hz.

2.2.3 Discontinuous measurements

In addition to the data acquired by the automatic stations, in situ measurements of vegetation properties were also carried out. Crop height and LAI were measured at various dates during the whole study, every week on average. They were

linearly interpolated between two successive dates to provide input data for ICARE. LAI was estimated from destructive measurements as well as hemispherical photographs for 23 fields representative of the various land use types. Gravimetric soil moisture profiles at each station were carried out each week. Those measurements allowed us to calibrate the CS616 TDR installed at each station. Soil texture was analyzed at each site at the beginning of the study period. Surface (0–5 cm) soil moisture was acquired spatially with ThetaProbe sensors (Delta-T) every week also at some 200 locations from December to May.

2.2.4 Remote sensing data

Seven ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer; http://asterweb.jpl.nasa.gov/) images were acquired in the thermal infrared domain from December 2007 to May 2008. The resolution of the land surface temperature product (AST08; https://lpdaac.usgs.gov/products/aster_productstable/ast08) is 90 m and is atmospherically corrected.

FORMOSAT-2 is an earth observation satellite launched in 2004 by the National Space Organization of Taiwan. It provides high-resolution (8 m) images of a particular area every day at 09:30 LT (local time) for four bands (blue, green, red and near infrared) and with the same view angle. More details can be found in Chern et al. (2008). For our study, 26 cloud-free images were obtained from 27 December 2007 to 13 May 2008, which represents a regular acquisition frequency of on average 1 image every 5 days (see Fieuzal and al., 2011, for more details). Remote Sensing products such as NDVI derived from FORMOSAT-2 data were linearly interpolated to the date of the ASTER acquisition date.

2.3 Data processing

2.3.1 Flux data

The turbulent fluxes from the EC stations were processed offline. A post-processing software (EC Pack) developed by the Meteorology and Air Quality group at Wageningen University (http://wwWmet.wau.nl/) was applied. A detailed explanation of the correction procedure is available in Van Dijk et al. (2004).

After rejection of inconsistent data due to instability or malfunction of the instruments, erroneous values of H and λE were still present. The closure of the energy balance was not achieved most of the time, with a residual at the half hourly time step comprised between 24 and 38 % of total available energy, depending on the station. This error is in the upper range of what can be found in the literature (Twine et al., 2000; Wilson et al., 2002). The gap in energy closure is often partly due to instrumental limitations. Indeed, although they acquire wind speed and specific humidity fluctuations at high frequencies, EC systems still miss part of the

energy fluctuations occurring at very high frequencies, leading to an underestimation of turbulent fluxes. Estimation of the soil heat flux can also be questioned for row crops since it is derived from a simple average of values measured at the top and bottom of the furrow, corrected for the soil energy storage through the use of a thermocouple buried at 5 cm. In addition, the malfunctioning of some instruments was identified during the Yaqui 2008 experiment, especially Krypton hygrometers of stations (2) and (4) and the CNR1 sensor of station (1). Finally, differences of energy closure issues between the stations result both from the variability of observed surface properties, like crop geometry, and differences between the various instrumental setups (brand, provenance, and maintenance history). At station (1), the CNR1, compared to those installed at other stations, proved to underestimate the incoming solar radiation component of the radiation budget. This term has been replaced with a mean value of measurements at other stations. Surface energy balance closure was forced using two different methods. In the case of stations (2) and (4) where the KH20 did not work, we chose to trust the estimation of H by the sonic anemometer and to discard the measurements of λE. The corrected λE was computed as the residual of the energy balance $R_n - G - H$. For stations (1), (3) and (6), the regression slope between available energy and the sum of turbulent fluxes was larger than 0.65, therefore both fluxes can be considered as consistent. The "Bowen-ratio closure" method is then used (Barr et al., 1994; Blanken et al., 1997). The estimation of the Bowen ratio $H/\lambda E$ is considered as correct and the fluxes are adjusted to close the energy balance. The safflower station (7) had both problems with the turbulent and soil heat fluxes so we excluded its data from the study. Finally, we did not succeed in integrating the drip irrigation data of the chili pepper station (5) into ICARE, therefore this station has not been used in the intercomparison. The dates and EC station numbers for which data were used in the intercomparison are reported in Table 2.

2.3.2 Remote sensing data

The observed distributed radiometric surface temperature used in this paper is the AST08 product from the ASTER sensor. The seven images were downloaded from the Earth Observing System Data Gateway. The overpass time is around 11:00 LT and the dates of the images are 30 December 2007, 23 February, 10 March, 11 April, 27 April, 6 and 13 May 2008. The resolution of this product is 90 m and the scenes are around 60 km × 60 km. The surface temperature is retrieved by the temperature and emissivity separation algorithm (Gillespie et al., 1998; Schmugge et al., 1998). The absolute registration of the images has been done based on a FORMOSAT-2 8 m resolution image (Merlin et al., 2010). The extracted sub-images of the 4 km × 4 km study zone were then resampled by bi-cubic interpolation at a resolution of 100 m for future model comparison.

Table 2. EC stations used to compute performance statistics at each date. Correspondence of numbers and stations : (1) wheat (east), (2) wheat (west), (3) broccoli/beans, (4) chickpea, (5) chili pepper, (6) potatoes/sorghum and (7) safflower.

Observed data	Date						
	30 Dec 2007	23 Feb 2008	10 Mar 2008	11 Apr 2008	27 Apr 2008	6 May 2008	13 May 2008
R_n	–	1–4	1–4	1–4, 6	1–4, 6	1–4, 6	1, 2, 4, 6
G	–	1, 2, 4	1–4	1–4, 6	1–4, 6	1–4, 6	2–4, 6
H	3	1–3	1–3	1–4, 6	1–4, 6	1–4, 6	1–4, 6
λE	3	1, 2, 4	1, 2	1–4, 6	1–4, 6	1–4, 6	1–4,
α	3	1, 3	1, 3, 7	1, 3, 7	1, 3, 7	1, 3, 7	1, 7
T_0	–	1–5	1–5, 7	1–5, 7	1–7	1–7	1, 2, 4–7

The error in the ASTER T_0 product is classically around 1.5 K (Jacob et al., 2008; Sabol et al., 2009).

The 26 FORMOSAT-2 images were registered using GPS ground control points and reprojected in the UTM WGS (Universal Transverse Mercator World Geodetic System) 1984 12 N coordinate system. Then an atmospheric correction was applied (Hagolle et al., 2008). Finally, the 4 km × 4 km study area was extracted and re sampled at 100 m resolution by aggregation (averaging) of FORMOSAT-2 pixels.

Although the ASTER platform provides more bands in the near (NIR) and shortwave infrared (SWIR) domain than FORMOSAT-2, which would suggest that a more consistent shortwave broadband albedo can be computed, a dysfunction of the SWIR sensor occurred on four of the seven available ASTER dates. This made calculus of an ASTER albedo impossible on those dates and FORMOSAT-2 data were chosen in order to keep a homogeneous albedo over the whole study and not multiply the sources of error. Albedo was computed as a linear combination of bands 3 (red) and 4 (near infrared) of FORMOSAT-2, according to Courault et al. (2008):

$$\alpha = 0.645\,\rho_{\mathrm{red}} + 0.382\,\rho_{\mathrm{NIR}}. \tag{20}$$

The NDVI was also calculated from bands 3 and 4 of FORMOSAT-2.

A remotely sensed LAI was computed from NDVI using a single relationship (Clevers, 1989) for the whole study area (Eq. 21):

$$\mathrm{LAI} = -\frac{1}{k}\ln\left(\frac{\mathrm{NDVI}_\infty - \mathrm{NDVI}}{\mathrm{NDVI}_\infty - \mathrm{NDVI}_{\mathrm{soil}}}\right). \tag{21}$$

This relationship between LAI and NDVI was calibrated using values of hemispherical LAI retrieved for all 23 fields during the growing season with a minimal RMSD criterion (Fieuzal et al., 2011). The calibrated extinction factor k is 1.13 and the asymptotical values of NDVI are $\mathrm{NDVI}_\infty = 0.97$ and $\mathrm{NDVI}_{\mathrm{soil}} = 0.05$. NDVI_∞ and $\mathrm{NDVI}_{\mathrm{soil}}$ are the NDVI values for a fully developed canopy and a bare soil respectively. We found that a single relationship was valid for all crops with a satisfying accuracy.

Table 3. Performance statistics of the models for the four components of the energy balance (in W m^{-2}).

		RMSD	CV (RMSD)	Bias
R_n	TSEB/SEBS/S-SEBI	42.5	7.1	−1.6
	ICARE	69.0	11.4	3.16
G	TSEB	60.4	55.0	−9.1
	SEBS/S-SEBI	57.4	52.2	−7.8
	ICARE	96.1	87.8	37.4
H	TSEB	131.4	62.2	−77.7
	SEBS	136.4	64.6	−41.1
	S-SEBI	133.6	63.2	−20.1
	ICARE	99.3	49.1	25.8
λE	TSEB	122.1	43.0	82.4
	SEBS	130.6	46.0	40.0
	S-SEBI	116.5	41.0	28.8
	ICARE	130.9	42.2	−72.1

Finally, vegetation height was interpolated spatially and temporally from discontinuous measurements based on NDVI time series to provide input data for the SEB models.

3 Results

First, the biophysical variables (surface temperature, albedo) extracted from the ASTER and FORMOSAT2 images (respectively) at each station are compared to the available measurements at the seven locations. Single-pixel values were selected since land use and most conditions (vegetation height etc) are fairly homogeneous within each field and EC stations are located at the center of those fields. Then, the available instantaneous energy components (net radiation and soil heat flux) and the turbulent fluxes at the time of the satellite overpass are compared to station values, with a special insight on water stress. The statistical results for the four models are summarized in Table 3. Finally, the patterns deduced from the models for the entire 4 km × 4 km square are analyzed
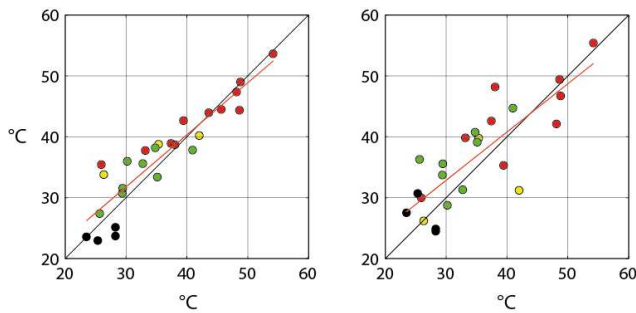
Fig. 4. ASTER (left panel) and ICARE (right panel) versus in situ radiometric surface temperature at each station at ASTER overpass dates. Color code: red: LAI < 0.4; yellow: 0.4 ≤ LAI < 0.8; green: 0.8 ≤ LAI < 1.2; navy-blue: 1.2 ≤ LAI < 2.0; black: LAI ≥ 2.0.

Fig. 5. Computed versus measured net radiation for the SEBS, TSEB and S-SEBI (left panel) and ICARE (right panel). Color code: red: LAI < 0.4; yellow: 0.4 ≤ LAI < 0.8; green: 0.8 ≤ LAI < 1.2; navy-blue: 1.2 ≤ LAI < 2.0; black: LAI ≥ 2.0.

by comparison with available information on moisture levels, vegetation type, etc.

3.1 Analysis of remotely sensed input data

Measurements of albedo were made on three sites ((1), (3) and (7), see Table 2 for dates) by CNR1 sensors. Comparison of FORMOSAT-2 against in situ albedo values shows a mean bias of around 4.4 % and a coefficient of variation of the RMSD, noted CV(RMSD) (ratio of the root mean square difference and the average value) of 10 %. Those results are quite acceptable with respect to previous studies (Bsaibes et al., 2009; Courault et al., 2008; Jacob and Olioso, 2005), which document CV(RMSD) values between 3 and 15 %.

Extraction of ASTER and ICARE surface temperature values at each EC station's pixel are shown in Fig. 4. We compared our ground data to ASTER values of temperature at the coordinates of the EC stations (Fig. 1). Brightness surface temperature of the Apogee sensors was corrected for emissivity and reflected atmospheric radiation using a classical emissivity–NDVI relationship (Van de Griend and Owe, 1993). Dates and stations at which surface temperature is available are displayed in Table 2. The absolute error on temperature is around 3.5K and clearly exceeds the values reported in the literature, but this can be explained by the difference between the footprint of the instruments (less than 1 m of diameter) and the pixel size (100×100 m) as well as the representativeness of the surface temperature when the surface is heterogeneous. Moreover, large temperature values are encountered in this semi-arid region. A mean bias around 0.9 °C appeared in the estimation of T_0. Despite those results, the mean bias was different for each station (from -0.2 to 3.9 °C), thus no global correction was applied to AST08. ICARE proves to be less accurate with a RMSD of 5.4 °C and a mean bias of 1.8 °C. ASTER seems to be less accurate for intermediate LAIs (between 0.8 and 1.2) where it seems to overestimate the surface temperature. This error can be attributed to the larger representative area of ASTER pixels than the Apogee footprint: ASTER can "see" more bare
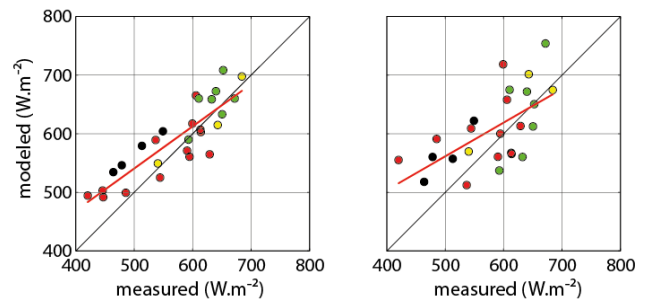
soil (which is hotter than the canopy) than the in situ instrument. Larger errors for ICARE radiometric temperatures can be interpreted as a consequence of model error in the energy balance resolution. The radiative surface temperature in ICARE is determined as a linear combination of the aerodynamic canopy and soil temperatures (Model RTM-TIR0 of Merlin and Chehbouni, 2014), which are all computed from the resolution of the energy budget at each source (soil and canopy).

3.2 Surface energy balance components

Scatter plots of modeled (SEB and ICARE) versus observed (measured) net radiation are displayed in Fig. 5. The estimation of the available energy by the three SEB models gives results often encountered in the literature with a RMSD of around 42 W m^{-2} (< 10 % of the mean value) for the computed R_n and of 57 W m^{-2} for G. Almost no bias is observed in net radiation.

Net radiation is better retrieved by the SEB models than ICARE. ICARE net radiation flux shows a coefficient of variation of the RSMD CV(RMSD) of 11 %, which, while still being a reasonable error, can be explained by the differences between the surface temperature and albedo used by the various algorithms. Indeed, in the SVAT model, albedo is computed using soil and dry/green vegetation albedos (respectively set at 0.15, 0.35 and 0.22 in our case, which correspond to rough estimates observed during the experiment). The ICARE albedo was not calibrated, thus it was not expected to obtain optimal results for each crop (~ 20 % CV(RMSD) on albedo for wheat and chickpea crops, ~ 55 % for beans). On the contrary, the formula used to calculate the broadband albedo from FORMOSAT-2 bands has been calibrated on a large area with a substantial variety of crops (Courault et al., 2008). It gets a CV(RMSD) lower than 10 % for wheat and beans and of around 17 % for chickpea, with a global CV(RMSD) of 9.6 % (against 35.7 % for ICARE). This gap can be explained by the fact that we did not use remotely
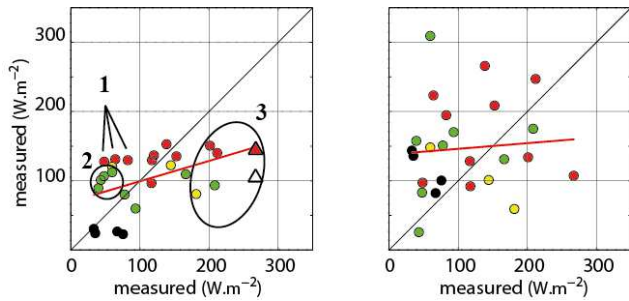
Fig. 6. Computed versus measured soil heat flux for SEB models (left panel) and ICARE (right panel). The only point differing between TSEB and the other SEB methods has been reported as a white-faced triangle, the Bastiaanssen (2000) estimation being reported as a red-faced triangle. Color code: red: $LAI < 0.4$; yellow: $0.4 \le LAI < 0.8$; green: $0.8 \le LAI < 1.2$; navy-blue: $1.2 \le LAI < 2.0$; black: $LAI \ge 2.0$.
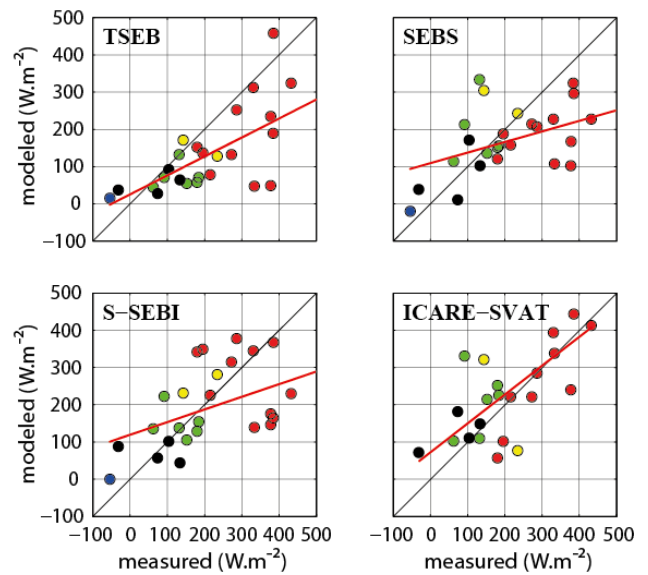

Fig. 7. Simulated versus measured sensible heat flux H for every station at ASTER overpass dates. Color code: red: $LAI < 0.4$; yellow: $0.4 \le LAI < 0.8$; green: $0.8 \le LAI < 1.2$; navy-blue: $1.2 \le LAI < 2.0$; black: $LAI \ge 2.0$.

sensed albedo data to calibrate the albedo parameterization of ICARE.

The results of the calculation of the soil heat flux for each model has been displayed in Fig. 6. Because of the structure of the TSEB algorithm, which computes G as the residual of the soil energy balance when the surface is totally stressed, estimation of G by TSEB differs from SEBS and S-SEBI only by one point. Thus, we did not plot results for TSEB on a separate graph and just reported that point on the left plot in Fig. 6 as a white triangle. For the SEB methods, an effect of saturation is clearly noticeable for bare soils and quasi-bare soils (red dots). This means that not only the maximum ratio of G over R_n but also the dynamic range for low LAI values should be larger than what is suggested by Bastiaanssen et al. (2000). The red dots group (group 1 in Fig. 6) shows an overestimation of the soil heat flux by the selected model. It corresponds to the end of the senescence period and the harvesting time of wheat. At that time, vegetation is still dense and standing, but dry. Since the formulation is based on NDVI, and therefore represents the amount of shading due to green vegetation only, shading is largely underestimated. A series of green and yellow dots (intermediate and low LAIs, group 2 in Fig. 6) corresponding to the same station (beans) are constantly overestimated with a bias around $50 \, W \, m^{-2}$. No particular overestimation of net radiation or surface temperature was observed at this location. We can assume that this bias comes in a large extent from the measurements. However, the models largely underestimate the flux for high values of measured G. The third group of outliers (Fig. 6) comes from the chickpea station. Even if the four points at the far right of the figure correspond to dates where R_n was greatly underestimated ($\sim 100 \, W \, m^{-2}$ at each date), it sums up with an underestimation of the $\Gamma = G/R_n$ factor by the models. At the station, Γ values estimated from observations ranged between 0.3 and 0.45 whereas the model computes values between 0.2 and 0.3. Although all in situ

values seem high for the study area (other stations give Γ values of around 0.3 for bare soil), the maximum of this factor encountered in the literature can be as high as 0.35 (Kustas and Daughtry, 1990; Monteith, 1973; Norman et al., 1995). In our case, the formulation of Γ cannot reach those values, and rarely exceeds 0.3.

Greater scattering is observed in ICARE with a larger RMSD (96 versus $57 \, W \, m^{-2}$ for the other models). Since the soil heat flux is determined by ICARE as the residual of the energy budget at the soil interface, those errors are mainly due to the absorption by G of the various errors made during the estimation of both net radiation and turbulent fluxes.

Scatter plots of simulated versus measured turbulent fluxes are presented in Figs. 7 and 8 for H and λE, respectively. TSEB has a systematic tendency to underestimate H (and thus overestimate λE) with a strong bias of $82 \, W \, m^{-2}$. In Fig. 7, strong underestimation of H at low LAIs can be observed. Underestimation of H for intermediate LAI values can be due to errors in the ASTER temperature. At high LAIs, we observe in most cases an overestimation of λE (see Fig. 8) certainly related to an overestimation of the canopy transpiration (Eq. 12), since TSEB's initial assumption is that the vegetation always transpires at the potential rate.

In terms of absolute error, SEBS shows a similar performance (Table 3). There is a great underestimation of H for quasi-bare soils. Since the remotely sensed temperature is rarely underestimated, and given the fact that vegetation height is well constrained by in situ measurements, H underestimation is likely to be due to an overestimation of the excess resistance kB^{-1} at low LAIs. This is in agreement
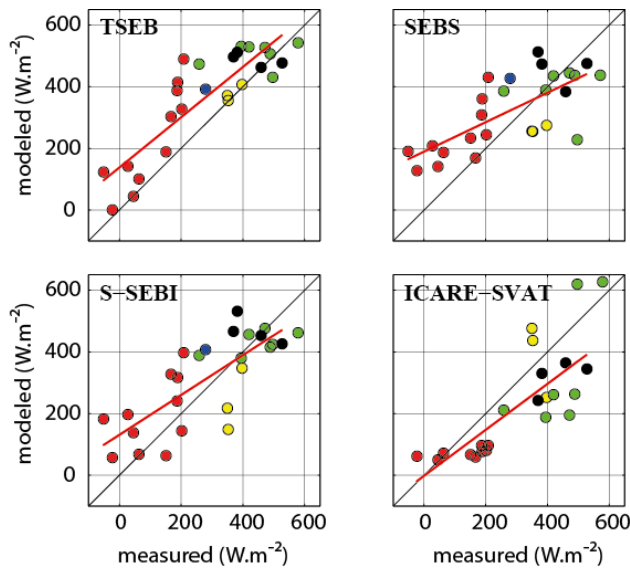
Fig. 8. Simulated versus measured latent heat flux λE for every station at ASTER overpass dates. Color code: red: LAI < 0.4; yellow: $0.4 \leq$ LAI < 0.8; green: $0.8 \leq$ LAI < 1.2; navy-blue: $1.2 \leq$ LAI < 2.0; black: LAI $\geq$ 2.0.

Fig. 9. Scatter plots of water stress at each station for TSEB, SEBS, S-SEBI, VIT method and ICARE. Values calculated for ASTER overpass dates. Color code: red: LAI < 0.4; yellow: $0.4 \leq$ LAI < 0.8; green: $0.8 \leq$ LAI < 1.2; navy-blue: $1.2 \leq$ LAI < 2.0; black: LAI $\geq$ 2.0.

with the current findings published in the literature (Boulet et al., 2012; Gokmen et al., 2012). The few overestimations of H observed for moderate LAIs (between 0.4 and 1) come from an underestimation of the aerodynamic resistance, in some cases combined with an overestimation of the surface temperature by a few degrees.

S-SEBI showed the best performance (among the SEB models) in the computation of λE with RMSD of 116 W m^{-2}. Opposite mean biases of -20 and 29 W m^{-2} have been observed on H and λE, respectively. A large underestimation of H occurs during senescence for the wheat plots. The vegetation is stressed but still dense and surface temperature and albedo are not as high as for bare-soil pixels. Senescence can be observed on the upper-right scatter plot of Fig. 2. Senescent wheat corresponds to yellow points with albedos between 0.2 and 0.3 and temperatures between 35 and 45 °C. They are located in the middle of the scatter plot because of the low surface temperatures induced by turbulent heat exchanges still going on above the canopy.

ICARE performs better than the other three models in the computation of H but falls behind in the estimation of λE with a RMSD of respectively 99 and 130 W m^{-2}. Contrary to SEB models, it tends to underestimate λE and overestimate H with biases of respectively -72 and $+26$ W m^{-2}.

3.3 Water stress

For hydrological applications, it is important to test not only the performance in estimating ETR (evapotranspiration) when water is not limiting (potential evapotranspiration being relatively easy to compute from NDVI and
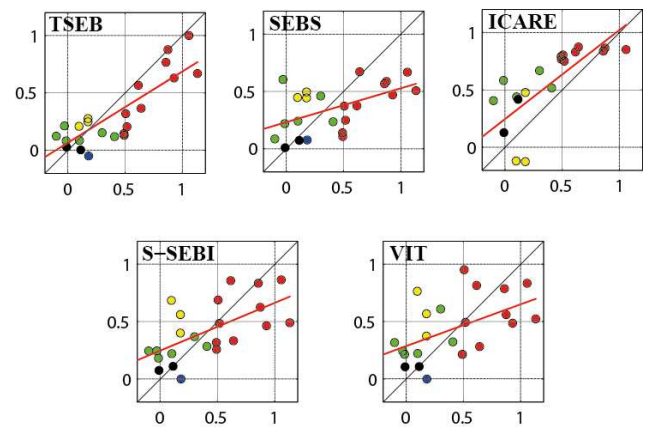
meteorological information) but during periods of water shortage as well. This is why we concentrate in this section on the assessment of the relative water stress of the surface. Of course, for agronomical applications, it is more interesting to get information about the water status of the plant. However, SEB models, TSEB excepted, only estimate evapotranspiration rates for the surface as a whole. Even in TSEB, the simulated vegetation component of the latent heat flux does not always represent transpiration. Indeed, when the vegetation is partially stressed, the soil evaporation term is set to zero and the transpiration term represents the evapotranspiration of the whole surface, excluding the case where the soil still evaporates at a very low rate. We computed a surface water stress at each measurement point for each date where ASTER data were available. Results are shown in Fig. 9. The definition of surface water stress reported here is the same as the one introduced as WDI for the VIT method (Eq. 17):

$$\text{water\_stress} = 1 - \frac{\lambda E}{\lambda E_{\text{max}}}. \tag{22}$$

In Eq. (22), λE_{max} is the maximum (potential) latent heat flux achievable for one particular pixel. Each SEB method uses its own potential conditions in order to compute actual evapotranspiration (Penman–Monteith in SEBS, Priestley–Taylor in TSEB and the available energy in S-SEBI). In order to assess the capacity of each model to retrieve water stress (which can be later on used as a proxy for total root zone soil moisture, cf. Hain et al., 2009), it appears only natural to consider water stress as an output of the models and to use their own version of potential conditions. In order to compute an in situ water stress, we used a two-source potential evapotranspiration model, which was the closest estimate amongst common models (Penman–Monteith, Priestley–Taylor), when compared to λE measurements at
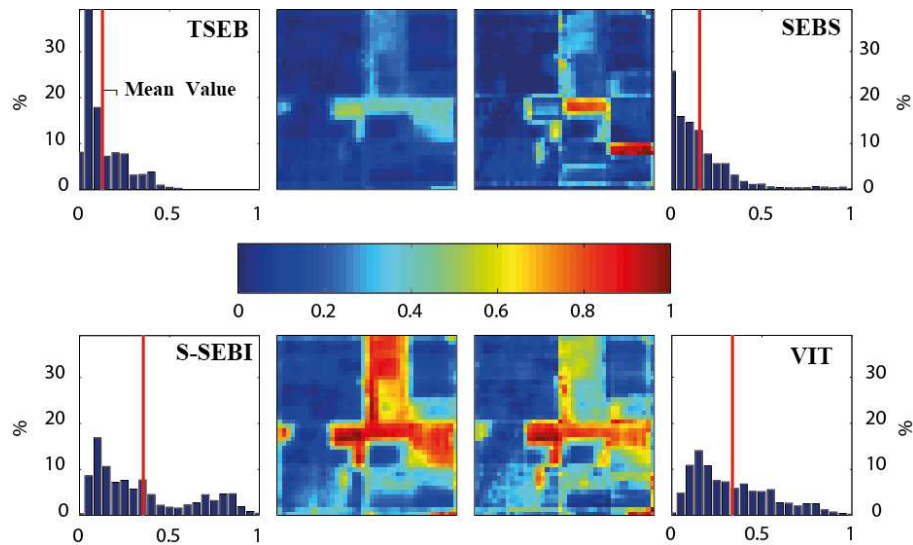
Fig. 10. Stress frequency histograms and maps for the whole area on 10 March.

each EC station for a short period following irrigation (not shown). Since potential evapotranspiration and average surface stress were not routine outputs of ICARE, we chose to run the SVAT model with a low continuous irrigation in order to obtain its own λE_{pot}.

In terms of water stress, ICARE would be expected to show good results, since it is the only model that is not forced with surface temperature (which is indirectly related to soil moisture) but directly with a time series of rain and irrigation, which control soil moisture and therefore stress levels. However, a significant overestimation of water stress is observable at medium and high LAIs. This seems to be due to the water balance module (force–restore), which tends to quickly dry the upper soil layer with bursts of λE_{s} on days immediately following irrigation.

TSEB globally underestimates stress, which is in agreement with the general overestimation of λE seen in Fig. 8, but it has a lower scattering and number of outliers than the other three SEB models. Additionally, it performs quite well for low stressed vegetation.

SEBS has a strong tendency to underestimate stress for low LAI, which is related to the overestimation of the surface aerodynamic resistance for bare soil discussed in Sect. 3.2. A group of points with large water stress values when stress is expected to be absent is, as explained in Sect. 3.2, mostly a consequence of an underestimation of the $k\,B^{-1}$ factor, and thus of the surface aerodynamic resistance.

The contextual models, as expected, dispatch all stress levels between extremes and produce a large spread of stress levels. Around observed medium stress values (0.5), S-SEBI and VIT simulate lower than observed stress values (red dots, LAI < 0.4). This is directly due to the observed surface temperatures at those locations, which are not high enough for the $T(\alpha)$ relation to detect a significant stress.

The corresponding points are in the lower part of the scatter plot so S-SEBI simulates evaporative fractions over 0.5 when measured values lie around 0.3 or 0.4. Two yellow points with intermediate LAIs (0.4–0.8) present overestimated stress values. This is due, for the one in the upper part of the diagram (chickpea, on 10 March), to an overestimation of the ASTER temperature of around 7 °C (one must keep in mind that the ASTER pixel is bigger than the EC instrument footprint and includes more bare soil with higher temperatures). The one in the lower part of the diagram (beans, 27 April) seems to be due to a high surface temperature of the whole bean field (around 40 °C), which is consistent with the station measurement, while the crop is still growing and well irrigated. Figure 10 shows an example of how contextual methods are dependent on the surface properties of the studied area. In the case of the VIT method, pixels with stressed and well developed vegetation (i.e., isolated pixels with high NDVI and high surface temperature values) seem to be lacking in the area, which is consistent with the agricultural practices: the whole zone being irrigated, stress is normally absent in the growing period. The boundaries drawn from the scatter of points shown in Fig. 2 are therefore misinterpreted as stressed for various pixels with high NDVI values located near the top-right edge of the trapezoid, while those points correspond to unstressed crops. However, stress is underestimated for some pixels classified as bare soils, which corresponds to pixels with low NDVI and a lower temperature than the observed maximum. If some straw is left on the ground after harvesting or if the bare-soil properties are different (e.g., higher albedo), evaporation is reduced. But since the temperature is in the lower range, the pixel is not located close enough to the bottom-right corner of the trapezoid to detect a strong reduction in evaporation.

Table 4. Mean values of turbulent flux and water stress on the whole area at two dates of ASTER overpass.

	10 March				6 May			
	TSEB	SEBS	S-SEBI	VIT	TSEB	SEBS	S-SEBI	VIT
H_{mean} (W m^{-2})	101.31	154.21	173.00	–	230.13	289.66	209.39	–
$\lambda E_{\mathrm{mean}}$ (W m^{-2})	425.57	372.67	353.88	–	272.70	207.76	288.02	–
Water stress	0.13	0.11	0.35	0.38	0.46	0.56	0.44	0.40

3.4 Spatial variability

In this section we move from the local to the spatial standpoint. Model intercomparison at a local scale allows us to assess their performance and draw specific behaviors but not to assess how they represent the spatial variability at the perimeter scale. In Fig. 11, we plotted frequency histograms for turbulent fluxes and remote sensing data (albedo, NDVI, surface temperature) on 10 March for the whole area. At this time of the year, most crops are well developed and green. On the histograms, TSEB and SEBS have a similar response: low H and high λE peaking around the potential rate. However, as expected for a contextual model, S-SEBI shows a large spread of values, which in this case does not seem to be representative of the real situation. As seen before, TSEB tends to simulate large values of latent heat flux for green landscapes. This is due to the starting hypothesis of the algorithm, which states that vegetation is transpiring at the potential rate given by the Priestley–Taylor formulation. This result is observed as well in Table 4, where arithmetical means of turbulent flux and water stress over the whole area are displayed. TSEB $\lambda E_{\mathrm{mean}}$ is higher than the other two (425 W m^{-2}), the smallest being S-SEBI with a $\lambda E_{\mathrm{mean}}$ of 354 W m^{-2}. The results in terms of stress are shown in Fig. 10. Both contextual models compute a larger amount of stressed areas than the other models. This is expected since they distribute stress values on the whole 0–1 interval. It is particularly true for the beans and chili pepper fields (heavily stressed fields in the northeastern 2 km × 2 km square, see Fig. 3). The surface temperature of those fields being high, the contextual models interpret them as stressed, whereas both crops are in an early growing stage and thus well irrigated. The reason of those high temperatures can be related to a large surface aerodynamic resistance (low LAIs and crop height) and to the effect of rows (large surfaces of bare soil are observed between crop rows by ASTER because the plant has yet to develop itself horizontally) rather than low water availability for plants. Single-pixel models compute surface water stress using theoretical potential rates taking into account LAI and thus tend to minimize this problem. This also shows the limitations of the use of a surface water stress to assess information about plant stress. Both models have mean water stress values of around 0.35 whereas single-pixel models compute low mean stress values (0.11 for SEBS, 0.13 for TSEB). Contrarily to TSEB, SEBS computes very high stress values (null
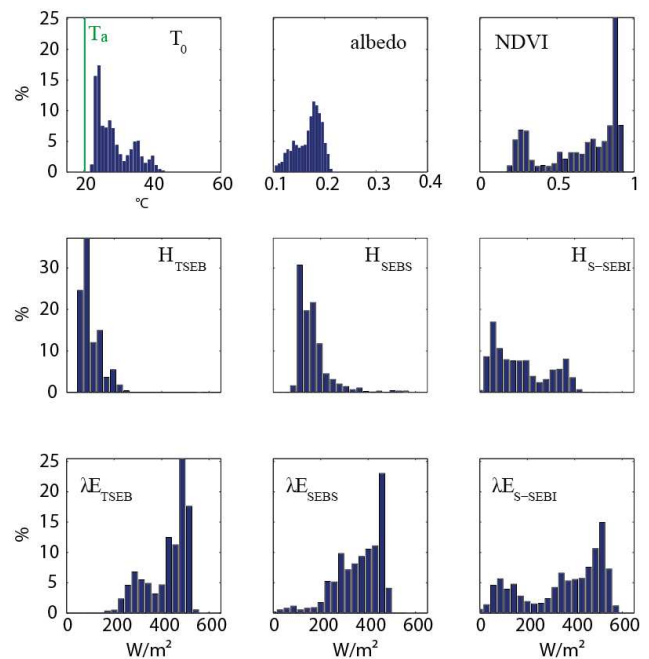
Fig. 11. Frequency histograms of simulated turbulent-calculated fluxes and remote sensing parameters on 10 March. On the T_0 histogram, the value of air temperature T_a is indicated as a green line.

evapotranspiration) for two orchard fields in the southeastern 2 km × 2 km part of the research area. Those sparse canopies have large amounts of intercropping bare soils with large temperatures. Both fields are irrigated, which means that the average surface stress is sensitive to fraction cover. For this land use type, one can question the validity of the de facto parameterizations of kB^{-1} and fraction cover from mean NDVI values, which have been built for herbaceous vegetation. TSEB seems to better accommodate the contrasted water status of those fields with its two-source energy balance.

In Fig. 10, as well, SEBS shows medium peaks of stress at the edge of many fields. Similar patterns are observed on NDVI (not shown) and thus LAI, and correspond to roads between the fields lowering the average NDVI of the surrounding pixels. Similar but more subtle edge effects are observed on VIT stress maps. Although it was expected that the VIT method would be the most sensitive to variations of NDVI,
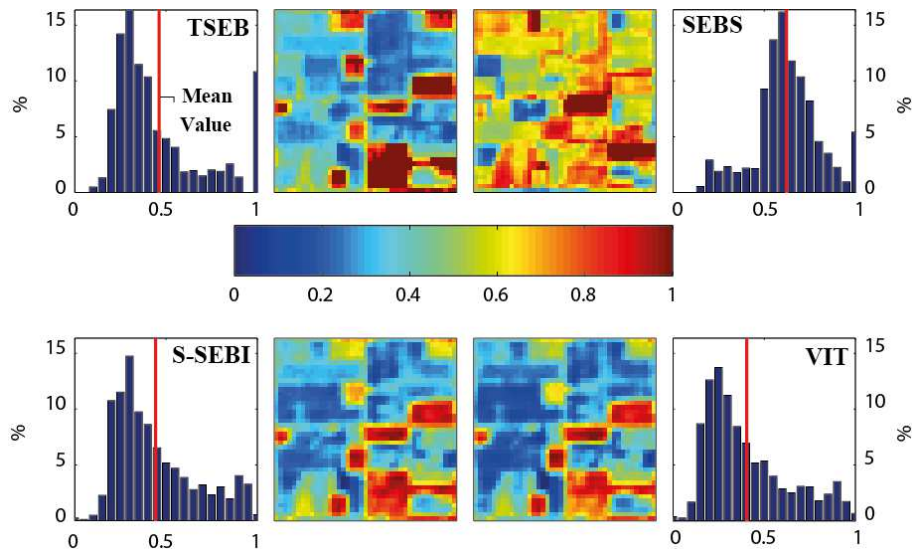
Fig. 12. Stress frequency histograms and maps for the whole area on 6 May.

it seems that the kB^{-1} parameterization in SEBS induces a high sensitivity of the model to vegetation cover.

On 6 May both well watered and stressed vegetation were present in the area. As expected, in heterogeneous conditions, contextual models react consistently to variations of the surface water status (see Fig. 12). TSEB, S-SEBI and VIT show very similar results in terms of patterns and stress distribution as well as mean stress value. Again, SEBS distinguishes itself from the others with a higher mean stress value (see Table 4) and a very different distribution of stress values over the area. Its difficulty to accommodate the hot bare-soil fractions is discussed below.

4 Discussion

Considering that we are looking for a possible assimilation of thermal data into SVAT or, more largely, hydrological models, SEB models – which provide information on the water status of the surface combining TIR, visible/NIR and meteorological data – seem like a decent lead. Understanding of their respective errors and robustness depending on surface and climatic conditions of the area is crucial.

4.1 Performance intercomparison

Models have shown fairly high errors in the computation of turbulent fluxes (RMSD over $100\,\mathrm{W\,m^{-2}}$ for the SEB methods), but it remains in the upper limit of what has already been published (see Table 1). Although part of those errors can be attributed to remote sensing uncertainties, most components of the energy budget are often large in semi-arid lands at low latitudes. In addition, fluxes are measured and simulated near midday, close to the peak in incoming solar radiation, and instantaneous energy flux values are also

large. It is therefore not surprising that RMSD values are in the upper range of the literature. Moreover, the Yaqui experiment was a "one shot" program carried out over a single cereal growing season and thus could not benefit from the experience in data understanding and correction that long-term projects like FLUXNET (http://fluxnet.ornl.gov/) can provide. Flux data quality is therefore also questionable.

In addition to the general performance of the models over the whole season, we had a closer look at their performance over different crop types and at different dates. ICARE performed better than the other methods over wheat and chick-pea fields, showing RMSD on turbulent fluxes lower by 20–$70\,\mathrm{W\,m^{-2}}$ than the others. It could be expected for wheat fields since the calibration of the soil resistance to evaporation and minimum stomatal resistance has been made over those crops. In addition, a fairly strong bias is present on the net radiation calculated by the three SEB methods, which tends to worsen their performance. However, over potatoes, sorghum, broccoli and beans, the three SEB models perform better, particularly over sorghum where the three models present RMSD of around 70–$90\,\mathrm{W\,m^{-2}}$. Despite the higher global performance of ICARE than the energy balance models, it seems that SEB models can give better results for some specific vegetation types. This can be partly due to their relative simplicity compared to ICARE, which needs a fair amount of surface parameters difficult to assess precisely. It could be expected that the SEB methods that use remotely observed surface temperature and albedo would be more consistent than the more complex SVAT model that internally calculates those parameters from user-defined properties of the soil and vegetation.

Another observation made during this experiment is that the relative error in surface temperature computation by ICARE seems to compensate the error made for the turbulent

fluxes. Indeed, the situation where ICARE is not accurate in surface temperature but has a very acceptable error in H and λE is not uncommon and it questions the relevance of assimilation of thermal data alone in SVAT models in order to improve its performance in turbulent flux determination. For instance, on 6 May, there is a 7 °C difference between the simulated and the observed surface temperatures for the wheat (1) station, while errors in turbulent heat flux are lower than 50 W m^{-2}. On 6 March, there is also a 7 °C temperature difference for the chickpea flux station while the sensible heat is correct (corresponding error is 10 W m^{-2}) and the latent heat is underestimated by more than 70 W m^{-2}. In those cases, correcting the model estimations based on thermal information may not be sufficient to correctly force ICARE into the right partition between the various components of the energy balance.

4.2 Models' structures and improvements

One big issue of the SEB models in their original formulation is that they have difficulties to account for the senescence phenomenon. For contextual methods, senescent vegetation has a lower temperature than stressed green vegetation (for VIT), due to a higher reflection of the incoming solar energy, and than bare, sandy soil (for S-SEBI), due to a higher crop height that enhances turbulent heat exchange at the surface. Since senescent vegetation pixels do not appear in extreme temperature conditions, those models do not detect very large stress values. RMSD on senescent pixels was found to be 40 W m^{-2} higher than for green ones.

SEBS tends to overestimate evapotranspiration for low LAIs. This seems to be mostly due to the bare-soil component of the $k\,B^{-1}$ factor that generates very high surface aerodynamic resistances and thus very low H values. RMSDs of 170 and 240 W m^{-2} are observed on latent and sensible heat fluxes (respectively) for pixels with LAI lower than 0.2. Solutions to this problem have been proposed. They consist in either using another empirical factor in order to compute turbulent fluxes (Boulet et al., 2012), including information about soil moisture in the $k\,B^{-1}$ (Gokmen et al., 2012) or lowering the soil component of the $k\,B^{-1}$ (Chen et al., 2013). Implementation of the Boulet et al. (2012) formulation leads to performance similar to the original SEBS but tends to overcorrect errors on λE with a lower bias of -70 W m^{-2} (against $+40$ W m^{-2} for Su, 2002). Overestimations of H at intermediate LAIs ($0.4 < \text{LAI} < 1$) seems to confirm the limitations of one-source methods and the advantage of using two-source approaches. Even if the new formulation of the $k\,B^{-1}$ factor introduced by Su (2002) was meant to add information about soil–plant interactions, the corresponding term in Eq. (8) seems to be underestimated. The generated error is lower than for bare soil but still needs to be noted, with a RMSD of around 130 W m^{-2} on H against 83 W m^{-2} for pixels with LAI larger than 1.

A major issue was encountered in TSEB during the senescence period in the partitioning of net radiation between soil and vegetation. The model has various new versions that modify the partitioning of net radiation between the soil and the vegetation for row crops, for instance, but we kept the original formulation by Norman et al. (1995). RMSD was found higher for senescent pixels than for bare soil and green ones (144, 106 and 109 W m^{-2}, respectively). Indeed, the LAI used in this study is computed from NDVI and is therefore a green LAI. However, during senescence, part of the canopy becomes yellow when drying. For crops like wheat, vegetation is fully yellow before harvesting. Our version of TSEB assimilates this part of vegetation as bare soil since it is not taken into account in the green LAI. It results in a great overestimation of the soil component of the net radiation for TSEB. Since the soil heat flux is bounded by a maximum fraction of $R_{\text{n,s}}$ and the soil sensible heat flux is well constrained by the soil temperature mostly (which is in turn well defined by T_0 and T_c), the residual soil latent heat flux is the most sensitive to big variations in $R_{\text{n,s}}$. It translates into a large overestimation of λE_s, and by consequence of λE. In order to integrate information about total and dry LAI into TSEB, we made the coarse assumption that during the senescence period, and until the harvesting date, the global LAI was constant and equal to its maximum value LAI_{max}. Global LAI extracted from the FORMOSAT-2 time series was used to compute the fraction cover during senescence and thus to better partition R_n. We then calculated the green fraction f_g according to Norman et al. (1995) with the new definition of global LAI:

$$f_g = \frac{\text{LAI}_{\text{green}}}{\text{LAI}_{\text{global}}} = \frac{\text{LAI}_{\text{green}}}{\text{LAI}_{\text{max}}}. \tag{23}$$

This approximation is not very accurate since the leaf's surface decreases while drying, leading to a gradual decrease of global LAI over the senescence period. However it is sufficient to greatly improve the repartition of net radiation and thus the estimation of λE and water stress. Results on λE and net radiation with this new formulation of LAI are compared to the initial version of the model in Fig. 13. Large variations are observed between the two formulations on soil and canopy net radiation, with differences as large as 500 W m^{-2}. This influences greatly the energy balance at the soil–atmosphere interface and produces smaller $R_{\text{n,s}}$ values. As a result, RMSD on λE is reduced by 10 W m^{-2} and the bias is lowered to 19 W m^{-2} (against 82 W m^{-2} with the initial formulation). The computation of water stress is also improved with a very low positive bias of 0.03 against -0.1 in the previous version (not shown). One must also stress that since we use global LAI in the computation of surface aerodynamic resistances, it tends to lower both r_a and r_s by adding more roughness to the surface and thus to favor sensible heat transfer at the soil interface over λE_s. Thus, integrating information about dry vegetation could greatly improve the performance of TSEB in senescent cases when vegetation
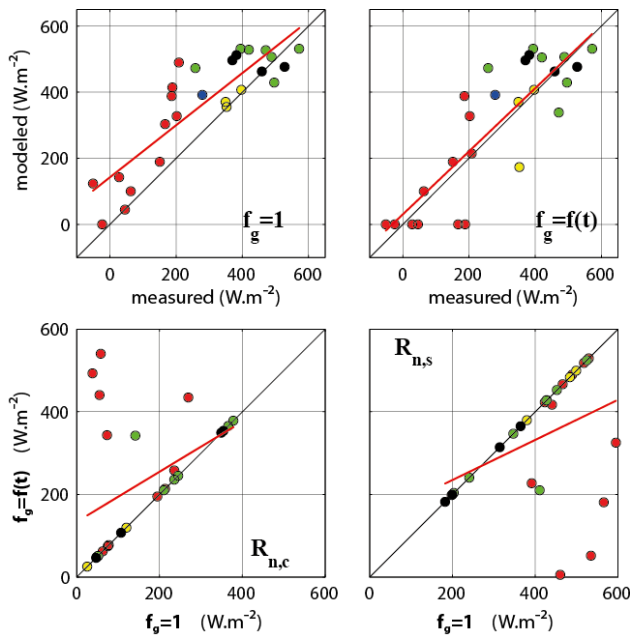
Fig. 13. Comparison of TSEB and observed λE fluxes with $f_g = 1$ (top left panel) and with f_g taking into account the drying of leaves (top, right), and corresponding values of the net radiation components for the vegetation ($R_{n,c}$, bottom left panel) and the soil ($R_{n,s}$, bottom right panel) with $f_g = 1$ (x axis) and with f_g taking into account the drying of leaves (y axis), respectively. Color code: red: LAI < 0.4; yellow: $0.4 \leq$ LAI < 0.8; green: $0.8 \leq$ LAI < 1.2; navyblue: $1.2 \leq$ LAI < 2.0; black: LAI $\geq$ 2.0.

height is reconstructed using LAI or NDVI time series. In our case, many cloud-free FORMOSAT-2 images were available, allowing us to follow each phenological stage of the vegetation and thus estimate the maximum LAI precisely. However, even if we hope that high-resolution images acquired frequently enough will be easier to afford in the near future, such clear-sky conditions over the whole growth season seem almost impossible to find in temperate or tropical regions and it would make the calibration of f_g much more delicate.

Comparing behaviors of TSEB and ICARE is also interesting since they are both dual-source models. They partition radiative energy between soil and vegetation in fairly different ways. In the original TSEB formulation, global net radiation is distributed between soil and vegetation based on fraction cover whereas ICARE computes two separate net radiations from LAI, soil and vegetation albedo, and emissivity, using a multireflection and transmission network, as proposed by Shuttleworth and Wallace (1985). The original TSEB method for net radiation can be criticized because it bypasses the effects of the vegetation's transmissivity and longwave radiation exchanges between soil and vegetation (emission of radiation from one layer to another). In Tang et al. (2011), a more physical way to calculate R_n is used, taking into account those two phenomena, and delivers better performance than a MODIS-based R_n (RMSD = 24 versus

44 W m^{-2} on a MODIS pixel). However, this method requires estimates of vegetation and soil albedos, as well as the vegetation transmission factor, which can be quite difficult if studying a very heterogeneous zone. The method used in ICARE considers all radiation exchanges between the two layers (vegetation and bare soil), including transmission and multiple reflections, but it assumes that the vegetation cover is homogeneous. It is not adapted to row crops (see Colaizzi et al., 2012 for recent formulations of net radiation for row crops). As long as FORMOSAT-2 images are available, the method used in this paper seems to be the most accurate for estimating R_n independently of the heterogeneity of the surface or regional topographic features. Furthermore, the assimilation of this net radiation into the ICARE SVAT model could be studied in order to calibrate vegetation and soil radiative properties.

In the partitioning of the soil and vegetation latent heat flux components, the two methods have also very different behaviors. Underestimation of λE by ICARE during the growing period is due to the quick drying of the first layer of soil, resulting in very low λE_s. However, since ICARE takes into account the dry part of the vegetation through green and dry LAIs, it performs better during senescence, with realistic low λE_c values, whereas TSEB maintains a potential rate for the vegetation. As a consequence, canopy temperature simulated by TSEB is almost equal to the air temperature while the soil temperature computed by the model rises towards ASTER surface temperature values. On the contrary, ICARE distributes both temperatures around the computed surface temperature. Although the partition of radiation sounds more realistic with TSEB (subject to some changes about dry LAI), the soil–vegetation partition between turbulent fluxes seems to make more sense for ICARE, mostly because of the strong initial hypothesis on vegetation water status in TSEB. After correction, TSEB seems to be the most accurate of all four SEB models after reconstitution of total fluxes, but its soil–canopy distribution of latent heat and estimates of component temperatures remains questionable.

4.3 Determination of distributed water stress

We focus here on the intercomparison of stress patterns simulated by all four SEB models. In order to remain as concise as possible, we chose not to discuss model results on every ASTER image but to focus on the dates that can represent the most significant cases encountered in this study. Differences were particularly important for two dates that represent contrasted vegetation cover and soil moisture patterns: one, 10 March, when the quasi-totality of the area is green and irrigated (and thus the spatial variability of water availability is relatively low) and the other, 6 May, when wheat has been harvested and thus the major part of the area is under water stress whereas many crops are still growing and irrigated (and thus the spatial variability of water availability is fairly high). We decided to focus on these two dates. In

Table 5. Double-entry table of coefficient of determination R^2 (lower-left part) and RMSD (upper-right part) between the SEB models and their standard deviation σ on 10 March.

		RMSD				σ
		TSEB	SEBS	S-SEBI	VIT	
R^2	TSEB		0.15	0.27	0.28	0.14
	SEBS	0.43			0.27	0.18
	S-SEBI	0.94	0.46		0.06	0.26
	VIT	0.91	0.58	0.95		0.27

Table 6. Double-entry table of coefficient of determination R^2 (lower-left part) and RMSD (upper-right part) between the SEB models and their standard deviation σ on 6 May.

		RMSD				σ
		TSEB	SEBS	S-SEBI	VIT	
R^2	TSEB		0.28	0.10	0.13	0.25
	SEBS	0.07		0.23	0.25	0.18
	S-SEBI	0.84	0.28		0.04	0.22
	VIT	0.78	0.32	0.95		0.23

Tables 5 and 6 coefficients of determination R^2 and RMSD values of the simulated water stress by the different models are displayed, as well as their standard deviation σ over the whole area. As expected for contextual models, which distribute stress over the whole 0–1 interval, S-SEBI and VIT show higher σ values than single-pixel models on 10 March. The quasi-totality of crops are green and well developed and stress should be almost absent of the scene. We enlarged the target area (from $4 \, \text{km} \times 4 \, \text{km}$ to the whole FORMOSAT-2 image) hoping to encounter truly stressed pixels, but the slight change in the resulting bounding relationships did not modify this finding (not shown). This tends to confirm that our sample zone is representative of the whole irrigation perimeter. For that date, the missing extremes should be inferred from theoretical boundaries and not derived from the images alone. For a later date in spring (on 6 May when wheat is mostly senescent), TSEB and contextual models are in good agreement with each other, showing equivalent σ values, and low RMSDs between models (a difference of around 0.1 between TSEB and contextual models, 0.04 between S-SEBI and VIT).

TSEB, S-SEBI and VIT show a strong correlation with each other on both dates but with a narrower interval of distribution than for contextual models, which can explain their differences in terms of mean values and standard deviation. In May, the correlation between TSEB and the contextual models is a bit lower than in March but the pattern shown in Figs. 10 and 12, as well as the mean value and standard deviation, are very similar. It shows that in the case of contrasted conditions, simpler contextual models reproduce quite faithfully the general behavior of a more complex model like TSEB, which is already used in operational algorithms at continental scales and thus whose performance is trusted. On the contrary, SEBS seems to behave very differently from the other three models during the whole season in terms of stress distribution, with a low standard deviation in winter as well as in spring. It has a very low correlation with the other three models and the patterns and histograms displayed in Figs. 10 and 12 are very distinct from the others. One possible explanation could be that SEBS is more sensitive to vegetation properties due to its use of the $k\,B^{-1}$ parameterization of the roughness length for heat exchange in the aerodynamic

resistance. The resulting roughness is very sensitive to the vegetation height, which is very difficult to retrieve spatially. Thus a lot of approximations are made using a priori values based on in situ qualitative knowledge in order to distribute crop heights over the whole area. The contextual models do not need this information and TSEB is strongly driven by fraction cover and surface temperature in his architecture, which enables it to by-pass some crop height determination issues.

5 Summary and conclusions

Performance and structure particularities of two contextual and two single-pixel methods to retrieve energy fluxes at the surface using thermal remote sensing data have been locally compared with in situ measurements and outputs of a complete SVAT model (ICARE) during a whole cereal-growth season. In terms of energy fluxes, TSEB, SEBS and S-SEBI showed comparable results with RMSD on λE ranging from 117 (S-SEBI) to 131 W m^{-2} (SEBS). These results are in the same range as the uncalibrated ICARE SVAT model forced with in situ data (RMSD = 131 W m^{-2} on λE) but with an opposite behavior in the repartition of turbulent fluxes. ICARE tends to underestimate the evapotranspiration whereas the SEB methods overestimate it. TSEB and ICARE are the two models that estimate water stress with the best accuracy. However, TSEB performs better at high LAIs (low stress) and shows difficulties in detecting stress during senescence, whereas ICARE has a strong tendency to overestimate stress for green vegetation but is more accurate than TSEB at low LAIs. SEBS performs poorly for senescent and bare-soil situations and contextual models present a lot of dispersion. Corrections for TSEB and SEBS have been suggested in order to account for their respective limitations in processing dry vegetation and low LAI cases.

From the spatial standpoint, general behaviors of the models have been described. SEBS distinguishes itself from the others in its way of incorporating and processing vegetation data, resulting in a singular distribution of stress. However, for dates with a strong contrast in surface soil moisture, contextual models have shown mean stress values and stress

patterns very similar to TSEB, whereas for dates with uniform surface moisture (winter and summer times mostly), they tend to accentuate extreme values of water stress. This was expected since they are self-calibrated and thus distribute the values between fully stressed and potential conditions for the whole image. In this paper, we used an automatic determination of the empirical laws driving contextual models. A manual selection was also tested but did not produce better or worse statistics (not shown). However, a user-monitored determination of the bounding curves might still be the only way to decide whether extremes correspond to target situations (for instance, wet bare soils or dense vegetation under water stress) or meaningless outliers.

This work has been carried out as a preliminary study of the SEB models in order to assess if thermal remote sensing data could yield valuable information to assimilate into a land surface or hydrological model. Results have shown that the single-pixel SEB models, after various modifications, can provide relevant information about water stress, especially during the growing period where ICARE tends to overestimate stress. By combining data assimilation of surface temperature, water stress index and broadband albedo in a Kalman ensemble assimilation scheme, both performance amelioration (by adjusting soil water content) and calibration of ICARE could be done (see Er-raki et al., 2008; Cammalleri and Ciraolo, 2012; Merlin et al., 2014). Furthermore, as current environmental and water use issues require a more regional point of view, distributed SVAT models could be an interesting tool, but very difficult to implement because of the lack of knowledge in subsurface characteristics of wide areas. Assimilation of thermal data could be a way to calibrate those models with a limited amount of information about the surface. Finally, it seems that contextual and single-pixel approaches could be merged in order to take advantage of the increased accuracy of the later in homogeneous conditions (winter and summer times) while keeping the uncertainty reduction of the former for very contrasted landscapes.

The publication of this article is financed by CNRS-INSU.

References

Allen, R. G., Tasumi, M., and Trezza, R.: Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC) – Model, J. Irrig. Drain. Eng., 133, 380–394, doi:10.1061/(ASCE)0733-9437(2007)133:4(380), 2007.

Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., González-Dugo, M. P., Cammalleri, C., d'Urso, G., Pimstein, A., and Gao, F.: Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery, Hydrol. Earth Syst. Sci., 15, 223–239, doi:10.5194/hess-15-223-2011, 2011.

Barr, A. G., King, K. M., Gillespie, T. J., Hartog, G., and Neumann, H. H.: A comparison of bowen ratio and eddy correlation sensible and latent heat flux measurements above deciduous forest, Bound.-Lay. Meteorol., 71, 21–41, doi:10.1007/BF00709218, 1994.

Bastiaanssen, W. G. M.: SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey, J. Hydrol., 229, 87–100, doi:10.1016/S0022-1694(99)00202-4, 2000.

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., and Holtslag, A. A. M.: A remote sensing surface energy balance algorithm for land (SEBAL), 1. Formulation, J. Hydrol., 212–213, 198–212, 1998.

Blanken, P. D., Black, T. A., Yang, P. C., Neumann, H. H., Nesic, Z., Staebler, R., den Hartog, G., Novak, M. D., and Lee, X.: Energy balance and canopy conductance of a boreal aspen forest: Partitioning overstory and understory components, Journal of Geophysical Research, 102, 28915, doi:10.1029/97JD00193, 1997.

Boulet, G., Chehbouni, A., Gentine, P., Duchemin, B., Ezzahar, J., and Hadria, R.: Monitoring water stress using time series of observed to unstressed surface temperature difference, Agr. Forest Meteorol., 146, 159–172, doi:10.1016/j.agrformet.2007.05.012, 2007.

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, doi:10.1016/j.agrformet.2012.03.008, 2012.

Brutsaert, W.: Aspects of bulk atmospheric boundary layer similarity under free-convective conditions, Rev. Geophys., 37, 439–451, doi:10.1029/1999RG900013, 1999.

Bsaibes, A., Courault, D., Baret, F., Weiss, M., Olioso, A., Jacob, F., Hagolle, O., Marloie, O., Bertrand, N., Desfond, V., and Kzemipour, F.: Albedo and LAI estimates from FORMOSAT-2 data for crop monitoring, Remote Sens. Environ., 113, 716–729, doi:10.1016/j.rse.2008.11.014, 2009.

Cammalleri, C. and Ciraolo, G.: State and parameter update in a coupled energy/hydrologic balance model using ensemble Kalman filtering, J. Hydrol., 416-417, 171–181, doi:10.1016/j.jhydrol.2011.11.049, 2012.

Cammalleri, C., Anderson, M. C., Ciraolo, G., D'Urso, G., Kustas, W. P., La Loggia, G., and Minacapilli, M.: Applications of a remote sensing-based two-source energy balance algorithm for mapping surface fluxes without in situ air temperature observations, Remote Sens. Environ., 124, 502–515, doi:10.1016/j.rse.2012.06.009, 2012.

Chehbouni, A., Escadafal, R., Duchemin, B., Boulet, G., Simonneaux, V., Dedieu, G., Mougenot, B., Khabba, S., Kharrou, H., Maisongrande, P., Merlin, O., Chaponnière, A., Ezzahar, J., Er-Raki, S., Hoedjes, J., Hadria, R., Abourida, A., Cheggour, A., Raibi, F., Boudhar, A., Benhadj, I., Hanich, L., Benkaddour, A., Guemouria, N., Chehbouni, A. H., Lahrouni, A., Olioso, A., Jacob, F., Williams, D. G., and Sobrino, J. A.: An integrated modelling and remote sensing approach for hydrological study in arid and semi-arid regions: the SUDMED Programme, Int. J. Remote Sens., 29, 5161–5181, doi:10.1080/01431160802036417, 2008a.

Chehbouni, A., Hoedjes, J., Rodriguez, J.-C., Watts, C., Garatuza-Payan, J., Jacob, F., and Kerr, Y.: Using remotely sensed data to estimate area-averaged daily surface fluxes over a semi-arid mixed agricultural land, Agr. Forest Meteorol., 148, 330–342, doi:10.1016/j.agrformet.2007.09.014, 2008b.

Chen, X., Su, Z., Ma, Y., Yang, K., Wen, J., and Zhang, Y.: An Improvement of Roughness Height Parameterization of the Surface Energy Balance System (SEBS) over the Tibetan Plateau, J. Appl. Meteorol. Clim., 52, 607–622, doi:10.1175/JAMC-D-12-056.1, 2013.

Chern, J.-S., Ling, J., and Weng, S.-L.: Taiwan's second remote sensing satellite, Acta Astronautica, 63, 1305–1311, doi:10.1016/j.actaastro.2008.05.022, 2008.

Choi, M., Kustas, W. P., Anderson, M. C., Allen, R. G., Li, F., and Kjaersgaard, J. H.: An intercomparison of three remote sensing-based surface energy balance algorithms over a corn and soybean production region (Iowa, U.S.) during SMACEX, Agr. Forest Meteorol., 149, 2082–2097, doi:10.1016/j.agrformet.2009.07.002, 2009.

Choudhury, B. and Monteith, J.: A four-layer model for the heat budget of homogeneous land surfaces, Q. J. Roy. Meteorol. Soc., 114, 373–398, 1988.

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W.: Regional evaporation estimates from flux tower and MODIS satellite data, Remote Sens. Environ., 106, 285–304, doi:10.1016/j.rse.2006.07.007, 2007.

Clevers, J.: The Application of a Weighted Infrared-Red Vegetation Index for Estimating Leaf-Area Index by Correcting for soil-moisture, Remote Sens. Environ., 29, 25–37, doi:10.1016/0034-4257(89)90076-x, 1989.

Colaizzi, P. D., Evett, S. R., Howell, T. A., Li, F., Kustas, W. P., and Anderson, M. C.: Radiation Model for Row Crops: I. Geometric View Factors and Parameter Optimization, Agronomy J., 104, 225–240, doi:10.2134/agronj2011.0082, 2012.

Coudert, B. and Ottlé, C.: An improved SVAT model calibration strategy based on the optimisation of surface temperature temporal dynamics, Geophys. Res. Lett., 34, L04402, doi:10.1029/2006GL028778, 2007.

Courault, D., Seguin, B. and Olioso, A.: Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches, Irrig. Drain. Syst., 19, 223–249, doi:10.1007/s10795-005-5186-0, 2005.

Courault, D., Bsaibes, A., Kpemlie, E., Hadria, R., Hagolle, O., Marloie, O., Hanocq, J.-F., Olioso, A., Bertrand, N., and Desfonds, V.: Assessing the Potentialities of FORMOSAT-2 Data for Water and Crop Monitoring at Small Regional Scale in South-Eastern France, Sensors, 8, 3460–3481, doi:10.3390/s8053460, 2008.

Delogu, E., Boulet, G., Olioso, A., Coudert, B., Chirouze, J., Ceschia, E., Le Dantec, V., Marloie, O., Chehbouni, G., and Lagouarde, J.-P.: Reconstruction of temporal variations of evapotranspiration using instantaneous estimates at the time of satellite overpass, Hydrol. Earth Syst. Sci., 16, 2995–3010, doi:10.5194/hess-16-2995-2012, 2012.

Er-Raki, S., Chehbouni, A., Hoedjes, J., Ezzahar, J., Duchemin, B., and Jacob, F.: Improvement of FAO-56 method for olive orchards through sequential assimilation of thermal infrared-based estimates of ET, Agr. Water Manage., 95, 309–321, doi:10.1016/j.agwat.2007.10.013, 2008.

Fieuzal, R., Duchemin, B., Jarlan, L., Zribi, M., Baup, F., Merlin, O., Hagolle, O., and Garatuza-Payan, J.: Combined use of optical and radar satellite data for the monitoring of irrigation and soil moisture of wheat crops, Hydrol. Earth Syst. Sci., 15, 1117–1129, doi:10.5194/hess-15-1117-2011, 2011.

French, a. N., Jacob, F., Anderson, M. C., Kustas, W. P., Timmermans, W., Gieske, a., Su, Z., Su, H., McCabe, M. F., Li, F., Prueger, J., and Brunsell, N.: Surface energy fluxes with the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) at the Iowa 2002 SMACEX site (USA), Remote Sens. Environ., 99, 55–65, doi:10.1016/j.rse.2005.05.015, 2005.

Galleguillos, M., Jacob, F., Prévot, L., French, A., and Lagacherie, P.: Comparison of two temperature differencing methods to estimate daily evapotranspiration over a Mediterranean vineyard watershed from ASTER data, Remote Sens. Environ., 115, 1326–1340, doi:10.1016/j.rse.2011.01.013, 2011.

Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., and Duchemin, B.: Analysis of evaporative fraction diurnal behaviour, Agr. Forest Meteorol., 143, 13–29, doi:10.1016/j.agrformet.2006.11.002, 2007.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J. S., Hook, S., and Kahle, A. B.: A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images, IEEE T. Geosci. Remote, 36, 1113–1126, doi:10.1109/36.700995, 1998.

Gokmen, M., Vekerdy, Z., Verhoef, A., Verhoef, W., Batelaan, O., and van der Tol, C.: Integration of soil moisture in SEBS for improving evapotranspiration estimation under water stress conditions, Remote Sens. Environ., 121, 261–274, doi:10.1016/j.rse.2012.02.003, 2012.

Gomez, M., Olioso, a, Sobrino, J., and Jacob, F.: Retrieval of evapotranspiration over the Alpilles/ReSeDA experimental site using airborne POLDER sensor and a thermal camera, Remote Sens. Environ., 96, 399–408, doi:10.1016/j.rse.2005.03.006, 2005.

Gonzalez-Dugo, M. P., Neale, C. M. U., Mateos, L., Kustas, W. P., Prueger, J. H., Anderson, M. C., and Li, F.: A comparison of operational remote sensing-based models for estimating crop evapotranspiration, Agr. Forest Meteorol., 149, 1843–1853, doi:10.1016/j.agrformet.2009.06.012, 2009.

Hagolle, O., Dedieu, G., Mougenot, B., Debaecker, V., Duchemin, B., and Meygret, A.: Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: Application to Formosat-2 images, Remote Sens. Environ., 112, 1689–1701, doi:10.1016/j.rse.2007.08.016, 2008.

Hain, C. R., Mecikalski, J. R., and Anderson, M. C.: Retrieval of an Available Water-Based Soil Moisture Proxy from Thermal Infrared Remote Sensing, Part I: Methodology and Validation, J. Hydrometeorol., 10, 665–683, doi:10.1175/2008JHM1024.1, 2009.

Jacob, F. and Olioso, A.: Derivation of diurnal courses of albedo and reflected solar irradiance from airborne POLDER data acquired near solar noon, J. Geophys. Res., 110, D10104, doi:10.1029/2004JD004888, 2005.

Jacob, F., Olioso, A., Gu, X. F., Su, Z., and Seguin, B.: Mapping surface fluxes using airborne visible, near infrared, thermal infrared remote sensing data and a spatialized surface energy balance model, Agronomie, 22, 669–680, doi:10.1051/agro:2002053, 2002.

Jacob, F., Schmugge, T., Olioso, A., French, A., Courault, D., Ogawa, K., Petitcolin, F., Chehbouni, G., Pinheiro, A., and Privette, J.: Modeling and Inversion in Thermal Infrared Remote Sensing over Vegetated Land Surfaces, in: Advances in Land Remote Sensing, edited by: Liang, S., Springer Netherlands, 245–291, 2008.

Jia, L., Su, Z., van den Hurk, B., Menenti, M., Moene, A., De Bruin, H. A., Yrisarry, J. J. B., Ibanez, M., and Cuesta, A.: Estimation of sensible heat flux using the Surface Energy Balance System (SEBS) and ATSR measurements, Phys. Chem. Earth A/B/C, 28, 75–88, doi:10.1016/S1474-7065(03)00009-3, 2003.

Jiménez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A., Fisher, J. B., Jung, M., Kanamitsu, M., Reichle, R. H., Reichstein, M., Rodell, M., Sheffield, J., Tu, K., and Wang, K.: Global intercomparison of 12 land surface heat flux estimates, J. Geophys. Res., 116, D02102, doi:10.1029/2010JD014545, 2011.

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, doi:10.1007/s10712-008-9037-z, 2008.

Kustas, W. and Anderson, M.: Advances in thermal infrared remote sensing for land surface modeling, Agr. Forest Meteorol., 149, 2071–2081, doi:10.1016/j.agrformet.2009.05.016, 2009.

Kustas, W. and Daughtry, C.: Estimation of the soil heat flux/net radiation ratio from spectral data, Agr. Forest Meteorol., 49, 205–223, 1990.

Li, F., Kustas, W., Prueger, J., Neale, C., and Jackson, T.: Utility of remote sensing-based two-source energy balance model under low-and high-vegetation cover conditions, J, Hydrometeorol., 6, 878–891, 2005.

Li, F., Kustas, W. P., Anderson, M. C., Prueger, J. H., and Scott, R. L.: Effect of remote sensing spatial resolution on interpreting tower-based flux observations, Remote Sens. Environ., 112, 337–349, doi:10.1016/j.rse.2006.11.032, 2008.

Long, D. and Singh, V. P.: A Two-source Trapezoid Model for Evapotranspiration (TTME) from satellite imagery, Remote Sens. Environ., 121, 370–388, doi:10.1016/j.rse.2012.02.015, 2012.

Ma, W., Ma, Y., Hu, Z., Su, Z., Wang, J., and Ishikawa, H.: Estimating surface fluxes over middle and upper streams of the Heihe River Basin with ASTER imagery, Hydrol. Earth Syst. Sci., 15, 1403–1413, doi:10.5194/hess-15-1403-2011, 2011.

Martin, M., Dickinson, R. E., and Yang, Z.-L.: Use of a Coupled Land Surface General Circulation Model to Examine the Impacts of Doubled Stomatal Resistance on the Water Resources of the American Southwest, J. Climate, 12, 3359–3375, doi:10.1175/1520-0442(1999)012<3359:UOACLS>2.0.CO;2, 1999.

McCabe, M. F. and Wood, E. F.: Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors, Remote Sens. Environ., 105, 271–285, doi:10.1016/j.rse.2006.07.006, 2006.

McCabe, M. F., Wood, E. F., Wójcik, R., Pan, M., Sheffield, J., Gao, H., and Su, H.: Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, Remote Sens. Environ., 112, 430–444, doi:10.1016/j.rse.2007.03.027, 2008.

Merlin, O. and Chehbouni, A.: Different approaches in estimating heat flux using dual angle observations of radiative surface temperature, Int. J. Remote Sens., 25, 275–289, doi:10.1080/0143116031000116408, 2004.

Merlin, O., Duchemin, B., Hagolle, O., Jacob, F., Coudert, B., Chehbouni, G., Dedieu, G., Garatuza, J., and Kerr, Y.: Disaggregation of MODIS surface temperature over an agricultural area using a time series of Formosat-2 images, Remote Sens. Environ., 114, 2500–2512, doi:10.1016/j.rse.2010.05.025, 2010.

Merlin, O., Chirouze, J., Olioso, A., Jarlan, L., Chehbouni, G., and Boulet, G.: An image-based four-source surface energy balance model to estimate crop evapotranspiration from solar reflectance/thermal emission data (SEB-4S), Agr. Forest Meteorol., 184, 188–203, doi:10.1016/j.agrformet.2013.10.002, 2014.

Minacapilli, M., Agnese, C., Blanda, F., Cammalleri, C., Ciraolo, G., D'Urso, G., Iovino, M., Pumo, D., Provenzano, G., and Rallo, G.: Estimation of actual evapotranspiration of Mediterranean perennial crops by means of remote-sensing based surface energy balance models, Hydrol. Earth Syst. Sci., 13, 1061–1074, doi:10.5194/hess-13-1061-2009, 2009.

Monteith, J.: Principles of environmental physics, Edward Arnold, London, 1973.

Moran, M., Clarke, T., Inoue, Y., and Vidal, A.: Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index, Remote Sens. Environ., 49, 246–263, doi:10.1016/0034-4257(94)90020-5, 1994.

Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterisation scheme, Global Planet. Change, 13, 145–159, doi:10.1016/0921-8181(95)00043-7, 1996.

Noilhan, J. and Planton, S.: A Simple Parameterization of Land Surface Processes for Meteorological Models, Mon. Weather Rev., 117, 536–549, doi:10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2, 1989.

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, doi:10.1016/0168-1923(95)02265-Y, 1995.

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

Oki, T., and Kanae, S.: Global Hydrological Cycles and
World Water Resources, Science, 313, 1068–1072,
doi:10.1126/science.1128845, 2006.

Oku, Y., Ishikawa, H., and Su, Z.: Estimation of Land Surface Heat
Fluxes over the Tibetan Plateau Using GMS Data, J. Appl. Me-
teorol. Clim., 46, 183–195, doi:10.1175/JAM2456.1, 2007.

Olioso, A., Inoue, Y., and Ortega-Farias, S.: Future directions for
advanced evapotranspiration modeling: Assimilation of remote
sensing data into crop simulation models and SVAT models, Ir-
rig. Drain., 19, 377–412, 2005.

Passerat de Silans, A.: Transferts de masse et de chaleur dans un sol
stratifié soumis à une excitation atmosphérique naturelle (com-
paraison modèles-expérience), Institut National Polytechnique
de Grenoble, Grenoble, 1986.

Roerink, G., Su, Z., and Menenti, M.: S-SEBI: A simple re-
mote sensing algorithm to estimate the surface energy bal-
ance, Phys. Chem. Earth B, 25, 147–157, doi:10.1016/S1464-
1909(99)00128-8, 2000.

Sabol, D. E., Gillespie, A. R., Abbott, E., and Yamada, G.: Remote
Sensing of Environment Field validation of the ASTER Temper-
ature – Emissivity Separation algorithm, Remote Sens. Environ.,
113, 2328–2344, doi:10.1016/j.rse.2009.06.008, 2009.

Santanello, J. A. and Friedl, M. A.: Diurnal Covariation in Soil
Heat Flux and Net Radiation, J. Appl. Meteorol., 42, 851–862,
doi:10.1175/1520-0450(2003)042<0851:DCISHF>2.0.CO;2,
2003.

Schmugge, T., Hook, S., and Coll, C.: Recovering surface tempera-
ture and emissivity from thermal infrared multispectral data, Re-
mote Sens. Environ., 65, 121–131, 1998.

Schuurmans, J. M., Troch, P. A., Veldhuizen, A. A., Bastiaanssen,
W. G., and Bierkens, M. F. P.: Assimilation of remotely sensed
latent heat flux in a distributed hydrological model, Adv. Water
Resour., 26, 151–159, 2003.

Shiklomanov, I: World water resources: a new appraisal and
assessment for the 21st century, UNESCO, Paris, available
from: http://www.vliz.be/imis/imis.php?refid=11667 (last ac-
cess: 23 April 2013), 1998.

Shuttleworth, W. and Wallace, J.: Evaporation from sparse crops –
An energy combination theory, Q. J. Roy. Meteorol. Soc., 111,
839–855, doi:10.1256/smsqj.46909, 1985.

Su, H., McCabe, M., and Wood, E.: Modeling evapotranspira-
tion during SMACEX: Comparing two approaches for local-and
regional-scale prediction, J. Hydrometeorol., 6, 910–922, 2005.

Su, H., Wood, E., McCabe, M., and Su, Z.: Evaluation of remotely
sensed evapotranspiration over the CEOP EOP-1 reference sites,
J. Meteorol. Soc. Jpn., 85, 439–459, 2007.

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

Tang, R., Li, Z.-L., Jia, Y., Li, C., Sun, X., Kustas, W. P.,
and Anderson, M. C.: An intercomparison of three remote
sensing-based energy balance models using Large Aperture
Scintillometer measurements over a wheat–corn production
region, Remote Sensing of Environment, 115, 3187–3202,
doi:10.1016/j.rse.2011.07.004, 2011.

Timmermans, W. J., Kustas, W. P., Anderson, M. C., and French,
A. N.: An intercomparison of the Surface Energy Balance Al-
gorithm for Land (SEBAL) and the Two-Source Energy Balance
(TSEB) modeling schemes, Remote Sens. Environ., 108, 369–
384, doi:10.1016/j.rse.2006.11.028, 2007.

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser,
P. R., Meyers, T. P., Prueger, J. H., Starks, P. J., and Wesely, M.
L.: Correcting eddy-covariance flux underestimates over a grass-
land, Agr. Forest Meteorol., 103, 279–300, doi:10.1016/S0168-
1923(00)00123-4, 2000.

Van de Griend, A. A. and Owe, M.: On the Relationship between
Thermal Emissivity and the Normalized Difference Vegetation
Index for Natural Surfaces, Int. J. Remote Sens., 14, 1119–1131,
1993.

van der Kwast, J., Timmermans, W., Gieske, A., Su, Z., Olioso,
A., Jia, L., Elbers, J., Karssenberg, D., and de Jong, S.: Eval-
uation of the Surface Energy Balance System (SEBS) applied
to ASTER imagery with flux-measurements at the SPARC 2004
site (Barrax, Spain), Hydrol. Earth Syst. Sci., 13, 1337–1347,
doi:10.5194/hess-13-1337-2009, 2009.

Van Dijk, A., Moene, A. and De Bruin, H.: The principles
of surface flux physics: theory, practice and descrip-
tion of the ECPACK library, Wageningen, the Nether-
lands, available from: http://wwWmaq.wur.nl/NR/rdonlyres/
3C2D48B3-2D23-4EE7-BA70-342863EBF822/60091/ir2004_
01v13.pdf (last access: 1 March 2013), 2004.

Venturini, V., Islam, S., and Rodriguez, L.: Estimation of evapora-
tive fraction and evapotranspiration from MODIS products us-
ing a complementary based model, Remote Sens. Environ., 112,
132–141, doi:10.1016/j.rse.2007.04.014, 2008.

Verstraeten, W. W., Veroustraete, F., and Feyen, J.: Estimating evap-
otranspiration of European forests from NOAA-imagery at satel-
lite overpass time: Towards an operational processing chain for
integrated optical and thermal sensor data products, Remote
Sens. Environ., 96, 256–276, doi:10.1016/j.rse.2005.03.004,
2005.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D.,
Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field,
C., Grelle, A., Ibrom, A., Law, B., Kowalski, A., Meyers, T.,
Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valen-
tini, R., and Verma, S.: Energy balance closure at FLUXNET
sites, Agr. Forest Meteorol., 113, 223–243, doi:10.1016/S0168-
1923(02)00109-0, 2002.

World Water Assessment Programme: The United Nations World
Water Development Report 4: Managing Water under Uncer-
tainty and Risk, United Nat., edited by: Andrzejwski, C., Cullen,
A., Dash, B., Franek, A., Kelty, M., McDonald, D., Smith, M.,
Wauchope, S., and P. Publishing, UNESCO, Paris, 2012.

Yang, J. and Wang, Y.: Estimating evapotranspiration frac-
tion by modeling two-dimensional space of NDVI/albedo
and day–night land surface temperature difference: A
comparative study, Adv. Water Resour., 34, 512–518,
doi:10.1016/j.advwatres.2011.01.006, 2011.