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

J. Chirouze\textsuperscript{1}, G. Boulet\textsuperscript{1}, L. Jarlan\textsuperscript{1}, R. Fieuzal\textsuperscript{1}, J. C. Rodríguez\textsuperscript{3}, J. Ezzahar\textsuperscript{4}, S. Er-Raki\textsuperscript{5}, G. Bigeard\textsuperscript{1}, O. Merlin\textsuperscript{1}, J. Garatuza-Payan\textsuperscript{2}, C. Watts\textsuperscript{3}, and G. Chehbouni\textsuperscript{1}

\textsuperscript{1}Centre d’Etudes Spatiales de la Biosphère, UMR5126, UPS, CNRS, CNES, IRD, Toulouse, France
\textsuperscript{2}Instituto Tecnológico de Sonora, Ciudad Obregón, Sonora, México
\textsuperscript{3}Universidad de Sonora, Hermosillo, Sonora, México
\textsuperscript{4}Centre National de l’Energie, des Sciences et des Techniques Nucléaires, Kénitra, Morocco
\textsuperscript{5}LP2M2E, Département de physique, Faculté des Sciences et Techniques, Université Cadi Ayyad, Marrakech, Morocco
Inter-comparison of four remote sensing based surface energy balance methods

J. Chirouze et al.

Received: 18 December 2012 – Accepted: 6 January 2013 – Published: 21 January 2013
Correspondence to: J. Chirouze (jonas.chirouze@cesbio.cnes.fr)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Remotely sensed surface temperature can provide a good proxy for water stress level and is therefore particularly useful to estimate spatially distributed evapotranspiration. Instantaneous stress levels or instantaneous latent heat flux are deduced from the surface energy balance equation constrained by this equilibrium temperature. Pixel average surface temperature depends on two main factors: stress and vegetation fraction cover. Methods estimating stress vary according to the way they treat each factor. 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 triangle method, inspired by Moran et al., 1994) and two single-pixel (TSEB, SEBS) are applied at seasonal scale over a four by four km irrigated agricultural area in semi-arid northern Mexico. Their performances, both at local and spatial standpoints, are compared relatively to energy balance data acquired at seven locations within the area, as well as a more complex soil-vegetation-atmosphere transfer model forced with true irrigation and rainfall data. Stress levels are not always well retrieved by most models, but S-SEBI as well as TSEB, although slightly biased, show good performances. Drop in model performances is observed 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 extreme hydric and vegetation conditions are encountered in the same image (therefore, esp. in spring and early summer) while they tend to exaggerate the spread in water status in more homogeneous conditions (esp. in winter).
1 Context and objectives

Evaporation is the largest water loss component of continental surfaces. In semi-arid areas, more than 80% of the annual available water is lost through evapotranspiration. In most countries, the largest water user is the irrigated agriculture, which represents more than 80% of all uses, with a low efficiency no greater than 50% in many cases. 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. 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.

Many methods exist to 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 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 (Cleugh et al., 2007). But these methods are of little help when vegetation suffers from water stress, which means that these methods 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 Infra-red (TIR) domain (3–15 μm) is an appropriate way to assess actual evaporation and soil moisture status at relevant space and time scales (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; Kustas and Anderson, 2009).

Those methods evaluate the instantaneous evaporation rate at the time of the satellite overpass which can be converted into daily values through the use of methods based on the diurnal self-preservation of the evaporative fraction (Delogu et al., 2012).

Soil-Vegetation-Atmosphere-Transfer (SVAT) models, on the other hand, are able to simulate directly the surface temperature. Because this latter is related to water stress (Hain et al., 2009), one can constraint model prediction through the assimilation of the observed surface temperature into SVAT models (Coudert and Ottle, 2007; Olioso et al., 2005). But for most hydrological models with daily time steps, which do not simulate the equilibrium surface temperature, a remotely-sensed evapotranspiration product could be used as well in an assimilation scheme (McCabe et al., 2008; Schuurmans et al., 2003).

2 State of the art

Methods to compute evapotranspiration from TIR data 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 latters are usually inferred in other optical wavelengths so that for a given vegetation type/extent or a given value of the scaling surface parameter one can associate securely contrasted temperature patterns with contrasted hydric
conditions. 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 an energy budget for each pixel independently from the others. Usually, latent heat flux is derived as a residual of the energy balance:

$$\lambda E = R_n(T_0) - G(T_0) - H(T_0)$$

Where $R_n$ is the net radiation at the surface, $G$ is the soil conduction flux and $H$ the sensible heat flux, all expressed in Wm$^{-2}$ hereafter. $T_0$ is the radiometric surface temperature.

Single-pixel models are more sensitive to absolute errors in surface temperature estimates, however, it’s usually expected that they are well adapted to homogenous landscapes or the use of low resolution data, since for the latter pixels are mixed and cover many individual plots with contrasted levels of NDVI and soil moisture. They are therefore applied to produce global maps of evapotranspiration (Jimenez et al., 2011).

Contextual models, on the other hand, are less sensitive to absolute errors in surface temperature estimates but the assumptions that all soil moisture conditions are present within one image for a large enough range of vegetation fraction cover 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, esp. in irrigated agricultural areas with sufficient water supply.

Many studies have tested the performance of these models at various scales, from very high (Jacob et al., 2002; Gomez et al., 2005; Su et al., 2005; Timmermans et al., 2007; Minacapilli et al., 2009) to low spatial resolution (Jia et al., 2003; Verstraeten et al., 2005; McCabe and Wood, 2006; Su et al., 2007; Yang and Wang, 2011). But in most cases, those studies lack temporal representativeness of a surface evolution during a growing season (Choi et al., 2009; French et al., 2005; Gonzalez-Dugo et al., 2009).
Inter-comparison of four remote sensing based surface energy balance methods

J. Chirouze et al.

Depending on the resolution and type of remote sensing data (number of bands, atmospheric correction algorithm, satellite versus airborne . . . ) and the model studied, Root Mean Square Differences (RMSD) for instantaneous retrievals of turbulent fluxes range from 40 (Li et al., 2008; McCabe and Wood, 2006; Verstraeten et al., 2005; Jia et al., 2003; Gonzalez-Dugo et al., 2009) to more than 150 Wm$^{-2}$ (Choi et al., 2009; Li et al., 2005; Oku et al., 2007). Understanding of the reasons for such a wide range in performances is crucial in order to implement thermal data in more operational calculation chains.

Amongst well known models, there are two important ones in the “single-pixel models” category: TSEB (Norman et al., 1995) and SEBS (Su, 2002); amongst the contextual approaches, one can cite the popular S-SEBI (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, such as SEBAL (Bastiaanssen et al., 1998) or METRIC (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 different inter- and intra-seasonal 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 turbulent fluxes observations 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.

From an assimilation point of view, it is equally important to retrieve the water stress level or to estimate the absolute evapotranspiration flux value. Only stress levels are
related to root zone soil moisture. Even if most SVAT models are able to assimilate directly the surface temperature, it is often hard to specify model errors, observation errors, and, especially spatial model error covariance when assimilating surface temperature images into distributed SVAT models. If contextual models show robust and reliable performances, the TIR-derived evapotranspiration products could be well adapted to be assimilated directly 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. The performance will be assessed through data collected at 7 flux stations and compared to outputs of a SVAT model forced with in-situ vegetation, climate and irrigation input data.

3 Material and methods

3.1 Site and data acquisition

3.1.1 Site description

This study was conducted in the Yaqui valley (27.4° N, 109.9° W), in the state of Sonora, north-west of Mexico. With an area of 225 000 ha, bordered on the south-west by the Sea of Cortez and on the north-east by the Sierra Madre Mountains, 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, http://smn.cna.gob.mx/observatorios/historica/obregon.pdf), 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 and water is provided by the Alvaro Obregon Reservoir on the Yaqui river, which has a capacity of 3 km$^3$ (Chehbouni et al., 2008). 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 by 4 km, located at the 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 rest is divided between broccoli, beans, chili pepper, potatoes, chickpea, safflower, orange and corn. In the framework of the MedMex project (Fieuzal et al., 2011) seven micro-meteorological stations equipped with eddy covariance flux measurement system were installed in different fields. Their positioning is shown on Fig. 1.

### 3.1.2 Automated data acquisition

The meteorological data are taken at a height of 10 m from a weather station installed at center of the zone. It provides us measurements of wind speed and direction (R.M. Young anemometer), air temperature and moisture (Vaisala humidity and temperature sensor). When the data at the station are not available, we replaced it with a combination of the meteorological data also available at each EC station. Data are acquired with half hour time step from 27 December 2007 at 02:00 p.m. to 17 May 2008 at 10:30 a.m.

At each of the seven sites, the net radiation was acquired using CNR1 (Kipp & Zonen) and Q7.1 (REBS) 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 and soil moisture was acquired at 0.05 and 0.3 m depths using CS616 TDR (Time Domain Reflectometer, Campbell Scientific Inc., UT, USA). Those data were acquired at a frequency of 10 s then averaged and recorded each 30 min.
Latent and sensible heat flux were measured with KH20 fast response hygrometers (Campbell) and Campbell CSAT3 or RM Young 81000 3-D Sonic Anemometer at a frequency of 10 Hz and converted to 30 min flux average, respectively.

3.1.3 Discontinuous measurements

In addition to the data acquired by the stations, measurements of vegetation properties were made in-situ. Crop height and Leaf Area Index (LAI) were measured at various dates during the whole study. The LAI was estimated from destructive measurements as well as hemispherical photographs of the different fields. Gravimetric soil moisture profiles at each station were carried out each week. Those measurements allowed 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) at different times from December to May.

3.1.4 Remote sensing data

Seven ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer, http://asterweb.jpl.nasa.gov/) images were acquired in the thermal infrared from December 2007 to May 2008. The resolution of the land surface temperature (AST08: https://lpdaac.usgs.gov/products/aster_products_table/ast_08) product is 90 m and is atmospherically corrected.

FORMOSAT-2 is an earth observation satellite launched in 2004 by the National Space Organization of Taiwan (China). It provides high resolution (8 m) images of a particular area every day at 09:30 a.m. (solar 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 the 27 December 2007 to the 13 May 2008.
3.2 Data processing

3.2.1 Flux data quality

The turbulent fluxes from the EC stations were processed offline. A post-processing software (EC Pack) developed by the Meteorology and Air Quality section of the Wageningen University in the Netherlands (http://www.met.wau.nl/) in the frame work of the Joint Eddy covariance Project was applied. A detailed explanation of the correction procedure applied by this program is available in Van Dijk et al. (2004).

After rejection of incoherent data because of instability or malfunction of the instruments, the processed fluxes still presented problems in the estimation of $H$ and $\lambda E$. The closure of the energy balance is not achieved most of the time, with a residual between 24 and 38% of available energy, depending on the station. This error is in the range of what can be found in the literature (Twine et al., 2000; Wilson et al., 2002). Two methods of correction exist for this problem. The first one is to trust the estimation of $H$ by the sonic anemometer and to discard the measurements of $\lambda E$, replacing it by the residual of the energy balance $R_n - G - H$. The other one is based on the assumption that both $H$ and $\lambda E$ of fluxes are under-estimated but that the Bowen ratio ($H/\lambda E$) is correctly measured. The fluxes are then adjusted to close the energy balance. It is referred to as the “Bowen-ratio closure”. Although the most widely used method is the second one, when the error on the available energy is well known, both closures give similar performances (Twine et al., 2000). The “Bowen-ratio closure” method is used here, except for the wheat [2] station and the chickpea [4] station, for which the KH20 did not seem to work well during the first half of the season and for the whole season, respectively. We used the residual method to estimate the latent heat flux for those two stations. The safflower station had both problems with the turbulent and soil heat fluxes so we excluded its data from the study. We did not achieve to correctly integrate the drip irrigation data of the chili pepper station into ICARE, therefore this station’s data has not been used in the inter-comparison. The dates at which each station’s data were used are available in Table 2.
3.2.2 Land surface temperature

The distributed radiometric surface temperature ($T_0$) used in this paper is the AST08 product from the ASTER project. The seven images were downloaded from the Earth Observing System Date Gateway. The overpass time is around 11:00 a.m. local hour and the dates of the images are 30 December, 23 February, 10 March, 11 April, 27 April, 6 May and 13 May. The resolution of this product is 90 m and the scenes are around 60 km by 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 $4 \times 4$ km study zone were then re-sampled by bi-cubic interpolation at a resolution of 100 m for future distributed model comparison.

The error on ASTER $T_0$ product has been estimated around 1.5 K (Jacob et al., 2008; Sabol et al., 2009). We compared our ground data to ASTER values of temperature at the coordinates of the EC stations (Fig. 1). The absolute error on temperature is around 3.5 K and clearly exceeds the literature’s values but this can be explained by the difference between field instruments footprints (a few meters of diameter) versus image’s pixel size ($100 \times 100$ m) as well as the representativeness of the surface temperature for the surface heterogeneity. A mean bias around 0.9 K appeared in the estimation of $T_0$. Despite those results, the mean bias was different for each station (from $-0.2$ to $3.9$ K), thus no global correction to ASTER product has been applied.

3.2.3 FORMOSAT-2 data

The 26 FORMOSAT-2 images were registered using GPS ground control points and re-projected in the UTM WGS 1984 12 N coordinate system. Then an atmospheric correction was applied (Hagolle et al., 2008). Finally, the $4 \times 4$ km studied square was extracted and re-sampled at 100 m resolution.
Albedo

Albedo was computed as a linear combination of band 3 and 4 of FORMOSAT, according to Courault et al. (2008) (see Eq. 2).

\[ \alpha = 0.645\rho_{\text{red}} + 0.382\rho_{\text{NIR}} \]  

(2)

Vegetation indices

The Normalized Differential Vegetation Index (NDVI) is calculated from bands 3 (red) and 4 (near infra-red) of FORMOSAT-2.

A remote sensing LAI was computed from NDVI (see Eq. 3). The relationship between LAI and NDVI was calibrated using values of hemispherical LAI retrieved in all the studied fields during the growing season with a minimal Root Mean Square Deviation (RMSD) criterion. The calibrated extinction factor \( k \) proves to be equal to 1.13 and the asymptotical values of NDVI are: \( \text{NDVI}_\infty = 0.97 \) and \( \text{NDVI}_{\text{soil}} = 0.05 \). \( \text{NDVI}_\infty \) and \( \text{NDVI}_{\text{soil}} \) are the NDVI values for a fully developed canopy and a bare soil respectively. Those values are slightly different but not far from what was estimated in Duchemin et al. (2006) on a similar study of winter wheat in Morocco (\( \text{NDVI}_\infty = 0.93 \), \( \text{NDVI}_{\text{soil}} = 0.14 \) and \( k = 0.54 \)).

\[ \text{LAI} = -\frac{1}{k} \ln \left( \frac{\text{NDVI}_\infty - \text{NDVI}}{\text{NDVI}_\infty - \text{NDVI}_{\text{soil}}} \right) \]  

(3)

3.3 The surface energy balance for estimating evapotranspiration

The first three models compute evapotranspiration as the residual of the energy balance according to Eq. (1).

Models differ primarily in the partitioning of available energy \( R_n - G \) into turbulent fluxes \( H \) and \( \lambda E \), and secondly on the way they compute the available energy.
3.3.1 The available energy

Different ways of estimating the net radiation $R_n$ and the soil heat flux $G$ and their components from remote sensing data are found in the literature. However, in our comparison of the models, we focused more precisely on the way available energy is partitioned between latent heat flux $\lambda E$ 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 using the available energy $R_n - G$ in the computation of evapotranspiration. Starting from the same basis, it will be easier to observe the different behaviors of the models.

The general formulation of net radiation is:

$$R_n = (1 - \alpha) \cdot R_{sw} + \varepsilon \cdot R_{lw} - \varepsilon \cdot \sigma \cdot T_0^4$$  \hspace{1cm} (4)

$R_{sw}$ and $R_{lw}$ are respectively the shortwave and longwave incoming radiation, $T_0$ the surface temperature, $\alpha$ and $\varepsilon$ are respectively the albedo and the emissivity of the surface and $\sigma$ the Stefan-Boltzmann constant. The emissivity value of the surface is fixed at 0.98 for the whole scene and for each date. The albedo is retrieved from FORMOSAT data (see Eq. 2).

The choice in the formulation for the soil heat flux is much larger than for the net radiation considering the variety of propositions that can be found in the literature (Norman, et al., 1995; Roerink, et al., 2000; Santanello & Friedl, 2003; Su, 2002; Bastiaanssen, 2000). It has been shown that for a fixed time around midday, the flux $G$ is directly 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 which were using soil flux conduction as well as the Bastiaanssen (2000) formulation. The latter proved to be the more accurate (in terms of RMSD) in our case, so we chose to implement it for all methods. Its formulation is detailed in Eqs. (5) and (6).

$$G = \Gamma \cdot R_n$$  \hspace{1cm} (5)
\[ \Gamma = T_0 \times (0.0038 + 0.0074\alpha) \times \left(1 - 0.98\text{NDVI}^4\right) \quad (6) \]

### 3.3.2 Single pixel models

Both single-pixel models are compared to field measurements at a height of just a few meters above the crop height. We used the Monin-Obukhov similarity theory in the Atmospheric Boundary Layer (ABL). As proposed by Brutsaert (1999), we used bulk ABL similarity functions in order to describe the wind and temperature profiles in the turbulent environment (Eqs. 7–8, see Brutsaert, 1999; Su, 2002).

\[
u = \frac{u_*}{k} \left[ \ln \left( \frac{z - d_0}{z_{0m}} \right) - \Psi_m \left( \frac{z - d_0}{L} \right) + \Psi_m \frac{z_{0m}}{L} \right] \quad (7)\]

\[
\theta_0 - \theta_a = \frac{H}{\rho C_p} \times \frac{1}{ku_*} \left[ \ln \left( \frac{z - d_0}{z_{0h}} \right) - \Psi_h \left( \frac{z - d_0}{L} \right) + \Psi_h \frac{z_{0h}}{L} \right] = \frac{H}{\rho C_p} \times r_a \quad (8)\]

\(r_a\) is the atmospheric resistance to heat transfer at the surface-atmosphere interface, \(z\) is the reference measurement height above the surface, \(u\) is the weed velocity at the level \(z\), \(k = 0.4\) is the von Karman’s constant, \(d_0\) is the displacement height \((d_0 \approx h_c \times 2/3, h_c\) being the crop height), \(z_{0m}\) is the roughness height for momentum transfer. \(\theta_a\) and \(\theta_0\) are respectively the potential temperature of the air at the height \(z\) and at the aerodynamic level, \(\rho\) is the density of the air, \(C_p\) is the heat capacity of air and \(z_{0h}\) is the roughness height for heat transfer. \(\Psi_m\) and \(\Psi_h\) are the stability correction functions for momentum and sensible heat transfer and \(L\) is the Monin-Obukhov length (see Eq. 9).

\[
L = -\frac{\rho u_*^3}{kg \left( \frac{H}{C_p \theta_0} + 0.61 \frac{\lambda E}{\lambda v \rho_w} \right)} \quad (9)
\]

In Eq. (9), \(g\) is the acceleration due to gravity.
If one can assume that Eq. (9) provides a fairly robust estimate of the atmospheric resistance to heat transfer \( r_a \), obtaining remote measurements of the aerodynamic temperature is a difficult issue. However, satellite imagery of the surface temperature is available. The difference between the two single-pixel models will thus lie in the approximation made in order to relate \( \theta_0 \) to the radiometric surface temperature \( T_0 \).

The Surface Energy Balance System (SEBS) model

SEBS (Su, 2000) 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 roughness height for the difference between the roughness for heat transfer \( z_{0h} \) (Eqs. 10–11) and momentum \( z_{0m} \) which accounts for the difference between \( \theta_0 \) and \( T_0 \):

\[
z_{0h} = \frac{z_{0m}}{\exp \left( kB^{-1} \right)}
\]

(10)

Where:

\[
kB^{-1} = A_1 \times f_c^2 + A_2 \times f_c f_s + kB_s^{-1} \times f_s^2
\]

(11)

The first term \( A_1 \) describes the full canopy aerodynamic properties (Choudhury and Monteith, 1988), \( kB_s^{-1} \) 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 is in the retrieval of the latent heat flux. It ensures that the retrieved latent heat flux is truly bounded by two extreme conditions (null and potential evapotranspiration rates resp.) and computes the relative evaporative fraction \( \Lambda_r \) according to:
$H_{dry} = R_n - G$  

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

$\Lambda_r = \frac{\lambda E}{\lambda E_{wet}} = 1 - \frac{H - H_{wet}}{H_{dry} - H_{wet}}$  

$H_{dry}$, $H_{wet}$ and $\lambda E_{wet}$ are the turbulent fluxes in extreme dry and wet conditions. $e$ and $e_s$ are respectively the actual and saturation vapor pressure, $\gamma$ is the psychrometric constant, $\Delta$ is the slope of saturation vapor pressure curve at temperature $T_a$ and $r_{ew}$ is the atmospheric resistance for wet conditions.

The components of the wet budget are obtained with Eq. (13) and by combining Eqs. (12), (13) and (14), $\lambda E$ is obtained.

**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. 15–16). Net radiation is computed according to Eq. (4). It is then partitioned according to 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) resp. Both energy balance equations read:

$R_{n,c} = H_c + \lambda E_c$  

$R_{n,s} = H_s + \lambda E_s + G$  

$H_s$ and $\lambda E_s$ are the sensible and latent heat flux at the soil/atmosphere interface and $H_c$ and $\lambda E_c$ the fluxes at the canopy/atmosphere interface (partitioning of the energy budget is summarized in Fig. 2).
The two main unknowns of these equations are the soil and vegetation surface temperatures. The “trick” to get two unknowns out of one single information (mixed pixel temperature) is to assume that in many cases the vegetation is unstressed and transpires at a potential rate $\lambda E_c$ which is obtained with a Priestley-Taylor formulation (see Eq. 17).

$$\lambda E_c = 1.3f_g \frac{\Delta}{\Delta + \gamma} R_{n,c}$$

(17)

Where $f_g$ is the green fraction of LAI. The LAI being retrieved from the NDVI, it is considered as green LAI and by consequence, $f_g$ is set to 1 in our case. This first computation of $\lambda E_c$ gives us a first guess of $H_c$, as a residual of the energy budget at the canopy. As shown in Eq. (18) (which describes the formulation of the global sensible heat flux), the canopy temperature $T_c$ can be deduced from $H_c$.

$$H = H_s + H_c = \rho C_p \frac{T_s - T_a}{r_a + r_s} + \rho C_p \frac{T_c - T_a}{r_a}$$

(18)

$T_s$ and $T_c$ being the temperatures of soil and canopy (resp.) and the resistance $r_s$ is added to take into consideration the resistance to heat transfer in the boundary layer immediately above the soil surface.

Equation (19) (Norman et al., 1995) links canopy and soil temperatures to the observed radiometric temperature $T_0$ and the fraction cover $f_c$ and allows to calculate $T_s$ (Eq. 19):

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

(19)

Combination of Eqs. (16) and (18) gives us $H_s$ and $\lambda E_s$. If $\lambda E_s$ is positive then a balance is reached. If $\lambda E_s < 0$, it means that the assumption that the vegetation transpires at a potential rate is no longer valid, thus the soil is considered as dry and $\lambda E_s$ set to zero and the other parameters are computed from Eqs. (16), (18) and (19). $\lambda E_c$ is obtained as a residual of Eq. (15). If $\lambda E_c$ is negative then soil and vegetation are dry and it is set to zero. Then $\lambda E = 0$ and $H = R_n - G$. 912
3.3.3 Contextual models

The Simplified Surface Energy Balance Index (S-SEBI, Roerink et al., 2000)

The model is based on the observation that for homogeneous atmospheric conditions over a scene, surface temperature and reflectance are strongly correlated and that for pixels in extreme hydric conditions, a linear regression can be established between the two parameters. For each albedo/$T_0$ image we infer manually the relation between extreme temperatures and albedo (see Fig. 3). The method then considers for each pixel two bounding extreme hydric conditions (with the same available energy) with the respective surface temperature $T_{\text{hot}}$ and $T_{\text{cold}}$. Those conditions correspond respectively to a dry surface ($\lambda E_{\text{min}} = 0, H = H_{\text{max}}$) and a saturated one where all the available energy is used to evaporate the water ($\lambda E = \lambda E_{\text{max}}, H_{\text{min}} = 0$). The evaporative fraction is then computed using those extreme temperatures according to:

$$\Lambda = \frac{\lambda E}{R_n - G - H} = \frac{R_n - G}{R_n - G} = \frac{H_{\text{max}} - H}{H_{\text{max}}}$$ (20)

From Eq. (8), we can deduce that in wet conditions ($H = 0$), $T_{\text{cold}} = T_a$, with $T_a$ the air temperature. So if we replace $T_a$ by $T_{\text{cold}}$, we can deduce from Eqs. (8) and (20) the following formulation of the evaporative fraction:

$$\Lambda = \frac{T_{\text{hot}} - T_0}{T_{\text{hot}} - T_{\text{cold}}}$$ (21)

The turbulent fluxes $H$ and $\lambda E$ are then deduced from the evaporative fraction and the available energy (Eq. 21).

A vegetation index-temperature trapezoid method (VIT)

Similarly to S-SEBI, Moran et al. (1994) proposed a method for retrieving $\lambda E$ on large scale areas combining relations between the temperature difference $T_0 - T_a$ and vegetation extent. The principle of the method is to use the Penman-Monteith equation to
retrieve \((T_s - T_a)\) in four extreme conditions of temperature and vegetation cover (full watered canopy, full stressed canopy, watered bare soil and dry bare soil). Four vertices of a trapezoid in the \(T_s/VI\) (VI for Vegetation Index) space are obtained this way. 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) SAVI (Soil-Adjusted Vegetation Index). This allows straight lines to be drawn between the vertexes 1 and 3 and the vertexes 2 and 4. The third assumption that permits to link this graphic representation to hydric state of the surface is that, for a given \(R_n\), vapor pressure deficit of the air and atmospheric resistance, \(T_s - T_a\) and \(T_c - T_a\) are linearly dependent on evaporation and transpiration, \(T_s\) and \(T_c\) being the temperature of soil and canopy. The land surface temperature \(T_0\) being directly linked to \(T_s\) and \(T_c\), the linear relation between \(T_0\) and evapotranspiration is established as follows:

\[
WDI = 1 - \frac{\lambda E}{\lambda E_p} = \frac{\Delta T - \Delta T_{\text{min}}}{\Delta T_{\text{max}} - \Delta T_{\text{min}}} \quad (22)
\]

WDI is the Water Deficit Index \(\Delta T\), \(\Delta T_{\text{min}}\) and \(\Delta T_{\text{max}}\) are respectively the difference of surface-air temperatures 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 (see Fig. 4). It can be assimilated to water stress.

In our case, a preliminary study had shown that the vertexes calculated with Moran et al. (1994) approach were not usable. Actually, the trapezoid calculated this way did not include all of the situations present in the images. Thus pixels with a negative or greater than 1 water stress were present, which is inconsistent with theory. We then chose to make an assumption similar to the one made in the S-SEBI method, which is that the variety of hydric conditions present in the image is sufficient to assess, if not the four extreme points of balance, at the very least the two extreme tendencies (lines 1–3 and 2–4) corresponding to well watered cold points and stressed hot ones. Other differences with the Moran et al. (1994) approach are that we use NDVI instead of SAVI (which is also linearly related to \(T_0 - T_a\)) and that, since we have a flat study zone, we
consider that the air temperature at 10 m height (our working height) is representative of the whole area. We can then use NDVI/$T_0$ scatter plots directly.

**Method to retrieve water status extremes in contextual models**

Various methods to retrieve linear relationships corresponding to bounding relationships between surface temperature and albedo for S-SEBI or NDVI for Ts-VI have been tested. 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.). This expertise allows to determine eventually whether some hydric conditions are missing on each image and determine more consistent albedo (resp. NDVI) – $T_0$ relationships. However in practical cases, detailed knowledge about the surface characteristics and human activities is not easily obtainable, all the more over a wide area like a watershed, so it can be quite difficult to interpret the scatter plots. Furthermore, 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. A classification of albedo values is done (in our case, ten classes of albedo values for each image). For each class, the median value of 5 % unique maximum and the median of 5 % unique minimum surface temperature values are identified, then the Least Square Method is used on those median values to retrieve each linear relation.

### 3.3.4 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 which to relate the SEB performances. ICARE is a classic dual-source SVAT model that solves the water 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 three points. First, the water
balance module, a two layers force restore model in our case, simulates the evolution of soil moisture and temperature for each soil layer (shallow and root zone). Thus, as a dynamic model, it is given initial conditions in surface and root zone temperature and moisture levels; therefore the surface temperature is not an input but an output. 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 albedo but computes its own broadband albedo, given fixed values of soil and vegetation albedos, with a multi-reflections model. It is likely to introduce differences in the computation of the net radiation. Finally, the soil heat flux $G$ is no more calculated as a fraction of the net radiation but with a temperature diffusion law between the soil layers, from a given soil thermal conductivity. In addition to the precedent point, the estimation of the available energy should differ significantly from the SEB models. A more technical explanation of the model is available in Gentine et al. (2007).

As a complex physical model, ICARE use 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 they are available, except for the soil resistance to evaporation $r_{ss}$, because soil texture and composition are almost uniform over the whole area. This choice was 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. Formulation of $r_{ss}$ is given in Eq. (23) 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 the two empirical constants that have been calibrated.

$$r_{ss} = \exp\left(A_{rss} - B_{rss} \frac{w_s}{w_{sat}}\right)$$  (23)
The calibration of $A_{rss}$ and $B_{rss}$ has been done simultaneously on a single wheat station when the surface was almost only bare soil using a multi-criteria approach. Conditions of minimum error were set for two model observations: the latent heat flux and the surface (5 cm) soil moisture. Results of this calibration gave a minimum mean error for values of 12 and 19 for $A_{rss}$ and $B_{rss}$ (resp.). The minimum stomatal resistance, which is a very sensitive parameter for the estimation of latent heat flux, has been set to 100 s.m$^{-1}$ (Martin et al., 1999; Gentine et al., 2007).

The SVAT model has been ran at each station at a half hour time step for the whole season, except for the chili pepper station which was irrigated by a drip system and covered by a white cap half of the season, which caused us problems in ICARE runs.

4 Results

The statistical results for the four models are summarized in Tables 3 to 6.

4.1 Albedo

Measurements of albedo were made on three sites ([1], [3] and [4], see Table 2 for dates) by CNR1 sensors. Comparison of FORMOSAT against in-situ albedo values shows a mean bias around 4.4 % and a relative RMSD of 10 %. Those results are quite satisfactory with respect to previous studies (Bsaibes et al., 2009; Courault et al., 2008; Jacob and Olioso, 2005) which show relative RMSDs between 3 and 15 %.

4.2 Surface temperature

Extraction of ASTER and ICARE surface temperature values at each station’s pixel are shown in Fig. 5. Dates and stations at which surface temperature is available are displayed in Table 2. ASTER temperature gives RMSD of 3.5 K and a mean bias of 0.9 K. 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 over-estimate the surface temperature. This error can be attributed to the representative area (footprint) of ASTER pixels that is bigger than the CNR1 footprint: ASTER should “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 temperature and the canopy and soil temperatures, which are all computed by resolving the energy budget at each source (soil and canopy). It is thus very dependent on the errors induced by the model.

4.3 Net radiation

Scatter plots of SEB models and ICARE versus in-situ measurements are displayed in Fig. 6. The estimation of the available energy by the three studied models gives results often encountered in the literature with RMSD around 44 Wm$^{-2}$ (< 10 % of the mean value) for the computed $R_n$ and 56 Wm$^{-2}$ for $G$. A notable positive bias (~24 Wm$^{-2}$, 4 % of the mean value) appears on the net radiation. A comparable bias is present on the values of FORMOSAT albedos (~4 %) and ASTER surface temperature (4.5 %), but the over-estimation of those parameters should lead to an under-estimation of the $R_n$. Therefore we can assume that this bias is not due to errors on surface temperature or albedo retrieval.

Compared to the performances of ICARE, SEB models are better in terms of statistics. The ICARE net radiation flux shows a coefficient of variation of the RMSD, defined as the RMSD normalized to the mean of observed values and noted CV(RMSD), of 11 % (in percentage of the mean value) which, while still being a reasonable error, can be explained by the differences between the surface temperature and albedo used by the models. Indeed, in the SVAT model, albedo is computed using soil and dry and green vegetation albedo set by the user (respectively set at 0.15, 0.19 and 0.22 in our case). 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,
5~55% for beans). On the contrary, the formula used to calculate the broadband albedo from FORMOSAT reflectances has been calibrated on a large area with a substantial variety of crops (Courault et al., 2008). It gets CV(RMSD) lower than 10% for wheat and beans and 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 favored physical meaning over model performances in our parameterization of ICARE.

4.4 Soil heat flux

The results of the calculation of the soil heat flux for each model has been displayed in Fig. 7. For SEBS and TSEB, an effect of saturation is clearly noticeable for bare soils and quasi bare soils (red dots). The over-estimated red dots (group 1 in Fig. 8) correspond to the end of the senescence period and the harvesting time of wheat. At that time and with those LAIs, we would expect higher values of $G$ but the LAI we use, being computed directly from the NDVI, is a green LAI. The expression of the factor $R_n/G$ uses NDVI too and does not take well into account the dry part of the vegetation. That counts for the senescence period when the plants are still high but dry and after the harvesting because farmers left the straw on the ground, which would lower $G$ values but still give low NDVI and high surface temperature. A series of green and yellow dots (intermediate and low LAIs, group 2 in Fig. 8) corresponding to the same station (beans) are constantly over-estimated with a bias around 50 Wm$^{-2}$. No particular over-estimation of net radiation or surface temperature was observed. We can assume then that this bias comes pretty much from the measurements at this station. However, for high values of measured $G$, the models under-estimate greatly the flux. The third group of outliers (Fig. 8) 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 ($\sim 100$ Wm$^{-2}$ at each date), it sums up with an under-estimation of the $\Gamma = G/R_n$ factor by the models. At the station, $\Gamma$ values estimated from observations ranged between 0.3 and 0.45 whereas the model computes values between 0.2 and 0.3. Although the in-situ values seem high for the studied zone (others stations gives
around 0.3 for bare soil), the maximum of this factor can be encountered as 0.35 in the literature (Monteith, 1973; Kustas and Daughtry, 1989; Norman et al., 1995) and in our case, it seems that the formulation of $\Gamma$ cannot reach those values, moreover rarely reach 0.3.

Greater scattering is observed in ICARE with bigger RMSD (96 Wm$^{-2}$ against 56 Wm$^{-2}$ for the other models). Those errors are mainly due to a poor estimation of the surface and deep temperatures by ICARE, which depend on the resolution of the energy balance for the surface and on the soil moisture for both.

4.5 Turbulent fluxes

Scatter plots of simulated versus measured turbulent fluxes are presented in Fig. 9 and Fig. 10 for $\lambda E$ and $H$, respectively. TSEB has a systematic tendency to over-estimate $\lambda E$ (and thus under-estimate $H$) with a strong bias of 99 Wm$^{-2}$. On Fig. 9, strong under-estimation of $H$ at low LAIs can be observed. Underestimations of $H$ for higher LAIs are mainly due to errors in the ASTER temperature. At high LAIs, we observe in most case an over-estimation of $\lambda E$ (see Fig. 10) certainly related to an over-estimation of the canopy transpiration (Eq. 17), as TSEB assumes that the vegetation always transpires at potential rate.

In terms of absolute error, SEBS shows similar performances (Tables 3 and 4). There is a great underestimation of $H$ for quasi-bare soils. The kB$^{-1}$ appears to be too big for low LAIs and this leads to an overestimation of the atmospheric resistance. The few overestimations of $H$ in moderate LAIs come from an overestimation of the temperature.

S-SEBI showed the best performance (among the SEB models) in the computation of both $H\lambda E$ with RMSD of 106 Wm$^{-2}$ and 120 Wm$^{-2}$, respectively. No bias was observed on $H$ and a positive mean bias of 40 Wm$^{-2}$ is obtained on latent heat flux. Error on the $H$ comes essentially from the error of the ASTER surface temperature. A large underestimation of $H$ occurs in period of senescence for the wheat. The vegetation is stressed but still dense and surface temperature and albedo are not as high as for
bare soil pixels. Consequently, those points are far from extreme temperatures and the evaporative fraction is overestimated.

ICARE performs better than the other three models in the computation of both $H$ and $\lambda E$ with RMSD of respectively 98 Wm$^{-2}$ and 116 Wm$^{-2}$ despite being less accurate in the calculation of the available energy. Contrary to SEB models, it tends to underestimate $\lambda E$ and overestimate $H$ with biases of respectively $-54$ Wm$^{-2}$ and $+34$ Wm$^{-2}$.

### 4.6 Water stress

Since one of the main purposes of estimating evapotranspiration is to get information about the water status of the plant, we computed the water stress (as defined in Eq. 24) at each measurement point for each date where ASTER data was available. Results are shown in Fig. 11.

\[
\text{water stress} = 1 - \frac{\lambda E}{\lambda E_{\text{max}}}
\]  

In Eq. (24), $\lambda E_{\text{max}}$ is the maximum (potential) latent heat flux achievable for a $p$. Indeed, 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). If we were to use a common potential evapotranspiration to compute water stress for each model, we would not get much more information than with the latent heat flux results. 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 seemed to be the most relevant amongst common models (Penman-Monteith, Priestley-Taylor), when compared to $\lambda E$ measurements at each EC station at irrigation dates (not shown). Since potential evapotranspiration and stress were not outputs of ICARE, we chose to run it with continuous irrigation in order to obtain its own $\lambda E_{\text{pot}}$.

The principal observation that can be made about those results is that although ICARE would be expected to show good results in terms of water stress, 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. A significant overestimation of water stress is observable at medium and high LAIs. It seems to be due to the water balance module (force restore) which tends to quickly dry the upper soil layer with bursts of $\lambda E_s$ on the days immediately following irrigation.

TSEB globally underestimates stress, which is in agreement with the general overestimation of $\lambda E$ seen in Fig. 10, but it has a lower scattering and number of outliers than the other three SEB models. The large underestimation of stress that appears for intermediate observed stress levels (0.5) seems to be due to difficulties of the model to take into account the senescence phenomenon. It will be discussed with more details in Sect. 5.2. However, 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 atmospheric resistance for bare soil discussed in Sect. 4.5. A group of overestimated stress points when it is expected to be near zero is a consequence of an overestimation of the temperature by the ASTER sensor and very turbulent atmospheric condition (low atmospheric resistance).

The contextual models, as expected, dispatch all stress levels between extremes and in turn produce a large spread of stress levels. Around medium observed stress values (0.5), S-SEBI simulates very low (to null) stress values. This is directly due to surface temperatures that are not high enough for the $T(\alpha)$ relation to detect stress. The relative 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 points which correspond to intermediate LAIs (0.4–0.8) present a greatly overestimated stress. This is due for the upper one (chickpea, on day of year 54) to an overestimation of the ASTER temperature of 7 K (one must keep in mind that the ASTER pixel is bigger than the EC instrument footprint and includes more bare soil with higher temperature). The lower one (beans, day of year 108) seems to be due to a high surface temperature of the whole bean field (around 313 K), which is consistent with the station measurement,
while the crop is still in growing period and well irrigated. Figure 11 shows an example of how contextual methods are dependent on the studied area surface properties. In the case of the VIT method, it seems that there is a lack of pixels with stressed and well developed vegetation (i.e. pixels with high NDVI and high surface temperature), which seems consistent in our area because, the whole zone being irrigated, stress shouldn’t appear in the growing period. It results in the simulation of medium to high stress for various pixels with high NDVI, which are located near the top-right edge of the trapezoid (see Fig. 4) but still represent non-stressed crop. On the other hand, the stress is underestimated for some quasi-bare soil pixels. It corresponds to pixels with low NDVI but that present a lower temperature than the maximum observed. If some straw is left on the ground after harvesting or if the bare-soil properties are different (e.g. higher albedo), the surface is fully stressed but, the temperature being lower, the pixel is not located close enough to the bottom-right corner of the trapezoid to detect strong stress.

4.7 Spatial variability

In this section we move from the local to the spatial standpoint. Local inter-comparison of the models allows us to assess their performances and to show some specific behavior but not to assess how they represent the spatial variability at the perimeter scale. In Fig. 12, we plotted frequency histograms for turbulent flux and remote sensing data (albedo, NDVI, surface temperature) on 10 March for the whole area. At this time of the year, most of the crops are well developed and green. On the histograms, TSEB and SEBS have a similar response: low $H$ and high $\lambda E$ peaking around the potential rate. On the other hand, 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 we have seen before, TSEB has a stronger tendency to overestimate the latent heat flux for green landscape because of the hypothesis of the vegetation always transpiring at potential rate given by PT formulation. This result is observed as well in Table 7, where arithmetical means of turbulent flux and water stress over the whole area are
displayed. TSEB $\lambda E_{\text{mean}}$ is higher than the other two (425 W m$^{-2}$), the smallest being S-SEBI with a $\lambda E_{\text{mean}}$ of 354 W m$^{-2}$. The results in terms of stress are shown in Fig. 13. The two contextual models compute a lot more stressed areas than the other models, which was expected since they distribute the stress values on the whole [0,1] interval. They both have mean water stress values around 0.35 whereas single-pixel models computed low mean stress value (0.11 for SEBS, 0.13 for TSEB). SEBS shows a singular behavior displaying heavily stressed areas but with the lowest mean stress value of all models (see Table 7).

On 6 May, both well watered and stressed vegetation are present in the area. As expected, in heterogeneous conditions, contextual models have a good response to variations of the surface water status (see Fig. 14). TSEB, S-SEBI and VIT showed very similar results in term of displayed 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 7) and a very different distribution of stress values over the area.

5 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 studied area is crucial.

5.1 Performances inter-comparison

Models have shown fairly high errors in the computation of turbulent fluxes (RMSD over 100 W m$^{-2}$ for the SEB methods), but it is still in the range of what has already been published (see Table 1). Most components of the energy budget are often large in semi-arid lands at low latitudes; 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 couldn’t beneﬁt of the experience in data understanding and correction that long-term projects like FLUXNET (http://fluxnet.ornl.gov/) can provide. All the more so as it involved a large acquisition setup which turned out very difﬁcult to monitor considering the available technical human resources and the shared knowledge of the in-situ instrumentation coming from each partner. Adding to that aspect the absence of real calibration and adjustment of the models to the experimental conditions, such results could have been expected.

In addition to general performances of the models over the whole season, we looked a bit deeper in their performances over diferent crop types and at diferent dates. ICARE performed better than the other methods over wheat and chickpea ﬁelds, showing RMSD on turbulent ﬂuxes lower from 20 to 70 W m$^{-2}$ to 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. Adding to that a fairly strong bias is present on the net radiation calculated by the three SEB methods, which tends to worsen their performances. However, over potatoes, sorghum, broccoli and beans, the three SEB models performs better and particularly over sorghum, where the three models present RMSD around 70–90 W m$^{-2}$. Despite his global performances that are higher than the energy balance models, it seems than SEB models can give better results, depending on the type of observed vegetation. 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 versatile than the more complex SVAT model that internally calculates those parameters from user-deﬁned properties of the soil and vegetation. Furthermore, although the net radiation calculation method in ICARE, a multiple reﬂection method, seems to be physically and well adapted to homogeneous crop types like wheat, it does not seems to be really
adapted to crops with a row geometry like chili pepper, broccoli, instrumented in this experiment.

Another observation made during this experiment is that the error made by ICARE in surface temperature computation is not necessarily representative of the error made on the turbulent fluxes. Indeed, the situation where ICARE is not accurate in surface temperature but has very acceptable error on $H$ and $\lambda E$ is not uncommon and it questions the relevance of assimilation of thermal data alone in SVAT models in order to improve its performances in turbulent fluxes determination. In those cases, correcting the model estimations based on thermal information may not be sufficient to correctly force ICARE into the right repartition of energy between the components of the energy balance.

5.2 Model’s structure and improvements

One big issue of the SEB models in their original form 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 enhance heat exchange at the surface, whereas the plant is likely to be fully stressed. Since senescent vegetation pixels do not appear in extreme temperature conditions, those models do not detect full stress.

SEBS tends to overestimates evapotranspiration for low LAIs and it seems to be mostly due to the bare soil component of the $k_B^{-1}$ factor that generates very high atmospheric resistances and thus very low $H$. Solutions to compensate this problem have been proposed by using another empirical factor in order to compute turbulent fluxes (Boulet et al., 2012) or by including information about soil moisture in the $k_B^{-1}$ (Gokmen et al., 2012). Implementation of the Boulet et al. (2012) formulation gives slightly better performances than the original Su (2002) with RMSD of 130 Wm$^{-2}$ (against 138 Wm$^{-2}$ for the original version) and a lower bias of $-27$ Wm$^{-2}$ (against $+70$ Wm$^{-2}$ for Su, 2002).
A major issue was encountered in TSEB during the senescence period in the partitioning of net radiation between soil and vegetation. Indeed, the LAI used in this study is computed from NDVI and thus is a green LAI. However, during senescence, a global drying of vegetation occurs and part of the canopy becomes yellow when drying. For crops like wheat it results in a fully yellow vegetation 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 and it results in a great overestimation of the soil component of net radiation. Since the soil heat flux is limited to a maximum fraction of $R_{n,s}$ 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 $R_{n,s}$. It translates in a large overestimation of $\lambda E_s$, and by consequence of $\lambda E$. 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 $\text{LAI}_{\text{max}}$. We then calculated the green fraction $f_g$ according to Norman et al. (1995) with the new definition of global LAI:

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

(25)

This approximation is not very accurate since during the drying of vegetation, the leave’s surface decreases and thus the global LAI should decrease smoothly over the senescence period. However it is sufficient to greatly influence the repartition of net radiation and thus the estimation of $\lambda E$ and water stress. Results on $\lambda E$ and net radiation with this new formulation of LAI are compared to the initial version of the model in Fig. 15. Big differences between the two formulations are observable on soil and canopy net radiation, with variations of fluxes magnitude up to 500 Wm$^{-2}$. This influences greatly the energy balance at the soil-atmosphere interface with smaller $R_{n,s}$. As a result, RMSD on $\lambda E$ is reduced by 30 Wm$^{-2}$ and the Bias is lowered to 33 Wm$^{-2}$ (against 99 Wm$^{-2}$ 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 more point to add is that since we use global LAI in the computation of atmospheric 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 and to lower $\lambda E_s$. Thus, integrating information about dry vegetation could greatly improve the performances of TSEB in senescent cases. In our case, many cloud-free FORMOSAT-2 images were available, allowing us to determine each phenological stage of the vegetation and thus calibrate the maximum LAI. 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 behavior of TSEB and ICARE is also interesting since they are both dual-source models. Their ways of partitioning radiative energy are different. TSEB calculates a global net radiation and then distributes it between soil and vegetation based on fraction cover whereas ICARE computes the two net radiations from LAI and soil and vegetation albedo and emissivity, using a multi-reflection and transmission network, as proposed by Shuttleworth and Wallace (1985). The TSEB method for net radiation can be criticized because it bypasses the effects of the vegetation’s transmissivity and long wave 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 performances than a MODIS-data based $R_n$ (RMSD = 24 Wm$^{-2}$ against 44 Wm$^{-2}$ on a MODIS pixel). However, this method requires estimates of vegetation and soil albedos and vegetation transmission factor, which can be quite difficult if studying a very heterogeneous zone. The method used in ICARE is the most physically based since it considers all radiation exchanges between two layers (vegetation and bare soil), including transmission and multiple reflections, but it assumes that the vegetation cover is homogeneous and thus is not adapted to row crops. As long as FORMOSAT 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 particularities. 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 between soil and vegetation latent heat flux, the two methods have also very different behaviors. Underestimation of $\lambda E$ by ICARE during the growing period is due to the quick drying of the first layer of soil, resulting in very low $\lambda E_s$. However, since ICARE takes into account the dry part of the vegetation, it performs better during senescence; with very low $\lambda E_c$ whereas TSEB still computes it as potential. As a consequence, TSEB computes canopy and soil temperatures as almost equal to respectively air and ASTER surface temperature. On the contrary, ICARE distributes both temperatures around the computed surface temperature, which seems to be closer to the reality. Although the partition of radiation seemed better with TSEB (subject to some changes about dry LAI), the soil/vegetation partition between turbulent fluxes and temperature seems to have more physical meaning in ICARE, mostly because of strong hypothesis on vegetation in TSEB. TSEB seems to be the most accurate of the SEB models after reconstitution of total fluxes, but his soil/canopy distribution of latent heat and estimates of component temperatures seems questionable.

### 5.3 Determination of distributed water stress

In order to carry out a spatial inter-comparison of the models, we chose to compute statistical operators allowing us to assess their relative behavior. In Table 8 and Table 9 are displayed coefficients of determination $R^2$ and RMSD of the simulated water stress by the different models, as well as their mean value and their standard deviation $\sigma$ over the whole area. As expected for contextual models, S-SEBI and VIT show higher $\sigma$ and mean 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. However, contextual models distribute stress over the whole $[0,1]$ interval since they suppose that extreme conditions are present on each image. For a later date in spring
(on 6 May, when wheat is mostly senescent), TSEB and the contextual models are in good agreement with each other, showing equivalent σ and mean values, and with low RMSDs (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 pattern showed in Figs. 13 and 14 as well as the mean value and standard deviation, are very similar. It shows than in the case of contrasted images, simpler contextual models reproduce quite faithfully the general behavior of a more complex like TSEB, which is already used in operational algorithms at continental scales and thus whose performances are trusted. On the contrary, SEBS seems to behave very differently from the other three models during the whole season in term of stress distribution with a low standard deviation in winter as well as in spring. It has very low correlation with the other three models and the patterns and histograms displayed in Figs. 13 and 14 are very distinct from the others. One possible explanation could be that SEBS is more sensitive to vegetation properties due to his use of the $k_B^{-1}$ factor in the determination of aerodynamic resistances. This parameter is very sensible 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 to by-pass some crop height determination issues.

6 Summary and conclusions

Performances 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 $\lambda E$ ranging from $119 \, \text{Wm}^{-2}$ (S-SEBI) to $138 \, \text{Wm}^{-2}$ (TSEB). Those results are in the same range as the ICARE SVAT model ($\text{RMSD} = 116 \, \text{Wm}^{-2}$ on $\lambda E$) 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 which estimate with the best accuracy the water stress. However, TSEB performs better at high LAIs (low stress), having difficulties to detect stress in senescent period, 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 points and contextual models present a lot of dispersion. Corrections for TSEB and SEBS have been proposed in order to account for their respective troubles in processing dry vegetation and low LAIs cases.

From the spatial point of view, general behaviors of the models have been described. SEBS distinguishes itself from the others in its way to incorporate and process vegetation data, resulting in a singular distribution of stress. On the other hand, for well contrasted dates in term of surface moisture, the contextual models have shown stress patterns, mean value and distribution very similar to TSEB, whereas at dates when the whole area is well watered (less surface moisture contrast), they tends to accentuate extremes values of water stress. This was expected since they are self-calibrated and thus distribute the values of the whole scene between totally stressed surface and potential conditions.

Different methods of determination of the empirical laws driving contextual models have been tested and were very similar in terms of statistics or global behavior but the manual method seems to be the most relevant to correctly process particular cases since it allows qualitative interpretation of each image.

This work has been carried out as a preliminary study of the SEB models in order to assess if thermal remote sensing data could bring valuable information to assimilate
into a land surface or hydrological model. Results have shown that the single-pixel SEB models, after various modifications, could 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 and Cammarelli and Ciraolo, 2012). 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 sub-surface characteristics of wide areas. Assimilation of thermal distributed data could be a way to calibrate those models with a limited amount of information about the surface.

Acknowledgements. Funding from the French space agency (Centre National d’Etudes Spatiales, CNES) for the MiSTIGRI (MicroSatellite for Thermal Infrared Ground surface Imaging) phase A study, as well as the FORMOSAT acquisition, the MISTRALS (Mediterranean Integrated STudies at Regional And Local Scales) SICMed (Continental Surfaces and Interfaces in the Mediterranean area) program, the European FP7 SIRIUS, the PLEIADES (Participatory multi-Level EO-assisted tools for Irrigation water management and Agricultural Decision-Support) program as well as the IRD (Institut de Recherche pour le Développement), the ITSON (Instituto Tecnológico de SONora), the University of Sonora and the Cadi Ayyad University (Morocco) in the setting up of the “Yaqui experiment” are gratefully acknowledged.

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


Table 1. Non-exhaustive review of validation exercises of instantaneous models. 1 site covers less than \(30 \times 30\) low-resolution pixels, 1 station covers less than 10 high-resolution pixels.

<table>
<thead>
<tr>
<th>Publication</th>
<th>Model(s)</th>
<th>Resolution</th>
<th>Precision on (LE) (Wm(^{-2}))</th>
<th>Spatial and temporal coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson et al. (2011)</td>
<td>TSEB</td>
<td>10 m to 10 km (airborne)</td>
<td>–</td>
<td>1 station on 1 site, 1 flight</td>
</tr>
<tr>
<td>Choi et al. (2009)</td>
<td>TSEB, METRIC,</td>
<td>60 m (LANDSAT 7)</td>
<td>RMSD 50–150</td>
<td>14 stations on 1 site, 2 dates</td>
</tr>
<tr>
<td>French et al. (2005)</td>
<td>TSEB, SEBAL</td>
<td>15–90 m (ASTER)</td>
<td>Bias 10–80</td>
<td>8 stations on 1 site, 1 date</td>
</tr>
<tr>
<td>Galleguilos et al. (2011)</td>
<td>S-SEBI, WDI</td>
<td>90 m (ASTER)</td>
<td>–</td>
<td>Comparison with SVAT (Hydrus 1-D) and 1 station on 1 site</td>
</tr>
<tr>
<td>Gomez et al. (2005)</td>
<td>S-SEBI</td>
<td>20 m (airborne)</td>
<td>RMSD ~ 90</td>
<td>7 stations on 1 site, 19 dates</td>
</tr>
<tr>
<td>Gonzalez-Dugo et al. (2009)</td>
<td>TSEB, METRIC</td>
<td>120–60 m (LANDSAT 5 and 7)</td>
<td>RMSD ~ 50</td>
<td>12 stations on 1 site, 3 dates</td>
</tr>
<tr>
<td>Jia et al. (2003)</td>
<td>SEBS, SEBI</td>
<td>1 km (ATSR-2)</td>
<td>RMSD 10–40 (BR)</td>
<td>11 stations on 1 site, 11 dates, 1 scintilometer</td>
</tr>
<tr>
<td>Li et al. (2005)</td>
<td>TSEB</td>
<td>120–60 m (Landsat 5 and 7, aircraft)</td>
<td>RMSD 40–120</td>
<td>5 stations on 1 site, 4 dates</td>
</tr>
<tr>
<td>Li et al. (2008)</td>
<td>TSEB</td>
<td>120 m (LANDSAT 5)</td>
<td>RMSD ~ 40</td>
<td>3 stations on 1 site, 3 dates</td>
</tr>
<tr>
<td>Ma et al. (2011)</td>
<td>SEBS</td>
<td>90 m (ASTER)</td>
<td>Bias &gt; 80</td>
<td>3 stations on 1 site, 4 dates</td>
</tr>
<tr>
<td>McCabe and Wood (2006)</td>
<td>SEBS</td>
<td>990–1020 m (aggregated ASTER and Landsat + MODIS)</td>
<td>RMSD 60 (HR)–80 (BR)</td>
<td>7 stations on 1 site, 1 date</td>
</tr>
<tr>
<td>Minacapilli et al. (2009)</td>
<td>TSEB, SEBAL</td>
<td>15 m (airborne)</td>
<td>–</td>
<td>1 date</td>
</tr>
<tr>
<td>Oku et al. (2007)</td>
<td>SEBS</td>
<td>5–7 km (GMS-5)</td>
<td>RMSD &gt; 100</td>
<td>2 dates, sites CEOP EOP-1</td>
</tr>
<tr>
<td>Su et al. (2007)</td>
<td>SEBS</td>
<td>1 km (MODIS)</td>
<td>RMSD 40–60</td>
<td>2 dates, sites CEOP EOP-1</td>
</tr>
<tr>
<td>Su et al. (2005)</td>
<td>SEBS</td>
<td>30 m (LANDSAT 7) and 20 km (GOES, MODIS)</td>
<td>RMSD 30 (HR)–140 (BR)</td>
<td>1 date, 8 stations on 8 sites</td>
</tr>
<tr>
<td>Timmermans et al. (2007)</td>
<td>TSEB, SEBAL</td>
<td>6–12 m (airborne)</td>
<td>RMSD 60–70</td>
<td>2 stations on 1 site, 5 airborne flights</td>
</tr>
<tr>
<td>van der Kwast et al. (2009)</td>
<td>SEBS</td>
<td>90 m (ASTER)</td>
<td>–</td>
<td>6 stations on 1 site, 1 date</td>
</tr>
<tr>
<td>Verstraeten et al. (2005*)</td>
<td>S-SEBI</td>
<td>1.1 km (NOAA/AVHRR)</td>
<td>RMSD ~ 40</td>
<td>13 stations on 13 sites, AVHRR series</td>
</tr>
<tr>
<td>Yang and Wang (2011)</td>
<td>S-SEBI</td>
<td>1 km (MODIS)</td>
<td>Only evaporative fraction evaluated</td>
<td>12 stations on 12 sites, 16 MODIS dates</td>
</tr>
<tr>
<td>Jacob et al. (2002)</td>
<td>SEBAL</td>
<td>20 m (airborne)</td>
<td>RMSD ~ 85</td>
<td>7 stations on 1 site, ~15 airborne flights</td>
</tr>
</tbody>
</table>
Table 2. EC stations used to compute performance statistics at each date. Correspondance of numbers and stations: [1]: wheat (east), [2]: wheat (west), [3]: broccoli/beans, [4]: chickpea, [5]: chili pepper, [6]: potatoes/sorghum and [7]: safflower.

<table>
<thead>
<tr>
<th>Observed data</th>
<th>364 (2007)</th>
<th>54</th>
<th>70</th>
<th>102</th>
<th>118</th>
<th>127</th>
<th>134</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_n$</td>
<td>–</td>
<td>[1–4]</td>
<td>[1–4]</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>1,2,4,6</td>
</tr>
<tr>
<td>$G$</td>
<td>–</td>
<td>1,2,4</td>
<td>[1–4]</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
</tr>
<tr>
<td>$H$</td>
<td>3</td>
<td>[1–3]</td>
<td>[1–3]</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
</tr>
<tr>
<td>$\lambda E$</td>
<td>3</td>
<td>1,2,4</td>
<td>1,2</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
<td>[1–4],6</td>
</tr>
<tr>
<td>albedo</td>
<td>3</td>
<td>1,3</td>
<td>1,3,4</td>
<td>1,3,4</td>
<td>1,3,4</td>
<td>1,3,4</td>
<td>1,4</td>
</tr>
<tr>
<td>$T_0$</td>
<td>–</td>
<td>[1–5]</td>
<td>[1–5],7</td>
<td>[1–5],7</td>
<td>[1–7]</td>
<td>[1–7]</td>
<td>1,2,[4–7]</td>
</tr>
</tbody>
</table>
**Table 3.** Performance statistics for the sensible heat flux $H$.

<table>
<thead>
<tr>
<th></th>
<th>TSEB</th>
<th>SEBS</th>
<th>S-SEBI</th>
<th>ICARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD</td>
<td>110.6</td>
<td>110.7</td>
<td>106.1</td>
<td>97.6</td>
</tr>
<tr>
<td>CV(RMSD)</td>
<td>56.1</td>
<td>56.2</td>
<td>53.8</td>
<td>49.8</td>
</tr>
<tr>
<td>Bias</td>
<td>−63.6</td>
<td>−40.4</td>
<td>0.13</td>
<td>32.34</td>
</tr>
</tbody>
</table>
Table 4. Performance statistics for the latent heat flux $\lambda E$ (in Wm$^{-2}$).

<table>
<thead>
<tr>
<th></th>
<th>TSEB</th>
<th>SEBS</th>
<th>S-SEBI</th>
<th>ICARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD</td>
<td>131.5</td>
<td>138.6</td>
<td>119.9</td>
<td>116.4</td>
</tr>
<tr>
<td>CV(RMSD)</td>
<td>49.2</td>
<td>51.9</td>
<td>44.9</td>
<td>40.0</td>
</tr>
<tr>
<td>Bias</td>
<td>99.2</td>
<td>70.5</td>
<td>40.5</td>
<td>−53.6</td>
</tr>
</tbody>
</table>
Table 5. Performance statistics in Wm\(^{-2}\) for the net radiation \(R_n\).

<table>
<thead>
<tr>
<th></th>
<th>TSEB/SEBS/S-SEBI</th>
<th>ICARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD</td>
<td>44.2</td>
<td>63.2</td>
</tr>
<tr>
<td>CV(RMSD)</td>
<td>7.7</td>
<td>10.9</td>
</tr>
<tr>
<td>Bias</td>
<td>23.9</td>
<td>26.4</td>
</tr>
</tbody>
</table>
Table 6. Performance statistics in Wm$^{-2}$ for the soil heat flux $G$.

<table>
<thead>
<tr>
<th></th>
<th>TSEB</th>
<th>SEBS/S-SEBI</th>
<th>ICARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD</td>
<td>59.4</td>
<td>56.4</td>
<td>96.1</td>
</tr>
<tr>
<td>CV(RMSD)</td>
<td>53.8</td>
<td>51.1</td>
<td>87.8</td>
</tr>
<tr>
<td>Bias</td>
<td>$-8.2$</td>
<td>$-7.0$</td>
<td>37.4</td>
</tr>
</tbody>
</table>
Table 7. Mean values of turbulent flux and water stress on the whole area at two dates of ASTER over-pass.

<table>
<thead>
<tr>
<th></th>
<th>10 Mar</th>
<th>6 May</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TSEB</td>
<td>SEBS</td>
</tr>
<tr>
<td>$H_{\text{mean}}$ (W m$^{-2}$)</td>
<td>101.31</td>
<td>154.21</td>
</tr>
<tr>
<td>$\lambda E_{\text{mean}}$ (W m$^{-2}$)</td>
<td>425.57</td>
<td>372.67</td>
</tr>
<tr>
<td>Water stress</td>
<td>0.13</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Table 8. Double-entry table of coefficient of determination $R^2$ (lower-left part) and RMSD (upper-right part) between the SEB models and their mean value and standard deviation $\sigma$ on 10 March.

<table>
<thead>
<tr>
<th></th>
<th>TSEB</th>
<th>SEBS</th>
<th>S-SEBI</th>
<th>VIT</th>
<th>Mean Value</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>TSEB</td>
<td>0.15</td>
<td>0.27</td>
<td>0.28</td>
<td>0.11</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>SEBS</td>
<td>0.43</td>
<td>0.27</td>
<td>0.26</td>
<td>0.17</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>S-SEBI</td>
<td>0.94</td>
<td>0.46</td>
<td>0.06</td>
<td>0.35</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>VIT</td>
<td>0.91</td>
<td>0.58</td>
<td>0.95</td>
<td>0.36</td>
<td>0.27</td>
</tr>
</tbody>
</table>
Table 9. Double-entry table of coefficient of determination $R^2$ (lower-left part) and RMSD (upper-right part) between the SEB models and their mean value and standard deviation $\sigma$ on 6 May.

<table>
<thead>
<tr>
<th></th>
<th>TSEB</th>
<th>SEBS</th>
<th>S-SEBI</th>
<th>VIT</th>
<th>Mean Value</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSEB</td>
<td>0.28</td>
<td>0.10</td>
<td>0.13</td>
<td>0.46</td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>SEBS</td>
<td>0.07</td>
<td>0.23</td>
<td>0.25</td>
<td>0.56</td>
<td></td>
<td>0.18</td>
</tr>
<tr>
<td>S-SEBI</td>
<td>0.84</td>
<td>0.28</td>
<td>0.04</td>
<td>0.44</td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td>VIT</td>
<td>0.78</td>
<td>0.32</td>
<td>0.95</td>
<td>0.40</td>
<td></td>
<td>0.23</td>
</tr>
</tbody>
</table>
Fig. 1. Satellite view of the studied zone with respective positions of the EC stations (Imagerie © 2012 Cnes/Spot Image, DigitalGlobe, Données cartographiques © 2012 Google, INEGI).
Fig. 2. Scheme of the resistance network of TSEB following Norman et al. (1995).
Fig. 3. ASTER surface temperature versus FORMOSAT albedo and the bounding linear regressions, on 27 April 2008.
Fig. 4. Schematic of the VIT (vegetation index-temperature) trapezoid method. Values from the 11 April.
Fig. 5. ASTER and ICARE versus in-situ radiometric surface temperature (in °C) at each station at ASTER over-pass dates. Color code: red: $0 \leq \text{LAI} < 0.4$; yellow: $0.4 \leq \text{LAI} < 0.8$; green: $0.8 \leq \text{LAI} < 1.2$; cyan: $1.2 \leq \text{LAI} < 1.6$; navy-blue: $1.6 \leq \text{LAI} < 2.0$; black: $2.0 \leq \text{LAI}$. 
Fig. 6. Computed versus measured net radiation (W m$^{-2}$) for the SEBS, TSEB and S-SEBI (left) and ICARE (right). Color code: red: 0 ≤ LAI < 0.4; yellow: 0.4 ≤ LAI < 0.8; green: 0.8 ≤ LAI < 1.2; cyan: 1.2 ≤ LAI < 1.6; navy-blue: 1.6 ≤ LAI < 2.0; black: 2.0 ≤ LAI.
Fig. 7. Computed versus measured soil heat flux ($\text{Wm}^{-2}$) for TSEB (left), SEBS (center) and ICARE (right). Color code: red: $0 \leq \text{LAI} < 0.4$; yellow: $0.4 \leq \text{LAI} < 0.8$; green: $0.8 \leq \text{LAI} < 1.2$; cyan: $1.2 \leq \text{LAI} < 1.6$; navy-blue: $1.6 \leq \text{LAI} < 2.0$; black: $2.0 \leq \text{LAI}$. 

955
Fig. 8. Detailed scatter plot of TSEB versus measured $G$ flux ($\text{Wm}^{-2}$). Color code: red: $0 \leq \text{LAI} < 0.4$; yellow: $0.4 \leq \text{LAI} < 0.8$; green: $0.8 \leq \text{LAI} < 1.2$; cyan: $1.2 \leq \text{LAI} < 1.6$; navy-blue: $1.6 \leq \text{LAI} < 2.0$; black: $2.0 \leq \text{LAI}$.
**Fig. 9.** Simulated versus measured sensible heat flux $H$ for every station at ASTER overpass dates. Color code: red: $0 \leq \text{LAI} < 0.4$; yellow: $0.4 \leq \text{LAI} < 0.8$; green: $0.8 \leq \text{LAI} < 1.2$; cyan: $1.2 \leq \text{LAI} < 1.6$; navy-blue: $1.6 \leq \text{LAI} < 2.0$; black: $2.0 \leq \text{LAI}$. 
**Fig. 10.** Simulated versus measured latent heat flux $\lambda E$ for every station at ASTER overpass dates. Color code: red: $0 \leq \text{LAI} < 0.4$; yellow: $0.4 \leq \text{LAI} < 0.8$; green: $0.8 \leq \text{LAI} < 1.2$; cyan: $1.2 \leq \text{LAI} < 1.6$; navy-blue: $1.6 \leq \text{LAI} < 2.0$; black: $2.0 \leq \text{LAI}$. 
Fig. 11. Scatter plots of water stress at each station for TSEB, SEBS, S-SEBI, VIT method and ICARE. Values calculated for ASTER over-pass dates. Color code: red: $0 \leq \text{LAI} < 0.4$; yellow: $0.4 \leq \text{LAI} < 0.8$; green: $0.8 \leq \text{LAI} < 1.2$; cyan: $1.2 \leq \text{LAI} < 1.6$; navy-blue: $1.6 \leq \text{LAI} < 2.0$; black: $2.0 \leq \text{LAI}$.
Fig. 12. Frequency histograms of calculated fluxes and remote sensing parameters on 10 March.
Fig. 13. Stress frequency histograms and maps for the whole area on 10 March.
Fig. 14. Stress frequency histograms and maps for the whole area on 6 May.
Fig. 15. Comparison of TSEB and observed \( \lambda E \) fluxes with \( f_g = 1 \) (top, left) 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) and the soil (\( R_{n,s} \), bottom, right) with \( f_g = 1 \) (x-axis) and with \( f_g \) taking into account the drying of leaves (y-axis), respectively. Color code: red: \( 0 \leq \text{LAI} < 0.4 \); yellow: \( 0.4 \leq \text{LAI} < 0.8 \); green: \( 0.8 \leq \text{LAI} < 1.2 \); cyan: \( 1.2 \leq \text{LAI} < 1.6 \); navy-blue: \( 1.6 \leq \text{LAI} < 2.0 \); black: \( 2.0 \leq \text{LAI} \).