Authors' response to referee #1 comment (hessd-10-C248-2013) on "Intercomparison of four remote sensing based surface energy balance methods to retrieve surface evapotranspiration and water stress of irrigated fields in semi-arid climate". MS No.: hess-2012-545, by Chirouze et al.

The authors would like to give their thanks to anonymous referee #1 for his advised comments that will greatly help to improve the manuscript's scientific quality.

In the following document, bold & italic text is extracted from the referee's comment file; normal text corresponds to authors' response.

### 2. Manuscript organization

The manuscript will be reorganized as suggested.

### 3. Concepts to be clarified

### 3.1 About the use of ICARE/SVAT model as reference

The performances of the four RS models were evaluated through a comparison with a set of micro-meteorological EC system and also using the spatially distributed outputs of ICARE/SVAT model. The latter, as described by Authors (pag. 915 lines 24,25 and pag. 916 lines 1-5), is ..."a classical dual-source SVAT model that solves the water balance of "a surface".....using two restore the lavers force model" ...that..."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".

In my opinion the "water balance" sub-model should be better described because in ICARE/SVAT the outputs of energy balance sub-model, i.e. evapotranspiration (used as reference for model's comparison) strongly depends on water balance sub-model, through the "force and restore" approach. I think that the SVAT approach (ICARE) can be usefully used for an inter-comparison exercise, but its characteristics should be clearly explained.

On the basis of this argument, I think that the Author should better explain in the revised introduction of paper (state of art) the characteristics of ICARE/SVAT and the reason of its use as reference.

One expects that a SVAT model forced with true irrigation and rainfall data will perform better than a sole energy budget model and can be somewhat considered as a "reference"; it is not, strictly speaking, a benchmarking tool (as originally stated in the paper) since the model has not been completely calibrated at each site; this will be clarified in the revised manuscript.

### Moreover in the "Material and Methods" the Authors should also describe and detail data and parameters used in WB sub-model of ICARE model

#### (hydraulic soil parameters, depths of soil layers, initial conditions in surface and root zone, conditions at the bottom of soil profile, root uptake functions, etc,).

A paragraph will be added to address this.

It will appear in the revised manuscript page 916, line 2 as:

"The evolution of soil moisture in those two layers is described following (Noilhan and Planton, 1989) :

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

Where  $w_s$  and  $w_2$  are respectively the surface layer and root zone 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 0.05m and 1m,  $\tau$  the diurnal time period (one day) and  $\rho_w$  the density of water.  $C_1$  and  $C_2$  are force restore coefficients depending on soil texture estimated following the ISBA model (Noilhan and Mahfouf, 1996) and  $w_{eq}$  an equilibrium soil surface moisture representing the case where capillarity and gravity processes compensate exactly."

Then, page 916, line 4:

"Initialization of soil moisture and temperature for each soil layer was done from in-situ measurements at each station."

#### Then, page 916, line 14:

"Soil texture has been determined from in-situ measurements at each station and main soil parameters were estimated following ISBA pedotransfer rules (Noilhan and Mahfouf, 1996)."

Another question that the Authors should clarify is the definition of the "main families" of Remote sensing method for ET estimation. The Authors used the following classification:

1. Contextual methods (pag. 899, lines 21-24) ..." 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".

2. Single-pixel models and (pag. 899, lines 3-4) ... "methods" .. that "solve an energy budget for each pixel independently from the others".

In my opinion this type of classification is not properly appropriate because, for example, in the group 1 can be included SEBS model which uses (as described by Authors at pag. 909, line 5) the concept of "hot/dry" and "cold/wet" pixels as boundary conditions.

In our definition "contextual" means "using simultaneously information from different pixels with contrasted vegetation and/or temperature values", this will be clarified and emphasized in the revised manuscript. Therefore, S-SEBi and the triangle method strictly belong to this category, contrarily to SEBS and TSEB.

Indeed, in SEBS, dry and wet conditions are idealized situations for the status of a given pixel: SEBS is not using several pixels of the same image to compute a relative EF.

# Moreover, following the classification proposed by Authors, in the group 2 could be partially considered the S-SEBI model, that computes for "each pixel, independently from the others", the evapotranspiration term by means of a simplified relationship for Evaporative Fraction calculation.

In S-SEBI, EF calculation depends on the endmember envelops in the observed temperature/albedo space (limiting edges corresponding to "radiation driven" and "evaporation driven" situations), therefore EF cannot be computed for one single pixel independently from the others, thus belongs to the "contextual" family according to our definition.

Note that the proposed classification has important impacts on precision (contextual methods are less sensitive to systematic absolute errors on surface temperatures) and data assimilation (variance/covariance matrices are not built in the same way when models are applied independently or not at each pixel location).

Therefore the Author should use a different way to classify the method for ET estimation from Remote Sensing data. I can suggest to use this type of classification (or similar):

1. Simplified energy balance index methods (for example S-SEBI and similar, that are all methods based on an analysis of the relationship between albedo or NDVI and surface temperature to obtain a simplified equation for the calculation of Evaporative Fraction or ET).

2. Direct energy balance methods (for example SEBAL, TSEB, SEBS, that are all methods based on the direct estimation of the latent heat flux,  $\lambda$ ET, as residual term from the surface energy balance equation). Within this family a distinction has to be done between:

2.1- Single Source approaches (as SEBS or SEBAL), where soil and vegetation are considered as a combined sole source;

2.2 Two Source approaches (TSEB), where soil and vegetation are treated separately.

Furthermore, a part the type of classification, considering that the Authors compared SEBS (single source) and TSEB (two source) models, in the next version of the paper they should explain the main differences between "single-source" and "two-source" approaches to estimate sensible heat flux H that is term with the largest uncertainty in estimating  $\lambda$ ET.

The proposed classification is perfectly valid and not far from ours, but provides little help to interpret the differences at local and spatial standpoints between the outputs of the various models. We believe that using or not the context (i.e. information obtained on other pixels) in a given image is an important difference in the way the various methods perform. For instance, SEBAL is a complex contextual model, which will behave differently from SEBS and TSEB. We agree though that within each category there is a second level of classification related to complexity, or the possibility to discriminate evaporation from transpiration, and we'll stress that in the revised manuscript.

The Authors used both ASTER and FORMOSAT-2 imagery data-set. ASTER data were used to exploit its Thermal band and FORMOSAT-2 for VIS-NIR bands. Indeed, ASTER provides also VIS-NIR and SWIR bands useful to compute albedo, so I don't understand the need to use FORMOSAT-2 data. Moreover using the greater number of VIS-NIR and SWIR ASTER bands the computation of albedo would have been improved respect to the method used by Author (eq. 2).

Although it's true that the number of available bands in ASTER is better suited to produce a broadband albedo than FORMOSAT, we had issues with ASTER SWIR bands. Indeed, some of the SWIR bands were not usable for 4 of the 7 available ASTER images. Although we could have used them on the valid days, we preferred to use the same product for the whole study in order to keep the same error sources for each date. That justifies the use of FORMOSAT-2 data to assess broadband surface albedo. It seems, from comparison with in-situ data, and by looking at S-SEBI's performance, that albedos produced by FORMOSAT are meaningful.

#### 3.3 About definition and estimation of Water Stress (Pag. 621 L:12-26). The Authors used as Water Stress index the term: $1 - \lambda E / \lambda E_{max}$ , where $\lambda E$ and $\lambda E_{max}$ are actual and potential latent heat flux from a plant, respectively (pag. 921, L:10-12). This definition is not properly correct, because the actual Stress of a plant should be related only to the transpiration term removing the evaporation term.

We agree with this, therefore we propose to replace the term "plant water stress" by "surface water stress" which is more consistent with its definition. Surface water stress is computed for the whole surface (total evaporation fluxes).

#### Moreover the term $\lambda E_{max}$ was computed using ICARE/SVAT model with the option of continuous irrigation. Sincerely, I don't well understand this choice and I have the doubt that the comparison between the Water Stress indexes derived from models (Fig. 11) could be not properly homogeneous.

In order to compute surface water stress levels, we had the choice to use a common maximum (non water limited) rate  $\lambda E_{max}$  for all methods (thus producing values outside the range [0-1]) or to use the maximum level derived from each model. Indeed, each model has a clear definition of  $\lambda E_{max}$ . Each actual evapotranspiration rate is computed with this inner  $\lambda E_{max}$  as reference. It seems more consistent, and, in a way, homogeneous, to use the latter, which can be translated into a relative soil moisture level. For ICARE, a simple way to derive this is to run the model with continuous irrigation. That warranties not only that both transpiration and evaporation are at a maximum level but as well that all fluxes (including soil heat flux) are consistent with this excess soil moisture.

The difference in max levels is part of the  $\lambda E/\lambda E_{max}$  comparison, we did not think useful to add a figure comparing the max levels which are very similar. Using, as proposed, a single common  $\lambda E_{max}$  (potential ET) for all estimates leads to the following Figure 1. The difference is not very significant and lead to values outside of the [0-1] range due to discrepancies in the various  $\lambda E_{max}$  estimates.



Figure 1: Scatter plot of stress computed from  $\lambda E$  ouput of each model with the same potential evapotranspiration model as the one used for calculating in-situ surface stress.

## About this problem my question is if a more homogeneous comparison could be conducted using only TSEB and ICARE (the only model that, being dual-source, are able to retrieve actual transpiration) and using reference ET<sub>0</sub> in place of $\lambda E_{max}$ .

Such a comparison would be possible in ICARE (which computes transpiration according to available water in the root zone) and in TSEB when there is no stress. However when stress appears, the vegetation part of the latent heat flux calculated by TSEB is more of an artifact used to assess surface evapotranspiration than a real vegetation transpiration. Indeed, in this case, soil evaporation is set to a fix value (either zero as in the original paper or 50 W/m2 as in more recent TSEB applications), which would mean that when the available water volume is not sufficient for the plant to transpire at potential rate, the soil stops to evaporate, which is not always true. Therefore, in this case, the  $\lambda E_c$  is in fact the total surface latent heat flux considering partially stressed vegetation and evaporating soil. Given this statement, the comparison of ICARE and TSEB transpiration terms would be strongly biased.

In order to estimate an "observed" surface water stress, we combined total evapotranspiration measurements for each flux station with a potential evaporation rate computed with observed meteorological forcing and LAI and vegetation height measured at each flux station (ET0 is not dependent on vegetation extent). It provides a satisfying lower and upper bounding condition for observed latent heat flux (the reconstructed "observed" water stress being rarely outside the range [0-1]).

### *4. Specific comment and technical corrections*

*P:897, L:1-4.* "Remotely sensed surface temperature can provide a good proxy for water stress level and is therefore particularly useful to estimate spatially distributed evapotranspiration". It is on the contrary: RS can provide a good proxy for ET estimation and is therefore useful to quantify water stress.

We agree that this sentence is not well formulated. Surface temperature provides information about stress level and evapotranspiration at the same time, therefore there is no causality link between those two variables. This will be reformulated as follow:

"Remotely sensed surface temperature can provide a good proxy for evapotranspiration and water stress level and is therefore particularly useful to estimate those variables at a regional scale."

### *P:897, L:5. Clarify the term "equilibrium temperature".*

We agree that this is not well formulated. The equilibrium temperature of the surface energy budget is not directly the remotely sensed surface temperature. The word "equilibrium" will be deleted.

### P:897, L:7-11. Reorganize following my previous comment (3.1).

We would like to keep our original classification (see response to comment 3.2).

### *P:898, L:1-4. Add some references about "water use" data.*

It will be added.

## *P:899, L:11-18. Clarify the nature of SVAT and the reason of its use as reference for Remote Sensing approaches. Clarify also the data-assimilation concept.*

The nature and reason for its use as reference for Remote Sensing approaches will be clarified as explained in the response to comment 3.1.

### *P:899, L:17. Check the reference "Schuurmans et al., 2003 (..or Schuumans ?).*

It is Schuumans et al., 2003, indeed.

### P:899, L:21. Reorganize following my previous comment (3.1).

We would like to keep our original classification (see response to comment 3.2).

### *P:900, L:28.* "Choi et al., 2009" is not reported in the Reference list.

This will be corrected.

### *P:901, L:2. Invert years in "Su et al., 2007,2005"*

This will be corrected.

**P:901**, L:2. "..in most cases, those studies.." ... "and are limited to two or three intercompared models". This comment is not useful, after all the Authors inter-compared four models !!.

Comment will be deleted.

### P:901, L:15-20. Reorganize following my previous comment (3.2).

See response to comment 3.2.

### *P:902, L:8-14.* These are the objectives of work. Move at the end of new introduction.

By merging Introduction and State of the Art, this paragraph will be placed at the end of the introduction.

### P:907, Eq. (2). Clarify following my previous comment (4).

A sentence will be added to justify the use of FORMOSAT albedo.

It will be added page 907, line 2:

"Although the ASTER platform provides more bands in near and shortwave infrared than FORMOSAT-2, which would suggest that a more consistent short-wave broadband albedo can be computed, a dysfunction of the acquisition instrument in the SWIR (shortwave infrared) domain occurred on four of the seven available ASTER dates. This made calculus of an ASTER albedo impossible on those dates and FORMOSAT-2 data was chosen in order to keep an homogeneous albedo over the whole study and to not multiply the sources of error."

*P:* 908,909,910,911. In my opinion these are the pages where the description of models could create confusion. On the basis of previous comments about the classification of Remote Sensing method (3.2), I suggest to explain how the models compute Evapotranspiration using this order:

1. Description of surface balance equation to obtain the instantaneous  $\lambda ET$  as residual term ( $\lambda ET = Rn - H - G$ ) and definition of Evaporative Fraction, EF.

2. Description of methods to compute Rn and G (as at pag. 908) that are common in TSEB and SEBS;

3. Description and discussion of the differences in H estimation between TSEB and SEBS (also following my previous comment 3.2). To

#### do this, I think that it is useful to describe, at first, the general equation for H (H=pcp∆T/Ra and its modifications in case of two-source approach, H=Hc+Hs); then, all terms used to describe wind and temperature profiles according Monin-Obukhov could be described in a synthesized form (Is it necessary to shown eqs. 7, 8, 9, 10, 11?) 4. Describe the simplified methods (S-SEBI and VIT).

We agree that the surface energy balance equation could be introduced at this point. However, only S-SEBI computes an evaporative fraction to determine the fluxes. SEBS computes an evaporation efficiency which is defined by  $\lambda E/\lambda E_{max}$  and thus is not the evaporative fraction defined as  $\lambda E/(R_n-G)$ . Instead of defining the evaporative fraction in the beginning of this section, we would like to make clearer both definitions of evaporative fraction and evaporation efficiency (which can be also found in the definition of WDI in the trapezoid method) in each of the models descriptions.

We agree that the description of Monin-Obukhov theory could be synthetized and will work to shorten it.

However, we think that differences between SEBS and TSEB do not simply summarize as different ways to compute H (simple or double source). For example, the loop used in TSEB to take into account vegetation stress of concepts of hot and dry conditions in SEBS are crucial in the interpretation of errors and understanding of the models. As a consequence, we think that grouping those two models in one single paragraph would tend to lose the reader and make the philosophy of the models more confuse.

### P:909, L:15. Clarify the concept of "potential temperature".

"potential temperature" is the term used in the original SEBS paper; it represents the temperature of the air corrected for pressure difference effects; in our case, the reference height used is low (10 m) therefore potential and real temperatures are similar.

P:916, L:15-20. "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"...." 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". This part is crucial (see my previous comment 3.1). As the ICARE is used a reference the Authors should be better detail this part of work.

This will be detailed.

### *P:920, L:19. The comment about the kB*-1 parameter is not clear, ..it is "too big" respect what ?

The kB<sup>-1</sup> is often overestimated (Gokmen et al., 2012, Boulet et al., 2012) due to the overestimation of kB<sup>-1</sup><sub>s</sub> by the Brutsaert (1976) formulation.

This will be reformulated as follow:

"The  $kB^{-1}$  appears to be overestimated, which is in concordance with literature (Boulet et al., 2012; Gokmen et al., 2012), and leads to an overestimation of the atmospheric resistance."

### *P:920, L:21. The comment about the about the overestimations of H should be more detailed.*

It will be detailed.

### *P*:921, *L*:13. "for a p."?

Correction  $\rightarrow$  "for a pixel".

P:921, L:13-24. Really, I did not understand this part. Please, clarify (See my previous comment, 3.3).

See response to comment 3.3.

*P:928, L:18.* "Shuttleworth and Wallace (1985)" is not reported in References list.

It will be reported.

P:941, Table 3. Specify unit for H.

*P*:953, Fig. 5. Insert labels and units in x and y axis and remove title.

*P:954, Fig. 6. Insert labels and units in x and y axis and remove title.* 

*P:955, Fig. 7. The size of figures is too small. Insert labels and units in x and y axis and remove title. Insert the same tics in both axis.* 

*P:956, Fig. 8. The resolution of captured figure seems too small. Insert labels and units in x and y axis and remove title. Insert the same tics in both axis.* 

*P:957, Fig. 9. Insert labels and units in x and y axis and remove title. Insert the same tics in both axis.* 

*P*:958, Fig. 10. Insert labels and units in x and y axis.

*P*:959, Fig. 11. Insert labels and units in x and y axis.

P:960, Fig. 12. The size of figures is too small.

P:961, Fig. 13. The size of figures is too small.

P:962, Fig. 14. The size of figures is too small.

### **P:963**, Fig. 13. The size of figures is too small. Insert the same tics in both axis

Figures and tables will be corrected according to comments.