Dear Editor,

Thank you so much for your kind attention to consider our work in HESS! With the helpful comments from the reviewers, we had done a lot of work to improve the paper. Please see the point-by-point response to the reviews below.

Interactive comment on "Mapping evapotranspiration with high resolution aircraft imagery over vineyards using one and two source modeling schemes" by T. Xia et al.

C. Jimenez (Referee)
carlos.jimenez@obspm.fr
Received and published: 19 December 2015

The paper presents an example of terrestrial evaporation (ET) estimation over 2 vineyard fields by 2 thermal-based methodologies. One is a relatively new and very simple one-source algorithm with a very complex name (DATTUTDUT, I'll shorten it as DAUT hereinafter) that only requires land surface radiometric temperature; the second one is a more established methodology (TSEBS) that requires also multi-spectra imagery at different bands and its algorithm is more complex to apply.
The paper is a good contribution to the field of deriving ET at the few meters resolution for agriculture and water management applications, illustrate our current capability to use thermal imagery to estimate ET over these type of fields, and discuss some of the remaining challenges. The paper is well written with the adequate level of detail, and the work carried out is easy to follow.

Reply: Thank you for your positive comments!

Some general comments about the paper are:
(1) It is expected that the more complex algorithm (TSEBS) will out perform the simpler DAUT, so I was wondering about the motivation of this comparison for a large

part of the article. The idea that they could be used together in an operational scheme with TSEBS only run when DAUT fails came later in the paper. It should perhaps be presented already in the introduction (and not at the end in the conclusions) to help understanding the motivation of the paper.

Reply: We agree with this comment and include the following text in the introduction on page 6, lines 8-13: "*However, more detailed comparisons between simple one-source contextual-based schemes versus more complex two-source models using high resolution imagery over different surfaces are still needed to fully understand the strengths and weaknesses of both modeling schemes. Such intercomparisons can facilitate development of hybrid schemes that leverage the strengths of different methodologies (e.g., Cammalleri et al., 2012), while incorporating simplications for routine application with airborne imagery.*"

(2) Although the spacecraft multispectral imagery is at sub-meter resolution, most of the data is spatially downgraded to a few meters resolution. This is required to allow TSEBS to operate, but my understanding of DAUT makes me think that it can operate at the sub-meter resolution. It would have been very interesting to have DAUT at sub-meter resolution, aggregate to meter-resolution, and compare with the TSEBS meter-resolution ET estimates. We could think that the one-source models are more adequate to deal with less "complex" pixels (in terms of soil canopy composition), so running DAUT at sub-meter resolution is likely to result in a larger number of "simpler" pixels (i.e., full canopy or full soil coverage) and a better performance. In the context of precision agriculture applications commented in the introduction that would have been interesting to see.

Reply: We agree with the reviewer that DATTUTDUT can be operated at the sub-meter resolution since in principle its parameterizations are not affected by the spatial resolution. In the original submission, DATTUTDUT was only run with high resolution thermal data in the sensitivity analysis section. In the current version, DATTUTDUT was also applied to the native resolution thermal-IR data and the results are described in the text, as well as in the related tables (Tables 2 and 3) and figures (Fig. 6 and 7).

In the text we now include the following (see page 17 line 6-11 "*At both instantaneous and daytime time scales, application of DATTUTDUT with the native (finer) pixel resolution thermal imagery yielded comparable (at Site 1) or significantly greater (at Site 2) discrepancies with the tower measurements than using the 5 m pixel resolution data (see Tables 2 and 3). Changes in the agreement with the tower measurements are mainly attributable to the new hot and cold temperature pixels selected by the DATTUTDUT procedure with the finer resolution TR data.*"

Using the finer thermal pixel resolution also had some effect on the spatial patterns/distributions of ET on some of days (see Fig. 9). This observation is included in the text (page 19, line 13-16) "*Use of the finer resolution data had generally a minor to moderate effect on the EF and ET distributions except for DOY 163 where the high resolution output indicates a bimodal distribution in EF and ET compared to the unimodal distributions using the 5 m resolution output from DATTUTDAT and TSEB.*"

(3) Thinking about real applications of these methods, the estimation of daily ET based on the ratio of instantaneous to daytime available energy at the tower seems a bit counterintuitive in the sense that in real life the tower will probably not be there. The ET methodology can still use the EF constant assumption, but it would require an estimation of the available energy not only at the instantaneous step but also integrated over the day. This could be perhaps discussed in the text.

Reply: Discussion referring to the method used to extrapolate instantaneous to daily ET is now included on page 10, line 19-22: "*In this study, the observed available energy from the two flux towers during the daytime period for all five days was used to extrapolate instantaneous model estimates to daytime ET totals. However, in practice tower measurements of A would not be available, so results using solar radiation to extrapolate to daytime ET will also be evaluated.*"

We computed daytime ET using solar radiation with the EF expressed as follows:

$$EF=\frac{LE_{ins}}{Rsolar_{ins}}=\frac{LE_{daytime}}{Rsolar_{daytime}} \quad (R1)$$

and the comparison between two methods were included in the revised manuscript (see Table 4). In Table 4, the MAE values for daytime ET using *A* versus S_d methods were similar for TSEB, but were larger for DATTUTDUT particularly at site 1 the north vineyard. A discussion of the results in Table 4 is on page 17 line 12-19: "*In practice, we will not have observations of available energy, A, from a flux tower for extrapolating the instantaneous ET from a single airborne observation to daytime ET, but instead are more likely to have weather station observations of incoming solar radiation, Sd. Results using Sd for extrapolating model estimates instead of flux tower measurements of A are listed in Table 4. In general, the differences between modeled and measured daytime ET (using RE method) increase, although not significantly for TSEB. On the other hand, discrepancies with the ET measurements for DATTUTDUT at the north vineyard (site 1) increase dramatically due to the large overestimation of instantaneous LE on DOY 162 and 219 (see Fig. 6b).*"

(4) The section about water consumption is just a comparison of TSEBS and DAUT field integrated ET with field integrated tower ET assuming that the tower ET fetch is representative of the whole field. It is of interest as it shows the variability that can exist at the field scale and the need of ET methodologies that could capture that variability. I imagine that for other type of crops growing in irrigated fields (i.e., water availability is not an issue) the variation would be smaller and micrometeorological methods still van be of utility. For this concrete example, based on the paper findings the variability seems related to a large extent to changes in LAI across the field. What about water availability, may it play a role here? The vineyards are drip irrigated, does this imply that water availability is constant through the field and ET response does not reflect water stress? I imagine that in practical applications there can be challenges in separating ET variability caused by water availability or biomass variability, and how this can be used to infer water stress and condition irrigation management.

Reply: Yes, we quite agree with the referee that both water availability and biomass will influence the spatial distribution of ET in field. Since the vineyards are drip irrigated, and the irrigation is replenishing available water in the vine root zone, the vine-rows with higher LAI are likely to have greater amount of root zone water. This suggests that either irrigated water was not evenly distributed across the vineyard and/or soil water holding capacity varied due to soil textural differences. The non-uniform distribution of irrigated water is illustrated in Fig. R1 below where some areas were over-irrigated or the water holding capacity of the soil was exceeded with the excess water causing re-growth of the senescent cover crop planted in the inter-row. This variation in root zone water availability is likely reflected in the spatial distribution in vine LAI.

This point is more clearly articulated in the revised manuscript on page 23 line 10-16. : "*The climate in this region is quite arid during the growing season with the drip irrigation being the only water source for the vines. As a result, the water availability (or soil water content) condition in the vine root zone plays a crucial role in the vegetation biomass. Therefore it is reasonable to assume there would be a strong correlation between ET and vine LAI as representative of the water availability in the root zone. The spatial variation in vine LAI is likely due to variation in the amount of irrigated water and/or variability in soil water holding capacity.*"

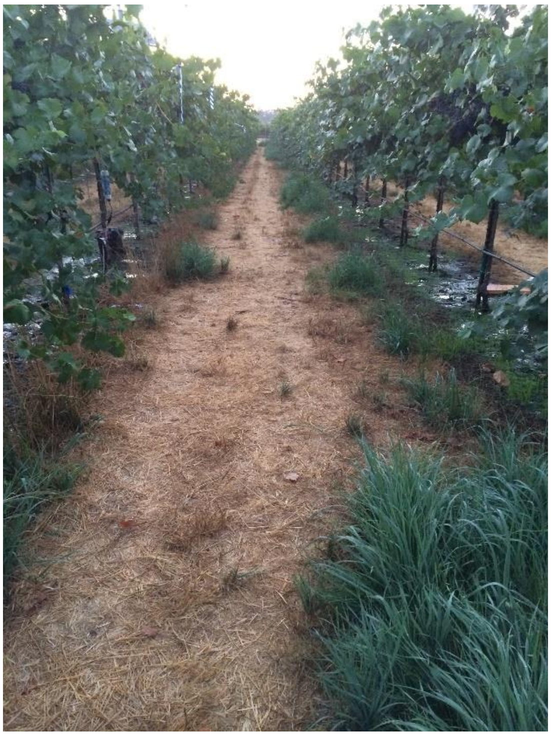

Figure R1. Example of inter-row cover crop re-growth due to excess irrigated water and/or soil water holding capacity.

Some more specific comments:

P11911.L10. I thought DAUT only required thermal images, no shortwave imagery as the downwelling radiation is based on astronomical calculations and the albedo scaled based on the thermal information.

Reply: Yes, the reviewer is correct that only a thermal image is required in the DATTUTDUT model.

P11913.L10. Given the importance of Rn for the ET estimation, it may be good to detail a bit how TSEBS operated here (as it is detailed for DAUT later in the text). My understanding is that it requires the downward SW and LW components (but these inputs are not listed in P11920.L8 as key inputs to TSEBS).

Reply: More details about the calculation of Rn in TSEB is now included (see Eq. (4) and (5) in the revised manuscript).

S_d can be from the measurement or calculated using astronomical information, and L_d can be calculated from meteorological data such as Brutsaert's (1975) longwave equation:

$$L_d = 1.24\left(e_a/T_a\right)^{1/7}\sigma T_a^4 \qquad \text{(R2)}$$

where e_a is actual water vapor pressure (kPa), T_a is air temperature (K), σ is the Stefan-Boltzmann constant (5.67×10^{-8} W m^{-2} K^{-4}).

Text is now included to describe how S_d and L_d are computed on page 7 line 8-12: "*Sd is either computed using sun-earth astronomical relationships under clear-sky conditions as done by DATTUTDUT (see below) or measured from a nearby weather station, and Ld is either measured or often computed using formulas based on weather station observations of air temperature and vapor pressure (i.e., Brutsaert, 1975).*"

P11917.L18. A map of LAI for one of the DOYs discussed my help illustrate the difference between both vineyards, and being useful also for the water consumption discussion later in the paper where ET variability within the field is linked to LAI variability.

Reply: An example of a LAI map for DOY 163 is now provided (see Fig. 4) and illustrates the variation in vine biomass/LAI for the north and south vineyards.

P11920.L1. See previous comment about imposing the scale where TSEBS can operate (few-meters) to DAUT.

Reply: We now run DATTUTDUT using the native (finer) pixel resolution of the thermal imagery as well as aggregated to 5 m used by the TSEB model.

P11920.L17. Ground-based reflectance was not collected for all the flights, but could the existing ground-based data have been used for a local (and possibly more accurate) calibration of the DN values? Was the existing ground-based data limited in terms of sampling the reflectance space? The airborne sub-meter and 30 meter LandSat resolutions are quite different.

Reply: We agree with the referee that using the ground-based data to fit the DN~reflectance relationship is a better way to obtain the reflectance maps. Unfortunately, the ground-based samples were very limited. The location of the sample sites are shown in Fig. R2 as the yellow stars. There are only a handful of sites where reflectance data were collected with most samples concentrated near the flux tower sites. Therefore it was felt that there was not a wide enough range in sampled reflectance values to provide a reliable conversion equation between DN and surface reflectance.

Figure R2. The location of the ground sampling sites for reflectance data.

P11921.L5. If we consider the 3 bands separately, the agreement to the ground based reflectance does not look that good, especially for the NIR. Any reason for the aircraft NIR being in worse agreement with the ground based NIR (if my impression of the NIR being worst than the others from looking at Figure 2 is correct)?

Reply: The reviewer is correct. The RMSD for NIR, green and red bands are 7.15%, 2.11% and 2.28%, respectively, and the error for NIR is relatively large. It is not known the source for the larger error with the NIR channel, although most likely it may be a drift in the sensor calibration so that the conversion of DN to reflectance was not constant for all flights.

P11922.L10. It may be of interest to give some details about the flux footprint model.

Reply: More details about the two-dimensional flux footprint model are included on page 15 line 14-15 although it is fairly standard, commonly used technique: "*This footprint model contains a lateral dispersion formulation to obtain a two-dimensional weighted source-area of flux from the upwind direction.*" The citations for this footprint model applied to imagery by Li et al. (2008) and the actual derivation of the model formulations given in Hsieh et al.(2000) provide all the necessary details.

P11923.L1. Wondering if bias and RMSD were similar for the estimates of Ts and Tc.

Reply: In Figure 5 of the revised manuscript the bias and RMSD for Ts and Tc are listed separately in the two plots. The Bias for T_s and T_c are both 0.5 ℃, and the RMSD values for T_s and T_c are 2.5 and 2.4 ℃, respectively indicating similar difference statistics for both soil and canopy temperature estimates.

P11923.L15. Some percentage figures may be nice to help judging the RMSD (e.g., the fraction of the expected instantaneous fluxes corresponding to the reported RMSD).

Reply: A percentage value is now listed in the text of the revised paper on page 16 line 13-16: "*Table 2 clearly shows that the RE closure adjustment method yields better overall agreement between measured and modeled fluxes with the average error computed as the ratio of RMSD and avereage observed flux value of ~27% for H and LE for the two sites, while the BR method has an error of ~37%.*"

P11924.L10. I would argue that the very simplistic determination of the Rn-G in DAUT is also key here, independent of the one-two source differences in how the fluxes are treated and the implications of the contextual scaling approach by DAUT. It would be curious to see how TSEBS and DAUT would score if their radiative inputs were exchanged (i.e., TSEBS using the downward SW and LW from DAUT, and DAUT using TSEBS available energy, if this makes any sense).

Reply: DATTUTDUT was run using Rn and G from TSEB, and the results are shown in Table 5 along with a discussion in the text on page 18 line 21-28: "*Using measured Sd from the towers instead of computing from the sun-earth astronomical relationships routinely applied by DATTUTDUT, there is only a minor reduction in the differences with the tower fluxes. An overall improvement in DATTUTDUT estimation of LE is achieved by adopting TSEB estimates of Rn and G (see Table 5). This is particularly true for the north vineyard (site 1). However, even with this better agreement in estimated LE, the discrepancies with observed LE from DATTUTDUT is still larger than with the output of TSEB. This indicates that the errors in available energy using the DATTUTDUT formulations are not the only significant source of error in estimating the LE flux.*" On average (the average value of Site 1 and 2) MAE and RMSD for LE_{RE} are reduced from ~65 and 84 W m^{-2} to ~52 and 68 W m^{-2}, which are still larger than ~37 and 44 W m^{-2} from TSEB. Therefore, using Rn and G from TSEB does have some value, however there still other sources of error for DATTUTDUT. Future applications will try to improve the Rn-G algorithm in DATTUTDAT.

Both the algorithm for energy and radiative exchange are mentioned as the potential strengths of the TSEB modeling scheme compared with DATTUTDUT on page 17 line 21-24: "*It is hypothesized that this likely results from a better physical representation of the* ***energy and radiative exchange*** *within TSEB, since it explicitly considers differences in soil and vegetation radiation and turbulent energy exchange and affects on the radiative temperature source*"

P11926.L25. For someone that works in evaluating satellite TR, this 3K bias is a relevant figure. Any ideas about the reasons suspected to be behind this bias? Out of curiosity, how are the ground-based TR measurements collected? TIR cameras?

Reply: As for possible reasons for the TR bias, the flying height for IOP 3 is 480 m above ground (see Table 1) which is higher than the other IOPs. It may be that the higher altitude caused greater atmosphere attenuation of the thermal signal at the sensor altitude. However, the aircraft-based T_R was higher than the ground-based T_R measured by the Apogee IRTs and CNR1 emitted longwave sensor converted to a T_R value.

P11927.L10. This is a nice exercise showing the relatively large sensitivity to the TR errors form thermal methods that depends on the absolute value of TR. I would argue that a more realistic determination of the TR uncertainties from available instruments is needed so this uncertainty can be properly propagated into the ET estimations errors.

Reply: We believe the 3K bias in T_R is a fairly large uncertainty and that in the other cases (days with aircraft imagery) the error in T_R is smaller based on comaprsions with ground-based T_R sensors. However, there is also the added complication of the pixel resolution which can give you a wide range in T_R values based on whether the sensor is viewing sunlit or shaded plant canopy or bare soil/cover crop inter-row.

P11928. L4. As discussed in this paragraph, perhaps a slightly more “complicated” but effective way of selecting the end members of the TR distribution could greatly improve DAUT estimates.

Reply: We agree with the reviewer. Modeled flux from DATTUTDUT model is largely dependent on the selection of the end members from the thermal imagery. There have been a number of published procedures to more reliably determine the temperature end members based on land cover considerations and vegetation cover information (such as using NDVI), although Timmermans et al. (2015) did not find more complex end member selection techniques consistently gave better results. However, as noted by Timmermans et al (2015) certain land cover conditions can cause significant errors in the automated end-member selection technique. Possibly adopting a technique similar to the one described by French et al. (2015) would improve reliability of DATTUTDUT flux estimates. This is a project for a future study.

P11931.L14. This section looks to me more just the "Conclusions" (the previous section was already labeled "Results and Discussions").

Reply: The title of this section is modified as "Conclusions" in the revised manuscript.

P11934. I imagine that the UAV technology is targeting the sub-meter resolution. Given the limitation of TSEBS to work beyond the plot/micrometeorological scale, any thoughts about possible candidates to complement DAUT for high precision agriculture (i.e., another thermal based ET methodology able to work at the sub-meter scale, operationally more demanding but able to deal with the cases where DAUT fails to provide decent ET estimates)?

Reply: This is a good question. We are indeed exploring this issue by adapting the TSEB to operate at the sub-meter resolution by developing a module that uses the coarser resolution output of TSEB at~5 m to define the key inputs to run TSEB in a parallel resistance network mode for plant canopy only pixels and bare soil/substrate only pixels. This will be the basis of a future paper on this topic.

Interactive comment on "Mapping evapotranspiration with high resolution aircraft imagery over vineyards using one and two source modeling schemes" by T. Xia et al.

Anonymous Referee #2

Received and published: 31 December 2015

This paper evaluates two energy balance models forced by very high resolution airborne thermal data over 2 vine fields. The simpler model (DATTUTDUT, or "rectangular" approach) only requires temperature data, while TSEB uses full climate forcing.

Main comments:

The originality of the paper lies in the use of very high resolution data obtained during 5 airborned overpasses, but this is not properly put forward in the paper: for TSEB, meter resolution data are used only to evaluate the component temperature retrieval, and for the rectangular approach running the model at both resolutions is only carried out in a sensitivity test. Model performances obtained when the rectangular approach is applied at the highest resolution could be brought forward or at least added to Table 4. Therefore, it's hard to grasp the added value of this intercomparison with respect to, say, Timmermans et al. 2015.

Reply: Results from DATTUTDUT using high resolution thermal data are now included in the revised manuscript (see the revised Tables 2-3 and Figures 6, 7, and 9). There is also discussion in the text on page 17 line 6-11: "*At both instantaneous and daytime time scales, application of DATTUTDUT with the native (finer) pixel resolution thermal imagery yielded comparable (at Site 1) or significantly greater (at Site 2) discrepancies with the tower measurements than using the 5 m pixel resolution data (see Tables 2 and 3). Changes in the agreement with the tower measurements are mainly attributable to the new hot and cold temperature pixels selected by the DATTUTDUT procedure with the finer resolution TR data.*"

The model inter-comparison is not the only motivation of this paper. Based on the performance of these two thermal-based ET models, one a simple contextual-based single-source approach versus a more physically-based two-source scheme using absolute surface-air temperature differences, we want to investigate their utility and ultimately look at ways of leveraging the strengths of both approaches in order to develop a hybrid modeling system that is more robust. We point out that the DATTUTDUT model is simple, readily operational with high resolution imagery and less sensitive to errors in absolute T_R, while TSEB is more physically-based and robust with greater accuracy in flux estimation under different environmental conditions, but it is more sensitive to uncertainty in absolute T_R. The paper then focuses on what attributes from these two modeling schemes are robust and would be important to incorporate in an operational ET modeling system applicable to high resolution imagery.

This motivation is now emphasized in the Introduction section in the revised paper on page 6 line 8-13: "*However, more detailed comparisons between simple one-source contextual-based schemes versus more complex two-source models using high resolution imagery over different surfaces are still needed to fully understand the strengths and weaknesses of both modeling schemes. Such intercomparisons can facilitate development of hybrid schemes that leverage the strengths of different methodologies (e.g., Cammalleri et al., 2012), while incorporating simplications for routine application with airborne imagery.*"

Also, the interest of either approaches in the context of precision agriculture should be further commented: why testing for precision agriculture a model that requires only temperature data? The cost of a met station is much smaller than that of acquiring airborne data, therefore I doubt that the low data requirement of the rectangular approach is a strong advantage for an application at local scale. Moreover, the rectangular approach provides total ET while TSEB targets E and T separately, which brings some added value in terms of water management and precision farming. Can't you evaluate E and T with the rectangular approach using two rectangles instead of one? (i.e. min/max values for soil and vegetated pixel groups separately ?)

Reply: We agree with the reviewer that the low data requirement of the DATTUTDUT model is indeed a strong advantage for local application, especially for near real time ET monitoring of agricultural fields without user expertise being required for specifying any model input variables. That is why we choose to test DATTUTDUT model in this paper. It is a very novel approach being recommended to model E and T with two separate rectangles (one for vegetation and the other for bare soil/cover crop), so that a simple rectangular approach can also obtain E and T more simply than TSEB. However, this requires having pure vegetation and soil temperature pixels, which there are very few such pure pixels with the current thermal imagery even at the finer pixel resolutions.

While it may be possible to estimate T_s and T_c using the visible bands which are at ~0.1 m resolution by sharpening the thermal images just like what is done in this study to compare with modeled T_s and T_c from TSEB in Section 4.1. These estimates only exist for some areas within the vineyards that are higher factional cover (greater LAI) where pure vegetated and bare soil/cover crop temperature pixels can be reliably extracted. Thus only a fraction of the north and south fields will have ET or E and T estimates using the rectangular approach.

The choice of the contextual model used in the study should be better commented: contextual models take advantage of the various land surface elements within a given landscape. Contextual models are meant to be applied for heterogeneous sites, here the area of interest is small, it is thus expected that contextual models won't perform very well (cf. your comment p11911L20). Here, there are essentially 2 fields, the full triangle or trapezoid method is therefore not necessary, and the rectangular model is preferred, with extreme values being mostly related to the irrigated vegetation and the dry inter-row bare soil. This could be expanded in the introduction to legitimate the methodology. One wonders what causes water status heterogeneity within the vineyards. Some indications about the irrigation system and scheduling are missing here.

Reply: The climate in the study area is arid during much of the vine growing season, and the vine vegetation relies on the drip irrigation to provide water to the root zone to meet atmospheric demand. In the vineyard, the amount of irrigated water infiltrates into the root zone varies as a function of the application rate and soil water holding capacity. In some areas in the vineyard excess irrigated water is transported laterally into the inter-row which results in re-growth of the cover crop. However, generally for most of the growing season the inter-row area is very dry consisting of bare soil and dry senescent cover crop stubble (usually mowed in mid to late Spring). Thus there is often a well irrigated vine crop alongside a very dry non-irrigated inter-row which permits the existence of a wet and dry end-member within the small area.

The sensitivity tests are interesting and should be better presented (at least explicitly) in the text and not only in Table 4.

Reply: The details about the sensitivity tests were presented in Section 4.3. In this Section, the first paragraph introduced the setting of the sensitivity tests, i.e., varying T_R for both two models, and varying the end member, study domain and resolution for DATTUTDUT. The second paragraph explicitly discusses the results from the sensitivity study based on the inputs and statistics reported in Table 6 and illustrated in Figure 10. Finally the last paragraph relates the results in this paper to previous findings in the literature.

Minor comments:

- P11907L28: the "rectangular" model (DATTUTDUT) is only insensitive to systematic errors, please correct.

Reply: In the revised paper, "systematic errors" is emphasized on page 2 line 17, thank you for the correction.

- P11910L18: it seems that the applicability of contextual methods for submeter data is more related to the size of the images analyses (I.e. where the extrememe pixels are identified) rather than the resolution itself. Please precise.

Reply: The size of the images had been stressed as the main reason in the revised version on page 4 line 22-25: "*With UAV imagery, the pixel resolution can be very fine (i.e., 10^0 cm – 10^0 m) in order to map the variability in crop condition within a field. This typically restricts the size of the area or field being monitored and hence reduces the likelihood of sampling the extremes in ET rates (i.e., ET ~0 and ET at potential).*"

- P11916L24: why didn't you use solar radiation as the main scaling factor?

Reply: Discussion referring to the method used to extrapolate instantaneous to daily ET is now included on page 10, line 19-22: "*In this study, the observed available energy* *from the two flux towers during the daytime period for all five days was used to extrapolate instantaneous model estimates to daytime ET totals. However, in practice tower measurements of A would not be available, so results using solar radiation to extrapolate to daytime ET will also be evaluated.*"

We computed daytime ET using solar radiation with the EF expressed as follows:

$$EF=\frac{LE_{ins}}{Rsolar_{ins}}=\frac{LE_{daytime}}{Rsolar_{daytime}} \quad (R1)$$

and the comparison between two methods were included in the revised manuscript (see Table 4). In Table 4, the MAE values for daytime ET using *A* versus S_d methods were similar for TSEB, but were larger for DATTUTDUT particularly at site 1 the north vineyard. A discussion of the results in Table 4 is on page 17 line 12-19: "*In practice, we will not have observations of available energy, A, from a flux tower for extrapolating the instantaneous ET from a single airborne observation to daytime ET, but instead are more likely to have weather station observations of incoming solar radiation, Sd. Results using Sd for extrapolating model estimates instead of flux tower measurements of A are listed in Table 4. In general, the differences between modeled and measured daytime ET (using RE method) increase, although not significantly for TSEB. On the other hand, discrepancies with the ET measurements for DATTUTDUT at the north vineyard (site 1) increase dramatically due to the large overestimation of instantaneous LE on DOY 162 and 219 (see Fig. 6b).*"

- P11923L3: you could mention more recent work by Colaizzi et al.

Reply: In Colaizzi et al. (2012), RMSD values comparing observed versus modeled T_c varied from 0.83-3.1 ℃. For T_s, RMSD values were from 2.4 to 5.0 ℃. These results are described on page 16 line 2-5: "*This accuracy was comparable with similar types of comparisons reported by Li et al. (2005), Kustas and Norman (1999, 2000), and Colaizzi et al. (2012a) which had RMSD values ranging from 2.4-5.0 ℃ for Ts and 0.83-6.4 ℃ for Tc when comparing observed to TSEB-derived component temperatures.*"

- P11924L13: It's probable that using the true solar radiation instead of the DATTUTDUT estimate is the prime source of error, please investigate or comment.

Reply: In the revised paper, DATTUTDUT is run using Rn and G from TSEB, and the results are shown in Table 5 with related discussion on page 18 line 21-28: "*Using measured Sd from the towers instead of computing from the sun-earth astronomical relationships routinely applied by DATTUTDUT, there is only a minor reduction in the differences with the tower fluxes. An overall improvement in DATTUTDUT estimation of LE is achieved by adopting TSEB estimates of Rn and G (see Table 5). This is particularly true for the north vineyard (site 1). However, even with this better agreement in estimated LE, the discrepancies with observed LE from DATTUTDUT is still larger than with the output of TSEB. This indicates that the errors in available energy using the DATTUTDUT formulations are not the only significant source of error in estimating the LE flux.*"

On average (the average value of Site 1 and 2) MAE and RMSD for LE_{RE} are reduced from ~65 and 84 W m^{-2} to ~52 and 68 W m^{-2}, which are still larger than ~37 and 44 W m^{-2} from TSEB. Therefore, using Rn and G from TSEB does have some value, however there still other source of error for DATTUTDUT. Future applications will try to improve the Rn-G algorithm in DATTUTDAT.

- Table 4: provide at least one performance meter (e.g. RMSD) for each case.

Reply: RMSD values for the daytime ET is now included in Table 6 (it was in Table 4 in the original manuscript), RMSD is calculated as:

$$\text{RMSD} = \left[\frac{(O_1 - M_1)^2 + (O_2 - M_2)^2}{2} \right] \qquad \text{(R2)}$$

where O1 and O2 are observed daytime ET at Site 1 and 2, M1 and M2 are modeled daytime ET at Site 1 and 2.

- P11929L10: provide an a priori estimate from the FAO56 method to show the added value of the thermal data compared to a classical crop coefficient method.

Reply: The crop coefficient (Kc) in calculating actual ET from the FAO56 method for vines will vary depending on numerous factors including vine type, growth stage/phenology, fractional vegetation cover, and soil water availability. The nominal Kc values recommended in FAO56 for vines are 0.30, 0.70 and 0.45 for initial, middle and late growing periods, respectively (See page 112 in FAO 56 manual). However, it was decided that given the lack of more detailed information about the Kc values for vineyards, it would be more interesting to evaluate Kc using model and measured ET.

Detailed discussion about the calculation of actual ET using FAO 56 are now provided on page 24 line 16-30 and page 25 line 1-13: "*Current operational techniques for estimating water use of crops primarily relies on the crop coefficient technique based on the FAO 56 publication (Allen et al., 1998). The actual ET of the crop is estimated by first computing a reference ET (ET_0) which is then multiplied by the crop coefficient (K_C). This single crop coefficient is often divided (called the dual crop coefficient) into a basal crop coefficient (K_{Cb}), which is associated with the crop transpiration and has been related to remotely sensed vegetation indicies (Neale et al., 1989) and a soil surface evaporation coefficient (K_e). There is also included a Ks coefficient to reduce crop transpiration for a deficit in water availability in the root zone so the expression has the form $ET=(K_{Cb}K_s + K_e)ET_0$. Determining K_e and K_s requires running a soil water balance model for the surface and root zone. A recent application of this methodology over corn and soybean croplands is given by Gonzalez-Dugo and Mateos (2008) where they find this reflectance-based crop coefficient technique can significantly overestimate ET during a prolonged dry down period. There also appears to be no consistent or universal relationship between crop coefficients and vegetation indices and so this approach is not readily transferable to different crops and climatic conditions (Gonzalez-Dugo et al., 2009).*

As an example, the spatial distribution of K_C was computed using FAO 56 estimated ET_0 and the ET map from TSEB from DOY 163 (Fig. 13). There is a *significant spatial variation in K_C due in part to the know effect of leaf area/fractional cover (Choudhury et al., 1994), which is seen in the correlation between the K_C map and LAI map of Fig. 4, but there are other factors including the vine variety and possibility of some level of stress in areas of the vineyard that cannot be reliably detected by this approach. Using the ET measurements from the flux towers and FAO 56 estimated ET_0, for the north vineyard site 1, the value of K_C ranged from 0.55 for DOY 100 to 0.76-0.82 for the other days. For the south vineyard (site 2), K_C values ranged from 0.59 for DOY 100 to 0.62-0.65 for the other days, indicating little variation in K_C with vine phenology. In contrast, the FAO 56 manual recommends K_C values for vineyards at early, peak and end of the growing season of 0.3, 0.7 and 0.45. Clearly, a calibration with this approach is required, which is not only dependent on vine variety but also on vine management (i.e., row orientation and spacing, pruning, irrigation scheduling, etc.)*"

- P11932-33: many typos errors, please review carefully.

Reply: We have checked the revised manuscript carefully before submitting it to HESS editorial system this time.

- P11932L26: in the rectangular approach, the max. LE corresponds to H=0 or LE=Rn-G so LE is different from a potential rate classically computed with a Penman-Monteith equation.

Reply: By "potential rate" in this paper, we actually mean maximum possible LE based on the surface energy balance LE=Rn-G, but not potential ET from the P-M equation. The "potential" was replaced by "maximum" ET in the sentence on page 26 line 16-21: *"The performance of DATTUTDUT model in computing reliable ET and generating distributions and patterns over the vineyards was similar to TSEB on some of the overpass dates, but for other times the DATTUTDUT model performance was less than satisfactory largely depending on whether there actually existed pixels in the scene that* *were representative of the extreme ET conditions, namely "maximum" ET (LE=Rn-G) and no ET (LE= 0)."*

Interactive comment on "Mapping evapotranspiration with high resolution aircraft imagery over vineyards using one and two source modeling schemes" by T. Xia et al.

Anonymous Referee #3

Received and published: 5 January 2016

In this paper, the comparison of two SEB approaches of different complexity based on airborne TIR observations over irrigated vineyards is carried out with a rigorous approach; significant details of the elaborations performed are given and the paper is generally more informative for the reader than other similar ones.

Reply: Thank you for your positive comments.

It should be noticed that the main concept (and the core) of DATTUTDUT model has been already published by Roerink et al., in Phys. Chem .Earth, Vol. 25 (2):147-157). This latter reference is not present in the paper, but it has been given instead in Timmermans et al. (2015), where the only addition is a simple definition of radiometric temperature end-members in order to easily extract them from the image.

Reply: The reference to Roerink et al. (2000) describing S-SEBI is sited in Introduction and Model overview Sections. The similarity and difference between two models are emphasized in the revised paper on page 5 line 11-13: "*The main concept of DATTUTDUT is similar to the S-SEBI (the Simplified Surface Energy Balance Index) proposed by Roerink et al. (2000); however, DATTUDDUT has a more simplified scheme to obtain radiometric temperature end-members and radiation-related factors.*"

Thanks for reminding!

The sensitivity tests presented for both models highlight some interesting feature of both models. According to the results presented in the Table 4, an uncertainty of +/-3°C is not acceptable in TSEB. However, in the present paper no atmospheric correction has been applied to airborne TIR data (P11919-L18), diversely from VIS-NIR data. Atmospheric effects on radiometric temperature are certainly in the order of magnitude of 2-5°C, but the authors do not comment on this issue, which is quite relevant for the correct application of TSEB in general. To this extent, it might still be useful to further explore the possibility of introducing contextual (image-based) information in the TSEB model, similarly to the approach proposed by Cammalleri et al. (2012, Remote Sensing of Environment, 124: 502–515).

Reply: In this study, the atmospheric effects appear to be insignificant since the aircraft altitude is less than 500 m above ground level. Ground-based T_R (calculated from upward longwave radiation measurements) was compared with T_R from aircraft imagery collected during IOP 3 in Table R1 (ground-based T_R was missing for the other IOPs). On DOY 218, T_R from aircraft was close to ground-based T_R (difference <1 °C). But on DOY 219, aircraft T_R was ~3 °C higher than ground-based T_R. In general, the atmospheric attenuation tends to reduce T_R observed at the sensor altitude, so aircraft T_R will increased after atmospheric correction is applied. The fact that no correction was applied suggests that the bias in aircraft T_R is due to sensor calibration. Therefore, the atmospheric affects are probably not significantly influencing the aircraft-based T_R observation.

Table R1. Comparison of ground-based T_R versus T_R from aircraft imagery from IOP 3 (°C).

DOY	Site	Ground-based T_R	Aircraft T_R
218	1	28.04	27.3
	2	30.06	30.9
219	1	27.29	29.8
	2	29.17	32.5

In the final part of the text, the latent heat flux is used for calculating the water consumption at plot scale, with the aim of emphasizing the impact of TIR observations in operational water management. This part raises some questions. Indeed, the proposed approaches (and the description given in the paper) do not give most relevant information on irrigation scheduling (i.e. occurrence of water stress, soil water deficit) but just a "one-shot" picture on the day of observation. It would have been interesting to highlight which threshold values of the evaporative fraction could be considered as an indication of crop water stress conditions, or to which extent the crop water requirements are met (accordingly to the "standard conditions" defined by FAO56). This element would have improved the paper rather than the simple water consumption calculation.

Reply: This reviewer makes a good point concerning vineyard water consumption and water stress. However the level of stress experienced by vine plants is likely to be dependent on a number of factors, most notably the vine variety and phenology. In addition, the estimated ET is comprised of transpiration and evaporation from the vine row and inter-row systems and so determining vine stress requires reliably partitioning the ET and T and E from these two systems. This is not easily validated without canopy level measurements of leaf conductance, water potential and photosynthesis to understand the relationship between vine T and stress.

Adopting reference ET from FAO56 as an indication of crop water stress condition is not advisable without *a priori* detailed information concerning the behavior of the crop coefficient for this particular vine variety. A detailed discussion about the calculation of actual ET and issues in using FAO 56 crop coefficient approach is now included on page 24 line 16-30 and page 25 line 1-13: "*Current operational techniques for estimating water use of crops primarily relies on the crop coefficient technique based on the FAO 56 publication (Allen et al., 1998). The actual ET of the crop is estimated by first computing a reference ET (ET_0) which is then multiplied by the crop coefficient (K_C). This single crop coefficient is often divided (called the dual crop* *coefficient) into a basal crop coefficient (K_{Cb}), which is associated with the crop transpiration and has been related to remotely sensed vegetation indicies (Neale et al., 1989) and a soil surface evaporation coefficient (K_e). There is also included a Ks coefficient to reduce crop transpiration for a deficit in water availability in the root zone so the expression has the form $ET=(K_{Cb}K_s + K_e)ET_0$. Determining K_e and K_s requires running a soil water balance model for the surface and root zone. A recent application of this methodology over corn and soybean croplands is given by Gonzalez-Dugo and Mateos (2008) where they find this reflectance-based crop coefficient technique can significantly overestimate ET during a prolonged dry down period. There also appears to be no consistent or universal relationship between crop coefficients and vegetation indices and so this approach is not readily transferable to different crops and climatic conditions (Gonzalez-Dugo et al., 2009).*

As an example, the spatial distribution of K_C was computed using FAO 56 estimated ET_0 and the ET map from TSEB from DOY 163 (Fig. 13). There is a significant spatial variation in K_C due in part to the know effect of leaf area/fractional cover (Choudhury et al., 1994), which is seen in the correlation between the K_C map and LAI map of Fig. 4, but there are other factors including the vine variety and possibility of some level of stress in areas of the vineyard that cannot be reliably detected by this approach. Using the ET measurements from the flux towers and FAO 56 estimated ET_0, for the north vineyard site 1, the value of K_C ranged from 0.55 for DOY 100 to 0.76-0.82 for the other days. For the south vineyard (site 2), K_C values ranged from 0.59 for DOY 100 to 0.62-0.65 for the other days, indicating little variation in K_C with vine phenology. In contrast, the FAO 56 manual recommends K_C values for vineyards at early, peak and end of the growing season of 0.3, 0.7 and 0.45. Clearly, a calibration with this approach is required, which is not only dependent on vine variety but also on vine management (i.e., row orientation and spacing, pruning, irrigation scheduling, etc.)"

Some other specific comments:

- It would be useful to give some comments about the influence of the flight acquisition time on the results. Were the flights time fixed in coincidence of Landsat overpass or there were other reasons?

Reply: Yes, an attempt was made to have at least one aircraft flight during each IOP \ center around Landsat overpass time in order to compare high resolution imagery from the airborne data with the courser resolution Landsat imagery. Since the visible and near-infrared reflectance bands on Landsat were used to develop a relationship with the aircraft sensor DN values this was very useful

- Why different equations are given for the LAI(NDVI) relationship on DOYs 163 and 218?

Reply: It was determined that the DN~reflectance calibration of the aircraft NIR sensor was not constant over all IOPs which may have contributed to this changing relationship, but also vineyard management (i.e., vine training) likely altered the structure of the vine canopy between the two IOPs, thus affecting the LAI distribution and hence the NDVI-LAI relationship.