RESPONSES TO THE REFEREES

We thank the reviewers for the comments. Below are our responses (in **blue** font) to the reviewers' comments and questions (in **black** font).

Anonymous Referee #1

Received and published: 9 February 2016

General Comments:

This manuscript compares two different approaches to retrieve energy budget components (including sensible and latent heat flux) at the land surface using satellite data from the Chinese HJ-1B. One approach (IPUS) uses information aggregated to the 300m resolution as given by the thermal channel; the second approach (TSFA) uses a temperature sharpening approach, making use of a NDVI – TS relationship and downscaling 300m Ts information to the 30m scale. Authors illustrate the differences between both approaches and demonstrate within a validation exercise the advantages and improved prediction capacities of the latter approach. I think this comparison and the results obtained are in principle worth publishing and will be of use for the readership of HESS. However, before a possible publication, author need to address and solve some significant concerns and questions that came up when working through the manuscript.

1. One of the major deficits of the manuscript is the following: The satellite data available are the 30m resolution data in the VIS/NIR spectral region and the 300m thermal information. As a "standard/normal" remote sensing user, I would try to make use of this available information. That means in a "reference application" (as it appears to me the IPUS scheme is meant to be) I would try to make use of the 30m data to derive NDVI and land use information (and all the other relevant parameters such as vegetation height, vegetation cover, roughness length etc., but also the simplification of individual fluxes for given LU type). Why are these parameters aggregated in the IPUS approach? Why don't use the high resolution information with an aggregated 300m Ts-signal. This should be compared to the TSFA approach in order to be able to evaluate the effect of purely temperature sharpening. Here actually the baseline situation is first worsened by aggregating information that is available in much higher resolution. In case the intention of the authors was to show what can happen when also in the VIS/NIR range only 300m resolution data were available, then all the (300m) average land surface parameters should have been derived from the aggregated reflectance information. So, I personally feel here are different aspects mixed and not properly separated.

Response: Thank you for your valuable comments. We assume that the ET estimation errors mainly come from inhomogeneity of surface landscape and variables. We aimed to reduce the uncertainties in estimated ET that caused by surface heterogeneities, and the TSFA is our final scheme. To evaluate the ability of the TSFA method to capture surface heterogeneity and reveal the scale effect, we used the IPUS method because it does not consider the effects of mixed pixels at all. According to your comments, we added the TRFA (temperature resampling and flux aggregation) method, which uses 30 m visible/near infrared band data with 300 m thermal infrared band data to estimate ET. In this method, simple spatial resampling (300 m to 30 m) of LST was used instead of spatial sharpening according to NDVI information. Comparisons between the TFSA and TRFA methods can be used to evaluate the effects of temperature sharpening on estimating ET, as well as the significance of separating inhomogeneity of landscape from that of surface variables (such as LST), and that would make our logic clearer.

2. The title of the manuscript suggests that the focus of the paper is on evapotranspiration – when looking through the manuscript and figures and tables, it seems to me that sensible heat flux is dominating the content and discussion. As a result, I would suggest to either change the title or put some more emphasis on ET in the presentation and discussion of results. As a result of my evaluation I would suggest major revisions of the manuscript before a possible publication in HESS.
Response: Thank you for your suggestion. We revised the manuscript by placing more emphasis on LE with a balance analysis and discussion of H.

Specific Comments/Questions

- In general, there are a very large number of abbreviations used in the manuscript – not all of them are intuitive and it is painful to always try and find the first position where they are explained. So I would suggest generating a list of abbreviations.
Response: Thank you for your suggestion. A table of abbreviations and the usage of input data were added in the appendix.

- Figure and table legends are not self-explaining throughout the manuscript and need extension!
Response: The figure and table legends were revised.

P2L14-22: While this paragraph is ok in principle, we as hydrologist all know how important ET – so in order to come quicker to the point it should be omitted.
Response: Thank you for your suggestion. We agree with your suggestion and have deleted this paragraph.

P6L17: Why choosing the 25% fractions having the lowest CV? Please explain in the text!
Response: We have added our justification for this decision in the manuscript.
According to the temperature sharpening method "DisTrad" proposed by Kustas et al. (2003), 25% of the pure pixels with the lowest CV are selected from each class. Regarding heterogeneity, lower CVs correspond with more homogeneous land surfaces. In addition, a fraction should guarantee that a sufficient number of pixels was obtained to fit a least-squares expression between $NDVI_{300}$ and T_{300}; thus, we choose 25% of the fractions with the lowest CVs.

P9L11: How is Ld calculated in the scheme?
Response: The Ld calculation method was introduced in Section 3.2.1.1. (P14L10).
Reference: Yu, S., Xin, X., and Liu, Q.: Estimation of clear-sky longwave downward radiation from HJ-1B thermal data, Sci. China Earth Sci., 56, 829-842, 10.1007/s11430-012-4507-z, 2013.

P12L21ff: It remains unclear how albedo is calculated
Response: The expression of the surface albedo computing method was modified as follows:
"According to the algorithm proposed by Liang et al. (2005) and Q. Liu et al. (2011), surface albedo

was obtained from the top of the atmosphere (TOA) reflectance by the HJ-1 satellite with a lookup table based on an angular bin regression relationship. The surface albedo and bidirectional reflectance distribution function (BRDF) of the HJ-1 satellite in the regression procedure were monitored by using POLDER-3/PARASOL BRDF datasets, and BRDF was used to obtain the TOA reflectance using the 6S (Second Simulation of a Satellite Signal in the Solar Spectrum) radiation transfer mode."

Reference: Liang, S., Stroeve, J., and Box, J. E.: Mapping daily snow/ice shortwave broadband albedo from Moderate Resolution Imaging Spectroradiometer (MODIS): The improved direct retrieval algorithm and validation with Greenland in situ measurement, Journal of Geophysical Research: Atmospheres, 110, D10109, 10.1029/2004JD005493, 2005.

Liu, Q., Qu, Y., Wang, L. Z., Liu, N. F., and Liang, S. L.: Glass-Global Land Surface Broadband Albedo Product: Algorithm Theoretical Basis Document. Version, 1, 1-50, College of Global Change and Earth System Science, Beijing Norman University, 2011.

P14L11: briefly describe how this is expressed/described (Ref)

Response: The top-of-atmosphere (TOA) brightness temperature of the HJ-1B thermal channel was used as the atmospheric effective temperature.

As shown in Yu et al. (2013), "To investigate the relation between TOA brightness temperature of the HJ-1B thermal channel and near-surface air temperature, TOA brightness temperature of HJ-1B is simulated using the Thermodynamic Initial Guess Retrieval (TIGR) atmospheric profile database TIGR2002 (http://ara.abct.lmd.Polytechnique.fr/index.php?page=tigr) and MODTRAN radiative transfer model; it has high correlation with near-surface air temperature."

Reference: Yu, S., Xin, X., and Liu, Q.: Estimation of clear-sky longwave downward radiation from HJ-1B thermal data, Sci. China Earth Sci., 56, 829-842, 10.1007/s11430-012-4507-z, 2013.

P16L2: What reliable methods? This needs to be more specific and with references.

Response: The methods were added as follows:

"Reliable methods were used to ensure the quality of the turbulent heat flux data. Before the main campaign, an intercomparison of all instruments was conducted in the Gobi Desert (Xu et al., 2013). After basic processing, including spike removal and corrections for density fluctuations (WPL-correction), a four-step procedure (data were rejected when (1) the sensor was malfunctioning, (2) precipitation occurred within 1 h before or after collection, (3) the missing ratio was greater than 3% in the 30-min raw record and (4) the friction velocity was below 0.1 ms-1 at night) was performed to control the quality of the EC data, and EC outputs were available every 30 min (for more details see Liu et al., 2011; Xu et al., 2013)."

P17ff: In the section 4.1 surface parameter and fluxes derived are evaluated against measurements. In order to put those results into a general context I think a discussion and comparison in relation to other international Remote sensing/Flux measurement campaigns should be given.

Response: Thank you for your comments. A discussion of the other ground campaigns was given.

Our surface variable retrieval methods were validated against other areas considered in remote sensing measurement campaigns. For example, the albedo algorithm was previously applied to retrieve Global Land Surface Satellite (GLASS) Products (Liang et al., 2014), the LST retrieval algorithm was validated in the Haihe River Basin in northern China (Li et al., 2011), and the soil heat flux correction algorithm was validated in the GAME-Tibet campaign (Yang and Wang, 2008). The Heihe River Basin has long served as a test bed for integrated watershed studies as well as land surface or hydrological experiments. Comprehensive experiments, such as Watershed Allied Telemetry Experimental Research (WATER) (Li et al., 2009), and an international experiment - the Heihe Basin Field Experiment (HEIFE) in World Climate Research Programme (WCRP) have taken place in the Heihe River Basin. One major objective of HiWATER is to capture the strong land surface heterogeneities and associated uncertainties within a watershed (Li et al., 2013). Because the surface of the Heihe River Basin is extreme heterogeneous, additional comparisons of our algorithm in other areas of research would be better.

References:

Hu Y Q, Gao Y X, Wang J M, Ji G L, Shen Z B, Chen L C, Chen J Y and Li S Q: Some achievements in scientific research during HEIFE, Plateau Meteorology, (03), 2-13, 1994.

Li, H., Liu, Q., Jiang, J., Wang, H., and Sun, L.: Validation of the land surface temperature derived from HJ-1B/IRS data with ground measurements, Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International, Vancouver, Canada, 293-296, 2011.

Li, X., Cheng, G. D., Liu, S. M., Xiao, Q., Ma, M. G., Jin, R., Che, T., Liu, Q. H., Wang, W. Z., Qi, Y., Wen, J. G., Li, H. Y., Zhu, G. F., Guo, J. W., Ran, Y. H., Wang, S. G., Zhu, Z. L., Zhou, J., Hu, X. L., and Xu, Z. W.: Heihe Watershed Allied Telemetry Experimental Research (HiWATER): Scientific Objectives and Experimental Design, Bulletin of the American Meteorological Society, 94, 1145-1160, 10.1175/BAMS-D-12-00154.1, 2013.

Liang, S. L., Zhang, X. T., Xiao, Z. Q., Cheng, J., Liu , Q., and Zhao, X.: Global LAnd Surface Satellite (GLASS) Products: Algorithms, Validation and Analysis, 1 ed., SpringerBriefs in Earth Sciences, Springer International Publishing, 2014.

Yang, K., and Wang, J.: A temperature prediction-correction method for estimating surface soil heat flux from soil temperature and moisture data, Sci. China Ser. D-Earth Sci., 51, 721-729, 10.1007/s11430-008-0036-1, 2008.

P20L9-10: This statement about errors is not very specific!

Response: We deleted this statement because it was repetitive with the sentence preceding it.

P22L5: How do you justify a ground heat flux of 0 for buildings?

Response: In our study area, ‘buildings’ contain residents and roads. Influenced by local climate situation, special materials with low heat conductance are used for residential buildings to maintain cool conditions during the summer and warm conditions during the winter. Thus, we justified using a ground heat flux of 0 for buildings. According to your comments, the buildings of these residents were not prevalent. Thus, we recalculated all the data, and $G = 0.4R_n$ for buildings during the summer (Kato, 2005).

Reference: Kato, S., and Yamaguchi, Y.: Analysis of urban heat-island effect using ASTER and ETM+ Data: Separation of anthropogenic heat discharge and natural heat radiation from sensible heat flux, Remote Sensing of Environment, 99, 44-54, http://dx.doi.org/10.1016/j.rse.2005.04.026, 2005.

P23L22ff: This statement is actually a result of what is summarized under point 1 in the general conclusion.

Response: We did not observe any relations between the statement at P23L22 and the general conclusion.

P24L25ff: Why do you use these specific day for calculating the sensitivities! In fig. 12 the x-axis shows variations in %. This makes it difficult to follow the interpretations of the curves in the section.

Response: Sensitivity analysis is a general mathematic analysis procedure, and the presented input data for a specific day can indicate the influences of surface variables. We calculated the sensitivity results using large amounts of data from different phenophases, and our input data illustrated the sensitivities of our ET algorithm. To make the sensitivity analysis results universal, we drew a figure with % on the x-axis. We revised the paragraph and the x-axis in Figure 12 (Figure 13 in latest revised version) to make it easier to understand, especially for LST, because the discussed manuscript did not describe the x axis of LST variation clear, as follows:

"Because LE is calculated as a residual item in energy balance equations, the sensitivity of H is analyzed first. Land surface variables (including LST, LAI, canopy height, and FVC) and meteorological variables (including wind speed, air temperature, air pressure and relative humidity) are needed to estimate H in this paper. To locate the error source when retrieving H, a sensitivity analysis was performed by adding errors at each 10% step (except LST). Fig. 13 presents the sensitivity analysis results: LST = 303.9 K (ranging from 298.4~309.4 K with a step size of 0.5 K), LAI=1.4 (ranging 0.14~2.66 with a step size of 0.14), canopy height equals 1 m (ranging 0.1~1.9 m with a step size of 0.1 m), FVC=0.5, wind speed u=2.48 m·s-1, air temperature Ta=297.9 K, air pressure = 97.2 kPa, and RH=40.29%. In addition, the land use type is maize, and the reference H is 230.2 $W \cdot m^{-2}$."

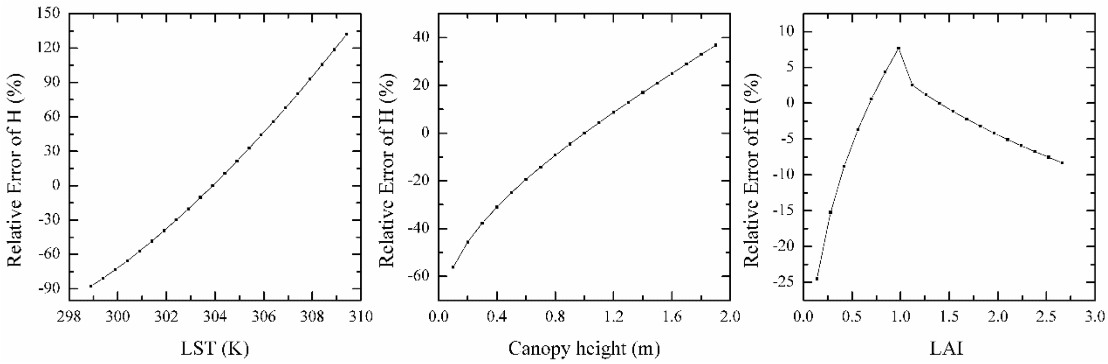

Figure 13. Sensitivity analysis of the surface variables for sensible heat flux

P29L5: Why do authors suddenly come up with the two source model – why didn't they use it initially?

Response: This sentence was deleted because it was not related to the objective of our study.

P30: While the difference between Ts and Taero has been mentioned in the introduction, why isn't that problem discussed here!

Response: Yes, we agree that the difference between Ts and Tareo should be discussed in this paper. "Excess" resistance r_{ex} was added to r_a to correct the discrepancy between Ts and Tareo in most

remotely sensed evapotranspiration models. Thus, the error caused by the difference between Ts and Tareo was shifted to the parameterization scheme error of “excess” resistance, which we discussed in the discussion. We revised this section as follows to clarify the discussion.
“In addition, to correct the discrepancy between remotely sensed radiative surface temperature and aerodynamic temperature at the source of heat transport, a brief and well-performed parameterization scheme (under uniformly flat plant surface) of “excess” resistance was used to calculate the aerodynamic resistance of heat transfer (Jiao et al., 2014). Because the objects of our study are mixed pixels, more parameterization methods should be compared to select the optimum method.”
Reference: Jiao, J. J, Xin, X. Z., Yu S. S., Zhou, T. and Peng, Z. Q.: Estimation of surface energy balance from HJ-1 satellite data. Journal of Remote Sensing, 18(5), 1048-1058, doi:10.11834/jrs.20143322, 2014

P50: Table 13 – there is an error in the definition of the relative error (twice the same expression in the difference)
Response: The mistake regarding the definition of relative error was corrected.

Minor Comments:

P3L4: Surface resistance is also needed for schemes classified under (1) because closure schemes need to calculate H where ra is required as well.
Response: We classified these remotely sensed models to discuss their drawbacks when used for heterogeneous surfaces. In addition, surface resistance is also needed for Penman-Monteith equations. Thus, we do not think surface resistance needs to be classified in this paper because it would disturb the flow of the manuscript and is not a focal point of our study.

P3L20: Which models? All those listed in (1) - (5) or only those in (5)
Response: All those models listed in (1) – (5).

P3L24-25: I do not understand “: : : inhomogeneity is a relative concept of homogeneity: : :!???
Response: We removed this sentence and have revised this paragraph as follows:
“However, heterogeneity is a natural attribute of the Earth’s surface. Studies have shown that different landscapes (Blyth and Harding, 1995; Moran et al., 1997; Bonan et al., 2002; McCabe and Wood, 2006) and the sub-pixel variations of surface variables, such as stomatal conductance (Bin and Roni, 1994), leaf area index (Bonan et al., 1993; Maayar and Chen, 2006), and land surface temperature (Ershadi et al., 2013), can cause large errors in turbulent heat flux estimations. Surface landscape inhomogeneity can be classified using two scenarios: nonlinear vegetation density variations between sub-pixels (e.g., different types of vegetation mixed with each other or with bare soil) and coarse pixels containing total different landscapes (e.g., vegetation or bare soil mixed with buildings or water). In mixed pixels, surface variables such as land surface temperature are set as singular to represent the entire pixel area in ET estimation models.”
References:
Bin, L., and Roni, A.: The Impact of Spatial Variability of Land-Surface Characteristics on Land-Surface Heat Fluxes, Journal of Climate, 7, 527-537, 10.1175/1520-0442(1994)007<0527:TIOSVO>2.0.CO;2,

1994.
Blyth, E. M., and Harding, R. J.: Application of aggregation models to surface heat flux from the Sahelian tiger bush, Agricultural and Forest Meteorology, 72, 213-235, http://dx.doi.org/10.1016/0168-1923(94)02164-F, 1995.
Bonan, G. B., Pollard, D., and Thompson, S. L.: Influence of Subgrid-Scale Heterogeneity in Leaf Area Index, Stomatal Resistance, and Soil Moisture on Grid-Scale Land–Atmosphere Interactions, Journal of Climate, 6, 1882-1897, 10.1175/1520-0442(1993)006<1882:IOSSHI>2.0.CO;2, 1993.
Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W.: Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, Global Biogeochemical Cycles, 16, 5-1-5-23, 10.1029/2000GB001360, 2002.
Ershadi, A., McCabe, M. F., Evans, J. P., and Walker, J. P.: Effects of spatial aggregation on the multi-scale estimation of evapotranspiration, Remote Sensing of Environment, 131, 51-62, http://dx.doi.org/10.1016/j.rse.2012.12.007, 2013.
Maayar, E. M., and Chen, J. M.: Spatial scaling of evapotranspiration as affected by heterogen eities in vegetation, topography, and soil texture, Remote Sensing of Environment, 102, 33-51, http://dx.doi.org/10.1016/j.rse.2006.01.017, 2006.
McCabe, M. F., and Wood, E. F.: Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors, Remote Sensing of Environment, 105, 271-285, 10.1016/j.rse.2006.07.006, 2006.
Moran, M. S., Humes, K. S., and Pinter Jr, P. J.: The scaling characteristics of remotely-sensed variables for sparsely-vegetated heterogeneous landscapes, Journal of Hydrology, 190, 337-362, http://dx.doi.org/10.1016/S0022-1694(96)03133-2, 1997.

P3L26: Density of what?
Response: The density of the vegetation variations. We revised this paragraph for clarity as follows:
“Surface landscape inhomogeneity can be classified using two scenarios: nonlinear vegetation density variations between sub-pixels (e.g., different types of vegetation mixed with each other or with bare soil) and coarse pixels containing total different landscapes (e.g., vegetation or bare soil mixed with buildings or water).”

P4L4ff: I do not understand that sentence/statement!
Response: We have revised this sentence as follows:
“However, it is difficult to develop linear operational models due to the complexity of mass and heat transfer processes between the atmosphere and land surface.”

P13L18: what is <dε> in equation (15)?
Response: <dε> is an effective value of the cavity effect of emissivity and is the mean dε of all vegetation species. In this paper, <dε>=0.015. The definition of <dε> was updated in the manuscript.

P14L1: Sentence (: : : H Li et al : : :) does not make sense.
Response: The names of the authors were located incorrectly due to typesetting. We corrected this problem as follows:
“A single-channel parametric model for retrieving LST based on HJ-1B/IRS TIR data developed by H. Li et al. (2010) was applied.”

P14L7: What is 6SLUT? Reference!

Response: 6SLUT is a look up table that was generated by the 6S (Second Simulation of a Satellite Signal in the Solar Spectrum) radiation transfer mode (Vermote et al., 2006). The following reference was added in the paper.

Reference: Vermote E F, Tanre D, Deuze J L, et al. Second Simulation of a Satellite Signal in the Solar Spectrum-Vector. 6S User Guide Version 3, 2006.

P28L1: Sentence (: : : greatly decreased the heterogeneity) does not make sense

Response: We have corrected this expression as follows:

"The temperature sharpening algorithm in TSFA uses the NDVI at 30 m to monitor the LST at 30 m and is capable of decreasing the influences of the heterogeneity of the LST."