A point-by-point response to the reviews

Firstly, we appreciate very much for the four referees' valuable and constructive comments and suggestions on "Area-averaged evapotranspiration over a heterogeneous land surface: Aggregation of multi-point EC flux measurements with high-resolution land-cover map and footprint analysis" by F. Xu et al. According to their comments, we have carefully revised all sections of the manuscript (revisions and corrections are marked in red). Our detailed responses to their worthwhile comments and suggestions (referee's comments in italics) are as follows:

Response to the comments from Referee #1

Main comments:

1. In Introduction section the authors argued that existing integration schemes often assume local flux measures area representative of an individual surface cover and thus result in errors. However, the paper did not explicitly address the issue with the data they used. The reviewer would like to see clearly to what extent and under what conditions the assumption may produce the error.

Response:

Thanks for your valuable comments. The three aggregation methods, particularly the simple arithmetic and/or the area-weighted method used before, are based on individual surface types, without high resolution land-use classification and fine footprint analysis. We have revised the relevant section in "Introduction" and believe a better result would be achieved based on present integration method.

2. While EC and LAS measures are valuable for large-scale ET estimation, they have measurement errors, either systematic or random. The errors are mixed or propagated into aggregated ET values. Are the measurement errors large or smaller than the ET differences due to spatial heterogeneity? A careful analysis between EC, LAS and spatial heterogeneity with the matrix flux data would provide valuable insights into the issue which is puzzling for many years.

Response:

Yes, both EC and LAS have measurement errors, either systematic or random. For the ECs used in our data analysis (the HiWATER) we have tried to reduce the systematic errors to a minimum with a pre-observation intercomparison (Xu et al., 2013 JGR) and careful maintenance during the observation. The random errors were also analyzed via a separate research (Wang et al., 2015, IEEE GRSL), which can be minimized in an ensemble average. As for the LASs used in HiWATER, major errors are from their data processing processes, for instance, the Bowen-ratio correction problem particularly for observations over Oasis. We have also tried our best to minimize them mainly through intercomparisons with fluxes from EC. All these uncertainties are minor when we compared with the spatial heterogeneity of our study area, especially, the large differences among the four kinds of land-use. We have added some descriptions on data quality and uncertainty of the EC and LAS measurements in our study in Section 2.2.1.

3. Section 4.4 seems not closely relevant to the aggregation topics addressed in the paper. Neither the ET estimates can be validated over the study area as a whole. It is better to remove it from the text.

Response:

We are sorry for our unclear statements in section 4.4. We have removed irrelevant information including the comparison with P-M ET products. The statements in the entire section have been re-written.

Specific comments:

Abstract: Page 1 line 20-23: it does not provide any new for audience.

Response:

Accepted: Line 20-23 are deleted.

1. Introduction

Page 2 line 13: earth - > Earth

Page 3 line 2-3: remove one of "remote sensing"

Page 3 line 8: there is - > there may be

Response:

Accepted.

Page 4 line 9: a nice statement on representativeness of flux measures over individual surface covers. However, the present version failed to explicitly address the assumption in Results and Discussion.

Response:

Thanks for your comments. We have added some statements in Section 4, Results and Discussion, to address the assumption explicitly.

Page 5 line 1-2: "disaggregation approach has not been fully investigated" does not absolutely mean it deserve investigation. Please state it more clearly.

Response:

We have improved the statements in this section more clearly.

2. Study sites and data

Page 6 line 14: two days only? Are they representative or enough to get conclusions that are general?

Response:

The two clear days we selected for analysis, 29 to 30 June 2012, are typical (due to the weather, surface status, and extended observations such as aircraft remote sensing) and representative for the general conclusions we got. Actually, we have applied the same flux aggregation method for other periods, such as the 6 days from 9 July to 14 July 2012. Figure S1 (attached below, similarly hereinafter) shows the daily ET for four land covers derived. Figure S2 describes the area-averaged daily ET over the study area. All the areal ET and its disaggregation to individual land types are similar to those of the two clear days analyzed.

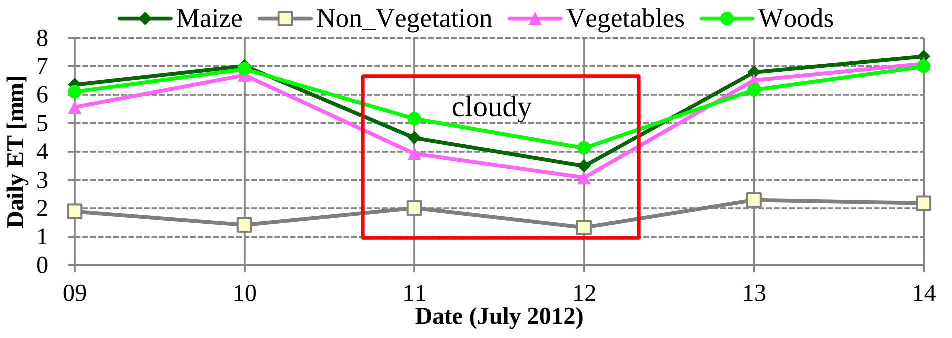

Figure S1: EC dis-aggregated daily ET for each land covers from 9 to 14 July 2012

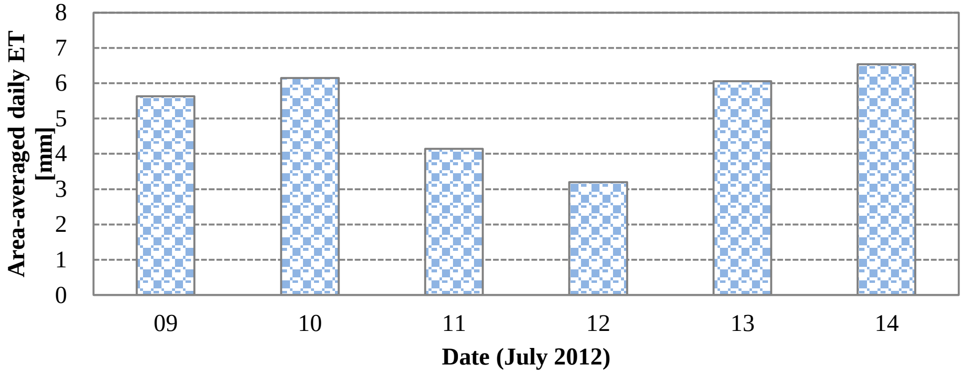

Figure S2: Area-averaged daily ET over the study area

Page 6 line 16: Please state the last time the irrigation done.

Response:

We have added the last irrigation time in the revised manuscript.

Page 7 line 16: It would be better to use local time. Otherwise, explicitly state the time difference to Beijing time.

Response:

The difference between local time and Beijing Standard Time is approximately +1 h 18 min. This has been added explicitly in the revision.

There are many places throughout the paper with mixture use of “remotely sensed”, ”remote-sensing”, ”remote-sensing based”, ”satellite-based”, etc.

Response:

We have unified the use of “remote sensing” throughout the revised paper.

3. Methodology

Page 9 line 8: add a reference here for footprint model.

Response:

A reference has been added:

Leclerc, M. Y., and Foken, T.: Footprints in Micrometeorology and Ecology, Springer, Heidelberg, New York, Dordrecht, London, XIX, 239 pp., 2014

Page 9 line 11: remove “The”. There are many places with misuse of “the” or “a/an”

Response:

Accepted. Other places with misuse of “the” or “a/an” have also been checked.

Page 11 line 10: what the mean of “footprint climatology function”?

Response:

We have revised the “footprint climatology function” as “the weighted footprint climatology”. Its meaning is clearly shown with Eq. (8).

Page12 line 23: framework -> data processing flow.

Response:

Accepted.

4. Results and discussion

Page 12 line 17-19: remove the paragraph. It is useless here.

Response:

Accepted.

Please use W/m^2 instead of mm/day throughout the paper.

Response:

We have unified the flux unit from mm d^{-1} to W m^{-2} in Section 4.1**.** However, in discussing the daily ET in Section 4.4, we still use mm d^{-1} as in usual applications.

Page 13 line 13: how about other period? Do the differences change with different periods?

Response:

The differences among all maize sites change slightly during the maize grown period. For example, Figure S3 shows the daily variation of the three major fluxes from 9 July to 14 July 2012. The standard deviations (SD) of all maize sites are also shown. All features are about the same as the two days analyzed. We have added some relevant statements in the revised paper.

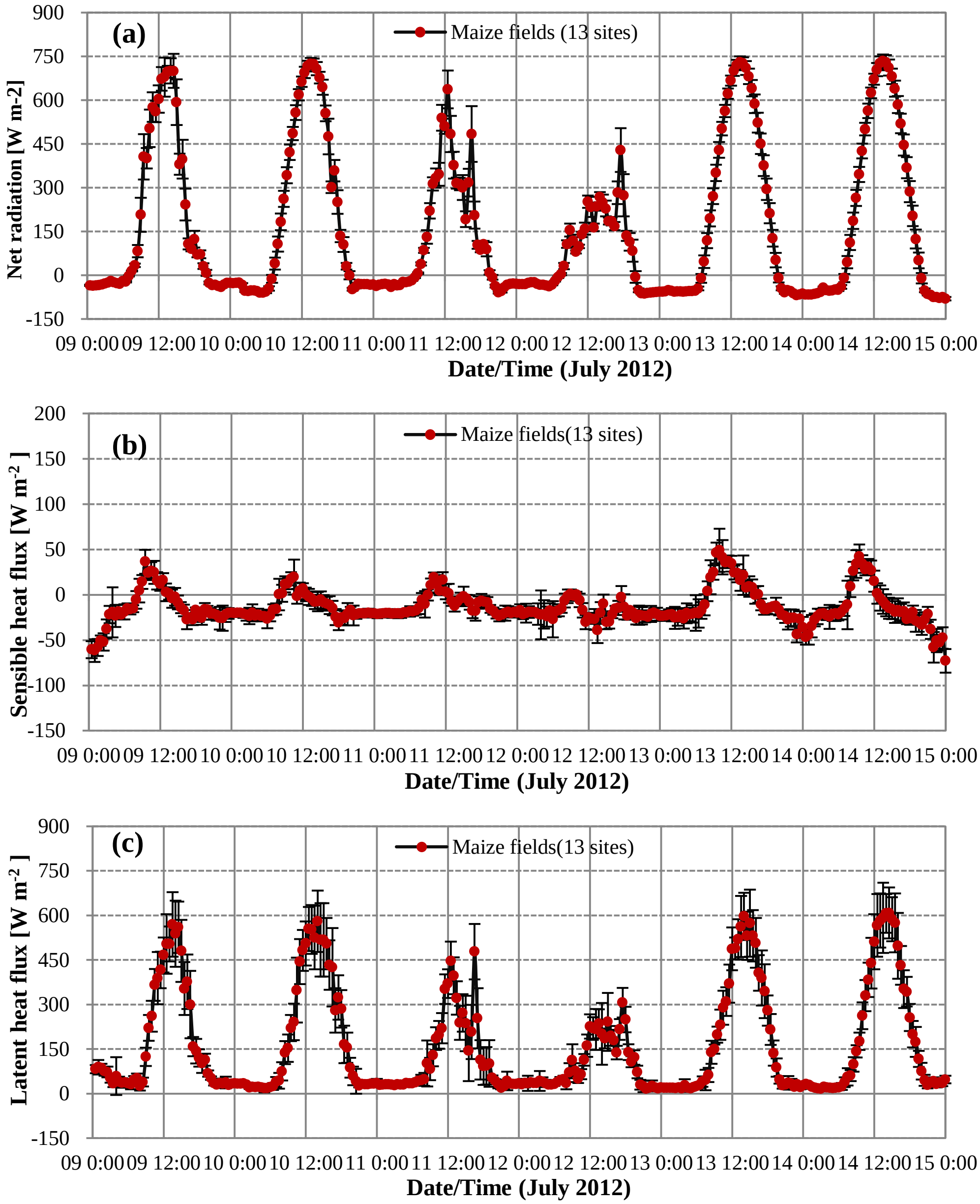

Figure S3: Diurnal cycle of the mean net radiation (a), sensible (b) and latent (c) heat fluxes for 13 maize field sites, the errors bar is the standard deviation.

Page 16 line 16-page 17 line 10: remove the text that describes regional ET over the study area as a whole. It provides no support for the scientific issues addressed.

Response:

The major objective of this study is to refine an aggregation method for area averaged fluxes based on our unique, comprehensive dataset of the HiWATER. The results are also useful for the water balance study extended to the whole Heihe River basin. So the results of regional ET over the study area are still kept in Section 4.4. Of course, some irrelevant parts are deleted according to the comments of yours and other two referees.

Response to the comments from Referee #2

Major comments:

1. While the authors' works are valuable, the size of paper is too large, with containing less-important information. I recommend authors to drop entire section 4.4 and the related descriptions available in other sections (section 3.3 etc.)

Response:

Thanks for your comments. The major objective of this study is to refine an aggregation method for area-averaged fluxes based on our comprehensive dataset of the HiWATER. The results are also useful for the water balance study extended to the whole Heihe River basin. Thus, the results of area-averaged ET over the study area are still kept in Section 4.4. But some irrelevant parts have been deleted, according to the comments from you and other referees.

2. I could not find a purpose of Fig 8, and comparison with "remotely sensed ET data" (Table 5 and 6).

Response:

The purpose of Fig. 8 is to show the spatial pattern of daily ET over the study area for readers. According to comments from yours and other referees, all the relevant statements about the comparison with "remotely sensed ET data" (including Table 5 and 6) and related descriptions in other sections have been deleted.

3. Also, including the comparison with remotely-sensed ET data in this paper might derive another problem on reviewing process because the procedure adopted in the paper is not well described in the paper, and the applied method may not be appropriate.

Response:

Thanks. According to your comments and similar comments from other referees, we have dropped all of the relevant parts on the comparison with remotely-sensed ET data in the revised paper.

Minor comments:

Page 8 Line 10: Authors manually revised land cover map using high-resolution CCD images and Google Earth imagery. Do those images applicable for year 2012?

Response:

Yes. The CCD images were acquired on 26 July 2012, while the Google Earth image used was collected on 3 September 2012. Both are in the HiWATER intensive observation period. We have added the acquisition dates of CDD images and the Google earth image in the revised manuscript.

By reviewing the results, the EC data used in the paper seems to be reliable. However, it is better to describe in the paper some more about the measurement accuracy of their EC data, for example, about the energy closure error.

Response:

Thank you very much for your suggestion. We have added the descriptions on the data quality of EC and LAS used in HiWATER, as well as the energy balance closure rate, in the revised paper (section 2.2.1).

It is author's preference and authors do not need to change, but I might recommend changing Fig 2-b (ET) from bar-graph with mm/d, to line graph with W/m^2 like Fig 2-a, so readers can understand the energy balance condition of the sites by directly comparing Fig 2-a and 2-b.

Response:

Accepted.

It's author's preference and authors do not need to change, but Fig.4 can be expressed not as figure but as table.

Response:

Thanks for your kind suggestion. However, here, a figure is probably more obvious than a table to show the spatial representativeness of all EC sites. So the original figure is still kept.

In Fig. 6, I recommend authors to show the character "a" "b" "c" and "d" in the figure, because authors are referring the figure such as "fig. 6c" in the text.

Response:

Accepted.

Response to the comments from Referee #3

General comments:

1. To address the scientific problem in this paper, 30-min flux might be sufficient, given the uncertainty in gap-filling method *(rainfall, fog etc). Daily ET (section 4.4) does not help a lot here. Indeed, it might be more necessary to clarify data quality and uncertainty of the EC and LAS measurements.*

Response:

Thanks for your comments. Yes, by giving the uncertainty in gap-filling, the 30-min flux might be sufficient to address the scientific problem in this paper. However, the major objective of this study, besides refining the aggregation method for area averaged fluxes (based on our unique and comprehensive dataset of the HiWATER), is also finally useful for the water balance study extended to the whole Heihe River basin. So the area-averaged daily ET over the study area is added for reference. We have also clarified the data quality and uncertainties of the EC and LAS measurements in Section 2.1.1 of the revised paper.

2. Besides, P-M estimated ET could be removed.

Response:

Yes. According to your comments (and those from other referees), we have removed the descriptions relevant to P-M estimated ET, such as those in Section 2.2.2, Section 3.3, and the paragraph on the comparison with P-M ET in Section 4.4.

Specific comments:

Page 5 Line 24 "following Fig.3" Since it is the first figure appearing in this article, it's better to change the number from 3 to 1.

Response:

Accepted.

Page 6 Line 15 "EC data from 16 towers…" According to section 4.2, in addition to site 3, sites 5/8/13/16 were also not used. It is better to use a consistent dataset throughout the paper.

Response:

We are sorry for the unclear statements about the data used in the paper. EC data from 16 towers were all used for analysis. We have improved our statements in Section 4.2 of the revised manuscript.

Page 6 Line 16 "no irrigation" and how was the weather during the period?

Response:

To choose "no irrigation" days is mainly for reducing the effect of local advection. The two days, 29 and 30 June 2012, were typical clear days.

Page 6 Line 22 "coordinate rotation" why not use Planar Fit?

Response:

Our study area, the oasis in the middle reaches of Heihe River basin, is relatively very flat. To use the common 2-D rotation method is not only simpler but also enough in this situation. Actually, we have compared the results of 2-D rotation and Planar Fit during previous works. The differences were very small.

Page 7 Line 13 "MOST" there are different solutions. Add either corresponding equations or references here.

Response:

A reference had been added:

Andreas, E. L.: Estimating Cn^2 over snow and sea ice from meteorological data, JOSA A, 5, 481-495, 1988.

And how were roughness height and zero-plane displacement estimated?

Response:

The roughness height and zero-plane displacement of the study area were obtained following Martano (2000). This has been added in the revised manuscript.

Martano P.: Estimation of Surface Roughness Length and Displacement Height from Single-Level Sonic Anemometer Data, Journal of Applied Meteorology, 39, 708-715, 2000.

Page 7 Line 16 'daytime…' It's a bit confusing. Local time is better.

Response:

We have stated the time difference between Beijing Standard Time and Local time in the revised paper.

For data quality control, what is the threshold value of signal strength?

Response:

For BLS series (BLS4500/BLS900), the threshold value of signal strength is 1000.

Section 2.2.2 This section could be abbreviated if the preliminary land cover has already been done by Liu and Bo (2015).

Response:

Accepted.

Page 8 Line 11, specify the date of the google earth image.

Response:

The Google Earth image used was collected on 3 September 2012. This has been added in the revision.

Page 8 Line 16-20, it might not be necessary to compare with PM-ET. The principle of that is the same as the comparison with LAS in terms of flux aggregation and there might uncertainty in PM-ET.

Response：

As mentioned in the beginning of this reply, we have accepted this comment and made revision.

Section 4.3 Page 15 &16. It's better to look into the details to figure out the factors contributing to the bias between EC and LAS, instead of just mentioning 'heterogeneous distribution of surface covers'.

Response：

Accepted.

Section 4.4 I didn't see the difference between Table 5 and 6 in terms of addressing the problem despite their different units.

Response：

We have removed all the related text on the comparison with PM-ET (including Table 6 and Table 5) in section 4.4 of the revised manuscript.

Response to the comments from Prof. Dr. Thomas Foken (as Referee #4)

Major comments:

1. Reading the manuscript, I found that the concept of the experimental design and the data analysis is very similar to the experiment LITFASS-2003, which was published in BAMS (Mengelkamp et al., 2006) and in a special issue of Boundary-Layer Meteorology (2006, vol. 121, issue 1). Some of these papers are quoted, but papers published later are missing (Foken et al., 2006; Foken et al., 2010; Charuchittipan et al., 2014).

Response：

Thanks. We have added the important references you specified.

2. Several parts in the paper are unclear, or information is missing that would enable the paper to be followed accurately:

(1) The area of investigation was very much dominated by maize fields. Only three stations had another dominant land cover (stations 1, 4, and 17). This is a significant limitation for the stated aim of the paper to determine area-averaged fluxes over a heterogeneous area. For the LITFASS-2003 experiment (and other experiments given as references), different land cover types were much better distributed. This deficit should be discussed.

Response:

As described in Section 2.1, even the dominant land-use type in the intensive observation area was maize field, the surface status of this oasis were actually very heterogeneous. The small square maize fields were all staggered with windbreak trees, roads, irrigation ditches, etc. We have classified four dominant types of the land-cover in the study area. The proportions of each land cover classes were 72 % (maize), 15% (non-vegetation), 8 % (woods) and 5 % (vegetable), respectively.

According to the crop planting structure and land cover, 13 sites were spatially distributed under the dominated maize cropland; while only three stations, namely site 1 (vegetable field) and site 4 (residential area) as well as site 17 (orchard), were separately installed in respectively rather small area of land-use. This has been discussed in more detail in the revised manuscript.

(2) The function of the LAS in the aggregation schema was not clear. I could not find a reason for the use of such data. In LITFASS-2003, LAS systems were also used with a specific function: It was assumed that LAS can also measure the fluxes of larger turbulence or circulation structures and that this is not affected by the non-closure of the energy balance (Foken, 2008). This information was used to discuss the unclosed energy balance of the flux measurements and to correct this. The problem of the unclosed energy balance is not mentioned in the whole paper, but it is a standard for the analysis of surface flux measurements (Foken et al., 2012).

Response:

The LAS measurements for this paper are an intermediate point in checking the established flux aggregation algorithm. The procedure is as follows: the sensible heat fluxes representative for LAS source area were firstly integrated from multiple EC flux measurements, and then compared with the sensible heat fluxes from the 4 paths of LAS systems, to test the reliability of the developed flux integration method. Finally, the latent heat fluxes (daily evapotranspiration) of EC systems were extended to the study area using the aggregation scheme.

The energy balance closure (EBC) is a significant problem we are concerning from the very beginning of HiWATER. Moreover, relevant research has just published in JAMC (Xu et al., 2017). Generally, the energy balance closure ratio (EBR) during the 3 and half months was good. For the 17 EC stations in the intensive observation area, the average EBR was about 0.92. Except the lowest (0.78) in orchard site (#17), values in other sites were scattered without clear relation to the surface status. Site #15 (Super-station) had 2 heights, 4.5 m and 34 m. The relevant EBR were 0.89 & 1.03 respectively. This is quite reasonable.

We have added the detailed description on the EBR for the EC data of the HiWATER flux matrix in Section 2.2.1 and inserted a referenced here. We have discussed the effect of the unclosed energy balance in EC flux measurements on the results of the flux aggregation method in the revised manuscript.

(3) Any information is missing as to why the footprint model by Kormann and Meixner (2001) was used in your study. Perhaps the textbook by Leclerc and Foken (2014) would give you the relevant information. Questionable is the exact location of the small non-maize-covered areas in the footprint of the EC and LAS measurements. A discussion of the accuracy of the footprint analysis combined with the accuracy of the EC and LAS measurements is urgently necessary.

Response:

The advantage of the analytical footprint model by Kormann and Meixner (2001) was referenced in the textbook by Leclerc and Foken (2014). Related descriptions have been added into our revised manuscript. Besides, as we have checked, the footprint estimates of the Kormann and Meixner (2001) were in good agreement with the results of sophisticated backward Lagrangian footprint models, such as the Kljun scheme (Kljun et al., 2002;Kljun et al., 2015). The results from the newest version of Kljun's scheme (October 2016) was used to compare with what from that of Kormann and Meixner (2001). The differences were really minor. We have added some statements and relevant references in the revised paper.

The land-cover map used in the study was initially derived from the aircraft remote sensing images with 1-m spatial resolution, and was then carefully post-processed. Thus, overlapping the accurate 1-m land-cover map with the footprint of EC and LAS with same resolution can determine the location of the small non-vegetation areas in the footprint.

Quality-control and uncertainty-estimation for the EC and LAS data of the HiWATER flux matrix were carefully done. For the EC systems used in the data analysis, we have tried to reduce the systematic errors to a minimum with a pre-observation inter-comparison and careful maintenances during the observation period (Xu et al., 2013). The random errors were also analyzed by a separate research, which can be minimized in an ensemble average (Wang et al., 2015). As for the eddy-covariance systems, flux data from the 4 groups of LAS were also quality controlled. The systematic errors from data processing, e.g. the larger effects of Bowen-ratio correction in this oasis area, were carefully minimized. We checked the sensible heat fluxes (H) from the 4 paths of LAS with that from the nearer ECs. Except LAS 3, under its path there are clearly some village buildings so the H_LAS is higher, others agreed very well with that of ECs. Relevant statements on this have been added into Section 2.1.1.

(4) The applied multiple-linear regression analysis needs more information. Did you aggregate the fluxes according to the land-cover type in different effect levels of the footprint? Compare your method with the methods presented by Leclerc and Foken (2014).

Response:

Yes, we aggregate the fluxes according to the land-cover type in different effect levels of the footprint. We have supplemented some statements on the applied multiple-linear regression method in the study into Section 3.1.

As mentioned above, we have compared carefully the footprint results from Korman and Meixner (2001) with those from Kljun's scheme (Kljun et al. 2002, 2015) to insure the quality of our footprint analysis.

(5) What is meant by "Remotely sensed ET products"? If I understood the paper correctly, only the land-cover type was determined by satellite measurements, but, as seems probable, did these also include the net radiation for use in the Penman-Monteith equation? But this would then be difficult for the heterogeneous land cover. It is impossible to discuss the underestimation of the fluxes by the Penman-Monteith equation without knowing the parameterizations used in this equation. E.g., the atmospheric resistance and the stomata resistance are extremely variable and should be included in any discussion.

Response:

We benefit a lot from your valuable comments. We have removed the parts on comparison with remotely sensed ET products derived by Penman-Monteith equation, according to the comments from you and other referees.

(6) Please also show in Fig. 2 the daily cycle of the evapotranspiration and not only the daily sum. This is necessary to indicate the energy exchange of the different sites, possible oasis effects, and the Bowen ratio. The latter may be a good indicator which to classify the sites.

Response:

Accepted. We have changed the Fig. 2(b) (Fig. 3(b) in revised paper) from bar-graph with mm/d to line graph with W m^{-2}, and also have re-stated the descriptions on the energy exchange of the different sites.

(7) Undoubtedly the authors have an interesting data set with a significant scientific potential. Such a data set should be published with a good scientific concept. Besides some deficits in the experimental design, the concept of area-averaged fluxes may be such a concept. But the paper needs significant improvements according to the points given above. Therefore I recommend major revisions.

Response:

Thanks for your constructive comments.

Minor remarks:

The numbering of the figures is confusing. Figure 3 should be renamed as Fig. 1.

Response:

Accepted.

Table 1: The instrumentation (sonic anemometer, gas analyzer) is missing.

Response:

Accepted.

Table 2: Do not mix LAS type and LAS producer, please give both for all sites.

Response:

Accepted.

p. 6, line 21: What do the flags mean?

Response:

The flag 0, 1 and 2 represent high-quality, intermediate-quality and poor-quality flux data (Mauder and Foken, 2015), respectively. We have added this reference in the revised manuscript.

p. 6, line 23: Why did you use 2D-rotation and not planar fit? Was the terrain absolutely even?

Response:

Our study area, the oasis in the middle reaches of Heihe River basin, is relatively very flat. To use the common 2-D rotation method is not only simpler but also enough in this situation. We have compared the results of 2-D rotation and Planar fit during previous data-processing works. The differences were very small.

p. 7, line 13: L can be easily misinterpreted as the Obukhov length in a micrometeorological paper.

Response:

We have changed the symbol L to R.

Fig. 4 and 5: Why did you use different names or land cover types in both figures?

Response:

We are sorry for our mistake. We have unified the use of land cover types in both figures.

Fig. 6: Probably y has a lower accuracy than given in the figure!

Response:

The different statistics (e.g. the root mean square error, RMSE) between x and y listed in Table 3 were calculated with data shown in the Fig. 7 of the revised manuscript.

p. 16, line 11: The reference should probably be Fig. 3!

Response:

Accepted.

p. 17, line 12: This is trivial; when maize dominates the land cover it is normal that maize also dominates the ET.

Response:

Removed.

Table 6: Give the units in the columns.

Response:

The relevant information on the comparison with P-M estimated ET throughout the paper has already been removed, including Table 6.

p. 19, line 16-25: Such a paper needs a well-written conclusion chapter and not only ten not very significant lines.

Response:

The conclusions have been re-written.

p. 22, line 13: Many authors are missing

Response:

Accepted.

p. 22, line 18: Print CO2.

Response:

Accepted.

References:

Kljun, N., Rotach, M., and Schmid, H.: A three-dimensional backward Lagrangian footprint model for a wide range of boundary-layer stratifications, Boundary-Layer Meteorology, 103, 205-226, 2002.

Kljun, N., Calanca, P., Rotach, M., and Schmid, H.: A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP), Geoscientific Model Development, 8, 3695-3713, 2015.

Kormann, R., and Meixner, F. X.: An analytical footprint model for non-neutral stratification, Boundary-Layer Meteorology, 99, 207-224, 2001.

Leclerc, M. Y., and Foken, T.: Footprints in Micrometeorology and Ecology, Springer, Heidelberg, New York, Dordrecht, London, XIX, 239 pp., 2014.

Mauder, M., and Foken, T.: Documentation and instruction manual of the eddy covariance software package TK3 (update), Arbeitsergebnisse, Universität Bayreuth, Abt. Mikrometeorologie (ISSN 1614-8916), 62, 64 pp, 2015.

Wang, J., Zhuang, J., Wang, W., Liu, S., and Xu, Z.: Assessment of Uncertainties in Eddy Covariance Flux Measurement Based on Intensive Flux Matrix of HiWATER-MUSOEXE, IEEE Geoscience and Remote Sensing Letters, 12, 259-263, 2015.

Xu, Z., Ma, Y., Liu, S., Shi, W., and Wang, J.: Assessment of the Energy balance closure under advective conditions and its impact using remote sensing data, Journal of Applied Meteorology and Climatology, 56, 127-140, 2017.

A list of all relevant changes made in the manuscript

All sections of the manuscript have been carefully checked and revised based on four referees' comments. And all the revisions and corrections are marked in red. The main modification and improvements in the manuscript are listed as follows:

1. In Section 2.2.1, the data quality and uncertainty of the EC and LAS measurements in our study, as well as the energy balance closure (EBC), have been added. Moreover, the statements in this section have been greatly improved.

2. In Section 2.2.2, the relevant descriptions on the collected land cover map have been abbreviated. Figure 3 has been inserted into this section and its number has also been changed from 3 to 1.

3. In Section 3, related descriptions on the advantage of the analytical footprint model have been added to explain our choice of the footprint model. Furthermore, some missing information has also been included. Meanwhile, less-important statements have been removed to make this section more concise.

4. In Section 4.1, the flux unit has been unified from mm d^{-1} to W m^{-2}, and the descriptions on the energy exchange of the different sites have been re-written. Besides, a figure, as Figure 4 of the revised manuscript, has been added into Section 4.1 to explain the differences in sensible and latent heat flux among all maize sites.

5. The statements in Section 4.3 have been carefully checked and improved, and some important information has been added to discuss the the bias between EC and LAS.

6. All of the information on the comparison with remotely-sensed ET data in Section 4.4, as well as the relevant parts in other sections (such as Section 2.2.2, Section 3.3 etc.), have been dropped. But the information on the results of area-averaged ET over our study area has still been kept. Because the results are useful for the water balance study extended to the whole Heihe River basin.

7. According to above revision, the abstract and the conclusions have been re-organized and re-written.

8. The manuscript has been prepared in a format compatible with MS Word, based on the Copernicus Publications Word template. And all the figures in the complete manuscript have been re-produced through professional drawing software.

A marked-up manuscript version

Area-averaged evapotranspiration over a heterogeneous land surface: Aggregation of multi-point EC flux measurements with high-resolution land-cover map and footprint analysis

Feinan Xu[1, 2], Weizhen Wang[1], Jiemin Wang[1], Ziwei Xu[3], Yuan Qi[1], Yueru Wu[1]

[1] Key Laboratory of Remote Sensing of Gansu Province, Heihe Remote Sensing Experimental Research Station, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, 730000, China

[2] University of Chinese Academy of Sciences, Beijing, 100049, China

[3] State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, 100875, China

Correspondence to: Weizhen Wang (weizhen@lzb.ac.cn)

Abstract. The determination of area-averaged evapotranspiration (ET) at the satellite pixel scale/model grid scale over a heterogeneous land surface plays a significant role in developing and improving the parameterization schemes of the remote sensing based ET estimation models and general hydro-meteorological models. The Heihe Watershed Allied Telemetry Experimental Research (HiWATER) flux matrix provided a unique opportunity to build an aggregation scheme for area-averaged fluxes. On the basis of HiWATER flux matrix dataset and high-resolution land-cover map, this study focused on estimating the area-averaged ET over a heterogeneous landscape with footprint analysis and multivariate regression. The procedure is as follows: Firstly, quality-control and uncertainty-estimation for the data of the flux matrix, including 17 eddy-covariance (EC) sites and 4 groups of large aperture scintillometer (LAS), were carefully done. Secondly, the representativeness of each EC site was quantitatively evaluated; footprint analysis was also performed for each LAS path. Thirdly, based on the high-resolution land-cover map derived from aircraft remote sensing, a flux aggregation method was established combining footprint analysis and multiple-linear regression. Then, the area-averaged sensible heat fluxes obtained from the EC flux matrix were validated by the LAS measurements. Finally, the area-averaged ET of the kernel experimental area of HiWATER was estimated. Compared with the formerly used and rather simple approaches, such as the arithmetic average and area-weighted methods etc., present scheme is not only with a much better database but also has a solid grounding in physics and mathematics in the integration of area-averaged fluxes over a heterogeneous surface. Results from this study, both instantaneous and daily ET at the satellite pixel scale, can be used for the validation of relevant remote

sensing models and land surface process models. Furthermore, this work will be extended to the water balance study of the whole Heihe River basin.

1 Introduction

Land surface evapotranspiration (ET) is not only a key component in the regional water circulation, but also essential in the surface energy balance and land surface process. Under the condition of increasing shortage of water resources, high precision estimation of ET at regional scale is essential for such applications, as the management of river basin water resources, regional planning and the sustainable development of agriculture etc. (Wang et al., 2003). Currently, the commonly used methods for acquisition of regional ET are ground-based observation, remote sensing based estimation and model simulation, respectively.

The Earth's surface is always characterized by spatial heterogeneity. Large land surface heterogeneity affects greatly the exchanges of momentum, heat, and water between the land surface and atmosphere (Mengelkamp et al., 2006). Indeed, the surface heterogeneity caused either by the contrast in soil moisture or vegetation type generates a large spatial variability of fluxes, which limit the use of the eddy-covariance (EC) system, unless one deploys a network of EC devices (Ezzahar et al., 2009b). Flux tower group can quantify the turbulent exchange of energy and mass between the atmosphere and a variety of surface types (Sellers et al., 1995), and these local point measurements need to be aggregated to provide a meaningful area averaged fluxes (André et al., 1986). If special aggregation rules for local flux measurements are applied, measurements can provide averaged fluxes at model grid scale (Beyrich et al., 2006; Mahrt et al., 2001). But given the EC network's high price and the requirement for their continuous maintenance, the large aperture scintillometer (LAS) is a useful alternative method for directly measurements of area-averaged sensible heat fluxes in the scale of 1 – 5 km (Ezzahar et al., 2009b; Ezzahar and Chehbouni, 2009).

Satellite has been considered as a promising data source for deriving regional ET with the development of remote sensing technique (Ezzahar et al., 2009a). In response to increasing demand for spatially distributed hydrologic information, many satellite-based approaches have been developed for routine monitoring of ET at a regional scale (Anderson et al., 2012). Nevertheless, the effectiveness of the remote sensing based methods for estimating ET must be fully assessed by ground-based area-averaged flux measurements, due to the uncertainties of model inputs and parameterization schemes etc. (Wang et al., 2003). Furthermore, there may be a bias in directly comparing a remote sensing based ET estimation with in situ measurements, because of their spatial-scale mismatch and spatial heterogeneity at the sub-pixel scale (Jia et al., 2012).

General atmospheric-hydrological models (e.g., Numerical Weather Prediction) can adequately describe the interaction between the atmosphere and the underlying surface using complex parameterization schemes. The development and validation of these models are usually based on measurements performed over homogeneous land surfaces. While the assumption of homogeneity might be justified at the local scale (10 m – 10^3 m), it is often violated at the scale of the grid resolution of current regional atmospheric models (about 10^4 m) (Beyrich et al., 2006; Beyrich and Mengelkamp, 2006). Therefore, it is significantly important to determine the area-averaged surface fluxes at the satellite pixel scale/model grid scale (10^3 m – 10^4 m) for the evaluation of general hydro-meteorological models and relevant remote sensing models.

A number of international field experiments have been performed over heterogeneous land surfaces in different geographical and climate regions of the earth in recent decades (Mengelkamp et al., 2006; Beyrich et al., 2006; Wang, 1999), such as HAPEX–MOBILHY (André et al., 1986), FIFE (Sellers et al., 1988), HAPEX-SAHEL (Goutorbe et al., 1994), BOREAS (Sellers et al., 1995), NOPEX (Halldin et al., 1998), LITFASS-2003 (Mengelkamp et al., 2006), etc. In these experiments, surface fluxes at the model grid scale, estimated from multi-point flux observations using various flux aggregation techniques, were compared with those obtained from LAS systems and remote sensing estimation methods. In former studies, the most commonly used and rather simple flux aggregation methods mainly include: arithmetic average method, the area-weighted method and the footprint-weighted method (Liu et al., 2016). These studies revealed, under careful data-processing and quality-control (Charuchittipan et al., 2014) as well as analysis of the energy balance closure for flux data (Foken et al., 2006; Foken et al., 2010), the integration of the multi-site EC flux measurements and area-averaged fluxes from scintillometers and aircraft observations etc. can provide reasonable estimates over a heterogeneous landscape (Mahrt et al., 2001; Beyrich et al., 2006; Liu et al., 2016).

However, the integration schemes of aforementioned methods are applicable for relative uniform sites, of which the local flux measurements are representative of the individual surface types. For the interpretation of tower flux measurements over a heterogeneous land surface, operational footprint analysis is an essential approach (Schmid, 2002). The development of footprint models provides diagnostic tools to quantify the representativeness of tower flux measurements for selected sites (Horst and Weil, 1992; Kim et al., 2006). Besides, it had been demonstrated that the footprint climatology can be combined with information on the spatial variability of vegetation types provided by satellite image (Kim et al., 2006; Chen et al., 2008). Land cover reflects the combined effects of vegetation, climate, soil and topography, some relationship could be found between land cover and measured surface fluxes (Ogunjemiyo et al., 2003). Ran et al. (2016) proposed four indicators with footprint analysis and land-cover map to improve the representativity of EC towers and correct the EC flux measurements. But this method did not obtain the surface fluxes of individual land cover types but just corrected the EC observations with some prior coefficients. Some previous studies have successfully related the aircraft observed fluxes to surface cover types with the integration of footprint analysis and satellite-based land-cover map (Ogunjemiyo et al., 2003; Kirby et al., 2008; Hutjes et al., 2010). Among these works, a flux dis-aggregation method (Hutjes et al., 2010), developed from former study presented by Ogunjemiyo et al. (2003), would be a promising method for integrate multiple tower-based flux measurements to satellite pixel or grid scale on account of its theoretical framework. The application of this method in attributing heterogeneous EC flux measurements to separate land over classes will be a hopeful way to have insight into the component fluxes from various land cover types. It also provides a chance to develop a flux aggregation scheme for exploring the extension of multiple EC flux observations to satellite pixel/gird scale.

A multi-scale observation experiment on evapotranspiration over a heterogeneous land surface was conducted in the middle reaches of Heihe River Basin during the Project of HiWATER (Heihe Watershed Allied Telemetry Experimental Research) in 2012 (Li et al., 2013; Liu et al., 2016). A comprehensive flux matrix, consisted of 17 EC sites and 4 groups of LAS systems within a 5 × 5 km^2 area, was specifically designed to capture the multi-scale characteristics of ET over a heterogeneous landscape during the experiment. HiWATER flux matrix, with an abundant of multi-scale flux data, provided a unique opportunity to build an aggregation scheme for area-averaged fluxes over a heterogeneous land surface. The objective of this study is to integrate multi-point EC flux measurements to area-average with high-resolution land-cover map and footprint analysis. The main issues were as follows: (1) the representativeness of EC flux matrix was quantitatively evaluated; (2) a flux aggregation scheme was established and used for estimating area-averaged sensible heat fluxes from EC flux matrix, taking LAS measurements as reference to check the integration algorithm; (3) the developed scheme was applied to determine the area-averaged evapotranspiration over a heterogeneous land surface.

2 Study sites and data

2.1 Site description

This study was based on ground-based observation dataset, collected from the multi-scale flux matrix of HiWATER from May to September 2012. The kernel experimental area (5 × 5 km^2) of the multi-scale observation experiment was located in the Yingke and Daman irrigation district within Zhangye oasis. The land-cover types were dominated by maize (72 %), vegetables (5 %), orchard and shelterbelt (8 %) as well as residential area and roads (15 %). As shown by the numbers 1 – 17 in Fig. 1, 17 sites were installed according to the distribution of crop planting structure and land cover. Each of them was equipped with an eddy covariance system (with two layers in site 15) and an automatic weather station (AWS) to capture the exchange process of surface water and energy budget at the local scale. Spatial distribution of EC/AWS systems is shown in Fig. 1, with site 1 of vegetable (pepper) field, site 4 of residential area, site 17 of apple orchard, and the others are in maize fields. Key micrometeorological observations at each AWS included four-component radiation, one or two levels wind / temperature / relative humidity, soil temperature / moisture and soil heat flux, etc. Among these sites, site 15 was a superstation equipped with two levels of EC system, seven-level wind speed/direction and air temperature/humidity. 4 groups of large aperture scintillometers were installed crossed over the experimental district (see Fig. 1). Details of the EC and LAS systems in HiWATER flux matrix are given in Table 1 and Table 2, respectively.

2.2 Data collection, processing and quality control

2.2.1 Flux data processing and quality control

Data in typical clear days of 29 to 30 June 2012 were selected for the following analysis, including EC data from 16 towers (except site 3 and the highest level (34 m) of site 15) and 4 groups of LAS data as well as multi-point micrometeorological data listed above. The last round of irrigation in each plot was done before 26 June; during the two days, there was almost no irrigation in the flux matrix. Firstly, AWS data sampled at 10min were averaged to 30min period. Careful data processing and quality control for EC and LAS raw data were then performed so as to insure a high quality flux dataset.

The EddyPro software developed by LI-COR (Lincoln, Nebraska USA, www.licor.com/eddypro) was used to process the 10 Hz raw EC data into a half-hourly averaged flux data, by procedures including spike removal, angle of attack correction (for Gill), time lag correction, coordinate rotation (2-D rotation), frequency response correction, sonic virtual temperature correction, and corrections for density fluctuation (Webb-Pearman-Leuning, WPL) etc. Data quality assessment was performed for the turbulent flux in each 30min using the flagging system with 3 different levels (0, 1 and 2) (Mauder and Foken, 2015). Detailed information on the processing steps can be found in Wang et al. (2015) and Xu et al. (2013). For this study, only the flux data of flag 0 (the best) were used. Flux data of flag 2, as well as the data at night when the friction velocity was below 0.1 m s^{-1} were discarded (Blanken, 1998; Liu et al., 2011). To obtain daily ET, at first, a gap-filling method, based on the nonlinear regression (establishing the relationship between the latent heat flux and net radiation), for the 30min latent heat fluxes (LE) was used. Then, the daily ET was calculated by summing the half-hourly gap-filled ET to 24h totals.

For the EC systems used in the data analysis, we have tried to reduce the systematic errors to a minimum with a pre-observation inter-comparison, and careful maintenances during the observation period (Xu et al., 2013). The random errors were also analyzed by a separate research, which can be minimized in an ensemble average (Wang et al., 2015). The energy balance closure ratio (EBR) for the EC data of the flux matrix was also carefully assessed. Generally, the EBR during the 3 and half months was good. For the 17 EC stations in the intensive observation area, the average EBR was about 0.92. Except the lowest EBR (0.78) in orchard site, values in other sites were scattered without clear relation to the surface status. For site 15 with two heights of EC system, the relevant EBR were 0.89 (at 4.5 m) and 1.03 (for 34 m), respectively (Xu et al., 2017).

The LAS system provided a measurement of the structure parameter for the refractive index of air (C_n^2) with an output period of 1 min. The raw data were firstly checked, mainly reject the saturated cases when $C_n^2 < 0.193R^{8/3}\lambda^{1/3}D^{5/3}$ (where R is the path length, D the optical aperture, and λ the wavelength) (Ochs and Wilson, 1993). Then, the data were averaged to 30min, and the path-average sensible heat fluxes were iteratively calculated based on Monin-Obukhov Similarity Theory (MOST) (Andreas, 1988). The parameters used in this calculation, like the roughness height and zero-plane displacement, were obtained following Martano (2000); other parameters, including wind speed, air pressure and temperature, Obukhov length, and Bowen ratio, were directly obtained from relevant EC measurements. For this study, the sensible heat fluxes from the LAS systems at daytime (8:30 am – 15:30 pm, Beijing Standard Time, BST; the time difference between Local time and BST is approximately +1 h 18 min) were selected.

As for the eddy-covariance systems, quality-control for the flux data from the 4 groups of LAS was also done. The systematic errors from data processing, e.g. the larger effects of Bowen-ratio correction in this oasis area, were carefully minimized. We checked the sensible heat fluxes (H) from the 4 groups of LAS with that from the nearer ECs. Except LAS 3, under its path there were clearly some village buildings so the H_LAS is higher, others agreed very well with that of ECs.

2.2.2 Collection and processing of remote sensing products

Based on the airborne hyper-spectral images acquired by the Compact Airborne Spectrographic Imager (CASI) on 29 June 2012 and the Canopy Height Model (CHM) data from the LiDAR data collected on 9 July 2012, a land cover classification map with 1m spatial resolution was derived using an object-based classification method. This was done mainly for the kernel experimental area. The classification accuracy of the land cover map is up to 90 %, and Kappa coefficient is approximately 0.9. The detailed classification process of the map can be found in Liu and Bo (2015).

Land cover misclassification was still occurred in this map because of spectral similarity, especially in the edges of different surface cover types. To obtain a more accurate land cover map, the misclassified patches of the land cover map were visually and manually revised, according to the high-resolution CCD images (acquired on 26 July) and the Google Earth imagery (on 3 September 2012). Finally, for the aim of this study, the refined 12 kinds of land classification types, of which most were different vegetables of small areas, were merged into 4 kinds (maize, woods, vegetables and non-vegetation types) in accordance with crop species and surface types, as shown in Fig. 1. Among the four land cover types, the non-vegetation types mainly contain two types of land surface cover, namely buildings and road; while the woods type consists of orchard and shelterbelt.

3 Methodology

3.1 Aggregation method combining footprint analysis and multivariate regression

It is generally accepted that an average flux equals the area-weighted sum of the component fluxes from individual land cover classes (Hutjes et al., 2010).

$$F = \sum_{k=1}^{n} A_k F_k \quad (1)$$

Where F is the total flux of any scalar (e.g. the sensible and latent heat flux in our case) for a specified area, A_k is the fractional coverage of an individual land cover class k within that area, F_k is the flux from the individual land cover class k; n is the number of land cover classes that is distinguished in the specified area.

The observed flux (F_{obs}) at height z_m can be closely related to the true surface flux from upwind measurement point through the footprint function, in continuous form (Leclerc and Foken, 2014):

$$F_{obs}(x_{obs}, y_{obs}, z_m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x, y, 0) w(x, y, z_m) dxdy \quad (2)$$

Here x_{obs}, y_{obs} are the site coordinates, z_m is the effective observation height, defined as $z_m = z - d$ (where z is the sensor height, d the zero-plane displacement). The footprint function $w(x, y, z_m)$ describes the flux portion seen at (x_{obs}, y_{obs}, z_m). Equation (2) can be discretized for a uniform grid over a landscape, as in a land cover classification map based on satellite image, leaving out the height dependence for simplification, and becomes:

$$F_{obs} = \sum_{k=1}^{n} F_k \sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij} \Delta x \Delta y \quad (3)$$

Where each pixel $\Delta x \Delta y$ of the map is assumed to be homogeneous, which is uniquely classified as belonging to class k. The fraction of the kth land cover type in the footprint (fp) is then defined as:

$$X_{fp,k} = \sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij} \Delta x \Delta y \quad (4)$$

Combing Eq. (3) and Eq. (4), the multi-linear model for the flux becomes:

$$F_{obs} = \sum_{k=1}^{n} F_k X_{fp,k} \quad (5)$$

A critical assumption under the flux aggregation method is that each land cover k (area A_k) is with a constant source strength (F_k). Thus, as Eq. (1), flux (F) for a specific area is a weighted aggregation of its various land cover classes. Based on multi-point tower flux measurements (F_{obs}), multiple linear regression equations can be formulated by overlaying flux footprint with land-cover map ($X_{fp,k}$), as follows Eq. (5). In this study, the multiple-linear regression method (using the 'regress' algorithm from the Matlab® statistical toolbox) determined the regression coefficients (estimates of the specific flux for each land cover class in the case, F_k) by minimizing the squared residuals. For each LAS path, the measured flux (e.g. sensible heat flux) can also be dis-aggregated into component flux by relevant footprint function as Eq. (5). This can be taken as a validation of the former step.

The accuracy of this method is highly dependent on four most important aspects: (1) better flux data for all EC sites; (2) better land cover classification map; (3) more precise flux footprint analysis; (4) good flux and footprint data for LAS. So properly processed flux data, accurate high-resolution land cover map and appropriate footprint models are the foundation of formulating a better multiple linear regression. Sometimes, the established multi-linear regression equations may not be solvable. When this problem happens, the classification accuracy of the used land-cover map should be carefully checked, and the selected footprint model should also be verified with an alternative method.

3.2 Footprint models

The Eulerian analytical footprint model, which developed by Kormann and Meixner (2001), was used for estimating the single time flux footprint of EC measurements, due to its ease of use and wide range of stability as well as its numerical stability (Leclerc and Foken, 2014). Besides, as we have checked, its footprint estimates were in good agreement with the calculations of more sophisticated backward Lagrangian footprint models, such as the Kljun scheme (Kljun et al., 2002; Kljun et al., 2015). The footprint function $w(x, y, z)$ can be expressed in terms of a crosswind integrated flux footprint function $f^y(x, z)$ and a Gaussian crosswind distribution function $D_y(x, y)$. The analytic solution of Kormann and Meixner

(2001) is depicted by Eq. (6). More details on the derivation of $f^y(x,z)$ and $D_y(x,y)$ as well as the relevant parameters can be seen in Kormann and Meixner (2001).

$$w(x,y,z) = f^y(x,z)D_y = \frac{1}{\Gamma(\mu)}\frac{\xi^\mu}{x^{1+\mu}}e^{-\xi/\mu}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}} \qquad (6)$$

The flux contribution source area of LAS measurements was estimated by combining the footprint function $w(x,y,z)$ for point flux measurement with the path-weighting function $W(x)$ of LAS (Meijninger et al., 2002). For equal transmitter and receiver apertures, this path-weighting function is symmetrical bell-shaped having a center maximum and tapering to zero at the transmitter and receiver end. For the LAS footprint calculation, the approach of Korman and Meixner (2001) was still used for the single-point footprint estimation. The equation of LAS footprint function is that:

$$f_{LAS} = \int_{x_2}^{x_1} W(x)w(x-x', y-y', z_{LAS})dx \qquad (7)$$

Where x_1, x_2 are the positions of LAS receiver and transmitter, respectively. x, y represent the locations of points along the path of LAS. x', y' are the coordinates of each of upwind points. z_{LAS} is the effective height of LAS measurements.

To obtain averaged footprint of EC flux measurements (daily, monthly etc.), the flux-weighted footprint climatology method was applied for each pixel (Liu et al., 2016). The expression of the weighted footprint climatology is as:

$$w_c(x,y,z) = \sum_i^N w_i(x,y,z)\,Flux(i) \Big/ \sum_i^N Flux(i) \qquad (8)$$

Here i denotes the timestep (e.g. 30min), N is the total number of 30min periods within the selected time frame (such as, daily scale), $Flux(i)$ is the EC observed flux at i time-step (half-hourly LE in our case), $w_i(x,y,z)$ represents half-hourly footprint estimate calculated via Eq. (6).

The inputs of the analytical footprint model mainly include the measurement height, wind direction, wind speed and the Obukhov length, which can be easily derived from flux tower measurements. The daily-averaged flux footprint of the EC observations was calculated by Eq. (8). The flux contribution of the total source area was set to 90 % for both EC and LAS measurements. The normalized daily-averaged footprint of ECs and half-hourly footprint estimates of LASs were separately overlaid with the land cover map to determine the footprint-weighted contribution of each land cover classes to the measured flux from EC and LAS systems.

3.3 Data processing flow for the determination of area-averaged fluxes

The overall data processing flow for determining the area-averaged evapotranspiration over a heterogeneous land surface mainly includes three steps (Fig. 2).

Firstly, the spatial representativeness of 16 EC sites within the 5 × 5 km^2 experimental area was quantitatively assessed by overlaying flux footprint climatology with high-resolution land-cover map. Detailed analyses on this aspect are going to be presented in the following section.

The second aspect was to evaluate the reliability of the established flux aggregation scheme through the area-averaged flux measured by LAS. Specifically speaking, based on footprint analysis and high-resolution land-cover map, the land cover specific flux was firstly dis-aggregated from multiple EC flux measurements by performing a multiple linear regression analysis (Eq. 5). To obtain area-averaged fluxes representative for LAS source area, the EC dis-aggregated fluxes for all land cover classes were aggregated again according to the fractional weight of each land cover class in the LAS footprint (Eq. 4). Finally, the EC-aggregated fluxes were compared with LAS observations.

At last, the area-averaged evapotranspiration over a heterogeneous land surface was estimated with the flux integration scheme that was developed and verified.

4 Results and Discussion

4.1 The characteristics of the surface heat and water vapor fluxes

Figure 3 depicts the diurnal cycle of the sensible and latent heat fluxes at different sites on two clear days. It not only reveals the energy exchange of different sites but also the significant differences in the magnitude of the sensible and latent heat fluxes between different surface types during the growing season.

The sensible heat flux over residential area reached a maximum of about 150 W m^{-2} at afternoon and was higher than over the vegetated surfaces (EC04, Fig. 3a), while the latent heat flux was smallest due to a certain fraction of sealed land surfaces, with maximum value of less than 300 W m^{-2} (Fig. 3b).

Over the vegetated surfaces (maize, orchard, vegetable), the sensible heat flux was nearly less than 100 W m^{-2} because of sufficient irrigation (Fig. 3a), and it was significantly different between different vegetated surfaces (Fig. 4a). Besides, deviations in latent heat fluxes were also found. The maize fields performed highly latent heat fluxes and lower sensible heat fluxes than the other two vegetated surfaces. One of the possible reasons is that both of the orchard area and the vegetable field are rather sparse compared with the maize cropland during the crop growing period. The sensible heat flux for maize field sites was relatively small and even negative in the midafternoon when the sensible heat was transported downward (known as the 'oasis effect') (Fig. 3a, Fig. 4a). The latent heat flux over maize cropland was quite large, with maximum value of up to 600 W m^{-2} (Fig. 3b, Fig. 4b).

There was also a difference in sensible and latent heat fluxes among maize sites (Fig. 4). The values of the standard deviation (SD) of LE and H for 13 maize sites were about 43 W m^{-2} and 8 W m^{-2}, respectively. This finding showed that the latent heat flux over maize cropland exhibited larger SD than the sensible heat flux, and it also indicated the LE differed between sites for same underlying surface. This can be partly explained by the discrepancy in plant physiology and crop growing stage.

The preliminary results revealed, for the HiWATER flux matrix sites, the variability and difference in the surface energy fluxes between different surface types were really significant during the crop growth period. The differences in sensible and latent heat fluxes between maize field sites could also be noticed.

4.2 Analysis of the representativeness of the multi-point EC flux measurements

To further understand the variability of surface energy fluxes between different sites in a heterogeneous landscape, footprint analyses for the spatial representativeness of 16 EC sites were performed by superimposing flux footprint with a land-cover map derived from aircraft images (Fig. 1). The fraction of all land cover classes present in the daily-averaged footprint is shown in Fig. 5. Given the source area (90 % flux contribution) of the 4 ECs (sites 5, 8, 13 and 16) on 30 June 2012 exceeded the extent of land cover map, the spatial representativeness of the 4 EC sites was not presented in Fig. 5b.

Due to the variations in the observation height, atmospheric stability, wind direction and wind speed, the exact shape and size of EC source area were distinctly different (Fig. 1). For each EC flux measurements, there was more than one type of land cover class in its footprint. The contribution of each land cover classes to the total measured flux was changed with the varying source area (Fig. 5).

At sites 1 and 17, the dominated surface types in the source area were vegetable and orchard, respectively. For site 4, however, there were mainly three types of land cover within its source area, namely non-vegetation, maize and woods type. Moreover, the fractional weight of the non-vegetation type and maize field in the footprint varied greatly, while the ratio of woods type was almost changeless.

At maize field sites, the relative contribution of maize field to the EC measured flux was approximately more than 0.9, except for sites 2, 9 and 10. At site 2, the percentage of non-vegetation type in the footprint was almost 0.18. For site 9, the proportion of maize and non-vegetation type present in footprint significantly varied. The contribution of vegetable type to the flux measurements at site 10 ranged from 0.15 to 0.1.

The above analysis shows that the tower flux measurements at the field scale are generally representative of multiple surface types. The result indicates that the sensible and latent heat fluxes measured by EC systems are representative of the averaged fluxes, which are determined by weighting the upwind surface flux from individual land cover classes with flux footprint. In general, it may be problematic to validate the model estimated fluxes by direct comparison with point flux measurements over a heterogeneous land surface. Thus, the extension of multiple tower-based flux observations to pixel/grid scale is urgently needed for the validation of model estimates of surface flux.

4.3 Evaluation of the EC aggregated fluxes

The determination of area-averaged fluxes from point measurements is usually not straightforward, especially for heterogeneous land surfaces. Based on multi-point EC flux observations and accurate land-cover map, a flux aggregation method for obtaining area averaging of fluxes was established with footprint analysis and multivariate regression, and its feasibility was assessed by the LAS measurements.

The first step was to dis-aggregate the specific flux for all land cover classes from EC flux observations via the established scheme. The diurnal cycle of the EC dis-aggregated sensible heat fluxes for each land cover classes is highly significant (Fig. 6). Then, the EC dis-aggregated fluxes for four land cover classes were aggregated again to obtain area-averaged fluxes representative for LAS source area. Fig. 7 illustrates a scatterplot of half-hourly sensible heat fluxes estimated from EC flux matrix (hereafter referred as H_ECagg) versus LAS measurements (H_LAS), as well as the linear regression parameters (including equations and R^2). And the different statistics between them are also listed in Table 3.

For LAS 1 (see Fig. 7a and Table 3), a good agreement is found between EC aggregated fluxes and LAS measurements, with high correlation coefficient and low RMSE value (R^2 = 0.79, RMSE = 0.96 W m^{-2}). The scatter points in the graph are nearly close to the 1:1 line. The MBE and MAPE values were 4.25 W m^{-2} and 9.93 %, respectively.

Compared with LAS 1, there was a little scatter between H_ECagg and H_LAS for LAS 2, but yielding a small mean bias error (MBE = 2.31 W m^{-2}) (Fig. 7b, Table 3). RMSE and MAPE values of LAS 2 were little higher than that of LAS 1, probably owing to slight part of urban areas distributing in its path center (Fig. 1).

For LAS 3 (Fig. 7c, Table 3), there was a slightly weak relationship between sensible heat fluxes derived from LAS and multi-point EC flux measurements, with correlation coefficient (R^2) of 0.57. RMSE, MAPE and MBE values were 17.63 W m^{-2}, 31.7 % and -18.01 W m^{-2}, respectively. The negative MBE indicates that the 30min estimates of flux were quite underestimated against the area-averaged fluxes of LAS 2. In Fig. 7c, the scatter points were overall below the 1:1 line. There was more large area of residential areas randomly distributing in its central path than other three paths (Fig. 1).

In Fig. 7d, the area-averaged sensible heat fluxes obtained using the established integration scheme were consistent with the LAS 4 measurements, with R^2 of 0.57. In contrast with LAS 3, the scatter points in this graph were almost above the 1:1 line (overestimate of H_ECagg, MBE > 10 W m^{-2}). RMSE value of LAS 4 relatively decreased by 4.88 W m^{-2}, but MAPE was up to 33.7 % (Table 3). Large proportion of area in the LAS 4 source area was occupied by urban area and woods types as well as vegetable (Fig. 1), which to a large extent ranged with the variation of wind direction.

The findings show, above the more homogeneous areas, there is a good agreement between EC aggregated fluxes and LAS measurements, while great discrepancy between them occurred in the more urban areas. For the maize dominated areas, the unclosure of the energy balance is surely low, but this is not the case for the more urban area and the orchard site. The EBR for site 4 and site 17 over heterogeneous areas exhibited low values. This may be the one factor attributing to the bias.

The energy balance closure of the HiWATER flux dataset was influenced by surface heterogeneity, which might cause the energy imbalance in EC flux measurements and result in large eddies or organized circulation structures (Xu et al., 2017).

The energy flux from large eddies or secondary circulations cannot be captured by single-point EC measurements but be measured via LAS system (Foken, 2008; Foken et al., 2010). Thus, the LAS observations might be able to close to the surface energy balance better than the EC method (Foken et al. 2010).

In the study, only three stations had another dominant land cover (site 1, 4 and 17), especially for urban area, which occupied much part of our study area. The sensible heat flux for non-vegetation type disaggregated from site 4 might be insufficiently representative for the flux from sealed buildings and roads that are part of non-vegetation type. The divergence between modeled and measured flux may partly be related to this deficit.

The results demonstrate that compared with LAS measured area-averaged fluxes, the fluxes obtained from multiple EC flux measurements are reliable. Therefore, the present scheme can be an effective way to estimate area-averaged fluxes over a heterogeneous land surface.

4.4 Estimation of area-averaged evapotranspiration

The flux aggregation scheme, which was established and evaluated in Sect. 4.3, was adopted to determine the area-averaged ET over the study area with multi-point EC flux measurements and high-resolution land-cover map. The EC dis-aggregated daily ET for four land cover classes on two typical clear days was shown in Fig. 8. As can be seen, the daily ET values for maize field were highest (7 mm – 8 mm) during the crop growing season. The value of daily ET was 6.4 mm for woods type, and it ranged from 6 mm to 7 mm for vegetable field. On the contrary, the daily ET for non-vegetation type varied largely, with values of 2.8 mm on 29 June and 1.5 mm on 30 June, respectively.

With the dis-aggregated daily ET for all land cover classes, the daily ET maps at 1m resolution were produced through a high-resolution land classification map. Fig. 9 depicts the spatial pattern of daily ET on 29 and 30 June 2012. It can be seen from the legend, the daily ET ranged from 1.56 to 7.95 mm during the two days, with higher values on 29 June (Fig. 9a) than on 30 June (Fig. 9b).

Table 4 lists the total ET for different land cover classes and their proportion of the total area ET. The total ET for our study area was approximately 169.62 10^3 m^3 per day on 29 June, while it was about 152.94 10^3 m^3 per day on 30 June. The results demonstrated that the ratio of ET for maize field to the total area ET was in excess of 80 %. In addition, the total rate of ET for both woods and vegetables types was approximately 13 %, and the ET value for non-vegetation type accounted for 4.83 % of daily totals on the average.

The area-averaged ET over the kernel experimental area of HiWATER was finally estimated, with values of approximately 7 mm d^{-1} on 29 June and 6 mm d^{-1} on 30 June 2012.

5 Summary and conclusions

On the basis of accurate high-resolution land-cover map and multi-point ground-based flux measurements from 16 EC systems and 4 groups of LAS systems during the intensive observation period of HiWATER, a flux aggregation method for determining area-averaged flux was established through the combination of footprint analysis and multi-linear regression. The method was applied to estimate the area-averaged surface fluxes over a heterogeneous surface from multi-point EC flux measurements, and its results were verified by the LAS measurements. Ultimately, the integration method was applied to estimate area-averaged ET over the study area.

Robust quality-control and uncertainty-estimation for the EC and LAS data, done through careful data-processing and inter-comparison as well as assessment of the energy balance closure, make sure the accuracy of the flux dataset used in data analysis. For the deep interpretation of the surface fluxes over different land surfaces, the combination of footprint analyses for the representativeness of EC flux measurements and high-resolution land-cover map can be a practical way, and it is also the foundation for the establishment of the flux aggregation algorithm.

With high-quality flux dataset (EC & LAS), precise flux footprint estimates and accurate land cover classification map, a flux aggregation method can be successfully established by multivariate regression, and it achieves the goal of determining the area-averaged fluxes over heterogeneous areas from the EC flux matrix, according to the LAS measured fluxes. However, the agreement between the results of the flux integration method and LAS observed fluxes partly relates to the heterogeneity of land surface. On the other hand, it may partly attribute to the insufficient distribution of flux stations under urban areas.

In spite of the limitations mentioned above, the current flux integration scheme provides a unique opportunity to disentangle the heterogeneous land surface fluxes in their single components, and the dis-aggregation process has the potential to scale up multiple EC measurements to an oasis landscape, even to a whole river basin. Besides, compared with the formerly used and rather simple approaches (e.g. the arithmetic average and area-weighted methods), present scheme is not only with a much better database but also has a solid grounding in physics and mathematics in the integration of area-averaged fluxes over a heterogeneous surface. Results from this study, such as daily ET at the satellite pixel scale, can be applied for the validation of flux estimates of meso-γ scale (1 ~ 20 km) models. Furthermore, this work will be extended to the water balance study of the whole Heihe River basin, which is quite practical for hydrological modeling and basin water resource management.

Data availability

The flux observation matrix datasets from the eddy covariance (EC) systems and large aperture scintillometer (LAS) systems and the meteorological data in this study are available at http://card.westgis.ac.cn/hiwater/mso on request. The revised land-cover data for this paper are available from the corresponding author on request.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements. This study was supported by the National Natural Science Foundation of China (Grant number: 41671373; 41271359; 41301363).

References

Anderson, M. C., Kustas, W. P., Alfieri, J. G., Gao, F., Hain, C., Prueger, J. H., Evett, S., Colaizzi, P., Howell, T., and Chávez, J. L.: Mapping daily evapotranspiration at Landsat spatial scales during the BEAREX’08 field campaign, Advances in Water Resources, 50, 162-177, 2012.

André, J.-C., Goutorbe, J.-P., and Perrier, A.: HAPEX-MOBLIHY: A Hydrologic Atmospheric Experiment for the Study of

Water Budget and Evaporation Flux at the Climatic Scale, Bulletin of the American Meteorological Society, 67, 138-144, 1986.

Andreas, E. L.: Estimating Cn^2 over snow and sea ice from meteorological data, JOSA A, 5, 481-495, 1988.

Beyrich, F., Leps, J.-P., Mauder, M., Bange, J., Foken, T., Huneke, S., Lohse, H., Lüdi, A., Meijninger, W. M., and Mironov, D.: Area-averaged surface fluxes over the LITFASS region based on eddy-covariance measurements, Boundary-layer meteorology, 121, 33-65, 2006.

Beyrich, F., and Mengelkamp, H.-T.: Evaporation over a heterogeneous land surface: EVA_GRIPS and the LITFASS-2003 experiment—an overview, Boundary-layer meteorology, 121, 5-32, 2006.

Blanken, P.: Turbulent flux measurements above and below the overstory of a boreal aspen forest, Boundary-Layer Meteorology, 89, 109-140, 1998.

Charuchittipan, D., Babel, W., Mauder, M., Leps, J. P., and Foken, T.: Extension of the Averaging Time in Eddy-Covariance Measurements and Its Effect on the Energy Balance Closure, Boundary-Layer Meteorology, 152, 303-327, 2014.

Chen, B., Black, T. A., Coops, N. C., Hilker, T., Trofymow, J. A., and Morgenstern, K.: Assessing Tower Flux Footprint Climatology and Scaling Between Remotely Sensed and Eddy Covariance Measurements, Boundary-Layer Meteorology, 130, 137-167, 10.1007/s10546-008-9339-1, 2008.

Ezzahar, J., and Chehbouni, A.: The use of scintillometry for validating aggregation schemes over heterogeneous grids, Agric. For. Meteorol., 149, 2098-2109, 2009.

Ezzahar, J., Chehbouni, A., Er-Raki, S., and Hanich, L.: Combining a large aperture scintillometer and estimates of available energy to derive evapotranspiration over several agricultural fields in a semi-arid region, Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology, 143, 209-221, 10.1080/11263500802710036, 2009a.

Ezzahar, J., Chehbouni, A., Hoedjes, J., Ramier, D., Boulain, N., Boubkraoui, S., Cappelaere, B., Descroix, L., Mougenot, B., and Timouk, F.: Combining scintillometer measurements and an aggregation scheme to estimate area-averaged latent heat flux during the AMMA experiment, Journal of hydrology, 375, 217-226, 2009b.

Foken, T., Wimmer, F., Mauder, M., Thomas, C., and Liebethal, C.: Some aspects of the energy balance closure problem, Atmospheric Chemistry and Physics, 6, 4395-4402, 2006.

Foken, T.: The energy balance closure problem: An overview, Ecological Applications, 18, 1351-1367, 2008.

Foken, T., Mauder, M., Liebethal, C., Wimmer, F., Beyrich, F., Leps, J.-P., Raasch, S., DeBruin, H. A., Meijninger, W. M., and Bange, J.: Energy balance closure for the LITFASS-2003 experiment, Theoretical and Applied Climatology, 101, 149-160, 2010.

Goutorbe, J., Lebel, T., Tinga, A., Bessemoulin, P., Brouwer, J., Dolman, A., Engman, E., Gash, J., Hoepffner, M., and Kabat, P.: HAPEX-Sahel: a large-scale study of land-atmosphere interactions in the semi-arid tropics, Annales Geophysicae, 12, 53-64, 1994.

Halldin, S., Gottschalk, L., van de Griend, A. A., Gryning, S.-E., Heikinheimo, M., Högström, U., Jochum, A., and Lundin, L.-C.: NOPEX—a northern hemisphere climate processes land surface experiment, Journal of Hydrology, 212, 172-187, 1998.

Horst, T., and Weil, J.: Footprint estimation for scalar flux measurements in the atmospheric surface layer, Boundary-Layer Meteorology, 59, 279-296, 1992.

Hutjes, R., Vellinga, O., Gioli, B., and Miglietta, F.: Dis-aggregation of airborne flux measurements using footprint analysis, Agric. For. Meteorol., 150, 966-983, 2010.

Jia, Z., Liu, S., Xu, Z., Chen, Y., and Zhu, M.: Validation of remotely sensed evapotranspiration over the Hai River Basin, China, Journal of Geophysical Research: Atmospheres, 117, D13113, doi:10.1029/2011JD017037., 2012.

Kim, J., Guo, Q., Baldocchi, D., Leclerc, M., Xu, L., and Schmid, H.: Upscaling fluxes from tower to landscape: Overlaying flux footprints on high-resolution (IKONOS) images of vegetation cover, Agric. For. Meteorol., 136, 132-146, 2006.

Kirby, S., Dobosy, R., Williamson, D., and Dumas, E.: An aircraft-based data analysis method for discerning individual fluxes in a heterogeneous agricultural landscape, Agric. For. Meteorol., 148, 481-489, 2008.

Kljun, N., Rotach, M., and Schmid, H.: A three-dimensional backward Lagrangian footprint model for a wide range of boundary-layer stratifications, Boundary-Layer Meteorology, 103, 205-226, 2002.

Kljun, N., Calanca, P., Rotach, M., and Schmid, H.: A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP), Geoscientific Model Development, 8, 3695-3713, 2015.

Kormann, R., and Meixner, F. X.: An analytical footprint model for non-neutral stratification, Boundary-Layer Meteorology, 99, 207-224, 2001.

Leclerc, M. Y., and Foken, T.: Footprints in Micrometeorology and Ecology, Springer, Heidelberg, New York, Dordrecht, London, XIX, 239 pp., 2014.

Li, X., Cheng, G., Liu, S., Xiao, Q., Ma, M., Jin, R., Che, T., Liu, Q., Wang, W., Qi, Y., Wen, J., Li, H., Zhu, G., Guo, J., Ran, Y., Wang, S., Zhu, Z., Zhou, J., Hu, X., and Xu, Z.: 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.

Liu, S., Xu, Z., Wang, W., Jia, Z., Zhu, M., Bai, J., and Wang, J.: A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem, Hydrology & Earth System Sciences, 15, 1291-1306, 2011.

Liu, S., Xu, Z., Song, L., Zhao, Q., Ge, Y., Xu, T., Ma, Y., Zhu, Z., Jia, Z., and Zhang, F.: Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces, Agric. For. Meteorol., 230, 97-113 2016.

Liu, X., and Bo, Y.: Object-Based Crop Species Classification Based on the Combination of Airborne Hyperspectral Images and LiDAR Data, Remote Sens., 7, 922-950, 2015.

Mahrt, L., Vickers, D., Sun, J., and McCaughey, J. H.: Calculation of area-averaged fluxes: Application to BOREAS, Journal of applied meteorology, 40, 915-920, 2001.

Martano, P.: Estimation of surface roughness length and displacement height from single-level sonic anemometer data, Journal of Applied Meteorology, 39, 708-715, 2000.

Mauder, M., and Foken, T.: Documentation and instruction manual of the eddy covariance software package TK3 (update), Arbeitsergebnisse, Universität Bayreuth, Abt. Mikrometeorologie (ISSN 1614-8916), 62, 64 pp, 2015.

Meijninger, W., Hartogensis, O., Kohsiek, W., Hoedjes, J., Zuurbier, R., and De Bruin, H.: Determination of area-averaged sensible heat fluxes with a large aperture scintillometer over a heterogeneous surface–Flevoland field experiment, Boundary-Layer Meteorology, 105, 37-62, 2002.

Mengelkamp, H.-T., Beyrich, F., Heinemann, G., and Ament, F., Bange, J., Berger, F. H., Bösenberg, J., Foken, T., Hennemuth, B., Heret, C., Huneke, S., Johnsen, K.-P., Kerschgens, M., Kohsiek, W., Leps, J.-P., Liebethal, C., Lohse, H., Mauder, M., Meijninger, W. M. L., Raasch, S., Simmer, C., Spieß, T., Tittebrand, A., Uhlenbrook, S., and Zittel,P.: Evaporation over a heterogeneous land surface: the EVA-GRIPS project, Bulletin of the American Meteorological Society, 87, 775-786, 2006.

Ochs, G., and Wilson, J.: A Second-generation Large-aperature Scintillometer, US Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Wave Propagation Laboratory, 1993.

Ogunjemiyo, S. O., Kaharabata, S. K., Schuepp, P. H., MacPherson, I. J., Desjardins, R. L., and Roberts, D. A.: Methods of estimating CO_2, latent heat and sensible heat fluxes from estimates of land cover fractions in the flux footprint, Agric. For. Meteorol., 117, 125-144, 2003.

Ran, Y., Li, X., Sun, R., Kljun, N., Zhang, L., Wang, X., and Zhu, G.: Spatial representativeness and uncertainty of eddy covariance carbon flux measurements for upscaling net ecosystem productivity to the grid scale, Agric. For. Meteorol., 230, 114-127, 2016.

Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: a review and perspective, Agric. For. Meteorol., 113, 159-183, 2002.

Sellers, P., Hall, F., Asrar, G., Strebel, D., and Murphy, R.: The first ISLSCP field experiment (FIFE), Bulletin of the American Meteorological Society, 69, 22-27, 1988.

Sellers, P., Hall, F., Ranson, K. J., Margolis, H., Kelly, B., Baldocchi, D., den Hartog, G., Cihlar, J., Ryan, M. G., and Goodison, B.: The boreal ecosystem-atmosphere study (BOREAS): an overview and early results from the 1994 field year, Bulletin of the American Meteorological Society, 76, 1549-1577, 1995.

Wang, J.: Land surface process experiments and interaction study in China: From HEIFE to IMGRASS and GAME-Tibet/TIPEX, Plateau Meteorol, 18, 280-294, 1999.

Wang, J., Gao, F., and Liu, S.: Remote sensing retrieval of evapotranspiration over the scale of drainage basin, Remote Sensing Technology and Application, 18, 332-338, 2003.

Wang, J., Zhuang, J., Wang, W., Liu, S., and Xu, Z.: Assessment of Uncertainties in Eddy Covariance Flux Measurement Based on Intensive Flux Matrix of HiWATER-MUSOEXE, IEEE Geoscience and Remote Sensing Letters, 12, 259-263, 2015.

Xu, Z., Liu, S., Li, X., Shi, S., Wang, J., Zhu, Z., Xu, T., Wang, W., and Ma, M.: Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE, Journal of Geophysical Research: Atmospheres, 118, 13140-13157, 2013.

Xu, Z., Ma, Y., Liu, S., Shi, W., and Wang, J.: Assessment of the Energy balance closure under advective conditions and its impact using remote sensing data, Journal of Applied Meteorology and Climatology, 56, 127-140, 2017.

Table 1 Details of the eddy covariance systems in the HiWATER flux matrix

Site No.	Longitude (°)	Latitude (°)	Elevation (m)	Turbulence sensors	Sensor height (m)	Surface type
1	100.3582	38.8932	1552.75	Gill/Li7500A	3.8	Vegetables
2	100.35406	38.88695	1559.09	CSAT3/Li7500	3.7	Maize
3	100.37634	38.89053	1543.05	Gill/Li7500A	3.8	Maize
4	100.35753	38.87752	1561.87	CSAT3/Li7500A	4.2/ 6.2 after 19 Aug.	Residential area
5	100.35068	38.87574	1567.65	CSAT3/Li7500	3.0	Maize
6	100.3597	38.87116	1562.97	CSAT3/Li7500A	4.6	Maize
7	100.36521	38.87676	1556.39	CSAT3/Li7500A	3.8	Maize
8	100.37649	38.87254	1550.06	CSAT3/Li7500	3.2	Maize
9	100.38546	38.87239	1543.34	Gill/Li7500A	3.9	Maize
10	100.39572	38.87567	1534.73	CSAT3/Li7500	4.8	Maize
11	100.34197	38.86991	1575.65	CSAT3/Li7500	3.5	Maize
12	100.36631	38.86515	1559.25	CSAT3/Li7500	3.5	Maize
13	100.37841	38.86076	1550.73	CSAT3/Li7500A	5.0	Maize
14	100.3531	38.85867	1570.23	CSAT3/Li7500	4.6	Maize
15	100.37223	38.85555	1556.06	CSAT3/Li7500A	4.5/ 34	Maize
16	100.36411	38.84931	1564.31	Gill/Li7500	4.9	Maize
17	100.36972	38.8451	1559.63	CSAT3/EC150	7.0	Orchard

Table 2 Details of the Large Aperture Scintillometers in the HiWATER flux matrix

Site	Longitude (°)	Latitude (°)	LAS type, Manufactures	Path length(m)	Effective height (m)
LAS 1	North 100.35090	38.88413	BLS900, Scintec, Germany	3256	33.45
	South 100.35285	38.85470	RR9340, Rainroot, China	3256	33.45
LAS 2	North 100.36236	38.88256	BLS900, Scintec, Germany	2841	33.45
	South 100.36171	38.85717	BLS450, Scintec, Germany	2841	33.45
LAS 3	North 100.37319	38.88338	BLS900, Scintec, Germany	3111	33.45
	South 100.37223	38.85555	LAS, Kipp&zonen, Netherland	3111	33.45
LAS 4	North 100.37841	38.86076	BLS450, Scintec, Germany	1854	22.45
	South 100.36840	38.84682	RR9340, Rainroot, China	1854	22.45

Table 3 Different statistics between LAS observed flux and EC aggregated flux at LAS sites

LAS sites	RMSE [W m^{-2}]	MBE [W m^{-2}]	MAPE [%]
LAS1	0.96	4.25	9.93
LAS2	6.91	2.31	16.39
LAS3	17.63	-18.01	31.70
LAS4	12.75	10.66	33.70

Remarks: $RMSE=\sqrt{\sum_{i=1}^{n}(P_i-O_i)^2/n}$, $MAPE=\frac{100}{n}\sum_{i=1}^{n}\frac{|P_i-O_i|}{\bar{O}}$, $MBE=\sum_{i=1}^{n}(P_i-O_i)/n$, P_i is EC aggregated value， O_i is LAS observed value， $\bar{O}$ is the mean measured value, n is the number of samples. RMSE is root mean square error, MAPE is mean absolute percentage error, MBE is the mean bias error.

Table 4 ET for each land cover classes and their proportion of the kernel experimental area ET

Land cover class	Area [km^2]	29 June 2012		30 June 2012	
		ET [10^3 m^3 d^{-1}]	ET proportion of total ET [%]	ET [10^3 m^3 d^{-1}]	ET proportion of total ET [%]
Maize	17.42	138.43	81.61	127.24	83.20
Woods	1.96	12.78	7.53	12.59	8.23
Vegetables	1.20	8.27	4.88	7.48	4.89
Non-vegetation	3.62	10.14	5.98	5.63	3.68

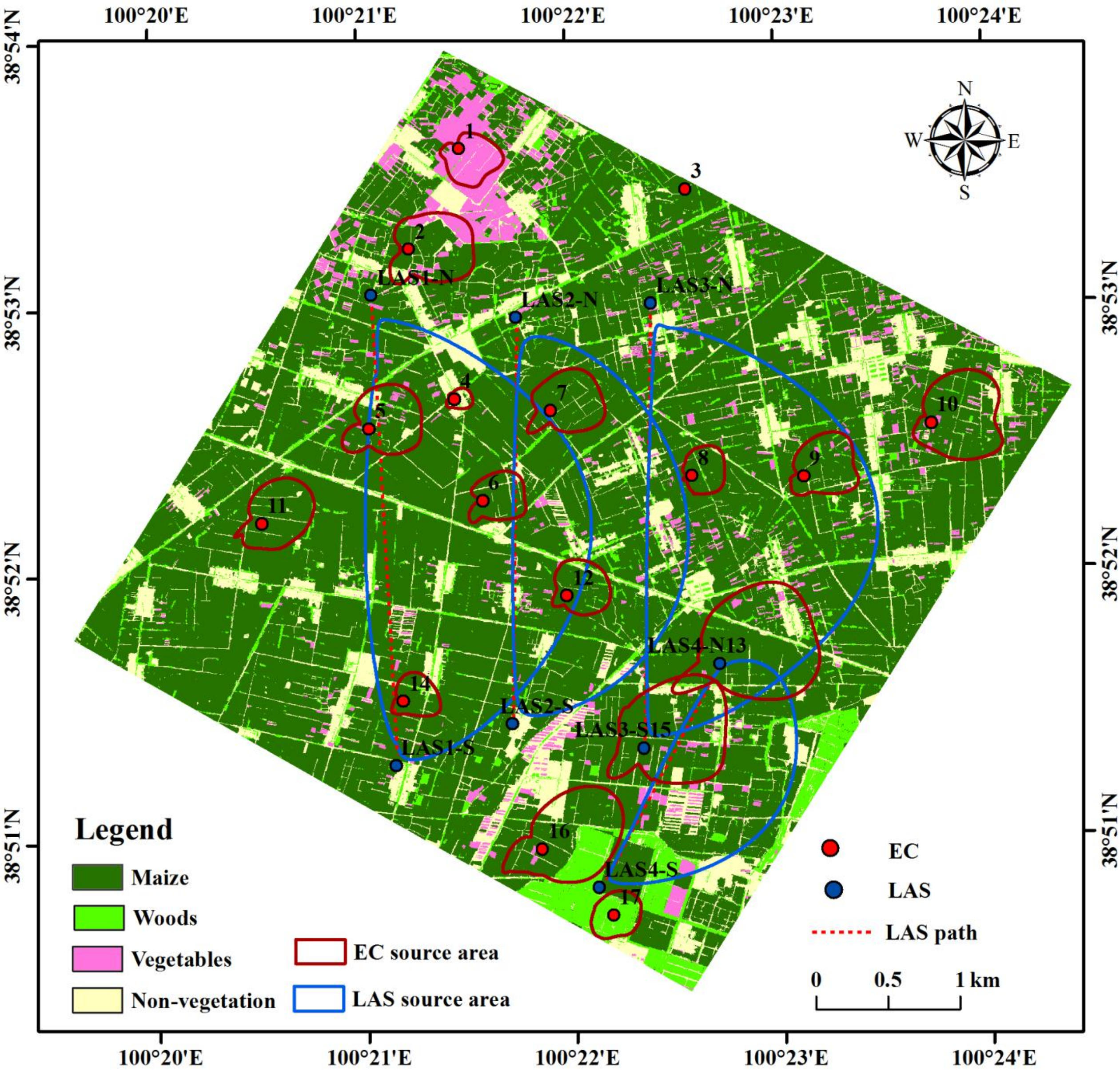

Figure 1: The land cover map of the kernel experiment area of HiWATER 2012. The small red circles represent the 90 % flux contribution source area of EC sites, and the large blue circles covering different land cover classes indicate the source area of LAS sites on 29 June 2012

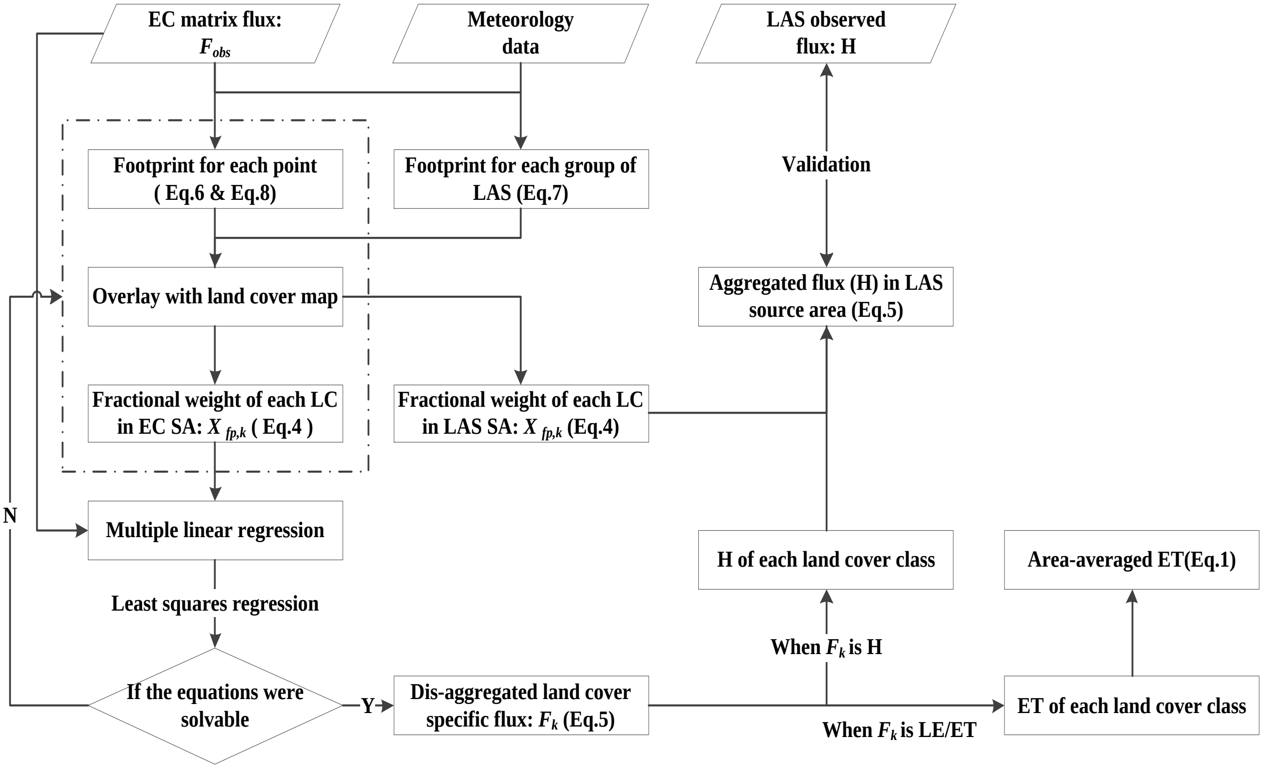

Figure 2: Schematic illustration of data processing steps; LC = land cover class; SA = source area; H = sensible heat flux; LE = latent heat flux; ET = evapotranspiration

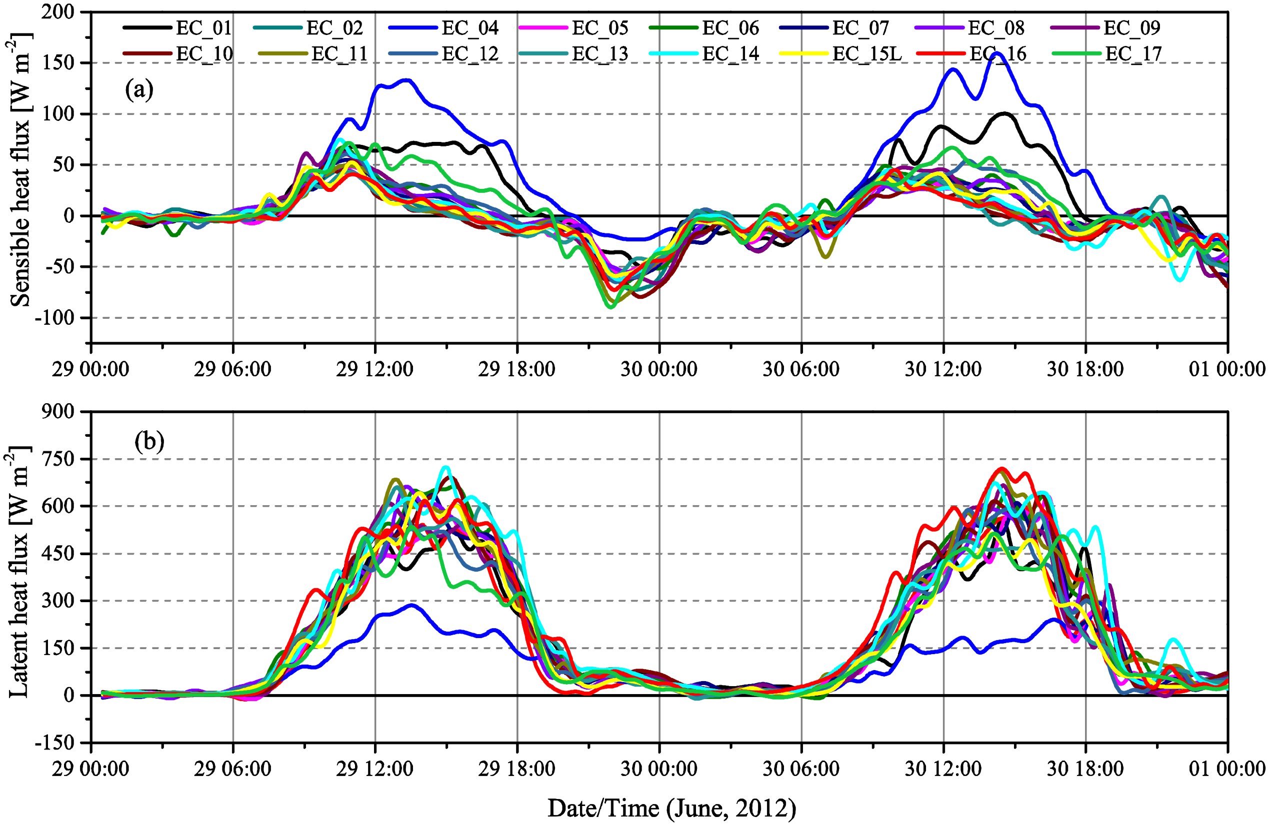

Figure 3: Diurnal cycle of the sensible heat fluxes (a) and latent heat fluxes (b) between different sites on 29 and 30 June 2012

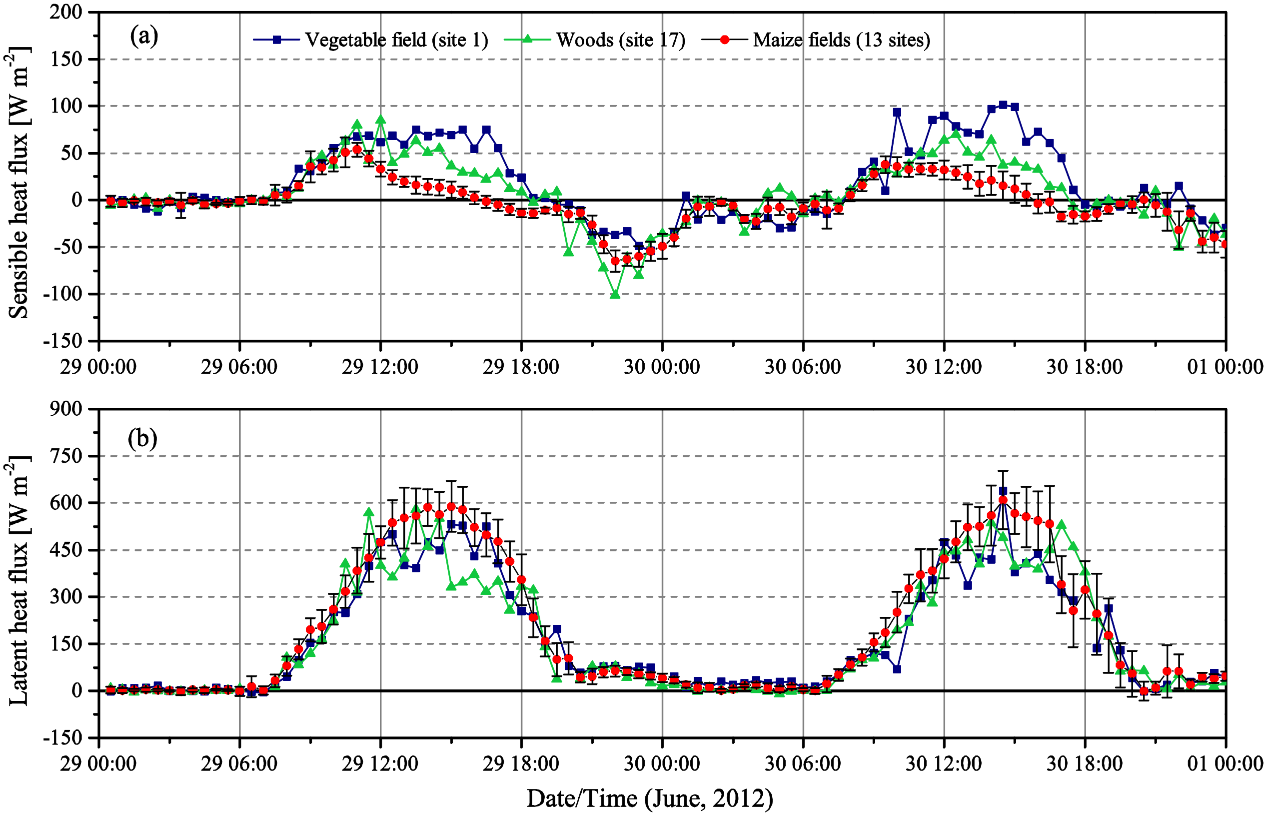

Figure 4: Diurnal cycle of the mean sensible (a) and latent (b) heat fluxes for 13 maize field sites and different types of vegetation, the errors bar is the standard deviation

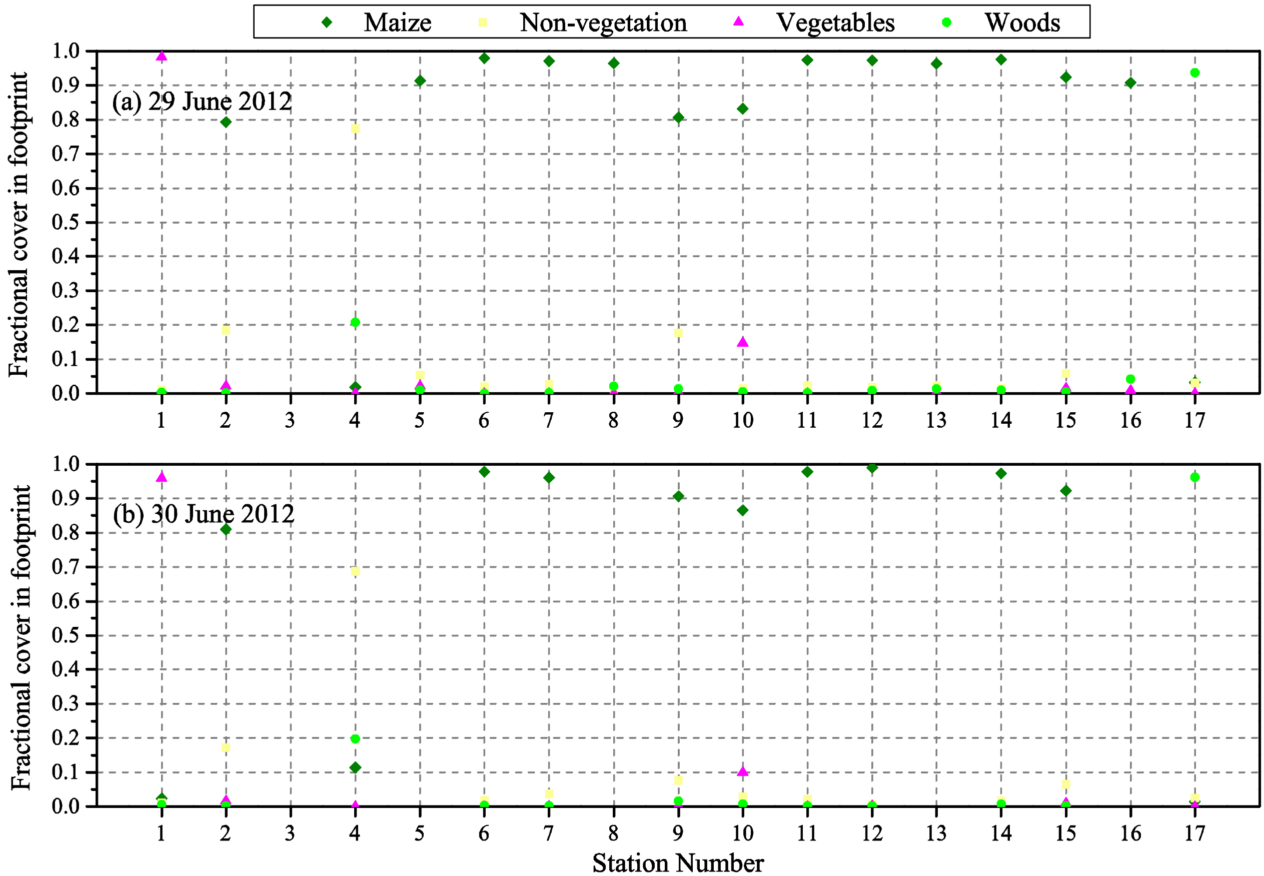

Figure 5: The fractional weight of each land cover classes in the daily averaged flux footprint of each EC flux measurements on 29 and 30 June 2012

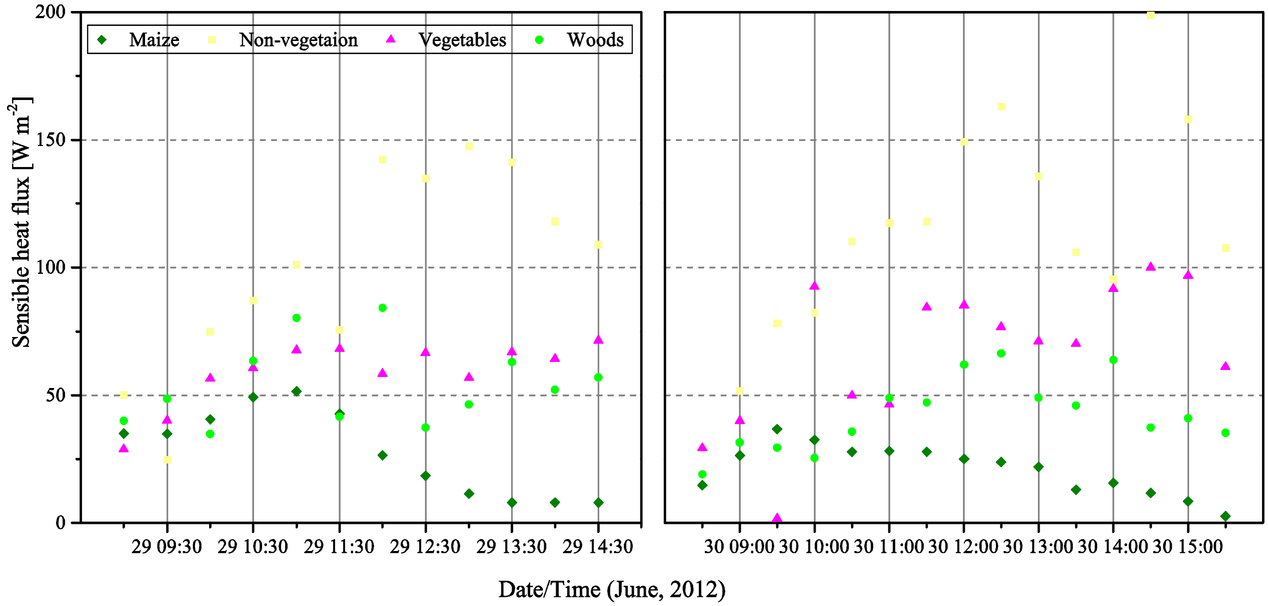

Figure 6: The diurnal cycle of the sensible heat flux for each land cover classes on 29 and 30 June 2012

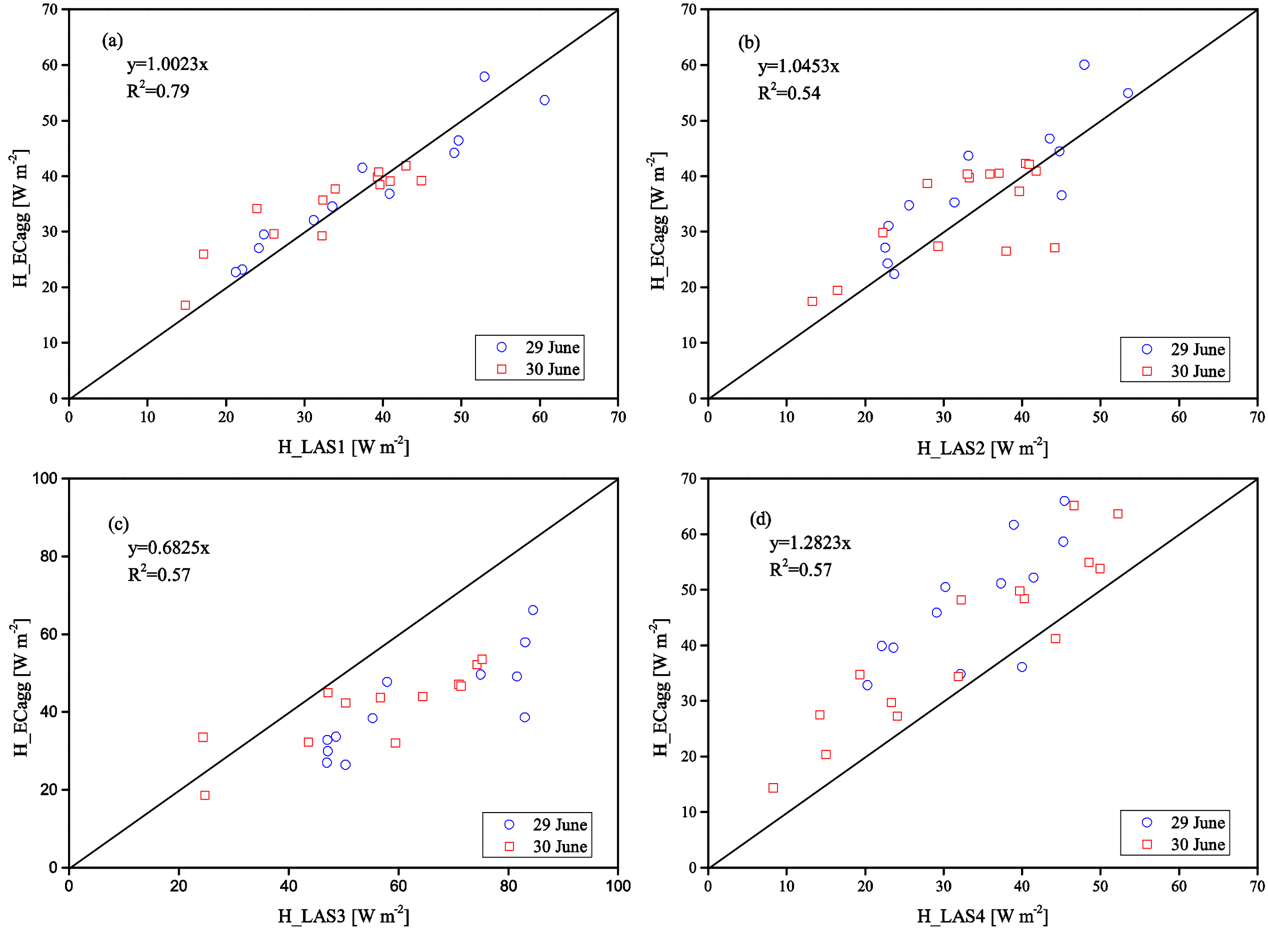

Figure 7: The comparison between LAS observed fluxes (X axis) and EC aggregated fluxes (Y axis)

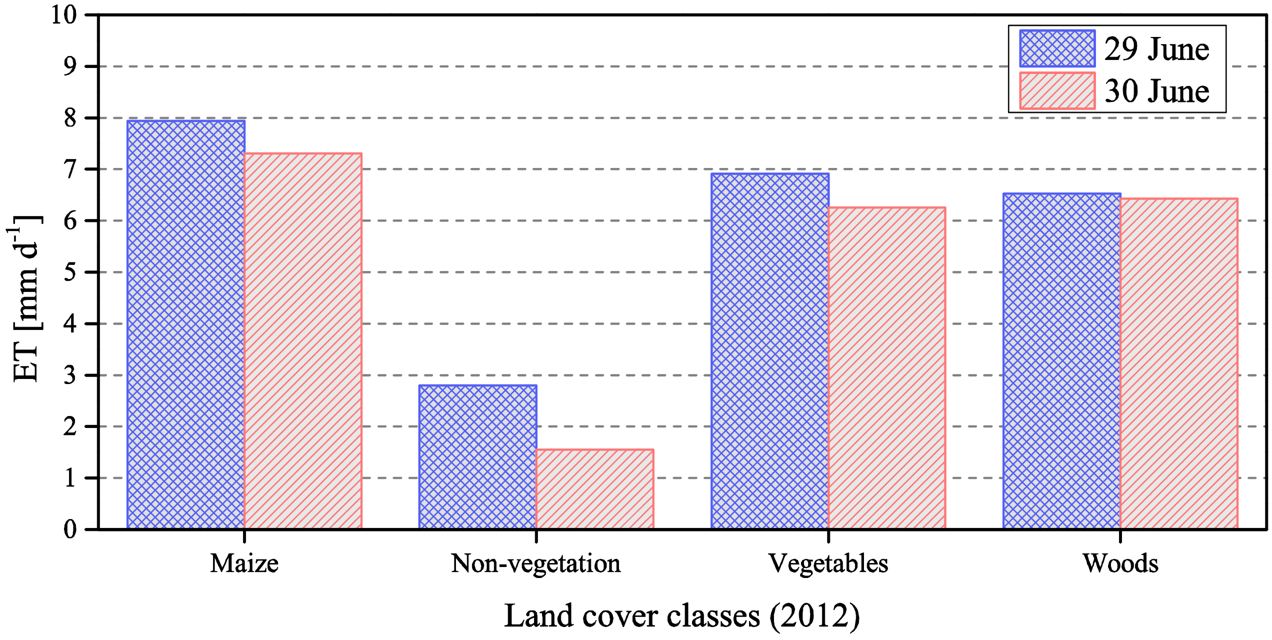

Figure 8: The dis-aggregated daily ET of each land covers in the kernel experimental area of HiWATER on 29 and 30 June 2012

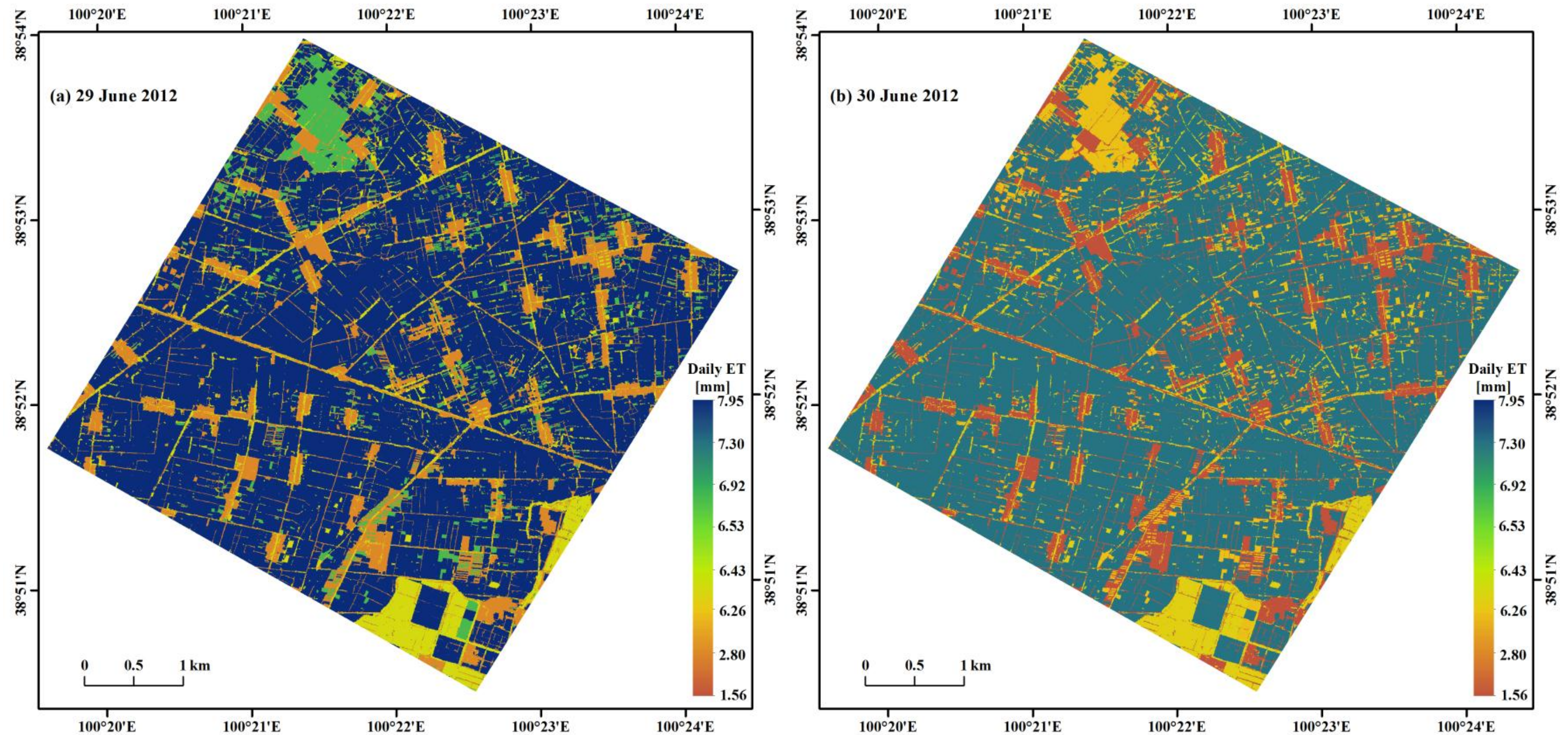

Figure 9: Spatial distribution of averaged daily ET over the kernel experimental area of HiWATER