Many thanks for handling the review process for our manuscript. The time and effort devoted to our manuscript by you and the reviewers are very much appreciated.

We have revised the manuscript carefully according to the reviewers' comments and suggestions. This includes a number of minor wording changes throughout the document to improve clarity and readability. In the following, we provide a point-by-point response to the reviewers' comments. The original comments are in black and blue regular font. Our responses are shown in blue italic font. Quotes from the revised paper are shown in green bold-face font. Additionally, there are a number of small grammatical and wording changes throughout the manuscript that are not specifically documented below.

REVIEWER COMMENTS

Reviewer (Raquel González Armas and Jordi Vilà-Guerau de Arellano):

We thank the authors for their responses and the changes. In our opinion, the manuscript has improved.

Below our reaction to their responses and the changes in the manuscript (in blue). If there is no reaction, we are satisfied with their answers and the revised manuscript.

Major comments

1. Although the processes of entrainment and boundary layer growth is acknowledged throughout the paper, we have the feeling that is played down in the research. We realized that with a surface data set is difficult to quantify, although the mixed-layer diagrams proposed by Santanello et al. (2009) could be an adequate tool to further quantify the relevance of entrainment of warm and dry air at the different sites. Could the authors elaborate and quantify more regarding the role of entrainment?

   Could they be more precise on the projects dominant after midday when entrainment becomes less relevant? The word obscure is vague in the revised manuscript.

   We also think that they are missing here a great opportunity in elaborating a bit more in the relevance of the processes at the sub-diurnal scales. Our recommendation is that based on their analysis and metrics they performed a more deep analysis on which processes are relevant or differ under the water and energy regimes.

   ➔ *We have tried to elaborate the impact of the entrainment as well as radiative cooling in the Summary section. This is discussed in the last paragraph of section 4.2. To characterize further the atmospheric entrainment in terms of heat and moist energy budget, we refer to a previous study that investigated the boundary layer budgets using radiosonde data (Barr and Betts, 1997). This contributes to the range of the diurnal variance in moist and thermal energy budgets in the mixing diagram. Text is added in Lines 537-544:*

   "The effect of atmospheric entrainment is greatest during the period of ML growth in the morning, when the entrained dry and high potential temperature air at the top of the PBL causes positive temperature and negative moisture tendencies in the ML. Entrainment weakens but continues after the ML reaches maximum depth until dissipation of the daytime boundary layer around sunset. The atmospheric entrainment characterizes the maximum of both tendencies around noon and the stronger negative moisture tendency (Barr and Betts, 1997). However, the impact of the entrainment is mainly from 7 AM to noon. Meanwhile, there is radiative cooling of the ML at all hours that there are no clouds above the ML trapping longwave radiation. The radiative cooling likely dominates afternoon when mean tendencies become negative (Fig. 4h)."

Please rephrase and elaborate more (in red there is typo):

"Although the effect of atmospheric entrainment continues until continues until dissipation of the daytime boundary layer around sunset, it is obscured by the other contributions after noon."

➔ *The red colored typo is only in the response to reviewers document and not in the main manuscript. We apologize for the confusion..*

5. Along the results section in part 4.2 Diurnal mixing diagrams and 4.3 Climate regime dependence, hysteresis of the thermal process chain versus the moist process chain is discussed. Regarding the discussion of hysteresis, we have three comments:

a) We highly encourage to define in this context the term hysteresis. Hysteresis is a word originally coined in science to describe systems which state depends on their history. The typical scientific example is the magnetic hysteresis. This refers to a magnet that is able to experience different magnetic moments when subject to the same magnetic field. Those magnetic moments depend on the previous states of the magnet. To us, using hysteresis in land atmospheric context may be misleading since the state of the system may be different between morning and afternoon because the external factors are also different. For instance, soil water content and vapor pressure deficit are generally different between morning and afternoon. Therefore, the sub-diurnal asymmetry may be attributed to it not because an inherent change on the interactions due to the previous history. Nonetheless, we acknowledge that hysteresis term is generally used in land-atmospheric interactions context. We recommend defining the term in this context. We already find a definition in conclusions section, line 417, the fact that "the evening path through the water-energy phase does not retrace the morning path". We would move or repeat the definition to results because there is where the hysteresis is widely discussed. In addition, we think it would be valuable to specify in which way we consider it a hysteresis. In essence, which system is subject to its previous history? Is it the vegetation, is it a vegetation-soil system? What are considered the external factors? Another simpler solution is to coin another term such as temporal asymmetry which does not imply previous history relations.

What do you mean by a kind of hysteresis? Is it not more appropriate to call it asymmetry? Please formulate with precision

We have changed the word "hysteresis" to "asymmetry" throughout the manuscript, and clarified what we mean by asymmetry at Lines 339-340:

"There is an asymmetry in the path of moisture and temperature across the diurnal cycle, in that the extremes in the thermal process chain lead the moist process chain by 2-3 hours."

b) We highly recommend discussing the hysteresis' possible causes both on the land and the atmospheric coupling. We argue that due to many processes that peak at different times (e.g., radiation peaks around noon, sensible heat flux peaks in the early afternoon and latent heat flux which with peaks later in the afternoon), morning-afternoon asymmetry can be expected. It is not clear to us what is the added value of assessing the asymmetry or if the aim of the research is simply to characterize it. We recommend clarifying either if the paper aims to characterize them as a general characteristic observed or if the asymmetry is seen as a possible option to evaluate land atmosphere interactions.

It is a pity that the authors do not develop and elaborate a bit more on the physics of this asymmetry based on the observational analysis

➔ *Based on the reviewers' comment, we have tried to explain the physics of the hysteresis (asymmetry) in the land-atmosphere coupling metrics. The impact of the diurnal peaks of land heat fluxes on the leading phase of land couplings. The asymmetric behavior in the atmospheric couplings is also deeply elaborated. This description is added in Lines 348-351, 357-358, 408-409, and 411-413:*

"H and LE peak in the early and later afternoon, respectively, each strongly controlled by gradients between the land surface and lower atmosphere. As the air warms in the afternoon and incoming solar radiation starts to decline, the thermal gradient weakens reducing H. At the same time, the warm air increases the potential evaporation by maintaining a large vapor pressure deficit, facilitating strong rates of LE."

"…, which reveals the phase of A(H,LCL) leading A(LE,LCL) by 2-3 hours, …"

"The closer to the water–limited regime, the higher the magnitude of the correlation between SWC1 and both surface fluxes."

"A(H,LCL) over the water–limited regime is stronger than over the energy–limited regime, which results from the larger LCL variability along with the marginal sensitivity of R(H,LCL) to the climate regime."

Other Comments

- In line 109, the lifting condensation level is used as the variable to understand the coupling of the land with the atmosphere. We think the reader would appreciate a short sentence in which it is stated why this variable is an important indicator of the coupling to the atmosphere (e.g., because its strong relation with cloud initiation or its importance in convection schemes in atmospheric models).

Perhaps here it is convenient to be more rigorous and stress that the condition h > LCL is a rough approximation. Majority of the situations in which shallow cumulus form are characterized by an opposite situation (h < LCL) (see for instance figure 7a at https://journals.ametsoc.org/view/journals/atsc/71/3/jas-d-13-0192.1.xml). Could they please elaborate a bit more here?

➔ *LCL is often used because it is easy to calculate from commonly available surface data, and that is also the case here. During the day when the surface inversion is broken and the mixed layer has formed, it is a reliable but rough approximation for the potential cloud base, subject to the limitations of parcel theory.*

We thank the reviewers for pointing out the paper of van Stratum et al. (2014); we also see relevance in the Heating Condensation Framework (HCF) of A. Tawfik to explain this difference. We remind the reviewers that the general issue of the limitations of parcel theory are discussed at the end of section 3.3, but we now elaborate on the specific points regarding LCL with modified text starting at Line 122:

"The LCL can be characterized as a potential level of cloud base formation based on parcel theory, and is easily calculated from surface meteorological measurements, but is an approximation subject to the limitations of parcel theory. In reality, the profile of temperature and moisture above the surface also determine the level of the cloud base (Tawfik and Dirmeyer, 2014). The LCL can be compared to the planetary boundary layer (PBL) height to define an LCL

deficit (PBL height minus LCL; Santanello et al., 2011). When the PBL grows to the height of the LCL (corresponding to positive values of the LCL deficit), water may condense from the air parcel, and cloud formation occurs, although clouds begin to form when scattered updrafts penetrate the condensation level before the entire ML reaches the LCL (Van Stratum et al., 2014)."

- 3.3 Mixing diagrams section. Along this section mixing diagrams are introduced. It is stated that for computing them, 2-m temperature and humidity or vapor pressure deficit are used. In the last paragraph of the section, some shortcomings of this approach are addressed. For instance, it is mentioned that embedded in this method it lies one hypothesis. The hypothesis that 2-m measurements reflect mixed-layer values. We find this hypothesis to be dubious for certain ecosystems. For instance, in vegetated areas whose trees are taller than 2-m, the measurements fall into the in-canopy range. Many forests have trees that surpasses this height. Therefore, unlike many of the observations in other land types, observations in forests lie inside the canopy. In the research 102 from 230 sites (approx., 44 %) are classified as forests. Consequently, for forests sites, we wonder how much sensitive the land and surface couplings are to the height in which the surface heat fluxes, temperature and humidity are measured. We would expect that using measurements located right above the canopy would reflect different land and atmospheric coupling. We do acknowledge the challenge of comparing the diverse land-types considered in the study within the same methodological framework. Nonetheless, we would appreciate a justification of using the 2-m height measurements for forests or at least addressing the special advantages and shortcomings of such approach for forests. In addition, we wonder how the inclusion of these observations affect the general conclusions for the land-atmospheric interactions. For instance, are patterns more easily generalizable (in figures 2, 3, 4 and 5) when forests are excluded?

  In the new manuscript it has been written that the observations of FLUXNET2015 data have been assumed to be taken at 2 m above the canopy. We do not fully understand what does "assume" means here. Does it mean that it is unknown the convention of height of FLUXNET2015? For instance, is it unknown if the documentation means 2m height from the surface or 2m height above the canopy? We would like some clarification or contacting FLUXNET for further details about the height of measurements with respect to height of the trees.

  ➔ *Canopy height and meteorological measurement height information for the flux tower observations (e.g., FLUXNET2015 and Ameriflux) is not always provided, The measurement height for such sites is assumed as the mean value across the sites that report this information. The average measurement height is 7.5 meters. The results do not change much as a result, but the component vectors in the mixing diagrams have been corrected in Fig. 4 and 5. The notation of '2-m' is replaced to 'near surface' across the manuscript and the formulation of Eq. 5 is also corrected to use the average measurement height. This description is added in Lines 133-136:*

    "As the instrument height varies among flux towers, this study computes the measurement height (h) as the difference between reported height of observation and average canopy height. When 83 observation sites (36% of the total) do not provide both heights, we assume the measurement height as the averaged value across the other available sites (7.5 m). All flux measurements are taken above the canopy while few meteorological sensors are below the canopy top."

<u>New comments connected to the revised manuscript</u>

- Regarding the addition of figure 6 and section "4.4. Canopy effects", we think it adds value to the manuscript. We think figure 6(b) could be omitted because the processes explained in the section can be visualized from figure 6(a), so we do not think it adds value to the analysis. In addition, it is a new way of visualization that has not been used before in the manuscript so it may disorientate the reader.

➔ *We agree with reviewer's comment that Fig. 6b could be omitted. Fig. 6a is now simply Fig. 6.*

- Section 5 is called "Conclusions" whereas section 6 is called "Discussion". We think it must be the other way around.

  ➔ *Based on the reviewer's comments, we changed the titles of sections 5 and 6 to "Summary" and "Conclusions", respectively.*