Reviewer 1:

General comments:

The paper explores a novel approach for estimating ET, or more specifically, several parameters required to model ET using satellite imagery, from remote sensing data. While the approach has limitations, which are acknowledged by the authors, it has the potential to be a useful tool for enhancing ET estimates over large scales by providing observation-based estimates of the parameters needed for ET modelling. Although the paper presents the initial evaluations of the approach and additional studies are needed to refine and confirm the utility of the methods presented, the potential benefits of the technique to the modelling community are sufficient to merit publication. Nonetheless, there are several aspects of the study that need clarification of further explanation (see below); the authors are also urged to work with a grammarian to improve the English syntax and structure of the paper.

We thank R1 for their supportive and helpful comments and will modify the text accordingly.

Specific

1. P2, L42: The "scales relevant to decision making" should be defined more rigorously. What are these scales?

The AIRS sounder scales are 100 km but geostationary sounders will come down to the 1 km scale. As a result, we speculate that this method could in principle furnish information from the field to the landscape scale and it is these scales we believe that are relevant to decision making on water resources.

2. P4, L89: These "counter arguments" need to be explained (justified) more fully. For example, the Bowen ratio is determined in terms of gradients in Eqs. 5 and 6. It is not clear how well measurements of temperature integrated over a volume of the atmosphere will represent the true gradient, particularly if these quantities vary nonlinearly with height? Also, relationships the authors present are based on assumptions of similarity (e.g. the eddy diffusivity terms are not shown in Eq. 5) given the course horizontal and vertical spatial scales of the soundings data, the authors need to justify the assumption that the measurements of the turbulent transport of heat and water are within the surface boundary layer and it conforms to similarity theory.

These same arguments can be made for towers in that any sensor measurement contains some element of vertical integration of the observation. So, although the scale at which we are applying the method is unusual, the principle of using observations as point estimates is no different. However, sounders integrate signal horizontally over scales of thousands of square kilometres and hence benefit from stronger horizontal averaging characteristics in the measurement relative to the vertical. We accept they also suffer from ambiguities in the vertical integration of signal but the effective vertical sensor separation is large. The point about nonlinearities in the vertical is important though. We correct for the environmental lapse rate using the geopotential height although we accept there could be other things we are not correcting for.

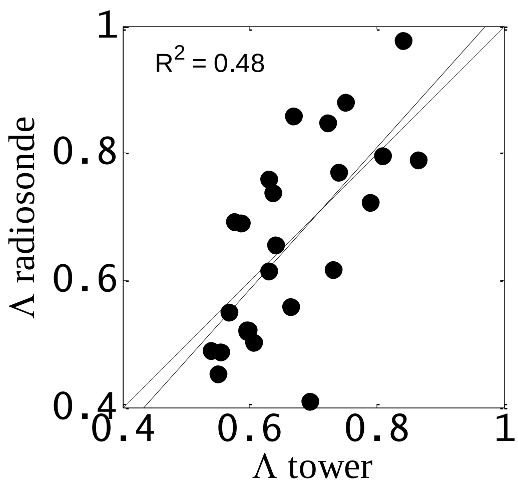
3. P6, L129: The authors use the 1000 mb and 925 mb levels to estimate the temperature and pressure gradients. Since especially the latter is unlikely to be within the constant flux layer, it is questionable that the underlying assumptions of similarity, etc. are valid. Or, why they should be expected to be valid? While the authors point to studies such as Swinbank and Dyer (1967), those analysis are based on the assumption that the sources and sinks for heat and moisture identical and uniformly distributed. The large spatial scales corresponding to the sounding measurements place that assumption in doubt.

[In response to R2 also on related points] We accept that sampling the boundary layer at this scale is likely to lead to non-idealised conditions in terms of the way the Bowen ratio was originally conceived (although such idealised conditions seldom exist at any scale), so the question is, to what extent do they impede the method? Obviously, this is a non-trivial question but we propose the following defense.

The soundings we utilize are for a 13:30 overpass time. Although not universally so, the turbulent boundary layer tends to be at its most mature around or soon after this time of day and the average depth of the turbulent boundary layer should extend well beyond the 925 mb level in all but high latitudinal environments (Fisch et al., 2004). Therefore, although the system we are sampling is not the constant flux region near the surface, in affect we have a surface source region (sampled by the 1000 mb sounding) exchanging with a well-mixed volume (sampled by the 925 mb sounding). The flux exchange between these two should be approximately linear and equivalent in the concentration differences between the two providing we are near dynamic equilibrium (i.e. the turbulent boundary layer isn't growing/contracting excessively) and that additional fluxes into and out of the boundary layer (including phase changes) are small relative to the surface sourced fluxes of heat and water vapour. The 1000 mb level would correspond to heights of approximately 10 m and the 925 mb level to approximately 600 m, thus Δz = 590 m and the effective measurement height is, therefore, at approximately 300 m. This contrasts with horizontal scales of the soundings of 100 km i.e. nearly three and a half orders of magnitude larger. Therefore, although advective fluxes affect the vertical profile, they should be small relative to the effects of vertical exchange on these scales. The principle difficulty as far as we can ascertain is the effect of phase changes associated with cloud formation, producing latent warming of the boundary layer whilst removing water vapour. Providing this happens above the 925 mb sounding we anticipate it being less of a problem, but if it happens below it then clearly this is problematic. Of course, this also impacts on the estimation of the net available energy.

In an attempt to reassure the reader about the validity of the assumptions we are making we have also tested the same methodology over a surface flux measurement site in the central United States where both the radiosonde measurements and eddy covariance flux observations were available. Bowen ratio was estimated from the air temperature and dewpoint temperature measurements of the radiosonde observations using the same methodology as described in the manuscript. The Bowen ratio was then converted to the evaporative fraction and these were then compared with the eddy covariance derived evaporative fractions. The figure of this comparison is given below and shows a fair degree of correspondence between the two. This analysis produces a modest correlation (R^2 = 0.48), reasonably low RMSE (0.11) and mean absolute percent deviation (14%) between radiosonde derived

evaporative fraction and tower observed evaporative fraction. This figure will be added and explained in the revised manuscript.



4. P8, L187: With the possible exception of tall tower data, the source area of EC flux measurements is typically much less than 10 km2 so the mismatch in the scale of the source areas between the tower and soundings data likely exceeds three orders of magnitude. The smaller source area of EC systems would tend to mask the impacts of spatial heterogeneity that would be seen at the coarser resolution of satellite data.

We agree. This shouldn't impact on the flux measurements themselves for the horizontal vs. vertical scales arguments made above. The most important implications for spatial heterogeneity in the present context is that, in addition to complicating comparison with tower data, relating these observations to unique surface characteristics is likely to be problematic.

5. P10, L219: A correlation of 0.34 may be statistically significant, but it suggests only a modest relationship between the satellite and tower-derived estimates of evaporative fraction. Moreover, from Fig. 2 it appears that the maximum tower measurements of latent heat flux are about 350 Wm-2; in that case, an error of 79 Wm-2 would be an error of 20% to 25%.

Point taken, although we are not claiming high accuracy and precision for the method at this stage. Indeed, such things may be unattainable with this method - time will tell.

Technical

1. P2, L36: It's less managing climate change than its effects.

We will change it as 'to manage the effects of this change'

2. P3, L53: The sentence beginning here is awkward and confusion. It needs to be rewritten.

Will change it to 'The most common approach centres on assuming a physical model of evaporation given many of the variables required to compute evaporation using these models are available directly as satellite products (e.g., land surface

temperature, vegetation index, albedo etc.) (Choudhury and Di Girolamo, 1998; Mu et al., 2007, 2011).'

3. P5: The authors use a number of non-standard symbols (e.g. using phi to represent available energy and P to represent water vapour pressure). Using more typical symbology would avoid confusion for the reader. Also, the authors need to confirm that all of the symbols are consistent. For example, the authors inter-mingle "z" and "Z" to represent height.

We are happy to try and align with the norms of the field and note some inconsistencies in our usage of terms, although our general notation is in line with some classic studies in this area (e.g., Raupach, 1995).

4. P6, L121: The word "plain" is misspelled.

Noted.

References:

Fisch, G., Tota, J., Machado, L. A. T., Silva Dias, M. A. F., da F Lyra, R. F., Nobre, C. A., Dolman, A. J., and Gash, J. H. C.: The convective boundary layer over pasture and forest in Amazonia. Theoretical and Applied Climatology, 78, 47 – 59, 2004.

Raupach, M.R. (1995). Vegetation-atmosphere interaction and surface conductance at leaf, canopy and regional scales. Agricultural and Forest Meteorology, 73, 151 – 179.