General comments

This manuscript describes 16 months of radiation and surface energy balance fluxes observed at a mixed land-use site near Beijing (Miyun). The highly complex site is classified as suburban, consisting of moderately dense residential buildings (6 storeys) to the west of the tower and a mixture of farmland, roads and buildings to the east. The observed radiative, sensible heat and latent heat fluxes are supplemented with estimations of the remaining energy balance terms using simple models. The methodology and tools used are similar to other urban climate studies, but the characteristics of the Miyun site are rather atypical which makes this a potentially interesting addition to the urban literature.

Analysis and interpretation of this complex site is challenging. To some extent the methodology takes these challenges into account (e.g. via source area analysis, consideration of storage fluxes for farmland as well as the built environment, estimation of anthropogenic heat emissions, measurements of soil moisture). However, the analysis is quite brief which limits the impact and usefulness of the study. I believe this is partly due to the broad approach used and missing further investigations which would add weight to the findings.

One of the main conclusions of this work seems to be the range of factors affecting the surface energy fluxes, yet the bulk of the paper only really assesses monthly and average monthly diurnal variations with the effects of different meteorological conditions, surface cover fractions, anthropogenic influences, and spatial and temporal variability all combined.
In addition, these various factors don’t seem to be dealt with consistently. For example, the prevailing wind is from the east (so the source area is predominantly farmland, small residential buildings and roads) but large anthropogenic heat fluxes are calculated for the study site as a whole. Also, despite the variation in building and tree heights around the site it seems a single value has been used for the roughness length and displacement height for all wind directions. The storage heat flux estimation attempts to account for the different source area characteristics but most of the analysis (except Fig 11 and 12 right at the end) is very broad (i.e. monthly differences without any stratification by wind direction). In my opinion, the manuscript needs to deal with these challenges in a more consistent way and refine the analysis to provide insight into these physical controls.

Substantial revisions are needed in order to develop this manuscript into a valuable contribution. The language is mostly understandable but often difficult to follow and there are many mistakes. Therefore the revised manuscript will need English language editing (both for small errors and to improve readability and structure). The main messages of the paper should be given more attention, in particular the comparisons with previous suburban studies. This is one of the purposes of this study but in many places the analysis is very brief and offers little new insight (see below for details). At the same time, unnecessary and unimportant details should be removed (e.g. overly detailed descriptions of building heights; Table 1) or conveyed more efficiently (e.g. with a figure, table or single sentence), or, if important, these details would be better consolidated with the main text (e.g. some of the supplementary information).

I have tried to make suggestions for where particular changes are needed. These range from major topics to more minor issues. I have also provided some language corrections for the first half of the paper but stopped after Section 3.1 as there are too many to expect a reviewer to detail them all.

**Specific comments - major**

**Single values used for roughness length and displacement height?**

L158-72: There is a lot of text given over to the description of this complex site – the variation in land cover, building height and vegetation height. Since there is considerable variation in building and vegetation height spatially (around the tower) and temporally (with season), can it really be justified to use single value for the roughness length and displacement height? Or are these single values only used as an initial estimate for the footprint model? Although the initial estimates for the iterative footprint calculation may not be crucial, the roughness length and displacement height used for the flux processing and to calculate the stability parameter could impact the results. It is not clear what has been done here.

**Further analysis of radiation would be useful**
L209-249: This part could be made more concise as there are not very many new or interesting results here that go beyond what is widely known. L237-239 should be deleted if it is irrelevant or rephrased so that the reason for including this statement is clear to the reader. The period with very little incoming shortwave radiation in early June 2013 is potentially interesting and could be investigated further (e.g. with a zoomed-in timeseries and further discussion). This would also help to show that the observed data realistically capture the conditions on these days and avoid any uncertainty in the mind of the reader that there could have been an instrument issue at this time. In L247-9 December is highlighted as having the lowest net radiation because of the snow cover, but other winter months have similarly low values and snow was also present for the January data. The text should be reconsidered to provide a more balanced and accurate assessment of what the data shows (i.e. does snow cover really play as important a role as is implied?).

Figure S1 could be moved to the main text and analysed in more detail, for example is this asymmetrical diurnal pattern in albedo also seen for days without snow cover? What is the physical explanation? Could this (and the need to set the shortwave components to zero at night) indicate a potential levelling issue with the radiometer?

**Further analysis of anthropogenic heat flux needed**

L259-263: Little evidence is provided in support of these QF values being judged reasonable and inclusion of these data is of little value without further analysis. What could be the reason for differences compared to the other suburban studies suggested here? How does the Miyun site compare to the Basel, Montreal and Swindon sites in terms of temperature (heating/cooling degree days), population density, use of air conditioning. Why not add a few sentences about the relative contributions to QF (i.e. what proportion is from buildings (for heating and for cooling), traffic and human metabolism)? How do these contributions compare to values reported for other sub/urban sites? In Appendix A it seems to be suggested that the traffic contribution may be overestimated (L530-1) – how does this fit with L263?

**Storage heat flux methodology and analysis**

It is good to see that a lot of consideration has been given to the estimation of this important term in the urban surface energy balance and the discussion includes potentially useful results. However, this section needs to be better structured so that the key findings come across. The language needs to be improved (in terms of the grammar of individual sentences as well as the overall story) and it the manuscript would be much more readable if the important results from the supplementary information were included in the main text.

The methodology (i.e. various approaches used to calculate QS_OHM) should be more clearly set out to guide the reader – at the moment the text jumps between the impact of seasonality, time of day, source area and methodology from sentence to sentence. More
explanation of Fig S2 is needed so that the reader immediately realises that the monthly and diurnal variation seen is due to the combination of (i) different seasonal coefficients and (ii) different source area composition. A couple of sentences explaining the results would also help (e.g. ‘a1 is larger in the afternoons because...’).

L275-7: Given the expectation that QS_RES is an upper limit for the storage heat flux and the well-known issue of energy balance under-closure, is the neglected biomass heat storage flux the main explanation here? The text implies that this is the message the reader should take away, but from the results shown and brief analysis I am not yet convinced that this is an important term or one of the main reasons for the differences. This paragraph needs to better introduce some of the other considerations that follow.

L278-301: It is interesting to see the consideration of other possible factors influencing the storage heat flux estimates such as energy balance under-closure and uncertainties in QF. Differences in footprint between radiometer and flux tower as well as the performance of the OHM model shown in Fig B2 (especially concerning the large values of observed soil heat flux that are not reproduced by the model) should also be discussed here.

L279-83: I am not sure the comparison with the Swindon site is especially helpful, especially if no physical explanation is given for the difference. It would be much more interesting to use the Miyun dataset to investigate some of the variability with rainfall or radiation (e.g. as described in L290-2, L295).

L284-7: The second calculation of QS_RES_v2 as a lower limit is helpful and it would be nice to see this brought into the main text to make the OHM vs RES comparison more robust.

L287-290: It is not at all clear how this conclusion has been reached.

L297-301: The sharp drop in QS_OHM that is seen in Fig 5b around sunset is seen all year round not just in summer (actually it seems to be even larger in winter than summer). This drop occurs for all the QS_OHM estimates but not for QS_RES. What could be the possible explanation for this and is it really likely to be a ‘true’ feature (as suggested by the text) or might it be an artefact of the estimation approach? I am not yet convinced. Was this drop seen in the observed soil heat flux data used in Appendix B? It could be worth checking the terms of the OHM equation individually to see what is producing this effect. If you used different timesteps for the OHM calculation does this feature still appear?

The implied connection with the valley/mountain wind is not clear (L300). Similarly, what should the reader learn from the final sentence about clear/rainy conditions (L301)? Can you provide a physical explanation for this observation?
**Description of turbulent heat fluxes**

It is interesting to see the differences in seasonal variability compared to many of the previous sub/urban studies. However, more analysis is required to understand why this occurs.

Are the seasonal patterns shown a result of the land cover (farmland versus urban) or the local climate (i.e. a rainy summer season)? More direct comparison with other studies in similar climates (e.g. Hong et al., 2020, the other Beijing studies mentioned in L53-55) would be very helpful in understanding the physical reasons for these observed differences and therefore transforming this study from only locally relevant to being more widely useful for interpreting observations in other locations, improving process understanding, developing model parameterisations, etc.

Also, it is not clear whether these results essentially describe the farmland (i.e. would very similar results have been obtained if the flux station was installed in the fields?) or whether there are also urban effects (e.g. are the sensible heat fluxes enhanced by anthropogenic heat emissions?). It might be worthwhile reproducing Fig 6 using only data when the source area was over the farmland and (if there is enough data) also when the source area was over the more built-up area (similar to Fig 11, 12). In general the text could also give a clearer impression of the importance of the non-vegetated areas when the source area is to the east of the tower (perhaps adding site photos to an earlier section would help).

Similar to the radiation section, there are a lot of statistics given here but not many new insights. Given the relevance of the farmland to the east, finer-scale temporal analysis linked to the state of the crops, irrigation schedule, soil moisture and meteorological conditions would be a useful addition. This seems to be the purpose of Section 3.4 but it is only done very broadly.

L358-64: This simple comparison with two very different sites is not particularly useful.

L381-2: This comparison is much more interesting and perhaps could be further developed by concisely incorporating the relevant information from the references so that the comparison is clearly extended beyond Tokyo and Oberhausen. Similarly in L412-6 – it is difficult for the reader to get the importance of the message if they have to go and look in references and basically make the comparison themselves.

L408-428: These paragraphs should really form the main contribution of the manuscript, where the reader learns how these factors relate to the observations set out in previous
sections. However, there is hardly any analysis here.

Fig 9 and 10 are briefly mentioned without much interpretation. The meaning of Fig 9 is unclear to me (not helped by the confusing symbol and colour choices). What is the reader supposed to compare with what? Adding two or three sentences summary in the text could be helpful, although these data don’t appear to show a consistent pattern. Since these data are at a different timescale to the monthly fluxes, it is very difficult for the reader to see the connections briefly mentioned in the text.

Despite crop growth being relevant (L423) I don’t see this information presented anywhere.

Similarly Fig 10 is referred to in single sentence (L421) and the reader has to do all the interpretation. Consider combining some of Figs 7-11 and adding additional information (such as crop growth and harvesting dates, irrigation times) or averaging the observations over relevant time periods so that the reader can see that connections suggested in the text are supported by the data – i.e. can see how the surface fluxes respond to the supposed controls.

L429-35: This short paragraph mentions two figures but offers no analysis. Fig 11 is useful and relevant and should be analysed in more detail in the text. Why are the fluxes normalised by $Q_{\text{down}}$ here and not $Q^*$ as above? Are similar or different patterns seen if $Q_{\text{down}}$ is used for normalisation in Fig 7 or $Q^*$ is used for normalisation in Fig 11. Might it be worthwhile to add another panel for QS to Fig 11 (e.g. to help provide evidence for or against the hypothesis in L378-9)? The division into wind sectors needs to be properly introduced (not just mentioned in the caption of Fig 12).

It’s not clear what the reader is supposed to learn from Fig 12. Don’t these two plots basically suggest very little dependence of Bowen ratio on time since last rainfall and little difference between the two sectors? The text L433-5 doesn’t really make sense and is more of a hypothesis which does not seem to be convincingly supported by the data.

**Conclusions need to offer something more than a repetition of results**

The conclusion section is simply a repetition of the results. Too many statistics are given here and there is no summary, synthesis or bigger picture. The final paragraph refers to the importance of these kind of observational studies, but unless the previous sections can be further developed to provide insight into these key factors (i.e. to improve understanding of processes) this study cannot claim to be useful for any of the purposes given in L478-81.
Specific comments – minor

L46: Would be good to point out that this extreme example from Basel is a car park site, and not a typical urban neighbourhood site.

L53-60: It's not very clear to the reader which of the sites are being referred to as ‘urban’ and ‘suburban’ here. Then in L60, it’s not clear if the suburban site (Miyun) is the same as introduced above or a new site that has not previously been mentioned in the text. Please rephrase this paragraph to avoid confusion.

Fig 1: The aerial image in (b) and land cover in (c) don’t seem to match up very well close to the tower – it looks as though the tower position is 100-150 m further north in (b) compared to (c) and possibly on a building in (b) but on grassland in (c). It would be good to state in the text the land cover directly at the flux station (was the tower mounted at ground level?) and for the nearby weather station. It would also be helpful to add the weather station to the maps in Fig 1. The colour bar in (a) needs an axis label and units (or a more complete description in the figure caption). Using consistent colours to represent the land cover in panels (c) and (d) would help the reader.

L72-77: A lot of heights, distances and directions are given here and it is not easy for the reader to get a clear impression of what is important for this site or match the descriptions to the aerial imagery or land cover map in Fig 1. If the building heights and distances in each direction are very relevant, consider adding a building height map as an extra panel in Fig 1; if this information is not especially relevant (which seems to be true) then less detail can be given here. In any case, the relevant building height information should be consolidated with similar information given in Section 2.4. The distance to the tallest building should be mentioned. It might also be helpful to show where each of the crops are (perhaps in the extra panel in Fig 1) as the reader has to guess where the wheat and maize mentioned in L166 are growing.

L170: Are these respective anemometric values for z0 and zd? It is not completely clear from the text.

L181-4: These seem unrealistically precise distances for a footprint model which can only provide an indication of the likely source area. Probably these should only be given to two significant figures at most.

L201-2: It’s not clear which definition has been used for daytime/night-time – the incoming shortwave radiation or the times given.
Table 1 is very difficult to interpret, the caption is difficult to follow and I’m not sure the information is really necessary in this much detail. Also the table does not include the information suggested by the text concerning quality control procedures (L102-3) or definitions (L445-6).

L267-70: It’s not clear to me how this conclusion was reached from looking at Fig 5a. What is the calculated annual gain/loss?

L270-72: The meaning here is unclear - the text needs to be rephrased and possibly further explanation added. It’ not clear what this conclusion about wind direction dependence is based on.

L273-5: Where are the mean and median values shown?

L363: How was this vegetation fraction of 34% calculated? Is it a source-area averaged value?

L533-8: It’s not clear in this paragraph which population data has been used for the QF estimation and what the implications for the analysis are.

**Technical corrections/language edits (not exhaustive!)**

L22: suggest changing to ‘effects of climate change’ since ‘urban’ is already mentioned several times in this sentence

L31: It’s not clear what ‘These’ refers to – please rephrase explicitly.

L35-6: Again it’s not clear what ‘these’ refers to – consider rephrasing so that the start of the next sentence (“The EC method sites...”) is also better expressed

L41: an ‘and’ appears to be missing here

L42-44: This sentence does not make sense. Please rephrase
L44: ‘The ratio of impervious...’

L56: I don’t understand what is meant by ‘both individually and concurrently’

L60-1: This sentence does not quite make sense. Suggest rephrasing as ‘We focus on the impact of site characteristics and environmental factors on energy partitioning.’ However, ‘environmental factors’ is quite vague so it would be good to be more explicit here.

L87: Suggest rephrasing as ‘and 90% eddy covariance source area’

L103-4: ‘the anemometer was installed with respect to magnetic north’

L111: I think you only need to mention the month-long gap here (the other periods of 1-2 days are probably not important).

L129: There is no bare soil fraction mentioned in the land cover description or shown in Fig 1(d).

L138: Delete ‘The’

L144-8: This doesn’t make sense grammatically – please rephrase.

L159-62: This is difficult to read and doesn’t quite make sense – please rephrase

L191: ‘difficulty’

L191: ‘use’ or ‘used’

L194-6: Please rephrase
L204-5: I don’t understand the meaning here. Please rephrase.

L220: ‘becomes increases’ à ‘to increase’, ‘decreases’ à ‘decrease’

L223: Is ‘variability’ really meant here, i.e. are the statistics given a measure of variability?

L386-7: This sentence doesn’t belong here without a ‘(see Section 3.5)’ or similar.

L453: I’m not sure it makes sense to call $Q^*$ a component of the radiation balance.

References

Hong J-W, Lee S-D, Lee K, Hong J (2020) Seasonal variations in the surface energy and CO2 flux over a high-rise, high-population, residential urban area in the East Asian monsoon region. *International Journal of Climatology* 40: 4384-4407
https://doi.org/https://doi.org/10.1002/joc.6463