Responses to Editor' Comments:

Thank you for submitting your manuscript entitled "The Effects of Surface Heterogeneity Scale on the Flux Imbalance under Free Convection" [Paper #2018JD029550R] to Journal of Geophysical Research - Atmospheres. I have received 3 reviews of your manuscript, which are included below and/or attached. Based on the reviews, your manuscript may be suitable for publication after minor revisions.

The feedback provided in the reviewer assessments of your manuscript is important and should be taken into account as you complete your revision.

Please submit a revised manuscript that responds to or addresses the comments by April 13, 2019.

In your revision, please follow our Checklist and use our Templates for the main file and any supporting information. Please provide the following:

1. A response to reviewer file that lists each major comment and describes how the manuscript has/has not been modified in response to those comments.

2. A copy of the manuscript with the changes noted (e.g., highlighted, "track changes," italics or bold changes).

3. The final revised manuscript with the changes incorporated and separate final figure files (figure parts should be combined into a single file), which will be used for publication if the manuscript is accepted. If final figures are already uploaded, they can be easily copied over to the next revision version.

4. If any, supporting information text figures, captions, and small tables in a single PDF file using AGU's template. Large data tables, and multimedia should be uploaded separately in their native format (preferably .xlsx or .csv for tables).

Response: We deeply appreciate your help in the review process and we thank the three reviewers for their constructive suggestions and detailed comments. Based on these suggestions and comments, we have revised the manuscript as detailed in the point-by-point responses below. The reviewer's comments are in *italics* and the texts in the revised manuscript are underlined.

Responses to Reviewer 1

I appreciate the work that the authors have put in revising their paper. It has certainly become much clearer. The comparison with existing flux corrections is adequate, and they have answered my major comments. It is now also clear to me why the authors equate the measured temporal flux by the high-pass spatial flux, and it's good that they spell out that they rely on Taylor's hypothesis (as is customary).

However, I think the method could gain some more clarity by comparing the approximation to the (LES) literature and say what the missing term would be according to the terminology in e.g. Steinfeld et al (2007). In my understanding, the missing flux would be the sum of the storage change and the TOS:

$$\int_{0}^{k_{ec}} F_{w\theta}(k, z_m) \, dk = \left[\overline{w}\overline{\theta}\right] + \left[\int_{0}^{z_m} \frac{\partial\theta}{\partial t} \, dz\right]. \tag{1}$$

This follows from partitioning the surface flux into the spatial covariance and the storage change (for the whole domain average there are no other terms), and then decomposing the spatial covariance (for a simulation with periodic boundary conditions) into the temporal covariance and the TOS. I think this is instructive, especially in the light that the authors later assume that

$$\int_0^\infty F_{w\theta}(k,z) \, dk = \left[\overline{w^{\prime\prime}\theta^{\prime\prime}} \right]. \tag{2}$$

without the storage term, which does not look entirely consistent to me. The latter term is usually small and it may make the analysis easier when it is dropped, but it would be good to spell it out. **Response:** We thank the reviewer for the overall positive reception of our manuscript and also his/her very constructive and valuable suggestions, which led to significant improvement of the quality of our manuscript.

The reviewer is correct that for a simulation with periodic boundary conditions, the spatially averaged heat budget equation is (Finnigan et al. 2003):

$$\left[\overline{HFX}\right] = \left[\overline{w'\theta'}\right] + \left[\overline{w\theta}\right] + \left[\int \frac{\overline{\partial\theta}}{\partial t} dz\right] + \sum_{1}^{4} \left[\overline{v'\theta'}\right] + \sum_{1}^{4} \left[\overline{v\theta}\right].$$

As a result, the storage term is also part of the missing flux (i.e., contributing to the flux imbalance). Given that $[\overline{w''\theta''}] = [\overline{w'\theta'}] + [\overline{w}\overline{\theta}]$, the derivation of which can be found in Steinfeld et al (2007) and Zhou et al. (2018), the effects of the storage item and the integrated horizontal flux divergence items (i.e., the last two items in the above equation) are represented by the f_1 , which is expressed as follows:

$$f_1 = \frac{\int_0^\infty F_{w\theta}(k, z) \, dk}{\left[\overline{HFX}\right]} = \frac{\left[\overline{w'\theta'}\right]}{\left[\overline{HFX}\right]} = \frac{\left[\overline{w'\theta'}\right] + \left[\overline{w\theta}\right]}{\left[\overline{HFX}\right]}.$$

To address the reviewer's concern, we add the following description for f_1 in Sect. 2.2 (Lines 146-148 on Page 5):

"As such, the term f_1 represents the effects of the integrated storage term and horizontal flux divergence term in the spatially averaged heat flux budget, leading to inequality between the flux at height z and the surface flux (Finnigan et al., 2003)."

Some very minor remarks

1. Missing preposition: velocity scale that can be represented by

Response: Revised as suggested.

2. Line 81: I would simply remove the ", x" from the exponent and use a "."

Response: Revised as suggested.

Responses to Reviewer 2

My comments have been addressed appropriately in the revised version. However, I still have one more major concern with the equation that is proposed to describe the relationship of flux imbalance with driving factors (see Eq.15 in the revised version). The equation shows that the flux imbalance follows a linear relationship with height z, but this is definitely not correct when the surface heterogeneity scale is larger than 1200m (see the case B1200 indicated by the solid red line with red circle in Figure 7a). This means that the authors should use a more appropriate function to describe the relationship of flux imbalance with height (z).

Response: We thank the reviewer for this comment. Indeed, as the reviewer correctly pointed out, the linear relation between the flux imbalance and z does not apply for all cases based on Figure 6a (we think the reviewer means Figure 6a instead of Figure 7a). However, due to the dependency of f_2 , or more precisely $l_w/(UT)$, on the z (see Fig. 10), the relation between the flux imbalance and z (see Eq. 15) is non-linear. To avoid further confusion, we add the following comment in the revised manuscript (Lines 499-502 on Page 14):

"It should be pointed out that while Eq. (15) may seemingly indicate that [I] changes with z linearly, the fact that $-z_i/L \times l_w/(UT)$ also depends on z (Fig. 10) makes it a non-linear function in terms of z. One can see from Fig. 6 (a) that [I] does not change with z in a linear fashion, especially for cases with large heterogeneity scales."

Several minor comments are given as follows.

1. The study is focused on the free convective conditions (i.e., geostrophic wind is set to be zero in all simulations), but U is still included in Eq.(15). What are the typical values of U for the simulations included in this study?

Response: We thank the reviewer for pointing out this. U in our manuscript is NOT the horizontal mean velocity since the horizontal mean velocity would be extremely close to zero under free convective conditions. As mentioned in our manuscript, all velocity scales are first calculated locally and then averaged over the horizontal plane. Hence, U is calculated locally as $\sqrt{U_x^2 + U_y^2}$ and then averaged over the domain. To distinguish it from the horizontal mean velocity used in shear

convective conditions, we call it the "mean wind speed U". The typical values of U are added to Table 2 in our revised manuscript.

"Table 2. Summary of the experiments along with spatially and temporally averaged ABL characteristics in case B and H with standard deviations. The z_i represents the boundary-layer height and the u_* and w_* indicate the friction velocity and convective velocity, respectively. *HFX* and w_e (i.e., $\Delta z_i / \Delta t$) are the surface heat flux (i.e., the "true flux" in our study) and the mean entrainment rate, respectively. *U* denotes the mean wind speed, which is first calculated locally and then averaged over the horizontal plane. \overline{e} is the square root of turbulent kinetic energy (*TKE*).

Case	Heterogeneity	Zi	U	ē	<i>u</i> _*	<i>W</i> *	HFX	w _e /w*
Name	scales (m)	(m)	(m/s)	(m/s)	(m/s)	(m/s)	(W/m ²)	
B2000	2000	1415±157	1.26±0.73	1.33±0.95	0.30±0.22	2.52±1.56	457±869	0.018
B1200	1200	1388±142	1.10±0.59	1.27±0.81	0.29±0.22	2.50±1.55	460±890	0.017
B550	550	1306±147	1.53±0.76	1.47±0.87	0.33±0.22	2.55±1.54	463±913	0.013
B240	240	1287±135	1.66±0.79	1.58 ± 0.80	0.37±0.21	2.52±1.51	487±937	0.012
H2000	2000	1444±189	1.59±0.82	1.47±0.98	0.34±0.22	3.05±1.51	512±795	0.016
H1200	1200	1436±178	1.50±0.74	1.44 ± 0.88	0.34±0.21	3.08±1.50	515±811	0.015
H550	550	1384±172	1.76±0.86	1.56±0.95	0.37±0.22	2.99±1.52	520±828	0.014
H240	240	1379±160	1.88±0.89	1.63±0.91	0.39±0.22	2.92±1.53	521±848	0.014

".

2. k_ec should be defined as it appears at the first time (see Line 206?) rather than in Line 429. **Response:** In our manuscript, k_{ec} first appears at Line 138, where we have defined it as follows: "we assume that the EC method can only sample eddies that are larger (smaller) than a critical wavenumber (wavelength), which is referred to as k_{ec} (Fig. 1a).".

3. Line 429, "represented" should be "represented by"?

Response: Revised as suggested.

4. "buffer zones" hasn't discussed in the preceding sections, why do the authors discuss this topic in "Conclusion" section? Should it be moved to "Discussions" part?

Response: We thank the reviewer for pointing out this. Following the reviewer's suggestion, in our revised manuscript, the discussion on the "buffer zones" has been moved to Discussions part (i.e., Sect. 5.3).

Responses to Reviewer 3

I thank the author for the very detailed answers to all the comments I did on the first version. The present version of the paper is really improved. The authors added:

- some useful information on the methodology;
- some important characteristics of the simulations which help to understand the differences and strengthen the results;
- analyses of the results.

The paper presents an interesting contribution to the energy balance closure problem and is almost ready for publication from my point of view. I still have very few points which could be clarified.

Response: We thank the reviewer for the overall positive impression of our manuscript. We also thank the reviewer for his/her very constructive suggestions and detailed comments. Below we detail how his/her concerns are addressed in the revised manuscript. The reviewer's comments are in italics and the texts in the revised manuscript are underlined.

Comments

1. Table 2: I thank the author for adding some useful information in Table 2. What is now very surprising in this table 2 is that the way the heterogeneities are generated (B or H) has more effect on the HFX for the same heterogeneity scale (55 W/m2 between H2000 and B2000), than the scale of the heterogeneity (11 W/m2 between H2000 and H240) itself. This point could be commented by the authors. They added a comment about the larger entrainment in B cases which explains many differences between B and H cases vertical profiles (Figure 5), but no explanation is given concerning why same scale heterogeneity can give such a different surface flux and entrainment rate when they are generated differently. The B cases behave really differently than H cases whereas the heterogeneity scales are supposed to be the same. Perhaps that would need a comment?

Response: We thank the reviewer for raising this important question, which led us to think deeply about the differences between B and H cases. As the reviewer points out, there are large differences between B and H cases even when the surface heterogeneity scales are the same. For example, the convective velocity, boundary layer height and friction velocity in B cases are smaller than those in

H cases. The smaller surface heat fluxes in B cases are mainly responsible for these differences.

As for the large differences in surface heat flux between B and H cases (e.g., the 55 W/m² difference between B2000 and H2000) than those between different surface heterogeneity scales in the same case (e.g., the 30 W/m² difference between B2000 and B240), the fundamental cause is the different surface pattern generation methods of B and H cases (see Sect. 2.2). For example, one can clearly see from figure 2 that the surface patterns are quite different between the two cases. As a result, the surface heterogeneity scales calculated by B07 (i.e., Eq. 3) are probably not directly comparable to those calculated by HM09 (i.e., Eq. 4).

We acknowledge that this was not well explained in our original manuscript. In our revised manuscript, we add the following comment in Section 3.2 (Lines 281-286 on Page 9) to make it clear:

"<u>As can be seen from Table 2, the B cases and H cases have very different spatially averaged surface</u> <u>heat fluxes even with the same heterogeneity length scale (note in our study the surface heat flux is</u> <u>not imposed but computed</u>). This suggests that the surface pattern generation method has an important effect on the results and that the heterogeneity length scale calculated by one method may not be directly comparable to that calculated by the other method.".

In addition, in Sect. 4.1.1 we add the following (Lines 374-388 on Page 11):

"As mentioned previously, due to the different surface pattern generation methods between B cases and H cases, the profiles in B cases are different from those in H cases. For example, compared to H cases, θ are smaller in B cases, which is caused by the smaller spatially averaged surface heat fluxes in B cases (Table 2). The larger variation of spatially averaged surface heat flux across B cases compared to that across H cases might be responsible for the larger variation of entrainment rate and hence the larger variation of σ_w^2 . In fact, the differences in spatially averaged surface heat flux between B and H cases with the same surface heterogeneity scale (e.g., the 55 W/m² difference between B2000 and H2000) are larger than those of the same case type but with different surface heterogeneity scales (e.g., the 30 W/m² difference between B2000 and B240). This implies that the two methods are not directly comparable and thus studies examining the effects of surface heterogeneity should be explicit about the method used to quantify the surface heterogeneity scale. Considering that our main goal is to analyze the effects of different heterogeneity scales instead of the effects of different heterogeneity generation methods on the flux imbalance, the differences between B and H cases are not focused in this paper.".

2. Figure 5 and Lines 329-367: I thank the other for the improvement of the analysis of the Figure 5. The effects of the surface heterogeneity on vertical profiles are still very impressive. I agree with the authors that Patton et al (2005) and HM09 also found a reduction of σw/W* with increasing scale of surface heterogeneity. Patton et al found a reduction of about 15% (from 0.7 to 0.6) which is also the case (~ 20%) for H cases. But the reduction is 40% for B cases whereas the ratio of the scale heterogeneity to Zi is even larger in Patton et al. (1.72 in Patton et al. and 1.42 in the present study for 2km heterogeneity).

Response: We thank the reviewer for raising this question. First of all, we note that it is actually not 40% as the reviewer pointed out, since we plotted $(\sigma_w/w^*)^2$ while Patton et al. (2005) plotted σ_w/w^* . In terms of σ_w/w^* , it is a reduction of 20% in B cases (from 0.7 to 0.56) and 10% in H cases (from 0.6 to 0.54), which is actually comparable to the 15% reduction in Patton et al. (2005).

3. Line 338-340: Patton et al. also studied a free convective case.

Response: We thank the reviewer for pointing out this. In our revised manuscript, this reference has been removed, as follows (Lines 346-350 on Page 10): "<u>Similarly, in 0 - 0.5 $z_i \sigma_u^2$ generally decreases with increasing heterogeneity scale, which is different from the results in HM09, where σ_u^2 increases with increasing surface heterogeneity scale. The free convective atmosphere instead of the shear convective atmosphere in HM09 may be responsible for this difference.".</u>

4. Conclusion, line 625: I would move the remark "It is also pointed out that this model is constructed with data in the range 0.03-0.1 Zi". As mentioned earlier in the text by the authors, this is a limitation of the model and perhaps it should be moved in the next paragraph dealing with the limitations of the model?

Response: We thank the reviewer for pointing out this. Following the reviewer's suggestion, this sentence has been moved to the paragraph discussing the limitations of the study, as follows:

"...Third, our model is constructed with data in the range $0.03 < z/z_i < 0.1$. Whether our model can capture the behavior of flux imbalance near the surface ($z/z_i < 0.03$) needs to be investigated using higher-resolution LESs in the future."

5. Appendix D: I would have done the tests on the domain size differently. I think that keeping the ratio of heterogeneity scale to the domain size constant is expected to give the same results. I would have change the domain size without changing the size of the heterogeneities.

Response: We thank the reviewer for raising this great question. Theoretically, there are two ways to examine the effect of domain size on our results. One method is to keep the ratio of surface heterogeneity scale to the domain size constant, which is adopted in our manuscript. The other one is to only increase the domain size without changing the surface heterogeneity scale, as the reviewer suggested. Due to the constant ratio of surface heterogeneity scale to the domain size in the first method, it does not need to redesign the surface heterogeneity pattern. However, in the second method, it needs to redesign the surface heterogeneity pattern, especially over the enlarged regions. Based on the observed differences between B and H cases, it can be imagined that how to set up the surface heterogeneity pattern over the enlarged regions will affect the results, which can be complicated and thus is left for future studies.

We acknowledge that this was not explained in our manuscript. In our revised manuscript, we change the statements in the Appendix C as follows (Lines 700-713 on Page 20):

"In order to examine the effect of domain size on our results, we add an additional case, i.e., case D where all parameters are identical to case B except with a larger domain (i.e., $10 \text{ km} \times 10 \text{ km}$). The patches are also enlarged concomitantly so the heterogeneity scale is doubled. Note that due to the constant ratio of surface heterogeneity scale to the domain size in this method, it does not need to redesign the surface heterogeneity pattern. However, if the ratio of surface heterogeneity scale to the domain size was not kept constant, e.g., only increasing the domain size without changing the surface heterogeneity scale, it needs to redesign the surface heterogeneity pattern, especially over the enlarged regions. Based on the observed differences between B and H cases, it can be imagined that how to set up the surface heterogeneity pattern over the enlarged regions will affect the results, which can be complicated and thus is left for future studies.

Base on Fig. C1 and Fig. C2, we can see that in the large domain, the results are similar to those in

the main paper, suggesting that the main findings of our paper are not sensitive to the domain size if the ratio of surface heterogeneity scale to the domain size remains unchanged.".

Minor comments

6. Line 134: change Figure 1a by Figure 1

Response: Revised as suggested.

7. Line 192: the integral scale is rather an estimate of the largest scale of the turbulent structure (largest eddy in the inertial subrange) than an estimate of the size of organized turbulent structure. Perhaps the word "organized" is misleading here. The author should be consistent in naming these different scales: see lines 483-484 in the Discussion section.

Response: We thank the reviewer for pointing out this. In our revised manuscript, this sentence has been revised as "<u>...which is an estimate of the maximum size of large-scale turbulent structures (i.e., large eddies) in the turbulent boundary layer.".</u>

Line 294: "1-hour" integration time instead of "3-hour" integration time. **Response:** Revised as suggested.

9. Line 429: "where u is a velocity scale that can represent..." remove the "be"? **Response:** We thank the reviewer for pointing out this. In the revised manuscript, this sentence has been modified as "...where u is a velocity scale that can be represented by U...".

10. Line 574: "...reason. First..." replace ":" by "."?Response: Revised as suggested.

11. Figure 2: Could the authors precise in the legend or as a title for each panel (like in Figure 3) what case the panels stand for?

Response: Revised as suggested. In our revised manuscript, Figure 2 is as follows:



Figure 2. Spatial patterns of surface temperature in B cases (a to d) and H cases (e to h).

12. Figure 3: Could the authors put the cases in the same order than in Figure 2? From larger, on the left panel, to smaller, on the right panel, surface heterogeneity.

Response: Revised as suggested. In our revised manuscript, Figure 3 is as follows:



Figure 3. The cross-sections (x-y) of temporally averaged vertical velocity during the final running hour in B cases and H cases at z = 45 m.

13. Figure C1: Perhaps figure C1 is not necessary and comment in the text is enough.Response: Revised as suggested.