

Response to the Referee

We thank the Referee for the careful and constructive comments on our manuscript.

In the following, the format is:

Referee's comment

Our reply (**revised text and place**)

Referee #1

The manuscript describes an updated version of the Building Energy Model coupled to SLUCM. The new model is tested against a series of indoor temperature data in London and building energy consumption in Tokyo.

This is a needed improvement and a welcome addition to SLUCM. My opinion is certainly positive. I believe the manuscript can be improved, clarifying some aspects of the implementation and adding some details, in particular about the sensitivity of the new model to different input parameters.

I describe below what I think must be addressed, following the order of the lines of the paper (not the importance).

We thank the Referee for the positive assessment and for noting that the model development is a useful addition to SLUCM. We have revised the manuscript to clarify the formulation of the new indoor-temperature scheme, its correspondence with the source code, the meaning of the WRF/SLUCM+BEM options, the assumptions used in the validation and what-if simulations, and the limitations related to urban morphology, building parameters, and omitted window solar gains. The details are as follows.

Line 69. "EC was qualitatively overestimated". Overestimation is a quantitative judgement; I do not understand how it can be qualitative.

We agree. The word "qualitatively" was inappropriate because overestimation is a quantitative assessment. We deleted this word.

Changes made in: Introduction, line 69, and Table 2.

Revised text:

"EC was overestimated (Oleson and Feddema, 2020; Takane et al., 2024)."

There are a few things missing (probably implied) in the equations that describe the new version. Among others:

- 1. In equations (9,10,11), Q_{Cr} and Q_{Cw} should be multiplied by the roof and wall surfaces. Same for eq(15-17). The floor is missing.*

We agree that the original manuscript did not clearly distinguish between surface heat-flux densities and area-integrated indoor heat loads. This was confusing because the source code performs the area conversion, whereas the manuscript equations did not show it explicitly.

In the revised manuscript, Q_{Cr} , Q_{Cw} , Q_{Rr} , and Q_{Rw} are defined as surface heat-flux densities (W m^{-2}). The total convective and radiative heat-transfer terms, Q_C and Q_R , are then defined as area-integrated heat loads (W) for a unit canyon building. We retained the original variables Q_C and Q_R and revised Eqs. (9) and (15) as follows:

Revised text:

$$Q_C = W_R Q_{Cr} + 2Z_R Q_{Cw} \quad (9)$$

Here Q_C is the area-integrated convective heat transfer for a unit canyon building (W for a 1-m along-canyon length), whereas Q_{Cr} and Q_{Cw} are surface heat-flux densities (W m^{-2}). W_R and Z_R are the roof/building width and building height, respectively. Because SLUCM uses a two-dimensional street-canyon geometry, W_R represents the roof/ceiling area per unit along-canyon length and $2Z_R$ represents the total indoor wall area of the two canyon walls per unit along-canyon length. The factor 2 accounts for the two walls in the canyon cross-section.

$$Q_R = W_R Q_{Rr} + 2Z_R Q_{Rw} \quad (15)$$

Here Q_R is the area-integrated radiative heat transfer (W), whereas Q_{Rr} and Q_{Rw} are surface heat-flux densities (W m^{-2}). These surface flux densities are defined as follows:

Before the building heat flux is coupled to WRF, the unit-canyon heat-load terms (H_{out}) are converted to land-area-normalised fluxes (W land-m^{-2}) using the canyon width. The land-area-normalised building heat flux and COP are then used to calculate Q_{FB} (Eqs. 5 or 6), and Q_{FB} is fed back to WRF through Q_H (Fig. 1b, bottom).

The floor is not treated as a separate prognostic indoor heat-exchange surface in the present single-layer formulation. We clarified this point. Its contribution enters the model through (i) the indoor heat capacity Q_B , which includes the effective indoor material heat capacity, and (ii) the total floor area A_f used to convert internal heat gains from W floor-m^{-2} to W .

Eqs. (10), (11), (16), and (17) define surface flux densities in W m^{-2} , so we did not multiply these equations by area. We also clarified the length scale used in the natural-convection coefficient. In Eqs. (12)–(14), d_z is the representative length scale for indoor natural convection and is taken as the representative floor height in the present implementation, $d_z = h_f = 3.0 \text{ m}$, rather than the total building height. The area conversion is applied when the surface flux densities are aggregated into the building-scale heat loads Q_C and Q_R in Eqs. (9) and (15).

We also clarified the subsequent normalization step. The heat-load terms in Eqs. (21) and (22) are first evaluated for a unit canyon building. Before coupling the building heat flux back to WRF, the corresponding load is normalised by the canyon width to obtain a land-area-

normalized flux (W land-m^{-2}). This is consistent with the source-code sequence in which H_{in_URB} is first calculated for the unit canyon building and then converted to H_{in_land} by dividing by $ROOF_WIDTH + ROAD_WIDTH$.

We also corrected the wording in step (i) from “roof and floor temperatures” to “**roof and wall temperatures**”, because Eqs. (7) and (8) prognose the innermost roof/ceiling and wall temperatures, not a floor temperature.

We have added W_R to Table 3 to clarify the geometric parameter used in the area conversion.

Changes made in: Section 2.1.2, step (i), Eqs. (9)–(17), Table 3, and the explanatory text defining Q_C , Q_R , Q_{Cr} , Q_{Cw} , Q_{Rr} , Q_{Rw} , W_R , Z_R , A_f , and h_f .

2. How is the total indoor heat capacity computed? In other words, what is the volume of indoor air? Is it an infinitely long building? Is it a quadrangular prism? With which dimensions?

We clarified the geometry used for the indoor heat capacity. SLUCM+BEM v2.0 is formulated per unit canyon length. The representative indoor volume is therefore a rectangular-prism volume with building height Z_R and roof/building width W_R per unit length in the along-canyon direction. Thus, the indoor air volume represented by the model is $V_{in} = Z_R W_R$ per unit canyon length, and the total indoor heat capacity is

Added new Eq. (23):

$$Q_B = (C_p \rho + R_M) Z_R W_R$$

Here $C_p \rho$ is the volumetric heat capacity of indoor air ($\text{J m}^{-3} \text{K}^{-1}$), and R_M is the effective volumetric heat capacity representing indoor materials and furniture, which is specified as a fixed model parameter, $2.0 \cdot 10^5$ ($\text{J m}^{-3} \text{K}^{-1}$).

The term R_M represents the effective volumetric heat capacity of indoor materials and furniture, so Q_B is the total effective indoor heat capacity, not only the heat capacity of air.

Changes made in: Section 2.1.2, Eq. (22)-(23), and the paragraph defining Q_B , V_{in} , W_R , Z_R , A_f , and h_f .

3. I am confused about eq.(16,17), which are not in Kikegawa et al. 2003. These represent the exchange of radiation between the indoor surface and the indoor air. But each indoor surface also receives radiation from the other indoor surfaces. For example, the indoor wall receives longwave radiation from the roof, from the floor, and from the opposite walls. This

is missing in eq. 7,8,16, and 17. With the formulas of the paper, the indoor surfaces would lose much more energy than in reality. View factors are also missing. To define them, we need to have a clear definition of the indoor volume considered.

Thank you for pointing out this ambiguity. We agree that the original description was too short and could be misread as if the indoor surfaces radiated only to the indoor air and did not receive longwave radiation from the other indoor surfaces.

We also added an explicit definition of T_{in} at the beginning of Section 2.1.2. In the revised manuscript, T_{in} denotes the representative indoor temperature in the BOX-type BEM. The model assumes that indoor air, indoor structures, furniture, and inner surfaces are represented by this single indoor temperature, while their heat storage is included in the effective indoor heat capacity Q_B . Thus, T_{in} should be interpreted as a representative indoor thermal state, not as a detailed multi-zone indoor air temperature.

The radiative formulation used in SLUCM+BEM v2.0 follows the simplified indoor-side longwave treatment in the original BEM formulation underlying Kikegawa et al. (2003). The full derivation is described in Kikegawa's doctoral thesis. Because this derivation is not shown explicitly in Kikegawa et al. (2003), we added the relevant explanation to the revised manuscript. In that derivation, the net longwave flux from an indoor-side surface i is first written for an indoor enclosure consisting of m surfaces, including the indoor-side wall surface, inner walls, furniture, and other indoor structures, using Gebhart radiation absorption coefficients a_{ki} :

$$Q_{R,i} = \frac{1}{A_i} \left(A_i \varepsilon_w \sigma T_{wi}^4 - \sum_{k=1}^m a_{ki} A_k \varepsilon_w \sigma T_{sk}^4 \right)$$

Here a_{ki} represents the fraction of longwave radiation emitted by surface k that is ultimately absorbed by surface i , including emission, absorption, and reflection among indoor surfaces. Thus, in the original derivation, the longwave radiation received from other indoor surfaces is not ignored.

The BEM then applies a BOX-model approximation. The indoor structures, furniture, and inner surfaces are assumed to share the representative indoor temperature T_{in} , which denotes the common BOX-model indoor thermal state assigned to indoor air and indoor internal surfaces. Their heat-storage effect is represented by the effective indoor heat capacity Q_B .

Under this approximation, T_{sk} is represented by the indoor temperature T_{in} . Using the Gebhart coefficient summation rule, the above expression reduces to the linearised effective form

$$Q_{R,i} \cong h_R (T_{wi} - T_{in}), \quad h_R = 4\varepsilon_w \sigma T_{in}^3$$

Eqs. (16) and (17) in the manuscript correspond to this simplified form, applied to the innermost roof/ceiling and wall surfaces:

$$Q_{Rr} = h_R (TRL(4) - T_{in}), \quad Q_{Rw} = h_R (TBL(4) - T_{in})$$

Therefore, these equations should not be interpreted as radiation from the indoor surface to a black indoor-air sink. Rather, they represent an effective, linearised net longwave exchange

between the innermost roof/wall surface and the representative indoor thermal environment used in the BOX-type BEM.

We revised the manuscript to make this point explicit. We also clarified the limitation of the present implementation: SLUCM+BEM v2.0 does not solve an explicit indoor radiosity matrix or a prognostic floor-surface temperature. The effects of indoor structures and furniture are represented through the effective indoor heat capacity Q_B , and the floor area is used for internal heat gains. The area integration of the roof and wall radiative flux densities is now made explicit in Eq. (15):

$$Q_R = W_R Q_{Rr} + 2Z_R Q_{Rw}$$

Revised text added after Eqs. (18):

“where ε_w is the indoor-side longwave emissivity (-), σ is the Stefan–Boltzmann constant ($\text{W m}^{-2} \text{K}^{-4}$), and T_{in} is the indoor temperature around which the longwave exchange is linearised. The radiative flux densities Q_{Rr} and Q_{Rw} in Eqs. (16) and (17) should be interpreted as effective linearised net longwave exchanges between the innermost roof/wall surface and the representative indoor thermal environment. In the original BEM derivation, the indoor-side longwave exchange is first written for an indoor enclosure using Gebhart radiation absorption coefficients, which account for emission, absorption, and reflection among indoor surfaces. The BOX-type BEM then assumes that indoor structures, furniture, and inner surfaces are represented by the indoor temperature T_{in} , while their heat storage is included in the effective indoor heat capacity Q_B . Under this approximation, the detailed indoor surface-to-surface exchange reduces to $Q_{R,i} \cong h_R(T_{wi} - T_{in})$, where $h_R = 4\varepsilon_w\sigma T_i^3$. Thus, Eqs. (16) and (17) do not represent radiation to a black indoor-air sink, but an effective net exchange with the representative indoor thermal environment. The current implementation does not solve an explicit indoor radiosity matrix or a separate prognostic floor-surface temperature.”

We also added the following limitation statement in Section 4.4:

“The indoor longwave exchange is also represented by the simplified BOX-model treatment described in Section 2.1.2; the current version does not solve an explicit indoor radiosity/view-factor matrix or a prognostic indoor floor-surface temperature.”

Changes made in: Section 2.1.2, Eqs. (15)–(18), the explanatory paragraph after Eq. (18) and the limitations paragraph in Section 4.4.

- 4. It is unclear how T_{in} is forced to the target temperature. If you have 100% of the buildings with A.C., then (from eq. 3) $H_{out}=H_{in}$. In this case, T_{in} will stay constant in time, and so the initial temperature must be equal to the target temperature. But what is going to happen if H_{in} is negative? Will H_{out} be negative, so heat will be extracted from the atmosphere to keep the indoor temperature constant? In other words, how is the situation treated where there is no need for A.C. to keep the indoor air below the target temperature in summer?*

Thank you for pointing out that the original description did not explain the thermostat logic sufficiently. We revised the manuscript to state explicitly that TARGTEMP is not used as a continuous hard constraint on T_{in} . Instead, it is used (i) to initialize T_{in} at the first time step and (ii) as the thermostat threshold that determines whether the HAC system is active.

$$\alpha = (1 - a)bc,$$

$$\text{cooling: } H_{out} = 0 \text{ if } T_{in}^n < TARGTEMP;$$

$$\text{heating: } H_{out} = 0 \text{ if } T_{in}^n > TARGTEMP$$

For the cooling season, if the indoor temperature at the previous time step is below the cooling target temperature, no cooling is needed and the processed HAC load H_{out} is set to zero. In that case, T_{in} is allowed to evolve freely according to Eq. (22). Conversely, if T_{in} reaches or exceeds the cooling target, the cooling load is calculated for the conditioned fraction of the building. The heating season is treated analogously, with the inequality reversed.

This clarification addresses the case where H_{in} is negative during summer: when the indoor state does not require cooling, the model does not impose cooling merely to keep T_{in} fixed at the target temperature. Instead, H_{out} is zero and the indoor temperature changes according to the net indoor heat load. We also revised the wording to avoid the impression that the target temperature always clamps T_{in} .

Changes made in: Section 2.1.2, paragraph following Eq. (22).

We added the following explanation after Eq. (22):

“At the first model time step, T_{in} is initialised using TARGTEMP. Thereafter, TARGTEMP is used as the thermostat threshold, not as a continuous hard constraint imposed on T_{in} . In the cooling season, the HAC system is activated only when $T_{in} \geq TARGTEMP$; otherwise, $H_{out}=0$, and T_{in} evolves freely according to Eq. (22). In the heating season, the inequality is reversed: the HAC system is activated only when $T_{in} \leq TARGTEMP$. Thus, when the indoor condition does not require cooling or heating, the model does not force T_{in} to remain equal to TARGTEMP.”

5. *On the same line of point d), if only a fraction of the buildings are with A.C., then H_{out} will be less than H_{in} (see eq. 3), and therefore the indoor temperature can change with time. How's the information on the target temperature used in this case?*

We agree that this needed clearer explanation. In the partial-HAC case, TARGTEMP is still used as the thermostat threshold for the conditioned fraction of the representative building. However, because only a fraction of the floor area is conditioned, the processed load H_{out} does not generally balance the total grid-representative indoor heat load H_{in} . Therefore, the representative indoor temperature can still change with time even when HAC is active.

We revised the text to clarify that the conditioned fraction is $\alpha = (1-a)bc$, where a is AB_BUILD_RATIO, b is AC_FLOOR_RATIO, and c is AC_USAGE_RATIO_CL or AC_USAGE_RATIO_HT depending on season. When $\alpha < 1$, the HAC system does not remove

or supply the entire grid-representative heat load. Consequently, Eq. (22) still prognostically updates T_{in} . TARGTEMP acts as the on/off trigger for the conditioned fraction, not as a direct prescription of the grid-mean indoor temperature.

Changes made in: Section 2.1.2, paragraph following Eq. (22), and Table 3 variable descriptions.

We added the following explanation after Eq. (22):

“For partial HAC use, TARGTEMP is still used as the thermostat threshold for the conditioned fraction $\alpha=(1-a)bc$. However, when $\alpha<1$, only part of the representative building load is processed by HAC, and the grid-representative T_{in} is still updated prognostically by Eq. (22). Therefore, TARGTEMP should be interpreted as the control threshold for the conditioned fraction, not as a prescribed grid-mean indoor temperature.”

Lines 247-248. AHOPTION and SBEMOPTION – these seem to be WRF’s flags, and should be explained.

We agree and added a clearer explanation of these model-control flags. AHOPTION is a WRF-Urban/SLUCM-related option for anthropogenic heat treatment. In the original SLUCM case, AHOPTION = 1 adds prescribed anthropogenic heat to the sensible heat flux. In the previous SLUCM+BEM v1.0 implementation, AHOPTION = 2 was used for the BEM-related building anthropogenic heat calculation. In v2.0, we introduced SBEMOPTION as the user-facing switch for SLUCM+BEM v2.0 to avoid confusion with the older AHOPTION setting. In the v2.0 configuration used here, AHOPTION = 1 is used together with SBEMOPTION = 1, and AH_TRAFFIC is used to prescribe traffic anthropogenic heat separately from the building component simulated by BEM.

Revised text added to Section 2.2.1:

“AHOPTION and SBEMOPTION are WRF/WRF-Urban control flags. AHOPTION controls the treatment of prescribed anthropogenic heat in SLUCM, whereas SBEMOPTION is the switch that activates the SLUCM+BEM v2.0 building-energy calculation. In the v2.0 simulations, AHOPTION = 1 and SBEMOPTION = 1 were used; traffic anthropogenic heat was specified through AH_TRAFFIC, while building anthropogenic heat was calculated dynamically by SLUCM+BEM. ”

Changes made in: Section 2.2.1.

Section 3.1. Are there differences in the measurements depending on the floor? Is the top floor hotter during the day and cooler during the night than the other floors?

This is an important point. We agree that indoor temperature can differ by floor, and that top-floor flats may be warmer during daytime because of stronger roof solar heat gain and may also cool differently at night. Unfortunately, the public metadata associated with the Southwark indoor-temperature dataset used in this study do not provide sufficient floor-level information for each sensor/flat to stratify the observations by floor.

We therefore added this point as a caveat in Section 3.1. The revised manuscript now explains that part of the observed spread among individual dwellings may reflect not only window-opening and other occupant behaviour but also unreported building-position factors such as floor level, roof exposure, orientation, and dwelling layout. This limitation is one reason why we evaluate the model against the observed range and mean rather than attempting to reproduce each individual room.

Revised text added to Section 3.1:

“The metadata do not include sufficient information on the floor level of each measurement. Therefore, we could not stratify the observations by floor. Top-floor dwellings may experience stronger daytime warming due to roof exposure and different nocturnal cooling, and such unreported differences may contribute to the spread among individual measurements.”

Changes made in: Section 3.1.

Lines 329. Is vent-10 equal to vent-rate=10 (of Eqn. 19)?

Yes. We clarified this in the text and figure caption. The label “vent-10” means the sensitivity experiment with $\text{VENT_RATE} = 10 \text{ h}^{-1}$ in Eqs. (19)-(20). In the source code, the corresponding input parameter is named `VENT_FREQ`, so VENT_RATE in the manuscript corresponds to `VENT_FREQ` in `module_sf_urban.F` and `URBPARAM.TBL`.

We also clarified the standard VENT_RATE values used in the model setup. For residential buildings, $\text{VENT_RATE} = 0.5 \text{ h}^{-1}$ was used, following the minimum continuous ventilation requirement for dwellings in the Japanese Building Standards Act sick-house countermeasure guidance. For business and commercial buildings, $\text{VENT_RATE} = 1.7 \text{ h}^{-1}$ follows the standard office-building setting used in the original CM-BEM framework; the outdoor air introduction rate of $5.0 \text{ m}^3 \text{ m}^{-2} \text{ h}^{-1}$ corresponds to $5.0/3.0 = 1.67 \text{ h}^{-1}$, rounded to 1.7 h^{-1} , assuming a representative floor-to-ceiling height of 3.0 m.

Revised text in Section 3.1:

“The case name vent-X denotes the experiment with $\text{VENT_RATE} = X \text{ h}^{-1}$; for example, vent-10 corresponds to $\text{VENT_RATE} = 10 \text{ h}^{-1}$.”

Changes made in: Section 3.1 and Fig. 4 caption.

4. I know there are already many lines, but knowing the value of the outdoor air temperature would help in the interpretation.

We agree. To make the interpretation of the indoor-temperature response clearer, we added the outdoor air temperature used to force the offline London simulations to Fig. 4. This allows the reader to compare the observed and simulated indoor-temperature variations directly with the outdoor-temperature forcing.

Revised Fig. 4:

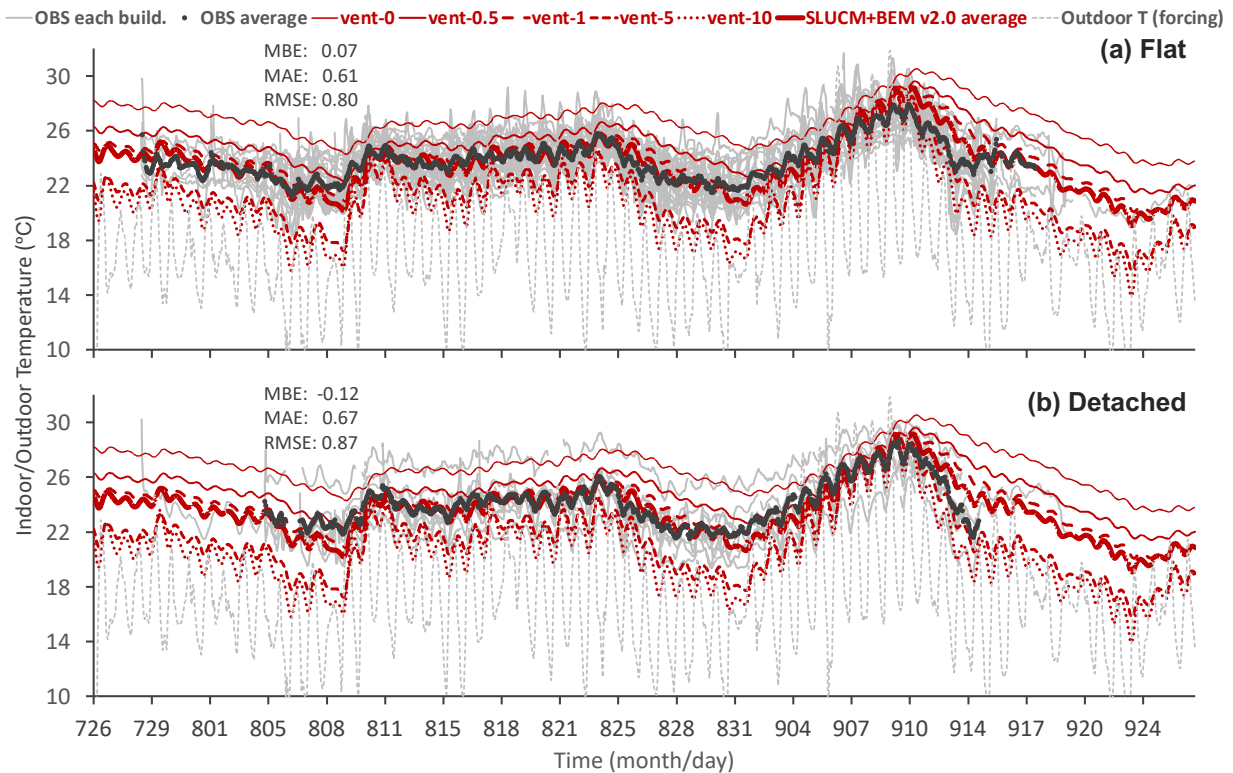


Figure 4: Diurnal changes in indoor temperatures (T_{in}) in (a) flats and (b) detached buildings in Southwark, London, UK (Fig. 2) over the period 26 July to 30 September 2023. Light grey lines are observations in each room. The dark grey line is the average value. The red lines are the simulated values for each case; vent-0, 0.5, 1, 5 and 10, and the average value determined by SLUCM+BEM. MAE is mean absolute error; MBE is mean bias error; RMSE is root mean square error. **The grey dashed line shows the outdoor air temperature from the ERA5 forcing used in the offline SLUCM+BEM simulations.**

Changes made in: Fig. 4 and its caption; Section 3.1 explanatory text.

Section 3.2. The values of electricity consumption are strongly affected by urban morphology. However, the values used in the simulations (Table 3) are based on only 3 urban classes, with fixed values of urban fraction and building height. How sensitive are the results to these values? Why didn't the authors use more detailed information (at least the 10 LCZs). Same for the building thermal properties – how realistic are they?

We fully agree that urban morphology and building thermal properties are important controls on simulated electricity consumption. We revised the manuscript to make the scope and limitation of the present evaluation clearer.

The main reason for using the three urban classes in the Tokyo evaluation was to allow a direct comparison with SLUCM+BEM v1.0, and the observation-based electricity-consumption evaluation framework used in our previous studies. Changing the urban classification system at the same time as changing the BEM formulation would have made it difficult to isolate the effect

of the new prognostic indoor-temperature and ventilation scheme. Therefore, the three-category setup was intentionally retained for the main evaluation.

At the same time, we agree that this setup limits the assessment of morphology sensitivity. We added text to clarify that the absolute EC values may be sensitive to urban fraction, building height, roof/building width, and thermal properties. We also clarified that SLUCM+BEM v2.0 is not restricted to the three-category setup: the implementation has been tested with LCZ and distributed urban-parameter input options, but a full evaluation of how LCZ- or distributed-parameter inputs affect EC skill requires additional observational constraints and is beyond the scope of this model-development paper.

Revised text added to Section 3.2:

“The three-category urban classification (Fig. 2a) was retained to enable a controlled comparison with SLUCM+BEM v1.0 under the same Tokyo evaluation framework. We acknowledge that the absolute magnitude and temperature sensitivity of EC_{HAC} can be affected by urban morphology and building thermal properties. Therefore, the thermal-property values in Table 3 should be interpreted as representative category-level parameters rather than building-specific values. SLUCM+BEM v2.0 has been confirmed to compile and run stably with LCZ and distributed urban-parameter input options (see section 4.3 and Table 5). However, evaluating how these more detailed inputs affect EC reproducibility requires a separate sensitivity study with appropriate observations.”

We also expanded the limitations paragraph to state that the thermal-property values should be interpreted as representative category-level parameters, not building-specific values.

Changes made in: Sections 3.2.

Lines 450-460. “What if” scenario. These results are with the assumption that the ventilation rate stays the same with and without A.C., which is doubtful.

We apologize for the unclear explanation. The what-if comparison did not assume the same ventilation rate for the HAC and no-HAC cases. In the control/HAC case, we used the standard ventilation-rate settings in Table 3: $VENT_RATE = 0.5 \text{ h}^{-1}$ for residential categories and 1.7 h^{-1} for business/commercial buildings. In the no-HAC case, AC_FLOOR_RATIO was set to zero and $VENT_RATE = 10 \text{ h}^{-1}$ was used to represent enhanced natural ventilation/window opening under free-running conditions (Table 3).

We revised Section 4.2 to state this explicitly and to clarify that the no-HAC case should be interpreted as “no mechanical heating/cooling with enhanced natural ventilation”, not as simply switching off HAC while keeping the same ventilation conditions.

Revised text in Section 4.2:

“In the no-HAC scenario, AC_FLOOR_RATIO was set to 0 for all hours and categories. The ventilation setting was also changed: the CTRL/HAC case used the standard Table 3 values (0.5 h^{-1} for residential categories and 1.7 h^{-1} for business/commercial buildings),

whereas the no-HAC case used $\text{VENT_RATE} = 10 \text{ h}^{-1}$ to represent enhanced natural ventilation associated with window opening.”

Changes made in: Section 4.2 and Table 3 caption/notes.

Line 515. The increase in computational time is strongly affected by the number of urban points. Please specify the fraction of urban points present in the domain.

We agree. We revised the computational-cost discussion to state the fraction of urban grid cells in the benchmark domains. Urban grid cells were identified from LU_INDEX in the wrfinput files, using LU_INDEX = 31, 32, or 33 as urban land use. In the innermost domain used for the Tokyo benchmark, the LU_INDEX mask contained $250 \times 250 = 62,500$ mass-grid cells. Of these, 7,100 cells were urban, corresponding to 11.36% of the grid cells. The urban fractions in d01, d02, and d03 were 0.20%, 3.93%, and 11.36%, respectively. We added these values to the revised manuscript and clarified that the reported computational-time ratios apply to this nested-domain configuration.

Revised text in Section 4.4:

“The innermost WRF domain used for the Tokyo benchmark had $e_{we} = e_{sn} = 251$ in namelist.input. The corresponding LU_INDEX land-use mask in wrfinput_d03 contains $250 \times 250 = 62,500$ mass-grid cells. Of these, 7,100 cells, or 11.36%, were classified as urban land use (LU_INDEX = 31, 32, or 33). The corresponding urban fractions were 0.20% in d01 and 3.93% in d02. The computational-time ratios reported here therefore apply to this nested-domain configuration, urban mask, simulation period, hardware, and parallel setting.”

Changes made in: Section 4.4 computational-cost paragraph.

Lines 533-535. I agree. On the relevance of the radiation transmitted through windows, you can refer to Pappaccogli et al. <https://www.sciencedirect.com/science/article/pii/S2212095519304213>, who made a sensitivity study on this and other parameters.

Thank you for this helpful suggestion. We added Pappaccogli et al. (2020) to the limitations/future-development discussion. This reference is highly relevant because their WRF-BEP+BEM sensitivity experiments examined the effects of urban geometry, building material properties, window fraction, target indoor temperature, ventilation rate, and internal gains on urban microclimate and building energy consumption.

Revised text in Section 4.4:

“The omission of solar heat gain through windows is a likely contributor to the remaining mismatch in residential EC_{HAC} sensitivity, particularly under sunny summer conditions. Pappaccogli et al. (2020) showed with WRF-BEP+BEM idealised simulations that building-energy consumption is sensitive to internal building parameters such as target

temperature, ventilation rate, internal gains, and window-related parameters. Introducing a simple treatment of transmitted solar radiation through windows is therefore an important target for future versions of SLUCM+BEM.”

Reference added:

“Pappaccogli, G., Giovannini, L., Zardi, D., and Martilli, A.: Sensitivity analysis of urban microclimatic conditions and building energy consumption on urban parameters by means of idealized numerical simulations, *Urban Climate*, 34, 100677, doi: 10.1016/j.uclim.2020.100677, 2020.”

Changes made in: Section 4.4 and References.

We again thank the Referee for the careful and constructive assessment. We believe that these revisions have substantially improved the clarity of the model formulation, the correspondence between the manuscript and the source code, and the discussion of input-parameter sensitivity and limitations.