This study proposed a simplified mathematical model to explain the generation mechanism of urban heat islands (UHIs). This is a new attempt to understand the UHI phenomenon, however several limitations should be addressed.

We thank the reviewer for carefully reviewing and evaluating our manuscript. We also recognize that our approach has several limitations, and in the revision we have clarified the scope and intent of the study. We have addressed the specific points raised by the reviewer in the following responses, and we believe that these revisions have improved the clarity and completeness of the manuscript.

The main conclusion of this study is that the UHIs occur due to the large heat capacity of urban areas. This can be misleading. It is true that the urban surface materials usually have large heat capacity values. However, besides the material properties, the large thermal inertia in urban areas can be attributed to the inefficient longwave cooling and also the increase in surface area. Due to the presence of high-rise buildings, the reduced sky view factors in cities greatly reduce the longwave cooling at night. The incease in surface areas (envelops) due to buildings also contribute to storing more heat in urban areas compared to a flat surface. It would be great if the authors can quantify the individual contributions that result in large thermal inertia. If this is not possible, I strongly suggest that the authors clarify these points; otherwise, readers would simply think the UHIs are due to the large heat capacity of urban surface materials.

We fully acknowledge the reviewer's point that large thermal inertia in urban areas is not only caused by the heat capacity of urban surface materials but also by other factors such as inefficient longwave cooling and the increased surface area due to buildings. Since our study employs a simplified model, it has limitations in quantifying the individual contributions of these factors. To avoid any misunderstanding, we have revised the manuscript to explicitly state that the large heat capacity of urban materials should be considered as only one of the contributing mechanisms. These clarifications have been added in lines 24–26 and 430–432 of the revised manuscript.

I wonder if the dR/dT used in the defination of lambda (eq. 10) shows a clear diurnal variation or not in urban and rural areas. In the paper, constant values (11.8 or 30.9 W m-2 K-1) are used, but the use of a time-invariance value shoud be justified. If the dR/dT does vary with time, how can this temporal variation be considered in the theoretical model?

We understand the reviewer's comment. As noted, the variable used in this study can exhibit diurnal variation because it is influenced by the diurnal changes in several surface energy fluxes. However, such variations have little impact on the results at the primary timescale considered in our simplified model, and their overall order of magnitude remains similar. Therefore, we used the daily mean value as a representative input to reflect the typical characteristics of urban and rural areas. This revision has been added in lines 286–289 of the revised manuscript.

The authors argue that the stronger UHI intensity at night than at day is due to the larger heat capacity. This may be partially true. However, the authours seem not to consider the difference in boundary layer height in the daytime and in the nighttime. The much shallower boundary layer at night both in urban and rural areas amplifies the effects of differential heat fluxes. For example, the difference in sensible heat flux between urban and rural areas is larger in the daytime (let's say 200 W m-2) than in the nighttime (e.g., 20 W m-2). However, the much deeper daytime boundary layer in the daytime (e.g., 1.5 km) than in the nighttime (e.g., 150 m) dilutes the differential heat flux effect. The authors are recommended considering this effect.

We understand the reviewer's concern. In our simplified model, the effect of boundary layer height was not explicitly considered. We agree that the shallower nighttime boundary layer amplifies the impact of differential heat fluxes, whereas the deeper daytime boundary layer tends to dilute them. To clarify this limitation, we have revised the manuscript to explicitly acknowledge the role of boundary layer height in modulating the diurnal variation of UHI intensity. These revisions have been added in lines 351–355 and 438–439 of the revised manuscript.

I am not sure if section 4.3 is necesseary for this study. I see a weak connection between section 4.3 and the previous sections. The authors would need to provide strong rationales or links of this section.

We acknowledge the reviewer's concern that Section 4.3 lacked a clear connection with the preceding sections. To strengthen the link, we have revised the Introduction (lines 96–99) to clarify the purpose of Section 4.3, stating that we utilize observational data in a megacity (Seoul) and a major city (Suwon) to compare the diurnal variation of the UHI intensity obtained from the theoretical model with that observed in reality, and to analyze long-term records to examine how UHI intensity evolves depending on the stage of urbanization.

Furthermore, we revised the beginning of Section 4.3 (lines 363–367) to explicitly build upon the previous subsection by emphasizing that the simplified day–night model qualitatively agrees with observations in Seoul and Yangpyeong, and that this agreement demonstrates the role of heat storage in strengthening nighttime UHI. Building on this insight, Section 4.3 then extends the analysis to long-term records, highlighting the contrasting UHI evolution in Seoul and Suwon with different urbanization rates.

We believe that these revisions provide a stronger rationale for Section 4.3 and clarify its connection to the earlier parts of the paper.