We thank the reviewer for the valuable and constructive feedback on our manuscript. Below we explain point-by-point how we adjusted our manuscript based on the reviewer's suggestions.

Additionally, we have implemented the following more substantial modifications:

- **We added a table that summarizes the representation of urban areas in RCMs, and we elaborate on the different parameterizations and their potential impacts on our results in the discussion section.**
- **We added a new figure to the supplementary (Figure S2), which shows a map with the EURO-CORDEX biases relative to ERA5-Land and E-OBS in the 36 investigated European cities.**

The paper presents a thorough analysis of ambient heat projections in major European cities using the EURO-CORDEX ensemble, comparing them with data from E-OBS, ERA5-Land, and weather stations. The study evaluates temperature biases, uncertainties, and factors influencing spatial patterns. It highlights variations in biases across cities and emphasizes the role of downscaling by regional climate models in shaping temperature estimates. The paper introduces a novel examination of nighttime ambient heat and compares projections with CMIP5 and CMIP6 ensembles. While the paper offers valuable insights, providing additional methodological details and discussing the limitations of the EURO-CORDEX data in the context of cities and urban areas could enhance its clarity and impact.

Urban Processes:

One noteworthy concern I have regarding the paper is the potential limitation arising from the EURO-CORDEX dataset's lack of representation of urban processes. The absence of urban-specific factors like the urban heat island (UHI) effect in the EURO-CORDEX models might lead to an incomplete understanding of how local urban conditions could influence temperature distributions. Addressing this limitation explicitly and discussing its potential implications for the reliability of the findings could benefit the readership.

We added additional information about the representation of urban areas in RCMs, as a response to a comment from the other reviewer. Table 1 now summarizes the representation of urban areas in the regional climate models, accompanied by a short description of the table in section 2.1.2:

"The representation of urban areas varies considerably across RCMs (Table 1). Some RCMs represent urban areas as rock surfaces, others assume reduced vegetation and adjusted surface parameters (such as albedo and roughness) for urban areas, and one RCM even includes a sophisticated urban model." (lines 144-146)

Based on this information, we adjusted and extended the discussion of our results in light of the different urban parameterization implemented in the models:

"The ~12.5 km spatial resolution of the EUR-11 simulations enables a much more detailed assessment of climate variability and climate change at the city-level compared to GCMs, which have a much coarser spatial resolution (~100 km). Yet, most land surface modules of models in the 0.11° EURO-CORDEX ensemble only employ a simplified representation of urban areas (Table 1), which prevents the full exploitation of their high spatial resolution for studies focusing on urban areas. A few models represent urban areas as rock surfaces, thus neglecting the influence of urban vegetation on the surface energy balance and the influence of urban buildings on turbulence,

radiation, and hydrology. Other models apply adjusted parameters (e.g., for albedo and roughness length) and a reduced vegetation cover in urban areas, and thus consider the characteristics of cities to some extent. One of the models uses a sophisticated urban land model, which includes various aspects of urban areas, such as urban canyons, different levels of urbanisation, and radiation and hydrology schemes specifically adapted for urban areas. Despite these substantial differences in how urban areas are represented, no direct link can be found between the general behaviour of the different models in the projection of ambient heat (e.g., comparatively high levels of ambient heat in HadREM3-GA7-05 and WRF381P, and comparatively low levels in HIRHAM5, RACMO22E, and COSMO-crCLIM-v1-1, with all of these models using the adjusted-parameter approach to represent urban areas) and their representation of urban areas (Figure 8, Table 1)." (lines 611-624)

Clarity of Methodology:

While the section on "Identifying factors influencing spatial patterns" is intriguing, the exact statistical methods used to establish the relationships between climate and location factors and heat metrics and the limitations of these methods should be explicitly stated.

Section 2.4.1 provides the description of the statistical methods used to calculate how much of the spatial variability of the different heat metrics can be explained by the climate factors and by the location factors, respectively. To make this clear, we added a reference to section 2.4.1 in the first sentence of section 3.3. In the same sentence, we now highlight better that we do not analyze the relationships between climate and location factors, but that we separately analyze how much of the spatial patterns of ambient heat can be related to climate factors and to location factors, respectively. The revised sentence now reads:

"To better understand the spatial patterns of ambient heat projected by the different heat metrics, we estimate how much of the spatial variance is explained 1) by different climate factors, representing each city's temperature climatology as well as its projected changes, and 2) by different location factors (Figure 5; see Section 2.4.1 for methodological details)." (lines 385-387)

The limitations of the applied method are stated in Section 2.4.1. One limitation are potential collinearities of the explanatory variables: "The explanatory variables (i.e., the climatological factors or the location factors) may be correlated, and their contributions cannot be strictly disentangled." (lines 285-286).

An estimate for the uncertainty introduced by the collinearity of different explanatory variables can be obtained from the variability of the squared semipartial correlation estimates. We use this variability as an uncertainty estimate for the contribution of each explanatory variable (see Section 2.4.1 and Figure 5). Additionally, we highlight now more in detail that the employed correlation analysis does not allow any statements about causality:

„The variability of the squared semipartial correlation estimates is a measure for collinearities between the explanatory variables and can be used as an uncertainty estimate for the contribution of each explanatory variable. The estimated contribution of each explanatory variable to the spatial variability of each heat metric does not permit statements about causality, as it is purely based on correlation analysis. Instead, the contributions should be interpreted as a measure of the extent to which the explained variables may be influenced by the location of each city or by the climatic conditions and climate change at the location of each city." (lines 293-298)

Comparative Analysis:

The comparison between EURO-CORDEX, CMIP5, and CMIP6 ensembles in Section 3.5 is a valuable addition. However, the paper could provide more insights into the potential reasons behind the differences in projections. Elaborating on the distinct characteristics of the GCMs and RCMs, such as spatial resolution and physical parameterizations, could enhance the understanding of the results.

To highlight the differences between RCMs and GCMs, we added a paragraph to section 2.1.2 where we introduce the climate model data:

"The GCMs and RCMs used in this study differ in several aspects. Most importantly, the RCMs have a much higher spatial resolution (~12.5 km) than the GCMs (~100 km), and orography and coastlines are thus represented much more accurately in RCMs. GCMs and RCMs also differ in their projections of atmospheric aerosols over the European domain, with GCMs using future scenarios with decreasing atmospheric aerosol concentrations while RCMs assume a constant atmospheric aerosol load (Boé et al., 2020; Gutiérrez et al., 2020; Nabat et al., 2020). Additionally, unlike GCMs, several RCMs do not consider plant physiological CO2 effects, which might cause an underestimation of temperature extremes (Schwingshackl et al., 2019)." (lines 158-163)

Additionally, we added a paragraph to the discussion, where we elaborate on how the differences between GCMs and RCMs might impact our results.

"In many of the investigated cities, CMIP5 and CMIP6 project higher increases in TXx and larger HWMId-TX values than EURO-CORDEX. This is likely caused by discrepancies in external forcing data and differences in process implementation (see Section 2.1.2). Specifically, the CMIP5 and CMIP6 simulations are based on future scenarios with decreasing atmospheric aerosol concentrations over the European domain, while the EURO-CORDEX simulations assume a constant atmospheric aerosol load (Boé et al., 2020). The RCMs of EURO-CORDEX may thus underestimate future warming in Europe as they do not consider the amplified warming from the additional solar radiation reaching and heating the Earth's surface in Europe because of the decreasing aerosol concentrations. In addition, unlike CMIP5 and CMIP6 GCMs, several RCMs do not consider plant physiological effects (Schwingshackl et al., 2019). The closing of plant stomata due to higher CO2 concentrations and the associated decrease in latent and increase in sensible heat fluxes, which lead to enhanced extreme temperatures, are thus not fully captured by RCMs. These differences between GCMs and RCMs suggest that RCMs likely underestimate future levels of ambient heat in European cities. Yet, for several southern European cities the EURO-CORDEX models project considerably more days exceeding 30 °C than CMIP5 and CMIP6. In coastal cities, such as Istanbul, Athens, and Lisbon, these differences are likely due to the higher spatial resolution of EURO-CORDEX, which enables a better distinction of land and ocean grid cells. In other cities, like Madrid or Rome, better resolved orography might be the reason for the more frequent exceedances in EURO-CORDEX. Yet the causes for some discrepancies remain unclear, for instance for the more frequent exceedances above 30 °C projected by EURO-CORDEX for Milan, which lies in the rather flat Po Valley, or for the coastal city Barcelona, where EURO-CORDEX shows much fewer exceedances above 30 °C than CMIP5 and CMIP6." (lines 563-579)