Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

How convection-permitting climate models improve the representation of urban temperatures in Europe

Benjamin Le Roy

benjamin.leroy@meteo.fr

Helmholtz Centre Hereon Climate Service Center Germany: Helmholtz-Zentrum Hereon GERICS - Deutsches Institut fur Klimaservices https://orcid.org/0000-0002-5496-2462

Rechid Diana

Helmholtz Centre Hereon Climate Service Center Germany: Helmholtz-Zentrum Hereon GERICS - Deutsches Institut fur Klimaservices

Sophie Bastin

LATMOS: Laboratoire Atmospheres Milieux Observations Spatiales

Gaby Langendijk

Deltares

Lilly Zerbst

University of Innsbruck: Universitat Innsbruck

Marianna Adinolfi

Euro-Mediterranean Center for Climate Change: Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici

Nikolina Ban

University of Innsbruck: Universitat Innsbruck

Ségolène Berthou

Met Office

May Meret Bohmann

Karlsruhe Institute of Technology: Karlsruher Institut fur Technologie

Cécile Caillaud

Centre National de Recherches Météorologiques: Centre National de Recherches Meteorologiques

Rita M. Cardoso

University of Lisbon Faculty of Sciences: Universidade de Lisboa Faculdade de Ciencias

Hylke De Vries

Royal Netherlands Meteorological Institute: Koninklijk Nederlands Meteorologisch Instituut

Andreas Dobler

Norwegian Meteorological Institute: Meteorologisk institutt

Jesus Fernandez
Instituto de Física de Cantabria: Instituto de Fisica de Cantabria
Thomas Frisius
Helmholtz Centre Hereon Climate Service Center Germany: Helmholtz-Zentrum Hereon GERICS - Deutsches Institut fur Klimaservices
Klaus Görgen
Forschungszentrum Jülich Institute of Bio- and Geosciences Agrosphere 3: Forschungszentrum Julich GmbH Institut fur Bio- und Geowissenschaften Agrosphare 3
Ozan Mert Göktürk
NORCE Norwegian Research Centre AS Forskningsomrade Helse
Øivind Hodnebrog
Center for International Climate Research: CICERO
Stergios Kartsios
Aristotle University of Thessaloniki: Aristoteleio Panepistemio Thessalonikes
Eleni Katragkou
Aristotle University of Thessaloniki: Aristoteleio Panepistemio Thessalonikes
Josipa Milovac
Instituto de Física de Cantabria: Instituto de Fisica de Cantabria
Mario Raffa
Euro-Mediterranean Center for Climate Change: Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici
Pedro M.M. Soares
University of Lisbon Faculty of Sciences: Universidade de Lisboa Faculdade de Ciencias
Stefan Sobolowski
NORCE Norwegian Research Centre AS Forskningsomrade Helse
Merja Tölle
University of Kassel: Universitat Kassel
Kirsten Warrach-Sagi
University of Hohenheim Faculty of Natural Sciences: Universitat Hohenheim Fakultat fur Naturwissenschaften
Natalia Zazulie
Abdus Salam Centro internazionale di fisica teorica: Abdus Salam International Centre for Theoretical Physics

---

Research Article

Keywords: Urban climate, Urban Heat Island, Regional climate model, Convection-permitting climate model, Urban parametrization, Multi-model ensemble, Added value, European cities

Posted Date: April 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9278593/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

How convection-permitting climate models improve the representation of urban temperatures in Europe

Le Roy, B.[1*], Rechid, D.[1], Bastin, S.[2], Langendijk, G. S.[3], Zerbst, L.[4], Adinolfi, M.[5], Ban, N.[4], Berthou, S.[6], Bohmann, M. M.[7], Caillaud, C.[8], Cardoso, R. M.[9], De Vries, H.[10], Dobler, A.[11], Fernandez, J.[12], Frisius, T.[1], Goergen, K.[13,14], Göktürk, O. M.[15,16], Hodnebrog, Ø.[17], Kartsios, S.[18], Katragkou, E.[18], Milovac, J.[12], Raffa, M.[5], Soares, P. M. M.[9], Sobolowski, S.[15,16,19], Tölle, M.[20], Warrach-Sagi, K.[21], Zazulie, N[22,23]

1) Climate Service Center Germany (GERICS), Helmholtz-Zentrum Hereon, Fischertwiete 1, 20095 Hamburg, Germany
2) Laboratoire ATmosphere Milieux Observations Spatiales/Institut Pierre Simon Laplace (LATMOS/IPSL), UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, CNES, Guyancourt, France
3) Climate Adaptation and Disaster Risk Department, Deltares, PO Box 177, 2600, MH, Delft, the Netherlands
4) Department of Atmospheric and Cryosphere Sciences, University of Innsbruck, Innrain 52f, 6020 Innsbruck, Austria
5) CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy
6) Met Office, Exeter, United Kingdom
7) Institute of Meteorology and Climate Research – Troposphere Research (IMKTRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
8) Météo-France, CNRS, Univ. Toulouse, CNRM, Toulouse, France
9) Instituto Dom Luiz, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal
10) Royal Netherlands Meteorological Institute, Utrechtseweg 297, 3731 GA De Bilt, the Netherlands
11) Norwegian Meteorological Institute, Oslo, Norway
12) Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain
13) Institute of Bio- and Geosciences (IBG-3, Agrosphere), Forschungszentrum Jülich, Jülich, Germany
14) Centre for High-Performance Scientific Computing in Terrestrial Systems, Geoverbund ABC/J, Jülich, Germany
15) NORCE Norwegian Research Centre, Bergen, Norway
16) Bjerknes Centre for Climate Research, Bergen, Norway
17) Center for International Climate Research (CICERO), Oslo, Norway
18) Department of Meteorology and Climatology, School of Geology, Aristotle University of Thessaloniki, Greece
19) Geophysical Institute, University of Bergen, Bergen, Norway
20) Center for Environmental Systems Research (CESR), University of Kassel, Kassel, Germany
21) Institute of Physics and Meteorology, Faculty of Natural Sciences, University of Hohenheim, Germany
22) The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy; Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy
23) Department of Atmospheric and Ocean Sciences, Faculty of Exact and Natural Sciences, University of Buenos Aires (DCAO-FCEN-UBA) - National Council of Scientific and Technical Research (CONICET), Buenos Aires, Argentina

* Le Roy, B. current affiliation is Météo-France, CNRS, Univ. Toulouse, CNRM, Toulouse, France; benjamin.leroy@meteo.fr

Abstract

This study presents the first multi-model ensemble evaluation of the added value of km-scale (2-3 km) Convection-Permitting climate Models (CPM) for representing urban temperatures and Urban Heat Islands (UHI) across multiple cities, compared to coarser-resolution (12 km) Regional Climate Models (RCM). To this end, we analyze 21 CPM simulations driven by the ERA-Interim reanalysis from 2000 to 2009. The ensemble spans six families of RCMs with urban parameterizations of varying complexity, ranging from simple surface/bulk schemes with modified surface parameters, to dedicated Urban Canopy Models (UCM). We evaluate the potential added value of the increased horizontal resolution and improved urban representation of the CPMs compared to their respective driving RCMs. This evaluation is performed across six European cities where data from urban and rural meteorological stations are available. The study focuses on how well the models represent daily minimum and maximum temperatures, as well as UHI intensities, in winter and summer. The results show that all models incorporating an urban parameterization capture some urban signal; however, their performance varies from city to city. CPMs with dedicated UCMs or simpler bulk schemes that represent key processes (e.g., anthropogenic heat in winter or solar radiation trapping in summer) can better simulate UHIs. The added value for temperatures outside urban areas is limited, as RCMs already match observations well. However, it can be significant over cities when the increased resolution is combined with the adoption of a more complex urban parameterization in the CPM. While all urban CPMs in the ensemble can represent average UHIs, even the best ones struggle to capture extreme UHI intensities.

Keywords

Urban climate; Urban Heat Island; Regional climate model; Convection-permitting climate model; Urban parametrization, Multi-model ensemble; Added value; European cities

1. Introduction

Cities already host more than half of the world population (UN-Habitat, 2022) and are hotspots of climate impacts and extreme weather hazards. Urban areas are particularly vulnerable to heat extremes (Mishra et al. 2015, Tuholske et al. 2021, Oliveira et al. 2022), whose intensity and frequency are projected to increase under ongoing climate change (Seneviratne et al. 2021, Marcotullio et al. 2022). Understanding and modeling their local climate conditions is therefore critical for impact assessments and adaptation planning. Urban areas cause important changes in natural landscapes that alter local climate conditions. They modify the surface energy balance of natural areas by altering the exchanges of heat, water and momentum between the surface and the atmosphere. Through their complex 3D geometry, they alter the amount of solar energy reaching the surface and trap energy inside street canyons, which is then absorbed by artificial materials and released at night (Nunez and Oke 1977; Oke 1988). Cities impact exchanges of water by reducing ground infiltration and evapotranspiration due to the prevalence of impervious surfaces and vegetation loss. Lastly, areas with a high concentration of human activity are sources of anthropogenic heat from transport, buildings, and industries. These processes collectively generate what is known as the urban climate (Oke et al. 2017) and its most well known phenomenon, the Urban Heat Island (UHI) effect, where urban areas are warmer than their rural surroundings, especially at night (Arnfield, 2003). UHI intensities are influenced by meteorological drivers (solar radiation, wind speed, air and soil humidity; McGregor et al., 2006; Yow et al. 2007; Oke et al. 2017), local surface properties, city geometry and materials. Therefore, they vary in time (during the day and between seasons), in space (between different neighbourhoods, [Kim and Brown, 2021], between different cities [Oke, 1973, Adinolfi et al. 2023, Campanale et al. 2025], and different surrounding climates [Zhang et al. 2022]).

Today, climate change projections are conducted using Global Climate Models (GCMs; Chen et al. 2021) that can be downscaled to the regional scale using statistical methods (Wilby et al. 1997; Maraun and Widmann, 2018). However, these approaches are based on the assumption that the statistical relationships observed in the past will remain valid in the future, and they do not take account of dynamics and feedback between the surface and the atmosphere. In order to better represent mesoscale dynamics and surface heterogeneities, dynamical downscaling methods with limited-area Regional Climate Models (RCMs) are applied (e.g. Rummukainen, 2010; Giorgi and Gutowski; 2015; Giorgi, 2019). But even with nowadays common grid spacings on the order of a dozen kilometers,

RCMs are still too coarse to adequately represent intra-urban variability and heterogeneity, except for the most extensive megapolis (Langendijk et al. 2025). In addition, the latest-generation RCMs do not include urban surfaces in detail and instead rely on simpler parametrizations (Langendijk et al. 2019; Eunice Lo et al. 2020; Langendijk et al. 2025). A new generation of higher resolution (1-3 km) RCMs called Convection-Permitting RCMs (CPRCMs or CPMs; Prein et al. 2015; Lucas-Picher et al. 2021) has become available in the past decade. Their kilometric horizontal grid spacing makes it possible to represent the processes linked to complex heterogeneous terrain, such as mountainous areas, coastlines and cities, with the latter especially true when CPMs are coupled with more sophisticated Urban Canopy Models (UCM; Grimmond et al. 2011, Lipson et al. 2024).

The added value of CPMs compared to RCMs has been extensively documented for different processes such as: heavy precipitation and convective systems (Ban et al. 2021; Pichelli et al. 2021; Ha et al. 2024, Fosser et al. 2024), near surface air temperature (Soares et al. 2024) and wind over complex terrain (Belušić Vozila et al. 2024; Molina et al. 2024), as well as heatwaves (Sangelantoni et al. 2024). Some studies have used individual CPMs to look at the urban climate of different cities, but because of their high computational costs few have done so on climatic time scales (i.e. at least several years). Yang et al. (2012) simulated the climate of Nanjing (China) over the summers between 2000 and 2009 with the WRF model at 1 km horizontal resolution. Kusaka et al. (2016) analysed the effect of climate change and urban scenario on Tokyo (Japan) between 1990-1999 and 2050-2059 using WRF at 4 km. Murata et al. (2017) used NHRCM02 to simulate the climate of Japan at 2 km for two time periods of 20 years (1980-2000 and 2076-2096). Using WRF, Marelle et al. (2020) studied the impact of four midlatitude megacities (Paris, France; New York City, USA; Tokyo, Japan; Shanghai, China) on extreme precipitation over 10 years in present climate (2008–2017; 20 years for Paris, 1998-2017); while Steensen et al. (2022), analysed future UHI influence on precipitation for Paris (France) and Shanghai (China), using the same regional model configuration (but different urban settings at 3 km). Keppas et al. (2021) used WRF at 2 km to investigate the future evolution of the climate of Rome (Italy) and Thessaloniki (Greece). The UM was used to study the urban climate of cities in the United Kingdom by Keat et al. (2021) (2.2 km) and the climate of Johannesburg (South Africa) by Keat et al. (2024) (4 km). The CNRM-AROME model at 2.5 km was used by Lemonsu et al. (2023) over Paris (France), and by Michau et al. (2023) and Michau et al. (2024) over other French cities. Most of these studies are single model experiments which limit robust assessment of uncertainty. Recently, ensembles of CPMs have been employed to study the future evolution of the

urban climate in different contexts e.g., during heatwave conditions over Paris (Michau et al. 2025) or for different seasons over several Alpines cities (Zerbst et al. 2026, submitted).

To the best of our knowledge, no work has assessed the potential added value of increased resolution and/or improved urban parameterizations in CPMs relative to their driving RCMs in an ensemble setting and covering multiple cities. This study aims to fill this gap by examining an ensemble of 21 CPM simulations from the Coordinated Regional Downscaling Experiment (CORDEX) Flagship Pilot Study (FPS) “Convective Phenomena over Europe and the Mediterranean” (FPS-CONV, Coppola et al. 2020) and their driving RCMs (available for 19 groups) over six European cities over the 2000-2009 period. The analysis focuses on the models' ability to represent observed seasonal UHI intensities. The model ensemble and observations are presented in Section 3, followed by a discussion of their performance in Section 4. More general comments and conclusions are presented in Section 5.

2. Urban representation in Regional Climate Models

2.1. Regional Climate Modeling

The CORDEX community brings together meteorological institutes, research centers, and academic institutions to coordinate continental scale downscaling experiments using RCMs (Giorgi 2006; Giorgi and Gutowski, 2015). For historical reasons – namely, the too coarse horizontal resolution of the first simulations and the fact that urban climates were not the focus – there was no coordination on the representation of urban areas in CORDEX. Some RCMs setups did not see cities at all, while others represented cities via an urban land use category and corresponding urban parameters (i.e. bulk approaches), and only a few (e.g. RegCM) had more complex urban parameterizations activated (Langendijk et al. 2019, Langendijk et al. 2025). This was still the case in the FPS-CONV, where the horizontal resolution (~3km) would allow for the cities to be better represented. Nevertheless, due to the growing interest in studying the impacts of climate change on cities, the CORDEX community has launched a new FPS on URBan environments and Regional Climate Change (FPS URB-RCC; Langendijk et al. 2024). As part of this effort, the representation of cities in existing multi-model ensembles of regional climate simulations is being investigated at standard RCM resolutions (10-25 km; Langendijk et al. 2025) and at convection-permitting scales (2-3 km; this study). This is an initial analysis based on FPS-CONV data to set the scene for the upcoming multi-model ensemble of

urbanized CPMs being developed within the FPS-URB-RCC. An overview of the different ways in which urban areas can be generally represented in RCMs is given in the following Section (2.2) and the FPS-CONV CPMs are presented next (Section 3.1) with further details regarding their land and urban representations.

2.2. Land surface models and urban representation

Land Surface Models (LSM) are the components of weather and climate models that simulate physical processes of mass, energy, and momentum transfers occurring at the Earth's surface in contact with the atmosphere (Blyth et al. 2021). They have evolved from simple surface energy balance models to models representing the evolution of vegetation, the water cycle, the dynamics of the carbon cycle, changes in land use, and complex two-way interactions and feedbacks with the overlying atmosphere. (Fisher and Koven, 2020). Over the years different approaches have been developed to represent urban areas in LSMs (Grimmond et al. 2011, Lipson et al. 2024). Here we focus on three types currently used in climate models (Figure 1).

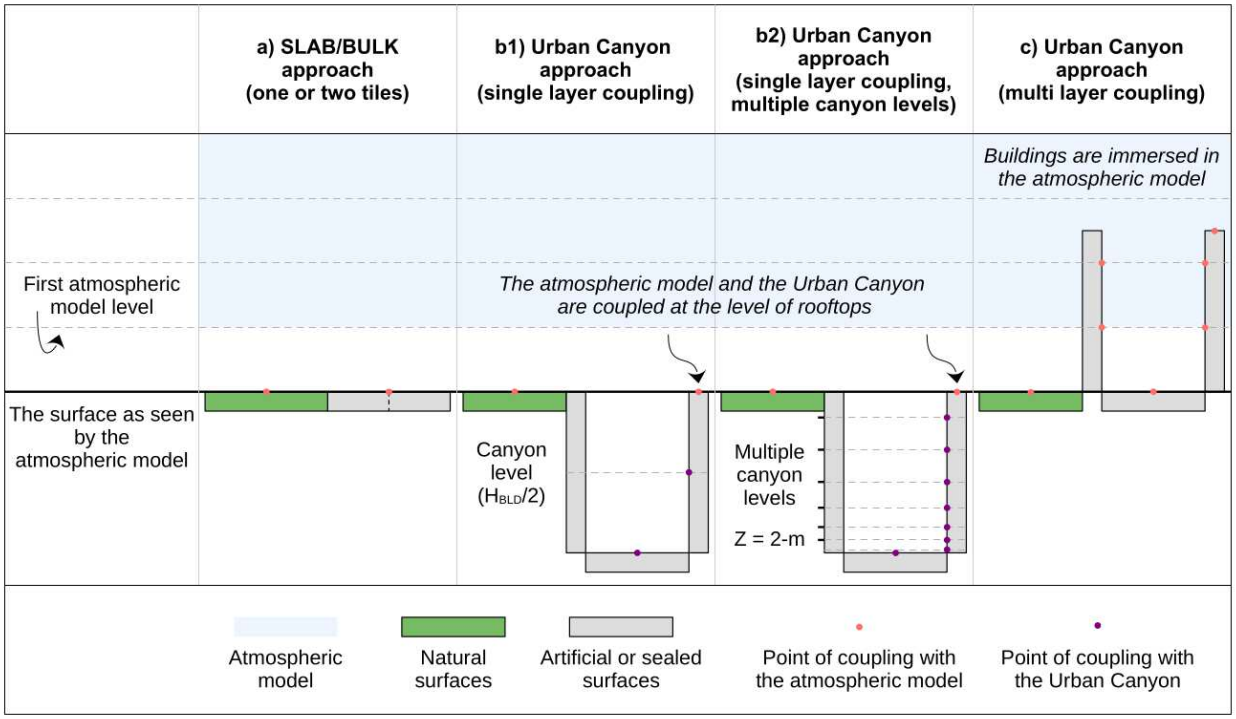


Figure 1. Types of urban parameterization schemes used in current-day RCMs: (a) In the Slab/Bulk approach, urban surfaces (grey block) are treated like natural surfaces (green block) with adjusted bulk parameters. No interior dynamics are resolved. (b) In

the single-layer urban canopy model (SLUCM), urban processes are represented within an idealized canyon (roof, walls, street; grey blocks) coupled to the lowest atmospheric level, with one or multiple canyon sublayers (b1 and b2, respectively). (c) As in SLUCM, but the urban canopy interacts with multiple atmospheric levels (red dots) and resolves several canyon layers.

Historically, cities were not represented in detail in LSMs. Instead, their interactions with the lowest atmospheric model level was approximated by using the non-vegetated/natural component of the LSM and by adjusting physical parameters such as surface roughness length, albedo and/or thermal capacity (Masson 2006; Martilli 2007). This approach, commonly referred to as the "slab", "bulk" or "rock" approach (Figure 1.a), is still used in most RCMs (Langendijk et al. 2019). BULK schemes can themselves be divided into two groups: one-tile schemes that represent urban areas as a homogeneous surface with adapted parameters (or sometimes just using the values of the "rock" cover from a soil-vegetation model; Daniel et al. 2019), and two-tile schemes that separate urban surfaces such as urban canyons and roofs (Lipson et al. 2024; Best et al. 2005).

A more complex approach developed in the 2000's is based on the Urban Canyon concept (Nunez and Oke, 1977) where the city is represented as one (Masson 2000) or several "canyons" (Kusaka et al. 2001; Kusaka and Kimura, 2004) of infinite length with a road, rooftops and one or two walls (sunlit and shaded). These UCMs can be divided into two separate categories depending on how they couple to the atmosphere. The first category is the single-layer (SL)UCM; these only interact with the atmosphere at the top of the urban canyon and roof (Figure 1.b1 and 1.b2; Masson 2006).
This means that the top of the urban canyon is located below the surface as seen by the atmospheric model. In their more simple form SLUCMs contain a single canyon layer (usually located at half the average canyon/building height; Figure 1.b1), which means that the air inside the canyon is assumed to be uniform. This approach makes it easy to couple the UCM to the atmospheric component of the model, but has the disadvantage of not directly representing the effect of building drag on the atmosphere. Because the air inside the canyon is assumed to be uniform, most SLUCMs are unable to provide prognostic variables at specified heights (e.g. air temperature at 2 m, T2M), which can make it challenging to compare a model grid cell to point-based observations, or two model grid cells with different building heights. To overcome this first issue, multiple levels can be

represented inside the canyon (Figure 1.b2) by parameterizing the impact of urban obstacles via introducing a drag force approach and resolving the surface boundary layer inside and above the canyon (Hamdi and Masson, 2008).

The last approach, called multi-layered (ML) UCMs (Kondo and Liu, 1998; Ca et al. 1999; Martilli et al. 2002), consists of coupling the atmospheric model and the UCM not only at the roof level but at different heights inside the urban canyon (Figure 1.d, Kondo et al. 2005; Chen et al. 2011). The main advantage of this approach is that the characteristics of the air within the urban canyon are no longer assumed to be uniform, but are simulated through the interaction between the atmosphere and urban surfaces, in addition to the representation of horizontal advection between urban canyons. The drag effect of buildings and horizontal advection are therefore represented, but at the cost of increased coupling complexity, with a modification of the atmospheric model equations (Masson, 2006; Martilli 2007). There are different levels of complexity in ML coupling; for example, depending on the number of wall surface temperatures used or whether the volume occupied by the buildings is explicitly represented in the atmosphere (Chen et al. 2011, Schoetter et al. 2020). It should be noted that the parameterization of multiple levels of urban canyons but with a single coupling layer (Figure 1.b1 and 1.b2) is sometimes referred to in the literature as MLUCM (Grimmond et al. 2011). In this study, however, when we refer to MLUCM we are referring to schemes that have multiple layers coupled with the atmosphere (Figure 1.c).

In addition to the differences highlighted above, the way near surface 2-m air temperature (T2M) is calculated is model-dependent. Differences can stem from multiple sources: the approach used to derive it (e.g. using Monin Obukhov Similarity Theory; MOST; Monin and Obukhov, 1954), the variables used in these approaches, and also the way these variables are aggregated over a grid point with different surface types. Note that this is not specific to urban areas: Breil et al. (2020) already highlighted the differences in the computation of the energy balance for LSMs representing tree canopies on several levels compared to a single one. Over grid points with different types of surfaces, some models will compute T2M by first averaging the fluxes of the different surfaces (e.g. urban surfaces, vegetation and inland water) and then computing the air temperature, while others will compute one T2M value for each type of surface and then average them based on the different tile fractions. Similarly to tree canopies, the addition of more sophisticated urban parametrization increases the complexity of the computation of T2M. SLUCMs with a single urban canyon level (Figure 1.b1) cannot output T2M, so the air temperature of the

point is an aggregate of T2M over natural surfaces and the canyon air temperature. SLUCMs with multiple urban canyon levels (Figure 1.b2) and MLUCMs (Figure 1.c) can output T2M inside the urban canyon and outside over natural areas, so the T2M of the model grid point is then a weighted average of the different temperatures.

3. Data and methods

3.1. Ensemble of Convection-Permitting Models

In this study, we analyzed an ensemble of 19 RCMs run within the scope of the EURO-CORDEX initiative, which has a focus on the European continent at a resolution of ~12km (Jacob et al. 2014, 2020). EURO-CORDEX climate simulations have been used for a large number of studies on a variety of subjects: temperature analysis (Dosio, 2016), heat waves (Vautard et al. 2013; Dosio, 2017; Molina et al. 2020), precipitation (Prein et al. 2016), agricultural impacts (Bathiany et al. 2023), winds (Moemken et al. 2018), land-atmosphere feedbacks (Knist et al. 2017), and over urban areas (Langendijk et al. 2019; Le Roy et al. 2021).

The ensemble served as a starting point for the first-ever downscaling of a set of CPMs (2–3 km) across a common pan-Alpine domain within the Flagship Pilot Study (FPS) “Convective Phenomena over Europe and the Mediterranean” (FPS-CONV; Coppola et al. 2020; Pichelli et al. 2021). Our study focuses on the evaluation runs (Ban et al 2021), driven by ERA-Interim reanalysis data (Dee et al. 2011) over the period 2000-2009. In addition to the increased resolution compared to their driving RCMs and the explicit resolution of deep convection and local scale dynamics, most CPMs in the FPS-CONV ensemble incorporate the urban parameterizations described in Section 2.2.

The ensemble used here consists of 6 RCM/CPM families: COSMO (Rockel et al. 2008, Sørland et al. 2021), WRF (Skamarock et al. 2008; Powers et al. 2017), REMO (Jacob and Podzun 1997, Pietikäinen et al. 2025), UM (Walters et al. 2017), RegCM (Giorgi et al. 2012) and AROME (HCLIM-AROME [Belušić et al. 2020] and CNRM-AROME [Déqué et al. 2016; Caillaud et al. 2021]). The evaluation study by Ban et al. 2021 provides details on the physics and parameterizations of these models, and we add information relevant to urban areas, namely the LSMs, the urban parameterizations, and the land use databases used. Models are listed by RCM-family, and in ascending order of urban parametrization complexity.

COSMO – Different versions of the climate branch of the COSMO model (Consortium for Small Scale Modeling) COSMO-CLM, conducted by the CLM community institutes, are used in this study (Table 1; Rockel et al. 2008, Sørland et al. 2021). To simulate land surface processes, all COSMO simulations are based on the TERRA LSM model (Schrodin and Heise, 2001; Schulz et al. 2016), with land covers being described by the GLC2000 database (Bartholomé and Belward, 2005). TERRA does not represent sub-grid surface heterogeneities (Avissar and Pielke, 1989), but instead relies on a dominant cover approach (Wouters et al. 2017). Only the 3 km CPM simulation carried out by the CMCC working group (CCLM5-0-9; Adinolfi et al. 2020) represents the effect of cities by activating the TERRA-URB urban parameterization (Wouters et al. 2015, 2016). Other simulations (including CMCC 12 km simulation) have replaced artificial areas with natural land covers with modified surface roughness length and reduced vegetation fraction and leaf area index (Doms et al. 2011; Garbero et al. 2021). The TERRA-URB urban parametrization used in the CMCC simulation, is a BULK scheme (Figure 1.a) based on the soil-vegetation module TERRA-ML with modified values of surface albedo, surface emissivity, heat capacity, heat conductivity, and surface roughness length. Additionally, other processes are represented/parametrized: prescribed time-varying (diurnal and seasonal) anthropogenic heat emissions (Flanner, 2009), thermal roughness length (Brutsaert, 1982; Kanda et al. 2007; Demuzere et al. 2008), impervious water storage (Wouters et al. 2015), a bare-soil evaporation resistance formulation (Schulz and Vogel, 2017), and a vegetation skin-temperature parametrization (Schulz and Vogel, 2017; Viterbo and Beljaars, 1995). One specificity of TERRA-URB compared with other BULK schemes is that it translates 3D urban canopy parameters into bulk parameters using the Semi-empirical URban canopY dependency parametrization (SURY, Wouters et al., 2016). Default values are adapted from the medium urban density class of Loridan and Grimmond (2012) and can be found in Wouters et al. (2016) and Wouters et al. (2017).

WRF – A sub-ensemble of simulations conducted with the WRF model (Weather Research and Forecasting; Skamarock et al. 2008) version 3.8.1 are analyzed to investigate the effect of different physical parameterizations. A detailed description of the differences can be found in Ban et al. (2021). Of particular interest for our study is the use of two different boundary layer schemes: MYNN2 (Nakanishi and Niino, 2009) for the CICERO, IDL, IPSL, UCAN and UHOH simulations and YSU (Hong and Dudhia 2006) for the AUTH-MC, BCCR, FZJ, and WEGC simulations. Also, two versions of the Noah LSM are used: Noah (Chen and Dudhia, 2001a, 2001b) for IDL, FZJ and WEGC and Noah with multiple physics

option (Noah-MP, Niu et al. 2011) for the other simulations (Table 1). WRF has several urban parametrizations to choose from: a BULK scheme, a SLUCM (Kusaka et al. 2001; Kusaka and Kimura, 2004) and a MLUCM (Martilli, 2002; Chen et al. 2011). In the FPS-CONV WRF simulations used the BULK approach to represent urban areas (Figure 1.a). WRF's BULK scheme modifies some of the LSM's physical parameters following Liu et al. (2006): surface roughness length, surface albedo, heat capacity, heat conductivity, and reduced green-vegetation fraction (see Chen et al. 2011 for a summary). While the representation of different land surface fractions inside a grid point is available in WRF (Li et al. 2013), it was not activated here. Instead, WRF relies on a dominant land cover approach, with land cover prescribed using the Corine Land Cover 2006 database (EEA, 2013).

REMO – One simulation is conducted by the GERICS working group, using the non-hydrostatic version of REMO (Regional Climate Model, Jacob and Podzun 1997, Pietikäinen et al. 2025). REMO uses its own internal land surface scheme with a sub-grid tile approach for land, water and sea ice (Semmler, 2002; Kotlarski, 2007). Over land, more than seventy land use / land cover (LULC) types are considered including "urban" surfaces, based on the 1 km resolution Global Land Cover Characterization (GLCC) dataset Version2, primarily derived from Advanced Very High Resolution Radiometer (AVHRR) data (reference years 1992 to 1993; Eidenshink and Faundeen, 1994). Land surface parameters are allocated to all LULC types based on Hagemann et al. (1999) and Hagemann (2002). Surface albedo is taken from Rechid et al. (2009). "Urban" surfaces refer to purely sealed/impervious surfaces, characterized by a relatively high roughness length of 2.5, zero vegetation cover and zero water holding capacity (Hagemann 2002, p.16). The parameter values of different LULC types including urban land cover are averaged for the land part of the grid cell, weighted by their fraction (see also Rechid 2009).

UM – The climate mode of the UM (Met Office Unified Model; Walters et al. 2017) is used for one simulation. UM is based on the UK Met Office operational model (Clark et al. 2016). Land surface processes are simulated by the JULES (Joint UK Land Environment Simulator) LSM (Best et al. 2011; Clark et al. 2011) itself based on the MOSES (Met Office Surface Exchange Scheme) LSM (Cox et al., 1999; Essery et al., 2003). In JULES, urban areas are either treated via a single tile BULK approach (Best 2005) or – as it is the case in this simulation – via a two tiles approach (Best 2006). Here, in addition to JULES, the MORUSES parameterization (Met Office Reading Urban Surface Exchange Scheme,

Porson et al. 2010a,b) is employed. The classification of JULES-MORUSES as a BULK scheme is trickier than for the previous models because it relies on two tiles. One tile represents urban rooftops as a flat surface and another tile represents the combined effects of roads and walls in an urban canyon as a simplified "bulk surface" (which was shown by Porson et al. 2009 to give similar results as treating them separately). Following Lipson et al. (2023) categories for the inter-comparison project Urban-PLUMBER, we also classify MORUSES as a BULK scheme since it only resolves two surface energy balances (Figure 1.a). Inside the canyon tile, radiative exchanges between the road, the walls and the sky are parametrized based on Sparrow and Cess (1970) and the simplified-geometry approach of Harman et al. (2004a,b), and Harman and Belcher (2006), with an effective albedo and effective emissivity (Harman et al., 2004a). Heat transfer is parametrized via the computation of effective roughness lengths for momentum and for heat (Harman et al., 2004b). Storage of heat inside the canopy is parameterized to represent the higher heat capacity of buildings and heat transfer to the soil is only represented in the canyon tile (Porson et al. 2010a). Anthropogenic heat release was implemented in MORUSES by Bohnenstengel et al. (2014) with realistic values for London (England) using an inventory made by Davies et al. (2008a), Davies et al. (2008b), and Hamilton et al. (2009). Unlike other CPMs analyzed in this study, UM is directly driven by ERA-Interim at its boundaries. More details of the configuration used here can be found in Berthou et al. (2020).

RegCM – Version 4.7 of the RegCM (Regional Climate Model; Giorgi et al. 2012; Coppola et al. 2021) is used by the ICTP working group for two simulations at 12 and 3 km horizontal resolution. Multiple LSMs can be used in RegCM, here the Community Land Model (CLM; Bonan 1996) version 4.5 (Oleson et al. 2010, Lawrence et al. 2011) was chosen. CLM4.5 represents the surface as a combination of up to 5 types: glacier, lake, urban, vegetated, and crop. The urban land unit is itself divided into urban density classes representing the building districts, high density urban areas, and medium density urban areas. Urban areas are modeled by the CLM Urban model (CLMU; Oleson et al. 2008a,b and Oleson et al. 2010), which is a SLUCM (Figure 1.b). Urban areas are composed of 5 different surfaces: roof, sunlit walls and shaded walls, and pervious and impervious canyon floor). The urban pervious floor has the role of representing the vegetation inside the canyon, but it is not explicitly modeled, instead evaporation is parameterized by a simplified bulk scheme. The global urban extent as well as the fraction of the different urban classes were developed by Jackson et al. (2010) using remote sensed satellite and population census data (Jackson et al. 2010; Oleson et al. 2010).

AROME – Simulations from two different AROME versions are used in this study: CNRM uses the climate version (Déqué et al. 2016) of the AROME weather prediction model (Seity et al. 2011; Brousseau et al. 2016), while the version used by HCLIMcom and KNMI originates from the HARMONIE-AROME modelling framework (Bengtsson et al., 2017). The specific model versions are CNRM-AROME41t1 (cycle 41; Termonia et al. 2018) for CNRM (Caillaud et al. 2021 updated from Fumière et al. 2020's cycle 38) and HCLIM38-AROME (cycle 38; Belušić et al. 2020) for HCLIMcom and KNMI (Lenderink et al. 2021; De Vries et al. 2024). The surface component of AROME is externalized in the SURFEX LSM (Masson et al. 2013). SURFEX uses a fractional approach to distinguish between vegetation, inland water, sea and ocean, and urban areas. Natural covers are simulated by the soil-vegetation model ISBA (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996) and artificial areas are handled by the SLUCM TEB (Masson 2000). The CNRM simulation uses SURFEX v7.3 while HCLIMcom and KNMI use a blend of SURFEX v7.3 and v8.1, but only for ISBA. All three models activate the TEB SL-UCM at 2.5 km or 3 km, but HCLIMcom and KNMI use the historical version of Masson (2000) where TEB only represents one urban canyon level (Figure 1.b). CNRM uses the TEB-SBL version of Hamdi and Masson (2008) where multiple canyon levels are represented by resolving the surface boundary layer inside and above the urban canopy thanks to the inclusion of a drag force approach taking into account the vertical effect of buildings (Figure 1.c). Even with the difference in the number of levels inside the canyon, both versions of TEB compute only one energy balance per surface (roof, wall and road) and are coupled to the atmosphere by a single level. Multi-level coupling of TEB (Figure 1.c, Schoetter et al. 2020) is not yet available in AROME. The physical surface parameters required to run SURFEX are provided by the ECOCLIMAP databases, version I for CNRM (Champeaux et al. 2005) and version II for HCLIMcom and KNMI (Faroux et al. 2013). In both versions of TEB used here, urban vegetation is not represented directly inside the canyon (as is the case in later versions of TEB, Lemonsu et al. 2012; Redon et al. 2017, 2020), instead the maximum possible urban fraction in a grid point is set to 90%, the rest being treated as natural cover by ISBA. This approach simulates part of the effect of urban vegetation when fluxes (as well as diagnostic variables) are aggregated on a grid point, but does not represent feedback between vegetation and the urban canyon (e.g. change in canyon geometry or radiation shading, Lemonsu et al. 2012). The CNRM simulation has been evaluated several times, and most notably Michau et al. (2023) evaluated its ability to simulate UHI over multiple cities in France.

In addition to the CPM simulations, the outputs of the RCMs used as an intermediate downscaling step between the ERA-Interim reanalysis and the CPMs are also available (Table 1). Among these, CMCC's CCLM and the three AROME simulations use different urban parametrisation between their RCM and their CPM, with more sophisticated schemes in the CPM (Table 1). The 12 km CCLM5-0-9 simulation (CMCC) does not represent urban areas at all. Two of the three CPMs from the AROME family (with the SLUCM TEB) are driven by different version of the ALADIN RCM: CNRM-ALADIN62 (Nabat et al. 2020) for CNRM and HCLIM38-ALADIN (Belušić et al., 2020) for HCLIMcom, both of these RCMs use a BULK approach with the vegetation model ISBA without modifying surface parameters. Instead, the ECOCLIMAP urban classes are converted into the rock class and treated by ISBA. The KNMI CPM is driven by a different RCM: RACMO23E (van Meijgaard et al., 2008). RACMO23E is a hydrostatic model based on the dynamical core of HIRLAM and physics packages of the ECMWF-IFS cycle Cy33r1 (ECMWF-IFS, 2008; van Meijgaard et al., 2008, 2012; Noël et al. 2015). It uses the ECMWF LSM HTESSEL to simulate land surfaces (Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land; Viterbo & Beljaars, 1995; Balsamo et al., 2009). By default HTESSEL follows a BULK approach to represent urban areas, treating them the same way as bare soil and desert (Luu et al. 2023). Lastly, MOHC didn't use a RCM to drive its CPM, but directly downscaled ERA-Interim. Their EURO-CORDEX simulations using the HadREM RCM (McCarthy et al. 2012) is used as a comparison.

Table 1. List of FPS-CONV CPM evaluation simulations and their respective driving RCM (nested into the ERA-Interim reanalysis). The corresponding Land Surface Model (LSM) and urban representation (URB) are also included. Table cells are colour-coded according to the urban representation: No urban areas (green), bulk (grey) or urban canopy model (UCM; blue). Table cells in dark grey represent RCMs with no available data. For full institute names see Ban et al. 2021 and Pichelli et al. 2021.

RCM/ CPM family	Institute	RCM (grid spacing km)	RCM LSM (URB)	CPM (grid spacing km)	CPM LSM (URB)	Fractional or dominant land cover approach
COSMO	BTU	CCLM5-0-14 (12)	TERRA-ML (NO URB)	CCLM5-0-14 (3)	TERRA-ML (NO)	Dominant
	ETH	COSMO-crCLIM (12)	TERRA-ML (NO URB)	COSMO-crCLIM (2.2)	TERRA-ML (NO)	Dominant
	JLU	-	-	CCLM5-0-15 (3)	TERRA-ML (NO)	Dominant
	KIT	-	-	CCLM5-0-14 (3)	TERRA-ML (NO)	Dominant
	CMCC	CCLM5-0-9 (12)	TERRA-ML (NO URB)	CCLM5-0-9 (3)	TERRA-ML (TERRA-URB; BULK)	Fractional*
WRF	AUTH-MC	WRF381BG (15)	NOAH-MP (BULK)	WRF381BG (3)	NOAH-MP (BULK)	Dominant
	BCCR	WRF381DA (15)	NOAH-MP (BULK)	WRF381DA (3)	NOAH-MP (BULK)	Dominant
	BCCR	WRF381BF (15)	NOAH-MP (BULK)	WRF381BF (3)	NOAH-MP (BULK)	Dominant
	CICERO	WRF381BJ (15)	NOAH-MP (BULK)	WRF381BJ (3)	NOAH-MP (BULK)	Dominant
	FZJ	WRF381BB (15)	NOAH (BULK)	WRF381BB (3)	NOAH (BULK)	Dominant
	IDL	WRF381BH (15)	NOAH (BULK)	WRF381BH (3)	NOAH (BULK)	Dominant
	IPSL	WRF381BE (15)	NOAH-MP (BULK)	WRF381BE (3)	NOAH-MP (BULK)	Dominant
	UCAN	WRF381BI (15)	NOAH-MP (BULK)	WRF381BI (3)	NOAH-MP (BULK)	Dominant
	UHOH	WRF381BD (15)	NOAH-MP (BULK)	WRF381BD (3)	NOAH-MP (BULK)	Dominant
	WEGC	WRF381BL (15)	NOAH (BULK)	WRF381BL (3)	NOAH (BULK)	Dominant
REMO	GERICS	REMO2015 (12)	REMO LSM (BULK)	REMO-NH (3)	REMO LSM (BULK)	Fractional
UM	MOHC	HadREM3-GA7-05** (12)	JULES (MORUSES; 2-TILES BULK)	UM10.1 (2.2)	JULES (MORUSES; 2-TILES BULK)	Fractional

RegCM	ICTP	RegCM4-7 (12)	CLM (CLMU;UCM)	RegCM4-7 (3)	CLM (CLMU; UCM)	Fractional
AROME	CNRM	CNRM-ALADIN62 (12)	SURFEX (BULK)	CNRM-AROME41t1 (2.5)	SURFEX (TEB-SBL; UCM)	Fractional
	HCLIMcom	HCLIM38-ALADIN (12)	SURFEX (BULK)	HCLIM38-AROME (3)	SURFEX (TEB; UCM)	Fractional
	KNMI	RACMO23E (12)	HTESSEL (BULK)	HCLIM38h1-AROME (2.5)	SURFEX (TEB; UCM)	Fractional

* Because the TERRA-URB parametrisation is activated in the CCLM5-0-9 simulation, urban and rural areas are separated using a “poor man’s tile approach” (Wouters et al. 2017)

** The MOHC UM simulation was directly driven by the ERA-Interim reanalysis, without the use of an intermediate RCM. The EURO-CORDEX HadREM simulation is used for comparison purposes.

3.2. Selection of cities, observations and model grid points

This study analyzes the urban climate of six cities within the FPS-CONV domain; they are selected on the basis of their spatial extent, their geographical/climatic context and the availability of urban temperature observations (Figure 2). They can be grouped in 3 different climate zones following the Köppen–Geiger climate classification (Köppen, 1884; Geiger, 1954). Paris (France) and Frankfurt (Germany) have temperate oceanic climates. Munich (Germany) has a temperate oceanic climate and some humid continental climate features. Milan (Italy) has a humid subtropical climate. Both Munich and Milan, are close to the Alps and are occasionally affected by warm and dry downhill winds (föhn). Lastly, Barcelona (Spain) and Marseille (France) are both coastal cities with hot-summer Mediterranean climates (Beck et al. 2018).

Due to the lack of urban measurements over long periods of time, previous studies commonly relied on the weather stations at the airports closest to the cities as the urban reference. Here we made the choice to select cities where urban observations were available inside the city’s limits. Consequently, some of the urban stations used aren’t located 2 meters above the ground (some are placed on rooftops or in streets). Therefore, observation uncertainty must be kept in mind for this analysis because the most standards stations are located in urban parks, which likely leads to an underestimation of urban temperatures (e.g. in Paris, Frankfurt and Munich), while the most urbanised stations might be contaminated by other urban processes (radiative effects of walls and roofs, anthropogenic heat released by traffic, etc.).

All stations used are shown in Figure 2 and the urban stations are listed in Table 2 with information regarding the characteristics of their surroundings showing the dominant Local Climate Zone (LCZ, Demuzere et al. 2022) and the Copernicus surface imperviousness fraction computed in a 250 m radius buffer (EEA, 2018). Rural stations (Table SM1) were selected in a 60 km radius surrounding the city and when possible with a difference in elevation smaller than 50 m relative to the city. The stations cover different geographic directions on land as much as possible, to not be affected by eventual regional temperature gradients. For each city, urban and rural stations are averaged respectively. The urban average is only computed if all urban stations of a city have data (only Milan, where there are multiple urban stations, satisfies this criterion), whereas the rural average allows one rural station to be missing at a time.

The closest grid-point to each station is extracted from the models and considered representative of the respective station characteristics; after checking that rural (urban) points do not have too high (low) urban fractions, especially for the models relying on a dominant cover approach. This approach was found more reliable than using averages of the 9 grid-points surrounding the stations (see sensitivity analysis in Supplementary Materials). In order to be as fair as possible for each model, especially the coarse RCMs, we made some adjustments. For coastal cities, we made sure to select model grid-points that are over land even if they are not the closest ones (this concerns the rural station of Barcelona for all RCMs and one rural station for Marseille). We use hourly model output of near-surface air temperature and compute daily minimum (TN) and maximum (TX) values - just as it is done in the observations. This can differ between meteorological institutes with Météo-France (Paris and Marseille) computing the minimum temperature between 18 UTC of the previous day and 18 UTC of the current day, and the maximum temperature between 6 UTC and 6 UTC of the following day; while other institutes compute them between 0 and 24 UTC. Before averaging each model points into one urban and one rural temperature, the days missing from the observations are also removed from the models. Due to data availability, the comparison with observations does not always cover the same time period across different cities; Marseille and Barcelona have the shortest time periods due to their urban stations (Table 2)

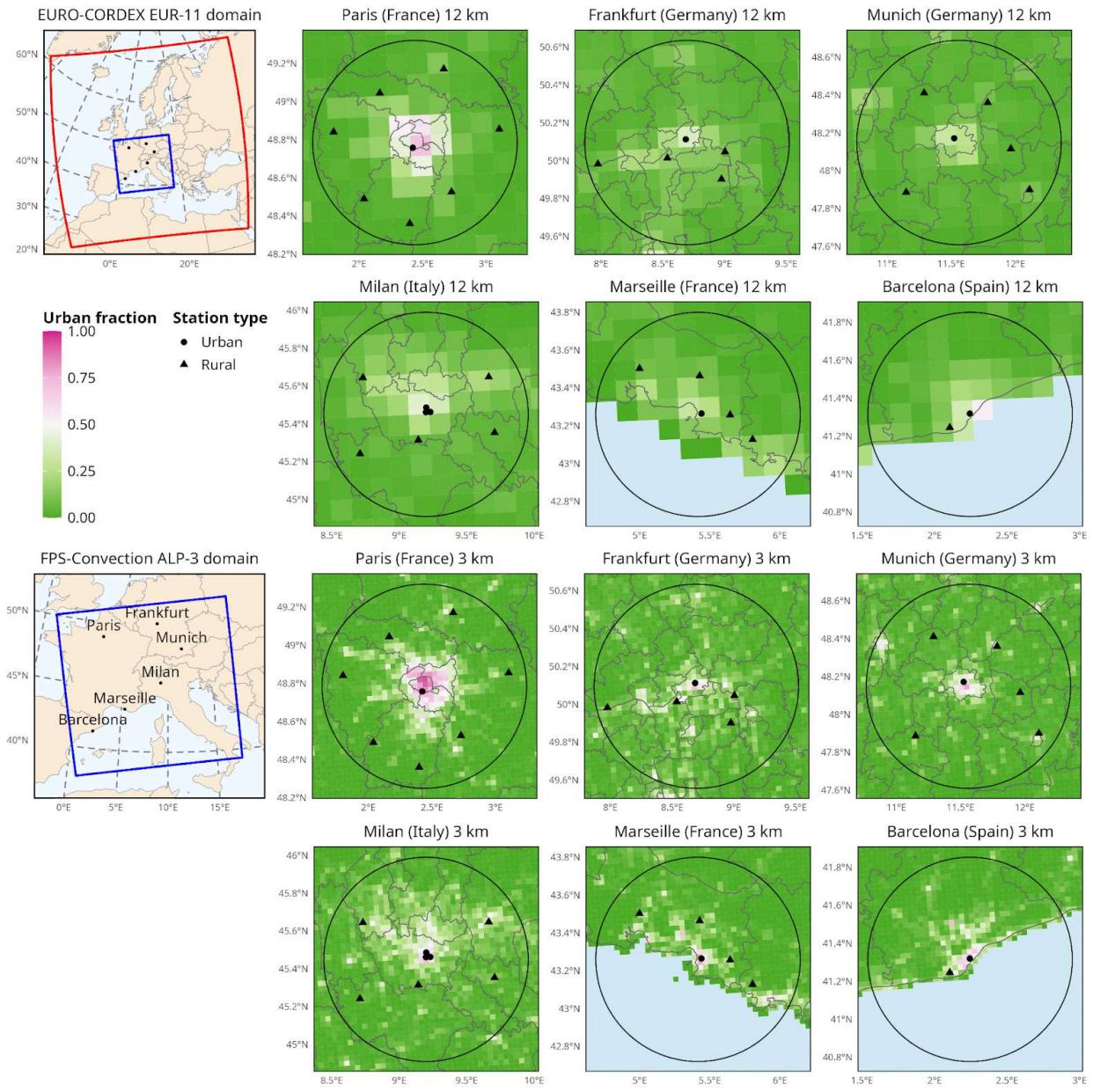


Figure 2. Simulation domains for the EURO-CORDEX and the FPS-CONV simulations (red and blue boxes in the left panels, respectively). To get an idea of urban density, gridded imperviousness (EEA, 2018) is interpolated at two grid resolutions: 12 km (representative for the RCMs) and 3 km (representative of the CPMs). Note not all models use the same grid at the same resolution, but all RCMs are around 12 km and all CPMs around 3 km resolution. The black circle marks a 60km radius around the city centers. Light grey lines delineate the European Union's Nomenclature of territorial units for statistics (NUTS) level 3 (small regions).

506

Table 2. List of the urban stations used. The complete list with rural stations can be found in Supplementary Materials (Table SM1).

City	Provider	Station Name	Availability period	Coordinates Latitude Longitude	Altitude (in m)	Dominant LCZ (and its fraction)* Imperviousness**	Location
Barcelona	Servei Meteorològic de Catalunya	Barcelona El Raval	2007-2009	41.384 2.168	33	2 (100%) 85%	Rooftop
Frankfurt	Deutscher Wetterdienst	Frankfurt/ Main-Westend	2000-2009	50.127 8.669	121	5 (90%) 46%	University campus with vegetation
Marseille	Météo France	Marseille-Obs	2000-2002	42.306 5.394	75	2 (74%) 60%	Rooftop of a building in a park
Milan	ARPA Lombardia	Milano v.Juvara	2000-2009	45.473 9.222	122	2 (71%) 56%	Street level
	ARPA Lombardia	Milano v.Brera	2000-2009	45.472 9.189	122	2 (96%) 79%	Rooftop
	ARPA Lombardia	Milano v.Marche	2000-2009	45.496 9.191	129	2 (89%) 70%	Street level
Munich	Deutscher Wetterdienst	München-Stadt	2000-2009	48.163 11.543	515	5 (56%) 67%	DWD campus backyard with vegetation
Paris	Météo France	Paris-Montsouris	2000-2009	48.822 2.338	75	5 (68%) 27%	Park

* The LCZ fractions are computed in a 500 m buffer around the stations using Demuzere et al. (2022) 100 m resolution global map. LCZ2 and LCZ5 refer to the “compact midrise” and “open midrise” built types, respectively.

** The imperviousness fraction is computed the same way using Copernicus' imperviousness density 2006 dataset at 100 m resolution (EEA, 2018).

3.3. Evaluation metrics

The ability of models to properly represent observed temperatures is evaluated on seasonal timescales by looking at average seasonal biases over winter (DJF) and summer (JJA). The TN and TX are evaluated over urban and rural points/stations (TN_{URB} and TN_{RUR}), respectively. UHI are defined as the difference between the urban and rural averages with UHIN for TN and UHIX for TX: $UHIN = TN_{URB} - TN_{RUR}$. The significance of the differences between the modeled and observed seasonal distributions is evaluated using a two-sample unpaired Wilcoxon–Mann–Whitney rank sum tests with continuity correction (Wilcoxon 1945; Mann and Whitney 1947). In addition to biases, the Pearson correlation coefficients (cor.) between simulated and observed UHI are also computed.

In addition to average biases, seasonal distributions are also analysed (between the 1st and 99th percentile). To summarise the potential added value of CPMs compared to RCMs in simulating seasonal distributions of UHIN, we use a similar approach as Soares and Cardoso (2018). They proposed the Distribution Added Value index (DAV), which is calculated based on the score developed by Perkins et al. (2007) representing the quality of the fit between two histograms. Perkins' score (PS) is defined as: $PS = \sum_{1}^{n} min(Z_{model}, Z_{observation})$ where n is the number of bins used to calculate the histograms, with Z_{model} and $Z_{observation}$ the frequency of values in a given bin in the modeled and observed data, respectively. The PS ranges between 0 and 1, with 1 representing a perfect match between the two histograms. The DAV is then the relative difference between the PSs of models of different resolutions: $DAV = \frac{PS_{HIGH-RES} - PS_{LOW-RES}}{PS_{LOW-RES}}$ (RCMs in Soares and Cardoso 2018; and RCMs and CPMs in Sangelantoni et al 2024). Because the DAV is relative to the PS of the RCM, it has no upper bound and can reach immense values when the latter is close to 0 (e.g. when applied to the UHI distribution of RCMs that do not represent urban areas at all), making its interpretation difficult. Therefore, we instead use the simple difference in the PSs between the two resolutions: $\Delta PS = PS_{CPM} - PS_{RCM}$, akin to Ciarlo et al. (2020)'s added value index. ΔPS ranges between -1 and 1, with positive values showing an added value of the CPM. The computation of the PS requires choosing a bin size for the histograms; for temperature a

bin size of 1 °C is common practice (although Perkins et al. 2007 originally chose 0.5 °C), but because UHI distributions cover narrower ranges than temperatures we rely on the Freedman–Diaconis rule to find more appropriate values: $Bin\,width = 2 \times \frac{IQR(x)}{\sqrt[3]{n}}$, where IQR is the interquartile range and n the number of days (Freedman & Diaconis, 1981).

4. Evaluation

4.1. Daily minimum and maximum urban and rural temperatures

The ability of the models to simulate daily minimum and maximum temperatures (Figure 3 and Figure 4, respectively) is evaluated by looking at average seasonal biases for winter (DJF) and summer (JJA) over the 2000-2009 period (depending on observations). We consider here that the biases of rural temperatures (TN_{RUR} and TX_{RUR}) should give an insight into the "*general biases*" of the models (i.e. not related to their urban parameterization) and even if they are calculated on single grid cells, they should be representative of surrounding non-urban areas (with the exception of Barcelona, where we only have one rural station).

Almost all CPMs tend to overestimate TN_{RUR} in DJF (Figure 3.a), up to 3 K, with two exceptions where biases are negatives: the WRF ensemble over Paris, Frankfurt and Munich, and the two HCLIM-AROME simulations for Paris, Frankfurt and Barcelona. Across all models and cities, absolute biases are almost always significant when above 0.5 K. When absolute biases are greater than 1 K, they seem to be systematic, with more than 2/3 of the season having days biased in the same direction as the average (colored text). This is not the case under 1 K (e.g. most models over Paris) where models are closer to the observations but also more variable. There is a general coherence between the same CPMs with different versions/parametrizations (e.g. CCLM overestimates TN_{RUR} for all cities, WRF is split between northern and southern cities and HCLIM-AROME tends more towards underestimation). For the CPMs that overestimate TN_{RUR} almost systematically (e.g. RegCM, CCLM, UM), a first hypothesis could be that the rural points are "contaminated" by an urban fraction leading to higher temperatures; but this is contradicted by the non-urban CCLM simulations that show warm biases similar to CMCC's. Most RCMs are colder than the CPMs they drive. Consequently, when RCMs overestimate TN_{RUR}, CPMs exacerbate these biases (e.g. WRF over Milan and Marseille;

RegCM, REMO and CCLM in general). Nevertheless, there is an overall coherence between RCMs and CPMs with the exception of the KNMI simulations because of their different models at the two resolutions. Another exception concerns Barcelona where TN_{RUR} is underestimated by almost all RCMs (albeit with small yet significant biases) and conversely overestimated almost all CPMs. This difference cannot be due to the presence of a sea fraction in the model's rural point because we avoid it in the RCMs (see Section 3.2) and it is naturally absent in the CPMs due to their resolution. Instead it might be explained by the difference in elevation between the two resolutions that is already important when selecting the same coordinates, and is even more exacerbated by our selection aimed at avoiding points covered by the sea.

Now focusing on TN_{URB} (Figure 3.c and Figure 3.d), the first thing to note are the higher temperatures in the observations than for TN_{RUR} (i.e. the presence of an UHI effect in DJF for all cities). Compared to TN_{RUR}, most models have cold biases, both at the CPM and RCM resolution. This is to be expected for the simulations that do not represent urban areas at all: the CCLM CPMs without an urban parametrization or KNMI and CMCC RCMs. For the simulations that do represent urban areas, this highlights potential limitations of the parametrizations. The CPMs showing warm biases in TN_{URB} (Figure 3.c) also showed even greater warm biases in TN_{RUR} (Figure 3.a), meaning that the urban parametrization is not necessarily simulating urban temperatures that are too high, but rather compensating for the general warm bias of the model (e.g. RegCM, CMCC-CCLM and UM). There are two notable situations where TN_{URB} are underestimated the most: models with no urban parametrizations such as CCLM, and RCMs that have either no urban parametrization (e.g. KNMI-RACMO and CMCC-CCLM) or a BULK combined with low urban fractions (see Barcelona for all RCMs but WRF).

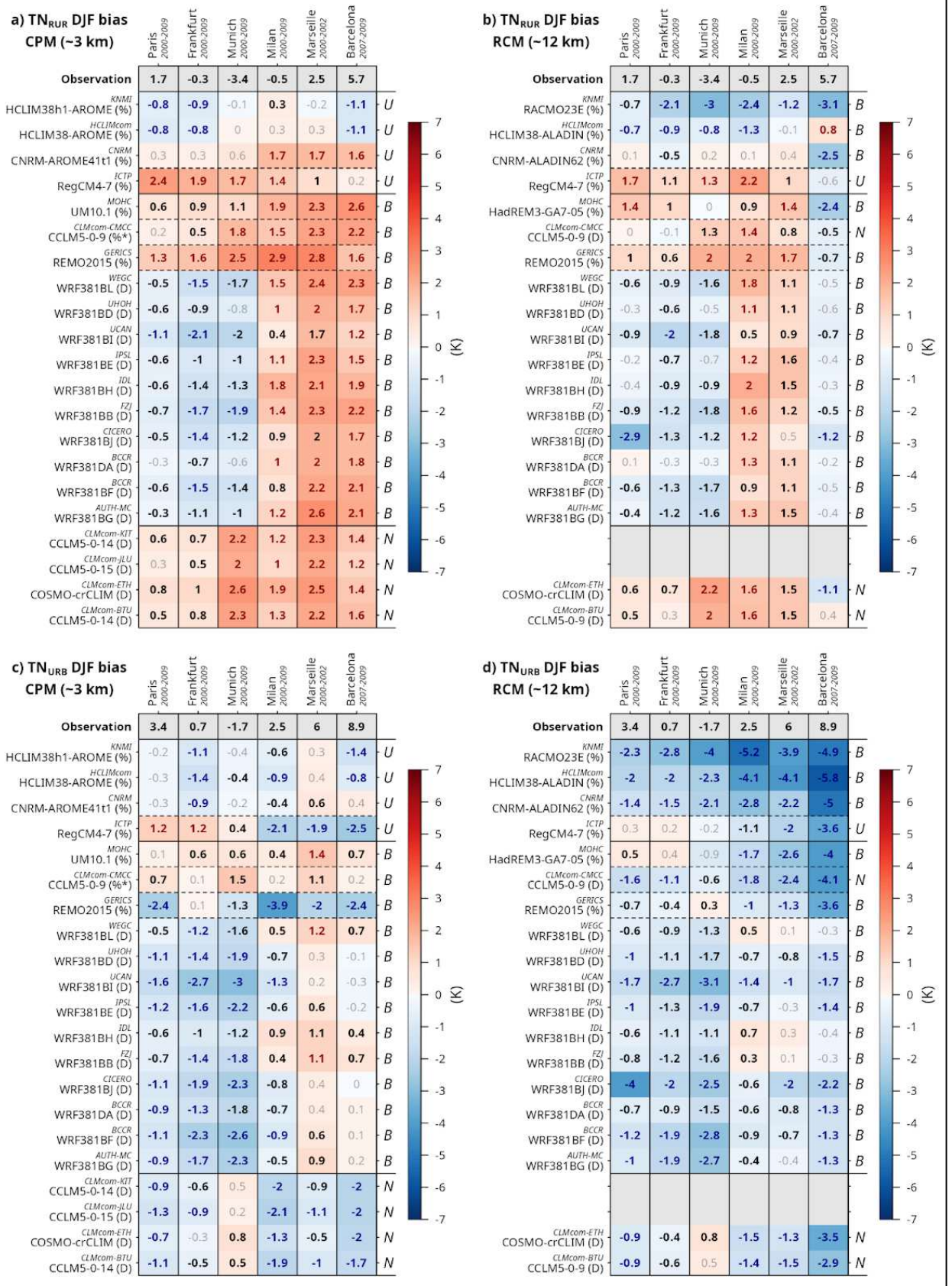


Figure 3. Average winter (DJF) biases for TN_{RUR} (a, b) and TN_{URB} (c, d) for the 21 CPMs (a, c) and their driving RCMs (b, d). Bold text highlights the model and city for which the difference is significant (see text). Bold red and blue text highlight the model and city for which the difference is significant and where more than 2/3 of the data is biased in the same direction as the average. Models are grouped based on the urban

parameterizations of the 3 km CPMs (N: no parametrization; B: BULK; U: Urban Canopy Model). The "%" and "D" signs next to the model names indicate the use of a fractional or dominant cover approach, respectively. Horizontal dashed lines separate different urban schemes of the same category. Note that due to missing data the metrics are not always computed over the same time period (shown by each city label) for the different cities; the missing days in the observation are also removed from the model for that city.

601

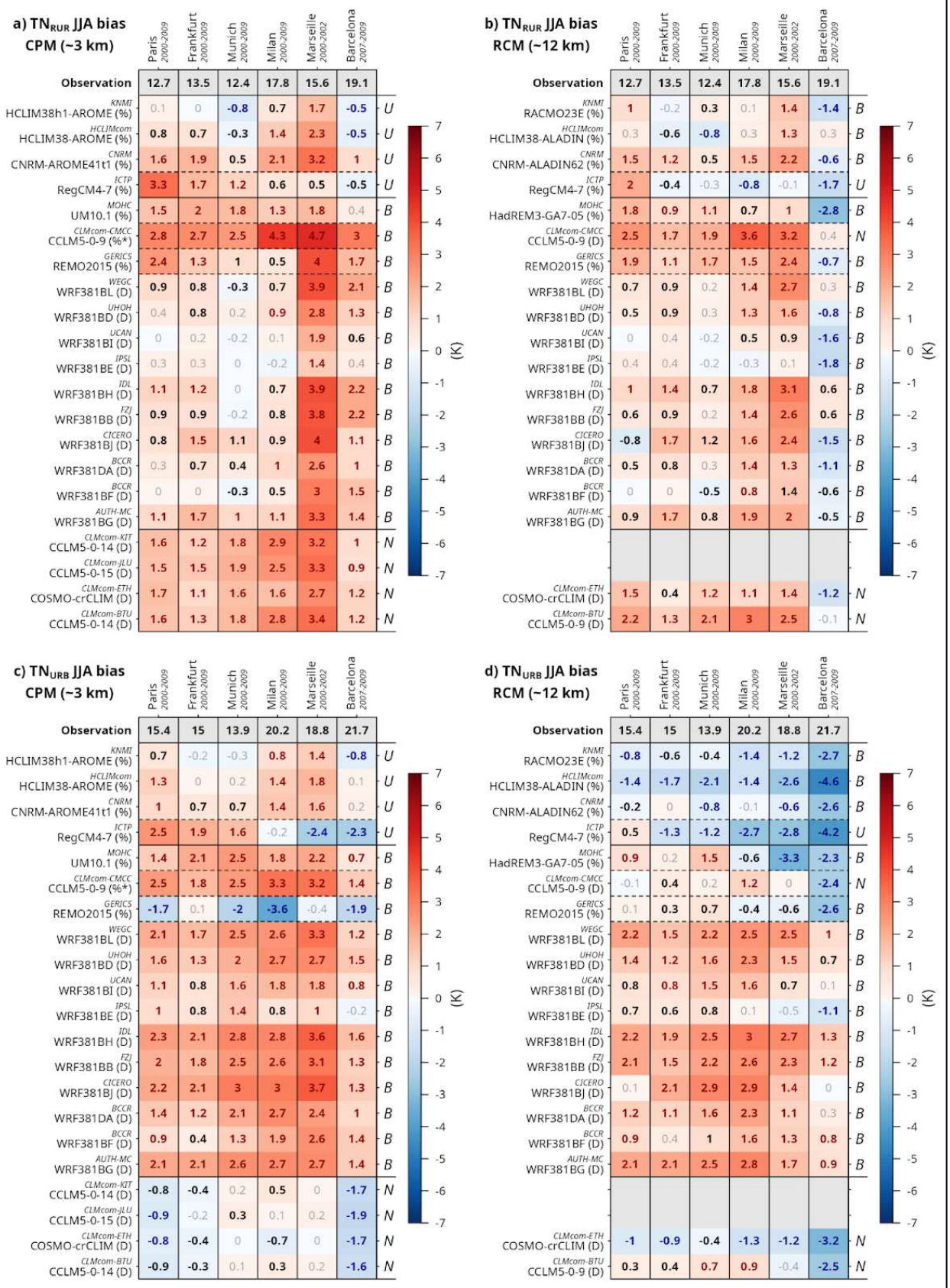


Figure 4. Same as Figure 3 but for TN during summer (JJA).

During JJA, almost all CPMs overestimate TN_{RUR} (Figure 4.a) between 1 and 3 K (and up to above 4 K for Marseille), with the exception of some WRF simulations for a few cities

(mostly Munich) where they show small biases. The only significant cold biases are smaller than 1 K (both HCLIM simulations over Munich and Barcelona). Absolute biases are greater in JJA than DJF for some models: e.g. CCLM, but this is not systematically the case for other models across cities (WRF). The greater biases noted in Marseille in DJF are even more apparent, reaching up to 4.7 K in CMCC's CCLM simulation. Besides Marseille, CMCC's CCLM simulation also stands out over Milan (4.3 K). Other CCLM CPMs also show large biases for Milan, but smaller (around 2 to 3 K). The differences in biases between the CCLM simulations might be due to an improvement of the skin conductivity made in the TERRA LSM by Schulz and Vogel (2020), that is absent in CMCC's model version.

When comparing RCMs and CPMs, we see that CPMs almost always simulate warmer TN_{RUR} than their driving counterpart, which most of the times increases warm biases; with the exception of Barcelona where RCMs sometimes show strong cold biases, and for WRF where the dominant cover approach most likely partially mitigates the added value of the increased resolution. The differences in biases between CPMs and RCMs are for the most part small (around 1 K), with the exception of CMCC's CCLM. This is consistent with Soares et al. (2024), who noted little to no added value on the distribution of TN in JJA (with the exception of coastlines) compared with the E-OBS dataset (Cornes et al. 2018), and to the findings of Cardoso and Soares (2022) for RCMs of different resolutions.

Over urban points, all CPMs with an urban parameterization overestimate TN_{URB} for all cities (Figure 4.c), with the exception of REMO and RegCM. Non urbanized CCLM CPMs show cold biases or no biases at all when compensating for their warm TN_{RUR} biases. All CPMs using a BULK approach overestimate TN_{URB} (up to 3.7 K). In the case of WRF, apart from the possible biases inherent to its BULK parameterization, biases could also be due to the dominant land cover approach which does not allow any vegetation in the urban grid points (while Table 2 shows that the station surroundings are never fully urbanized). It should be stressed that the biases in TN_{URB} ought to be put into perspective with those noted previously in TN_{RUR}: for example RegCM shows warm biases in TN_{URB}, which are not due to its urban parameterization because they are lower than those in TN_{RUR}, i.e. the model overestimates TN in general but underestimates the effect of the city.

Urban Heat Islands are mainly a nocturnal phenomenon (Oke et al. 2017), which is why differences between urban and rural temperatures and between models with varying resolutions and parameterizations should be less pronounced during the day (Lemonsu et

al. 2023) - i.e. when looking at maximum temperatures (Figures 5 and 6). On the other hand, higher daytime temperatures may be of particular interest because of their impact on human health (Basu et al. 2002; Cheng et al. 2014; Liu et al. 2022), especially during extreme events such as heat waves (Ebi et al. 2021; Romanello et al. 2023). Figures 5 and 6 show the average biases for TX_{URB} and TX_{RUR}, respectively for DJF and JJA. Overall, there are fewer differences between CPMs in TX than TN as most of them underestimate TX_{RUR}. Very few CPMs overestimate the maxima in DJF, and only for one or two cities (4 WRF simulations for Milan, and REMO for Milan and Barcelona; Figure 5.a). In general, the biases are around -1 K and -2 K with some models performing quite well (HCLIM-AROME and CNRM-AROME for all cities except Marseille, and RegCM but only for continental cities). Once again, driving RCMs are for the most part colder than their higher-resolution counterparts, with the exception of a few pairs of WRF simulations and cities where the biases are similar (Figure 5.b). The same picture is seen for TX_{URB} (Figure 5.c and 5.d), most likely because - as mentioned earlier - urban effects are expected to be less pronounced during the day. As with rural temperatures, CPMs perform better than their driving RCMs due to the aforementioned reduced cold bias and also because they seem to simulate some small urban effects, particularly the CPMs incorporating UCMs.

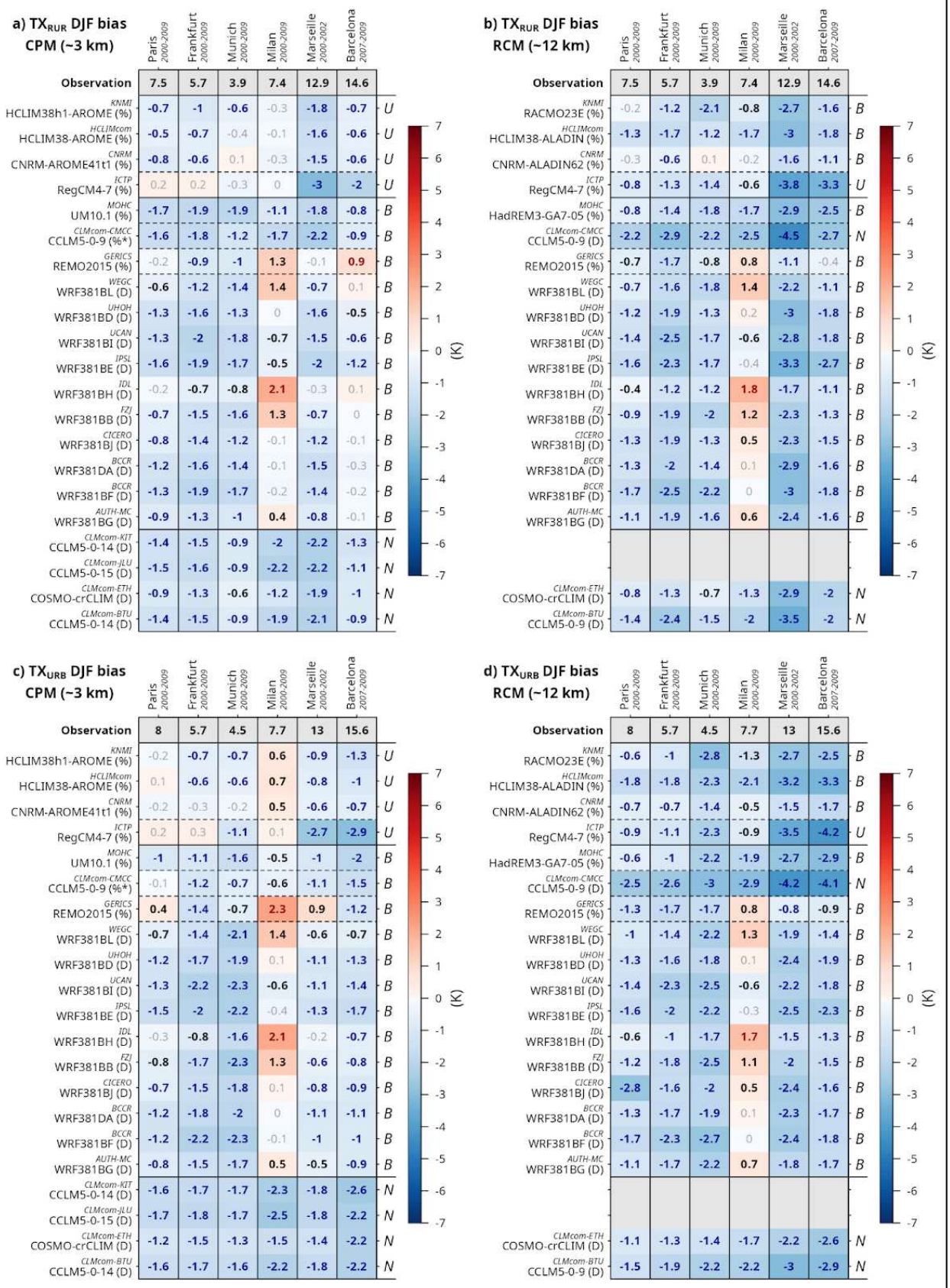


Figure 5. Same as Figure 3 but for TX DJF

In JJA, the average TX biases (Figure 6) are much more erratic than (DJF; Figure 5) or for TN (DJF and JJA; Figures 3 and 4). The same model configuration may show divergent biases between different cities that cannot be attributed to a simple difference between inland and coastal cities (e.g. BCCR's WRF381BF; Figure 6.a). Different versions of the same model may differ in the sign of the biases for the same city, (e.g. the CMCC and JLU CCLMs compared to the other CCLM simulations, or the HCLIM-AROME and CNRM-AROME simulations). These two examples seem to highlight the greater sensitivity of TX to small changes in model parameterizations and differences in local surface properties. Overall, with the exception of two special cases, biases rarely exceed 1.5 K. The first exception concerns the IPSL WRF CPM, which systematically underestimates TX_{RUR} (by up to -3.8 K for Marseille; Figure 6.a). The second exception and most prominent point is that all models (except for the IPSL WRF mentioned above) significantly overestimate TX in Milan, often by more than 2 K and up to 4.6 K (CICERO-WRF), with the smallest biases coming from models with a systematic cold bias for other cities (e.g. ETH-CCLM and RegCM; Figure 6.a). Similar patterns are found in the RCMs (Figure 6.b) and for TX_{URB} (Figure 6.c) with the same effects mentioned earlier: RCMs are colder on average which is improved (degraded) by CPMs when biases are cold (warm). The only important difference between RCM and CPM is noted for RegCM where the RCM shows a systematic cold bias.

The general warm bias in TX_{RUR} for most CPMs was already noted by Sangelantoni et al. (2024) in JJA in general and also during more extreme events such as heat waves (Soares et al. 2024; Sangelantoni et al. 2024) especially over low-elevation areas such as southern France or northern Italy (although we do not observe it for Marseille here). Ban et al. 2021 have pointed out that RCMs tend to overestimate the frequency of precipitation, which is improved by the ability of CPMs to resolve deep convection. However, as a result, and combined with the simpler representation of surface hydrology in current models (Barlage et al. 2021; Picher et al. 2021), this leads to drier surface conditions which in turn have an impact on soil water content and the partitioning of surface heat fluxes, with more sensible heat and less latent heat (Sangelantoni et al. 2024). This could explain larger biases in CPMs but not the ones at both resolutions over Milan. Other studies have also proposed that models are unable to represent temperature inversions (Bucchignani et al. 2016a; Adinolfi et al., 2023) present over flat terrain such as the Po Valley (Cacciamani et al. 1995; Rossa et al., 2012) or also that these biases may be due to the absence of crop irrigation in both the RCM and the CPM they drive (Asmus et al. 2023).

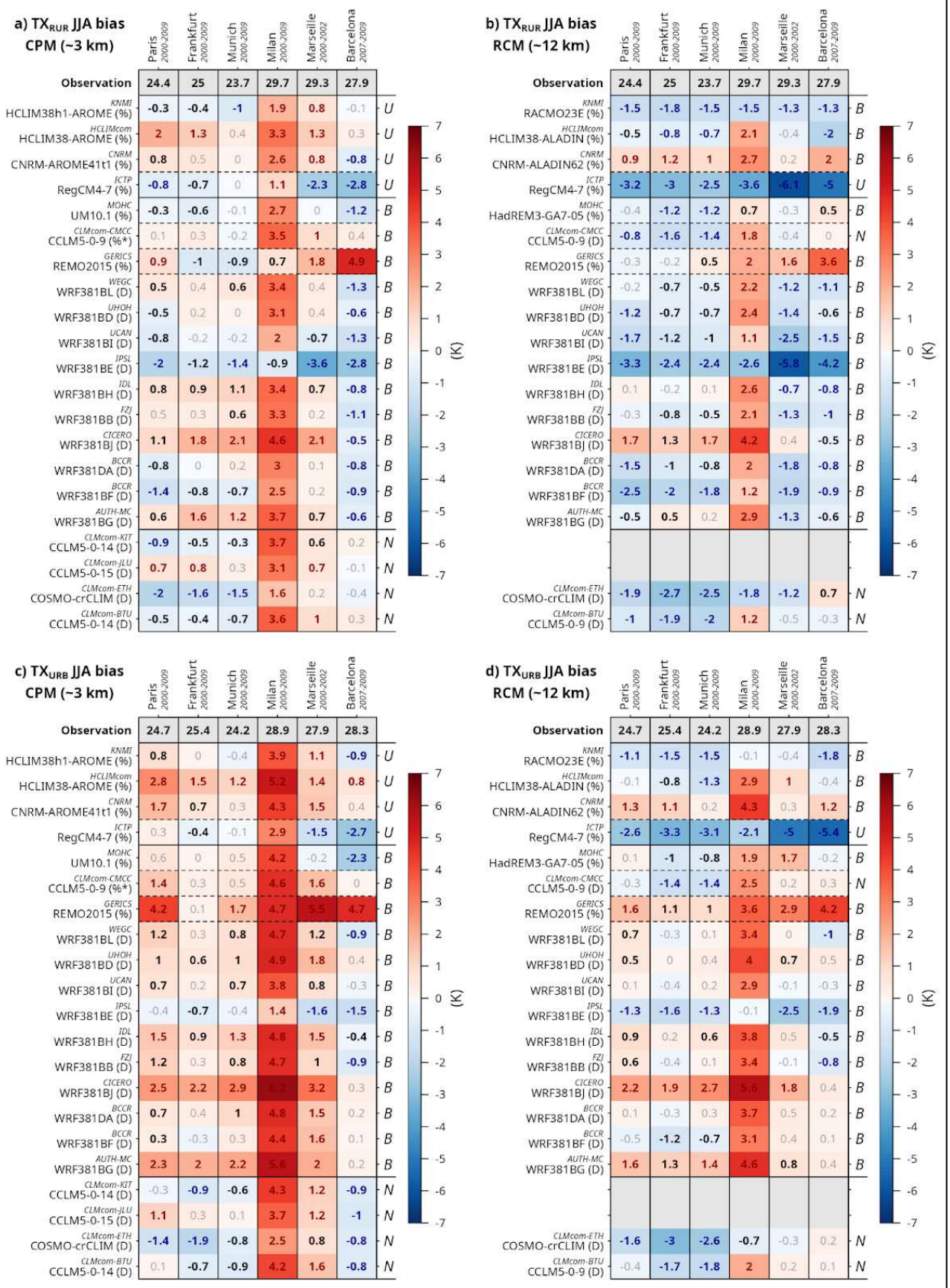


Figure 6. Same as Figure 7 but for TX JJA

This subsection highlighted that CPMs generally perform well, with relatively low seasonal biases (less than 1 or 2 K). However, there are a few exceptions, such as Milan during JJA

(particularly for TX). Performance varied substantially from city to city and between models. Nevertheless, several general trends were observed: TN_{RUR} is generally overestimated by CPMs (except for a few in DJF); TX_{RUR} shows greater discrepancies between models and cities during both seasons; and CPMs are, on average, warmer than their driving RCMs. This leads to an improvement or exacerbation of biases already present in the coarser-resolution RCMs.

4.2. Urban Heat Islands seasonality

We examine the representation of average UHI intensities and their evolution throughout the year. To this end, Figure 7 shows the evolution of monthly averages of UHIN (y-axis) relative to TN_{RUR} (x-axis) for observations (black line and labels), CPMs (solid coloured lines) and RCMs (coloured lines with dots). For clarity, CPMs and RCMs are divided into different panels for each city, and models are also grouped according to the urban parameterization of the CPM. The first two columns group the CPMs with no urban parameterization or BULK schemes (first column) and their driving RCMs (second column). The third and fourth columns show the CPMs with a UCM (third column) and their driving RCMs (fourth column). Each RCM is placed in the panel next to its associated CPM and not according to its own urban parameterization.

Focusing first on observations, three types of seasonality emerge, which correspond in some ways to the group of cities proposed in Section 3.2. All cities but Munich show distinct seasonal patterns already noted in observed surface UHI and modeled (canopy) UHI (Zhou et al 2013, Zhou et al. 2016, Manoli et al. 2020, Yang and Zhao 2023, Sismanidis et al. 2022, Li et al. 2024, Shen et al. 2024). Paris and Frankfurt clearly show higher UHINs during JJA and lower UHINs during DJF. UHIN intensities are highest between late spring and early JJA due to the combination of greater amounts of solar radiation stored in urban materials and active vegetation enhancing rural cooling. Munich's UHIN shows less pronounced seasonality, with maximum intensities in late winter and early spring. Milan, Marseille and Barcelona have similar annual UHIN cycles that differ significantly from those of the previous cities. They are as variable as that of Paris, but with maximum intensities in winter-spring and minimum intensities in summer-autumn. The Marseille cycle appears more irregular, which could be explained by the shorter observation period (only 3 annual cycles are averaged), but the models show consistent

730 patterns (not shown) when the entire period is taken into account (i.e. without removing
731 days in the models for which observations are missing).
732

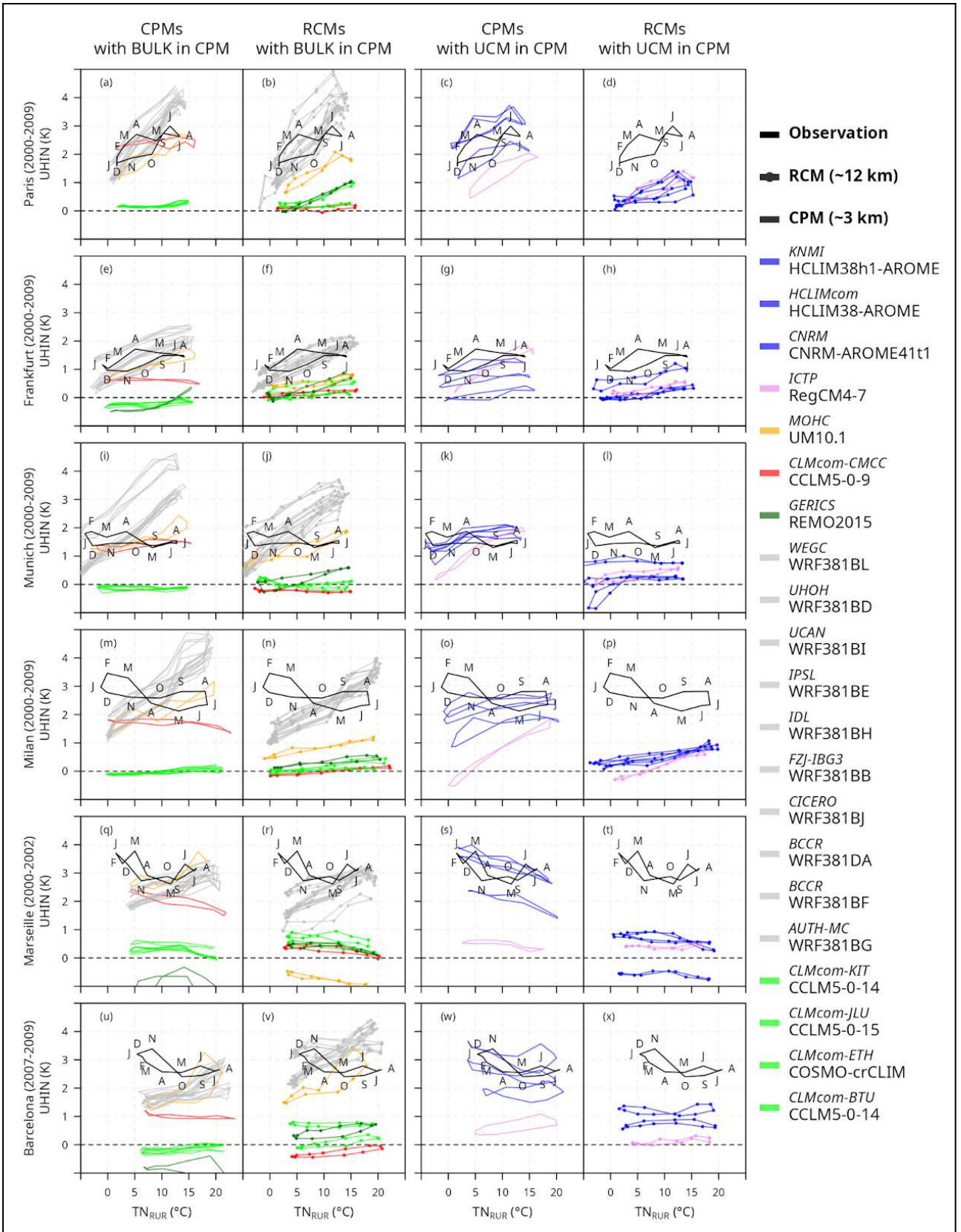


Figure 7. Monthly evolution of night time Urban Heat Islands (UHIN) and daily minimum rural

temperature (TN_{RUR}) observed and simulated by CPMs and RCMs. Monthly average of the periods 2000-2009 for all the cities but Barcelona (2006-2009) and Marseille (2000-2002).

In addition to Figure 7, DJF and JJA seasonal biases for UHIN intensities are presented in Figure 8 and Figure 9. The values should be identical to the differences highlighted previously (Figure 3 and Figure 4) between urban and rural biases (with some rounding differences), but the significance of the differences might vary. The Pearson correlation coefficients (cor.) between simulated and observed UHIN are also displayed.

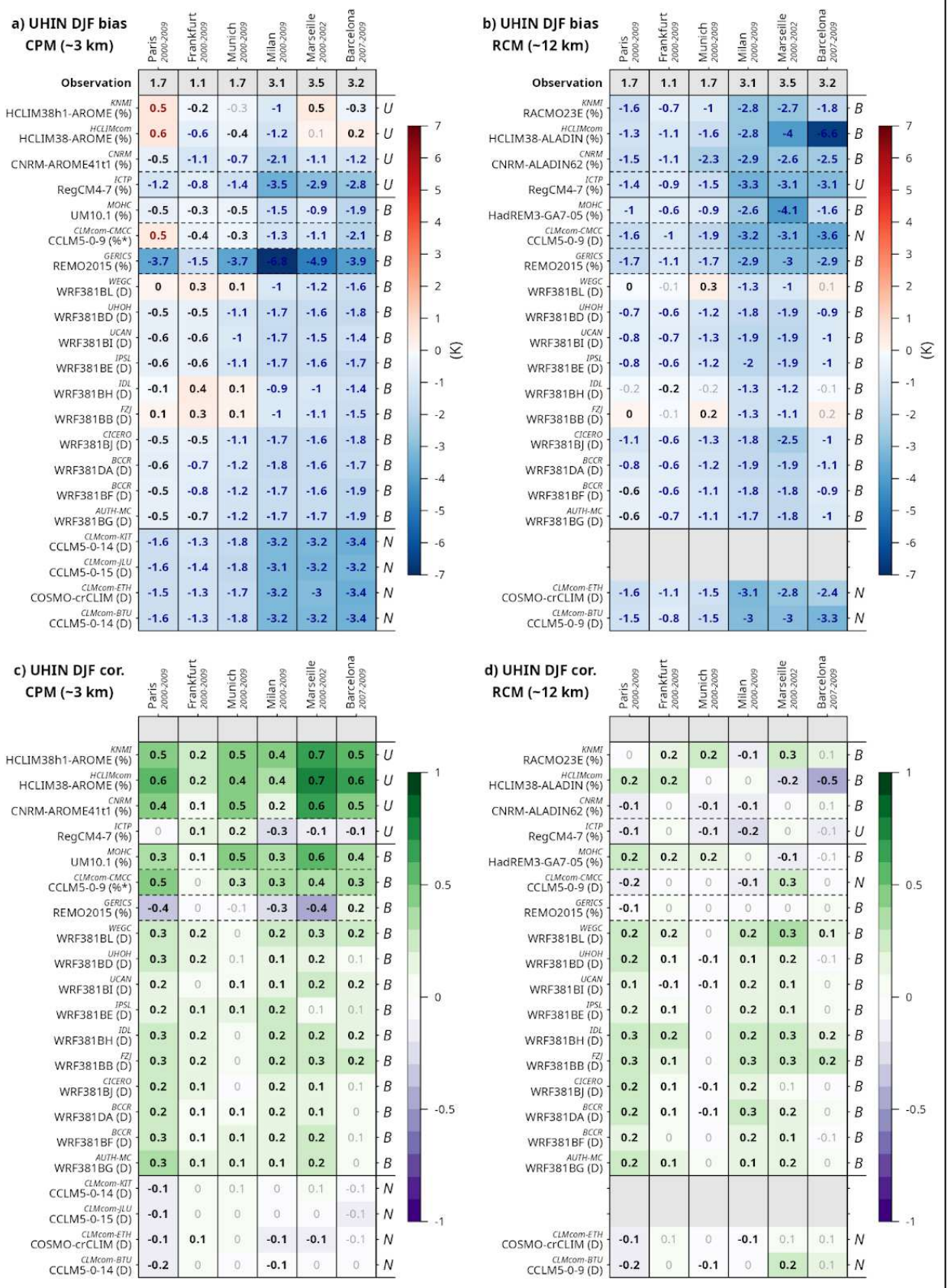


Figure 8. Same as Figure 3 but for UHIN DJF. In addition to average seasonal biases, the Pearson correlation coefficients are also displayed. Correlation coefficients are highlighted (bold) when they are significant at the 5% level.

740

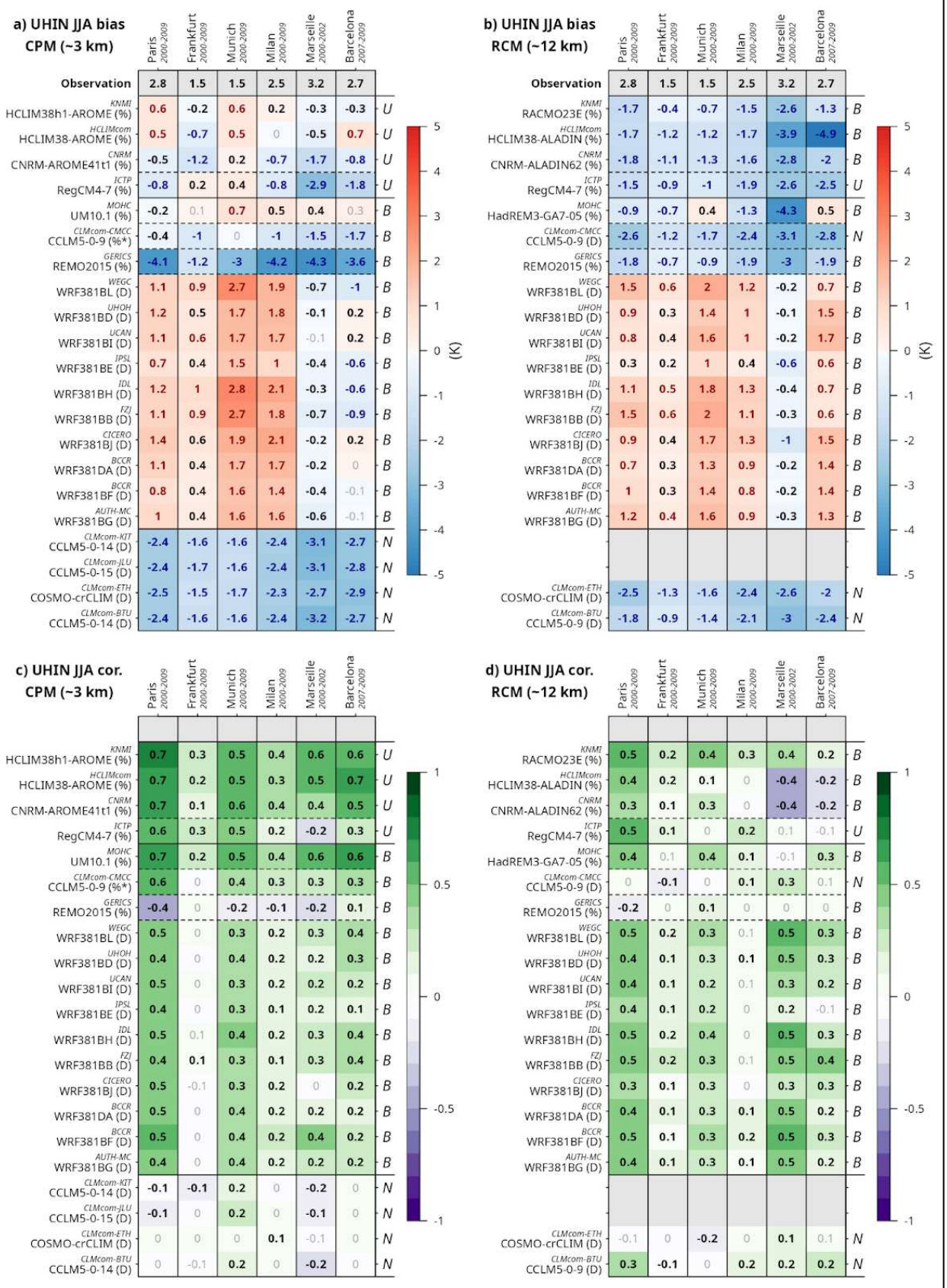


Figure 9. Same as Figure 5. but for UHIN JJA.

As expected, Figure 7 shows that non urbanised models simulate UHI intensities close to zero throughout the year, with the exception of Marseille, and Barcelona (for RCMs) where low intensities are visible. This could be explained by the fact that, even when urban parametrization is not activated in the CCLM simulations, points located at urban stations have a lower vegetation fraction and lower leaf area index values than those at rural points. Another reason could be the different location of the points along the coast, resulting in a different response to the sea. Urbanised models exhibit widely varying UHIN seasonal cycles. For BULK schemes, WRF tends to simulate a similar pattern regardless of the city (although slightly different for Marseille and Barcelona), namely intensities that are smaller (and too low) in DJF and greater (and too high) in JJA (Figure 8 and Figure 9). The shapes of the cycles are similar between WRF's CPMs and RCMs, with differences mainly in magnitude, except for CICERO's RCM simulation for Marseille. The reason why this simulation is atypical remains unclear and cannot be attributed to a difference in the selection of points as it shares the same grid as the other WRF simulations, or to a particular parameterization as listed in Ban et al. 2021. The overestimation of UHIN in JJA could be due to a combination of the particular sensitivity of NOAH's BULK approach and the absence of vegetation in urban points (due to the dominant cover approach), although its effect should be at least partially parameterised (Chen et al. 2011). On the other hand, the excessively low intensities in DJF could be due to the absence of anthropogenic heat fluxes, that are particularly significant during this season due to the heating of buildings. Another striking feature of the WRF ensemble, particularly visible for Munich (Figure 7.i), is that the subgroup of three models using NOAH LSM (FZJ, IDL, WEGC) simulates significantly higher UHI intensities due to a combination of too high TN_{URB} and better simulated TN_{RUR}, whereas other WRF simulations with NOAH-MP with similar urban biases (AUTH-MC and CICERO) also overestimate TN_{RUR}. Salamanca et al (2018) highlighted differences in the way in which the surface exchange coefficients used to compute T2M between the two LSMs (Niu et al. 2011).

At 3 km, the REMO simulation strongly underestimates UHIN in both seasons, which correspond to negative intensities (Figure 8.a and Figure 9.a). A closer look at temperature maps (not shown) seems to indicate that the model has a deficiency over fully urbanized points, with temperatures that are significantly lower than over adjacent points.

CMCC simulation, which uses the TERRA-URB BULK scheme, tends to simulate the same UHIN intensities throughout the year, which leads to small biases for cities where they do not vary much and are low (e.g. in Munich between 1 and 2 °C), but opposite errors when this is not the case (e.g. 0.5 K too warm in DJF and -0.4 K too cold in JJA for

Paris). Unlike WRF's LSMs, TERRA-URB incorporates anthropogenic heat fluxes as sensible heat (Wouters et al. 2015) using an inventory compiled by Flanner (2009) based on a combined annual average value and an idealised diurnal and seasonal cycle that takes latitude into account (Sailor and Lu, 2004). The absence of seasonality could then be explained by a better representation of DJF UHIN (higher intensities), but certain limitations appear during JJA. Campanale et al. (2025) evaluated TERRA-URB's ability to simulate Milan's UHI (in COSMO v6 and ICON v2.6.6 for August 2017) and obtained better results. In addition to the fact that the model was different, they did not use exactly the same stations and applied a lapse rate correction. MOHC's UM model simulates more seasonal variations than other CPMs with BULK schemes and gives a good shape for Paris, but the intensities are mostly underestimated (especially during DJF). Like WRF, but to a lesser extent, it simulates lower intensities in DJF and higher intensities in JJA. Even though anthropogenic heat is implemented in the JULES LSM via MORUSES (Bohnenstengel et al. 2014), the values are based on an inventory compiled for London (England; Davies et al. 2008a; Davies et al. 2008b; Hamilton et al. 2009; Keat et al. 2021) and may therefore not be the most appropriate for more southern cities.

For model configurations implementing UCMs, RegCM simulates UHIN intensities that are too low throughout the year (with the exception of Frankfurt and Munich in JJA). We noted earlier that RegCM has a systematic warm bias in TN_{RUR} and that UHIN intensities are underestimated, as this bias is weaker in TN_{URB}. In addition, RegCM simulates a seasonal cycle similar to that of WRF's BULK (although less pronounced) with lower intensities in DJF and higher intensities in JJA for all cities except Marseille and Barcelona. We don't expect this bias to come from a lack of anthropogenic heat fluxes (as in WRF) because they are represented in RegCM's urban scheme as sensible heat (Oleson et al. 2008a; Oleson et al. 2008b), with heat from buildings and air conditioning waste calculated from the internal temperature of buildings and setpoint temperatures (the internal temperature of buildings itself varying according to a weighted combination of the temperatures of the inner layer wall and roof; Oleson et al. 2010).

The CPMs incorporating the TEB UCM, namely HCLIM-AROME and CNRM-AROME, show the lowest UHIN biases in DJF and JJA, on a par with UM BULK performances (Figures 8 and 9). More generally, outside these seasons, they also better represent the distinct cycles of each city, with the exception of Milan around DJF, and with differences in intensity for certain cycles (e.g. Paris), but less than one degree. Figure 7 shows that the two HCLIM-AROME simulations simulate fairly similar UHIN values that differ in intensity from those of CNRM-AROME (even though the patterns are similar for the three models).

Even with these differences in UHIN, a closer look at TN biases (Figures 5 and 6) shows that TN_{URB} is quite similar for the three models and that the differences come from TN_{RUR}, with CNRM-AROME being warmer (cold bias reduced in DJF, exacerbated in JJA), which could be explained by the fact that HCLIM-AROME uses a more recent version of the ISBA soil-vegetation model. A final point that could explain CNRM-AROME's poorer performance in Frankfurt compared to other cities is that it is the northernmost city and is located in the model's relaxation zone (only for that model).

In addition to the average biases, Figures 5 and 6 show the Pearson correlation coefficients between each UHIN time series from the model and the observed series for each city. With the exception of Paris during JJA (Figure 7.c), the correlations are often relatively weak (e.g., less than 0.5), suggesting that while the models are capable of representing the average intensities of the UHI, they simulate them at dates different from those observed. The performance of the CPMs appears to be aligned with the mean biases, with the UCMs (and UM's BULK scheme) showing higher correlations than the BULK approaches. For a similar reason, the UHIN correlations of the CPMs are higher than those of their driving RCMs.

This subsection highlighted that all urbanized CPMs are capable of representing certain levels of the UHI, and do so better than their driving RCMs. However, their performance varies throughout the year and from one city to another. It appeared that CPMs using simpler BULK schemes appeared unable to represent the diversity of observed seasonal cycles between cities and, on the contrary, simulated the same seasonality for all cities. CPMs incorporating UCMs seemed better able to capture the differences between cities. These results were linked to the different processes represented in each model, which have varying seasonal influences. For example anthropogenic heat during DJF and the city's effect on solar radiation trapping during JJA.

4.3. Distribution evaluation

In order to assess the models' ability to simulate the full range of UHI intensities beyond their seasonal averages (for example, their most extreme intensities), Figure 10 and Figure 11 show a quantile-quantile comparison of models to observation for DJF and JJA, respectively.

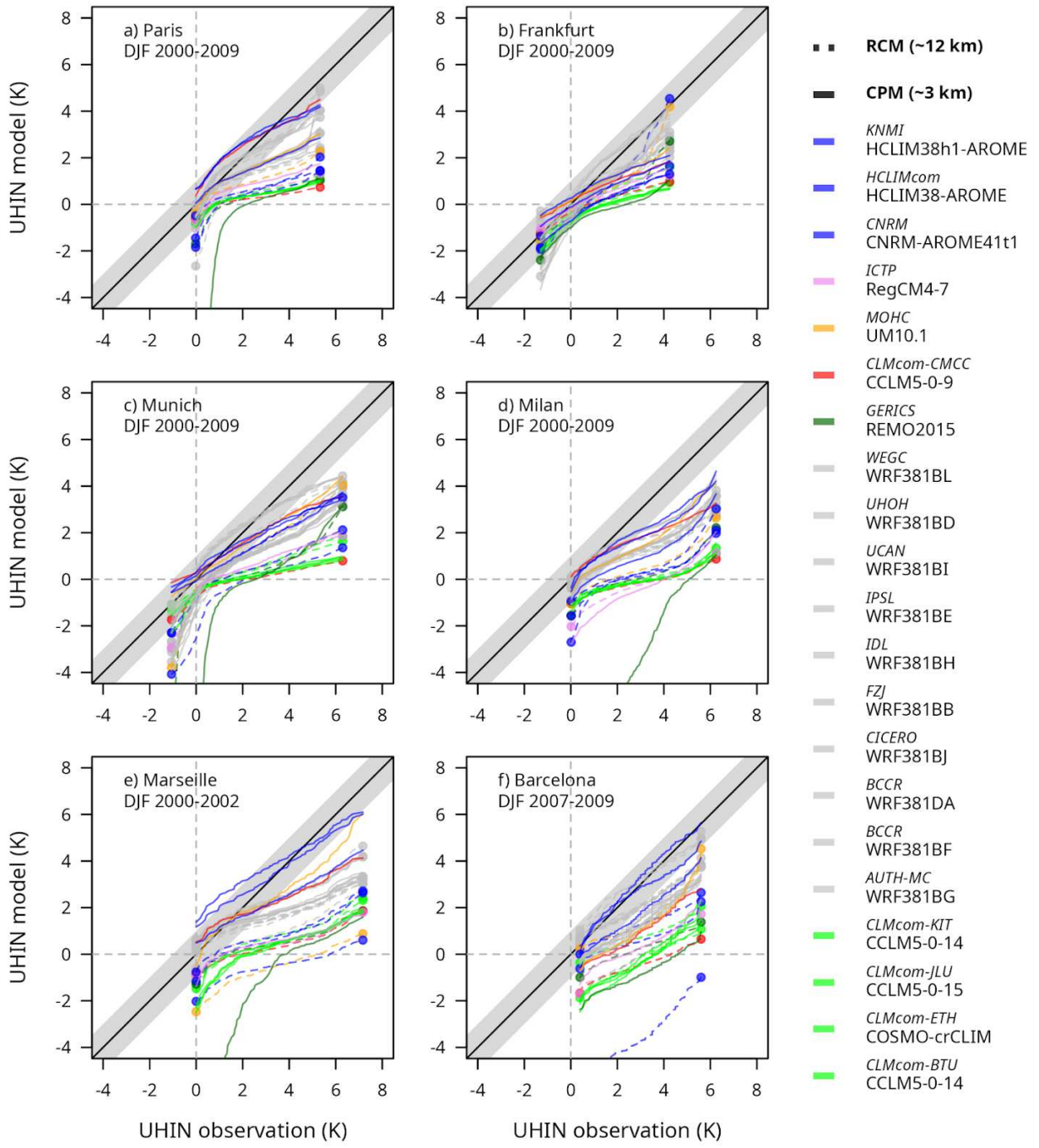


Figure 10. Quantile-quantile plot of observed and simulated UHIN for DJF. The grey shaded area represents a ±1 K difference from observed quantiles. Solid lines represent the CPMs and dashed lines the RCMs. RCMs are also highlighted with dots at the edges of their distribution.

844

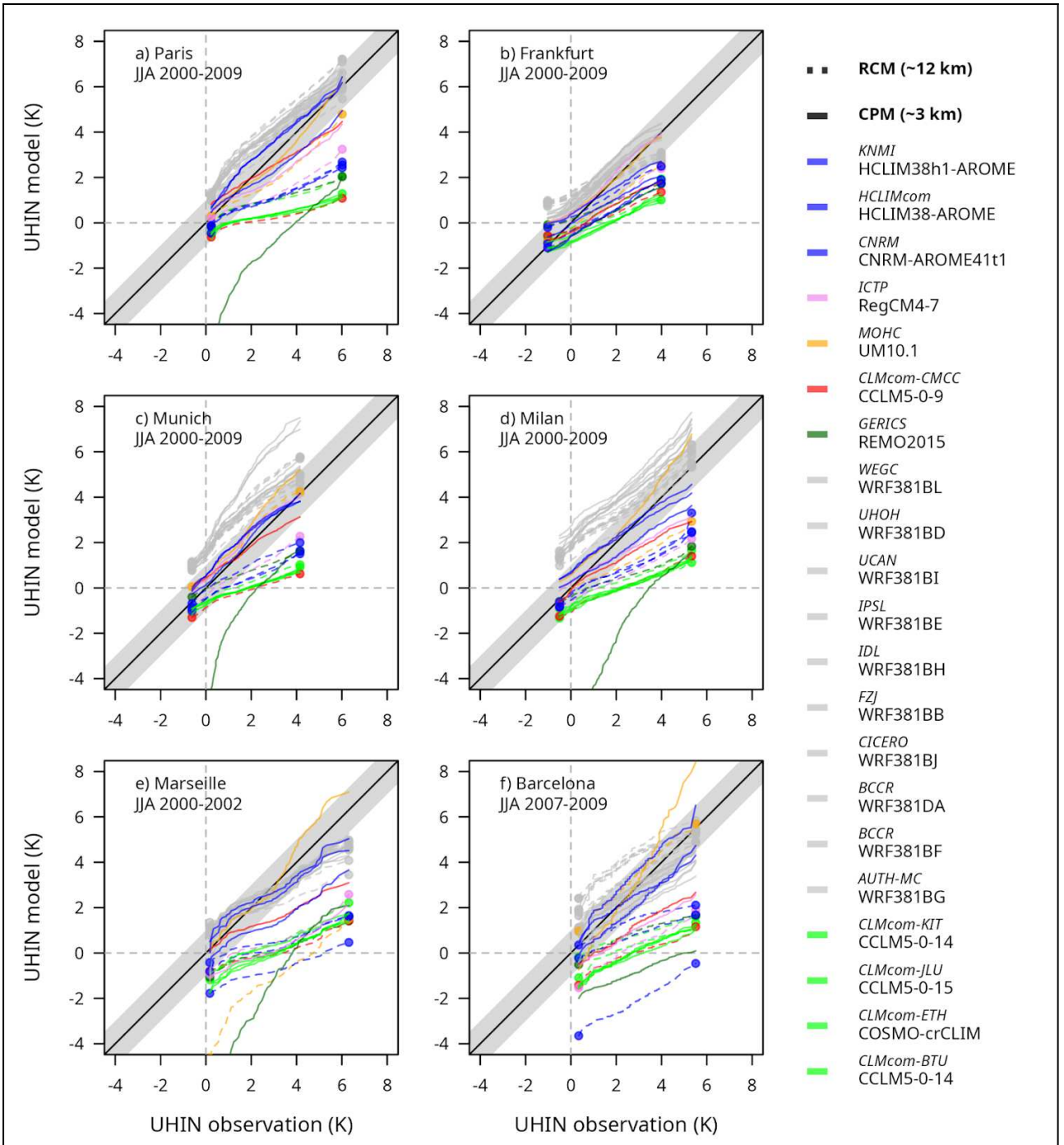


Figure 11. Same as Figure 10 but for UHIN JJA.

The first thing to note for both seasons is a coherence between the average biases identified previously: models that tend to be too warm (cold) on average often have their whole distribution shifted towards warmer (colder) values (see Figures SM9 and SM10 in Supplementary Materials for each percentile bias). This is apparent for almost all CPMs in DJF with their distributions being underestimated, with a few exceptions such as the two HCLIM-AROME simulations and CMCC's CLM for Paris that have low-to-medium UHIN intensities slightly too high. This way, in JJA, the sign of the average bias indicates the direction of the bias of the whole distribution e.g. the WRF UHIN distributions over

continental cities are systematically shifted towards warmer intensities (except for Frankfurt at 12 km). Besides their common shift toward lower (higher) values in DJF (JJA), the shape of the seasonal distributions of WRF members are also different from the observed ones (Figure 7). For almost all combinations of models and cities (with the exception of WRF with NOAH in Frankfurt during JJA) the highest UHIN intensities are underestimated; or have smaller biases than the average for distribution that are systematically shifted towards high values, like WRF. For model-city combinations where UHIN is better simulated (the three AROME simulations over Paris) this also corresponds to an overestimation of the more moderate intensities; models are able to simulate the average UHIN conditions but fail to reproduce the more intense cases. While this seems expected at first due to the fact that even in models with UCMs, the cities are described by average characteristics (building density, height), looking at UHIN values above the 90th percentile (not shown), we see that TN_{URB} is indeed underestimated, but at the same time TN_{RUR} is also overestimated. Thus, the issue is not only due to the urban parametrization or the city's description, but also comes from the model's ability to simulate the proper rural background conditions for these specific days.

RCMs share the same issue of underestimating UHIN intensities on average (Figures 8 and 9) and for their whole distribution (Figures 10 and 11), only more pronounced. Only WRF stands out as an exception in JJA, most likely because it suffers less from the 12 km coarser resolution due to its dominant land cover approach.

Observed IQR and corresponding bin values are listed in Table SM3 and Table SM4 in supplementary materials for TN_{UHI} and TN_{RUR} respectively. Between DJF and JJA, bin width averages across cities vary between 1.2 and 0.9 °C for TN_{RUR}, and 0.5 and 0.4 °C for TN_{UHI}, therefore we chose 1 °C for TN_{RUR} and 0.5 °C for TN_{UHI}.

Figure 12 compares the ΔPS between TN_{RUR} and TN_{UHI} for the six cities in JJA, while Figure 13 gives additional context by showing the original PS of RCMs and CPMs. The corresponding figures for DJF can be found in SM (Figures SM9 and SM10). UHIN is shown instead of TN_{URB} in order to avoid the general biases highlighted previously and to focus on the urban aspect (for example, if a model has a rural distribution that is completely different from the observations, it will have the same problem for its urban distribution). Most of the ensemble is grouped around 0 for ΔPS TN_{RUR} (with the exception of Barcelona) highlighting little to no added value for TN_{RUR} in JJA. In some cases the PS of CPMs is even lower than in RCMs, because of an exacerbation of an already existing warm bias as noted previously. Indeed, if we shift the CPM distribution to the RCM's so

that their averages match, then ΔPS can become slightly positive (not shown). However, this is not always the case: for some model-city pairs, the CPM actually degrades the distribution shape. However, these slight variations in ΔPS must be viewed in the context of the already high PS values in the RCMs (Figure 13). Most of the RCMs have PS values greater than 0.7 in TN_{RUR} JJA with the remaining percentage of the distribution not overlapping coming from a shift in TN_{RUR} distribution towards higher values (with medians a few degrees higher for most RCMs but WRF, not shown) and highest TN_{RUR} values overestimated by all RCMs (opposite for Barcelona). RCMs have lower PS values over Marseille and to a lesser extent Barcelona which can be explained by our selection of points that are fully inland and therefore represent less the potential sea influences that the observations might capture. Nevertheless not doing this selection would give even worse RCM results because of TN being greater over the sea in RCMs.

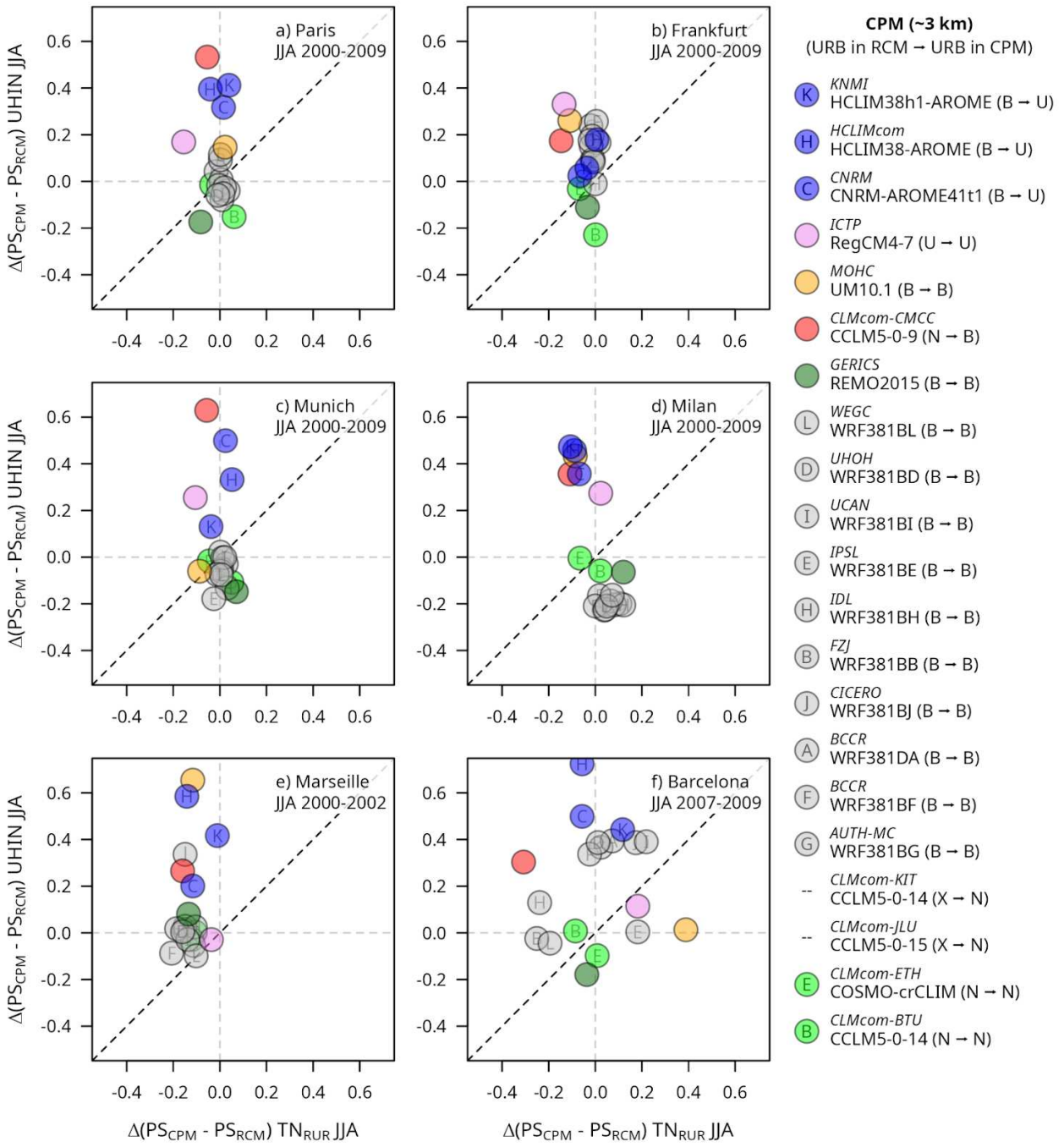


Figure 12. Average JJA added value for the distributions of TN_{RUR} and UHIN between a CPM and an RCM. The added value is defined as the difference in Perkins score between the CPM and the RCM. The B, U, N, X letters in the caption stand for Bulk, UCM, No Urban scheme, and No data respectively.

Over urban areas, the results are considerably different. The ΔPS UHIN are in fact no longer clustered around zero, and a systematic pattern emerges between the model groups, which appears consistent across all cities. Once again, as expected, COSMO simulations (except CMCC) that do not represent urban areas show similar ΔPS for UHIN and TN_{RUR}. The WRF results depend more on the cities, with sometimes very little added value (Paris), sometimes real improvements with better representation of both tails of the

distribution (Frankfurt, Figure 11), and sometimes deterioration due to a recurring and exacerbated warm bias (Milan). The highest ΔPS values are common to the remaining models and may be associated with a change in the level of detail of the urban parameterization between the two resolutions (with the exception of RegCM): CPMs incorporating a UCM driven by RCMs with a BULK (HCLIM-AROME and CNRM-AROME) and CPMs with a BULK driven by an RCM without any urban representation (CMCC's CLM). When examining the PS values directly (Figure 13), the CPMs with UCMs appear to have the highest scores (with the CMCC's CLM slightly lower), with the exception of UM.

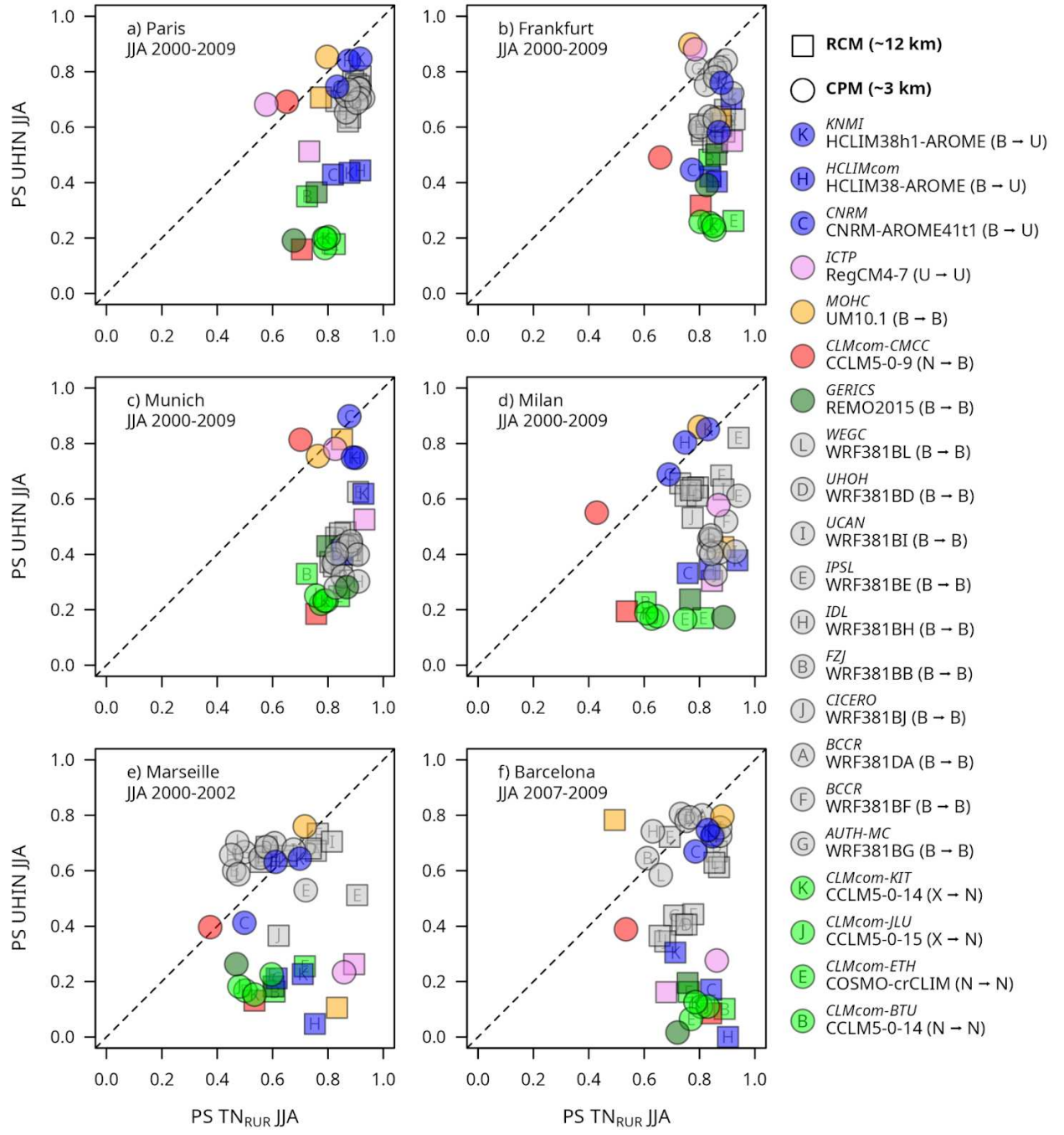


Figure 13. Average Perkins Score (PS) in JJA for TNRUR and UHIN. RCMs are represented by squares and CPMs by circles.

5. Discussion and conclusion

5.1. Added value

In light of the comparison carried out in this study, there is clear added value in using CPMs to represent urban temperatures, provided by the models that are combined with adequate urban parameterization (Table 3). CPMs show less added value in rural areas because (1) 12 km models already give good results in these areas, (2) in some cases, CPMs exacerbate existing warm biases (Sangelantoni et al. 2024), and (3) the analysis conducted here focused on a comparison between grid cells and stations and therefore did not fully reflect the effect of the increased resolution (see discussion below).

Table 3. Summary of model performance stratified by horizontal resolution and type of urban parametrization.

Resolution	Urban parametrization	UHI representation	Added value (vs RCM)	Limitations
12 km (RCM)	No urban	Not represented	-	-
12 km (RCM)	BULK scheme	Weak / underestimated	Reference (baseline)	No explicit urban processes; smoothed urban signal
3 km (CPM)	BULK scheme	Slight improvement but still weak	Limited	Resolution alone insufficient; missing urban physics
3 km (CPM)	BULK scheme + key processes	Better UHI seasonality	Good	Depends on included processes (e.g. anthropogenic heat, solar radiation trapping)
3 km (CPM)	Urban Canopy Models (UCM)	Best representation of seasonality and variability	Strongest (model and city dependent)	Still underestimate extreme UHI intensities

Initially, CPMs coupled with UCMs seemed to consistently yield better results, but closer examination shows that the original choice to distinguish between UCMs and BULK schemes based on the number of surface energy balances they simulate (Lipson et al. 2024) may not be the most appropriate for our ensemble. Rather, it seems that the most important thing is to focus on the model representation of the main physical processes that determine the UHI during the different seasons, for example anthropogenic heat in DJF and the effect of the city on solar radiation trapping in JJA (Arnfield, 2003). Indeed, models that simulate these processes tend to give better results in most cities. This is particularly

evident for the UM simulation, which was initially classified as a BULK scheme because it only solves two surface energy balances (although one of them covers the entire canyon), but which gives excellent results in JJA, comparable to those of CPMs coupled with UCMs, most likely because it represents the effect of the city on solar radiation via the MORUSES scheme.

However, biases persist in CPMs, both in urban and rural areas, and the UHI of smaller cities is not always fully represented. In addition, climate models rely on good urban cover datasets that accurately describe surface properties, which may be outdated and, even when up to date, must rely on average assumptions, i.e. even the best land cover maps may only include a limited number of urban classes with average characteristics suitable for the whole world (or at least a continent). Furthermore, these classes focus on describing the geometric aspect of cities, but rarely provide detailed information on the physical characteristics of the materials used, as these often depend on the city, and finally, urban station observations are sparse. Two main shortcomings appear to be common to most model-city pairs. Firstly, neither RCMs nor CPMs are able to reproduce the maximum UHI intensities, due to urban temperatures being too low and due to an overestimation of rural temperatures; as these models appear unable to simulate the appropriate surface conditions during these specific events. Similarly, even models with good average results are not always able to capture the daily variability of UHI, suggesting that they do not respond perfectly to background climatic conditions, which calls into question their ability to represent key relationships between UHI and relevant meteorological variables (e.g., wind speed, humidity, solar radiation, etc.).

These remaining issues should not obscure all the improvements made by CPMs in representing the UHI compared to their driving RCMs. It should also be noted that this selection of models was an ensemble of opportunity that was not originally designed for urban areas. The representation of cities is now a clear focus of research in the regional climate modelling community, particularly within CORDEX FPS URB-RCC (Langendijk et al. 2024). Climate models often lag behind in terms of development compared to the state-of-the-art parameterization available in research or weather forecasting models, and the latest versions of CPMs increasingly incorporate more recent developments in land surface and urban area parameterization.

5.2. Limitations of our analysis

Certain limitations have been highlighted throughout this article. A first point that has already been mentioned several times is the decision to focus the evaluation on a point based analysis that does not necessarily perfectly represent the climate in the areas surrounding cities. We have done our best, given the available observations, to cover as many areas around cities as possible, particularly in more complex terrains such as coastlines and mountains (e.g. in Barcelona). The main result of this choice is probably an underestimation of the added value provided by the increased resolution and by the difference between the dominant and the fractional land cover approaches. Figure 14 illustrates this issue by comparing the spatial UHI patterns of models with a dominant cover approach and a fractional approach, at two resolutions (the way in which the UHI maps were computed is detailed in Supplementary Materials).

Furthermore, this type of assessment did not take advantage of the spatial aspect of climate models, for example to calculate spatial UHI indicators (as done by Lemonsu et al. 2023). Ongoing and future work could use urban and rural masks to compare models with each other (Diez-Sierra et al. 2025; Langendijk et al. 2025), but the results are likely to differ from observations. Nevertheless, this should further highlight the added value of CPMs, based solely on their increased resolution.

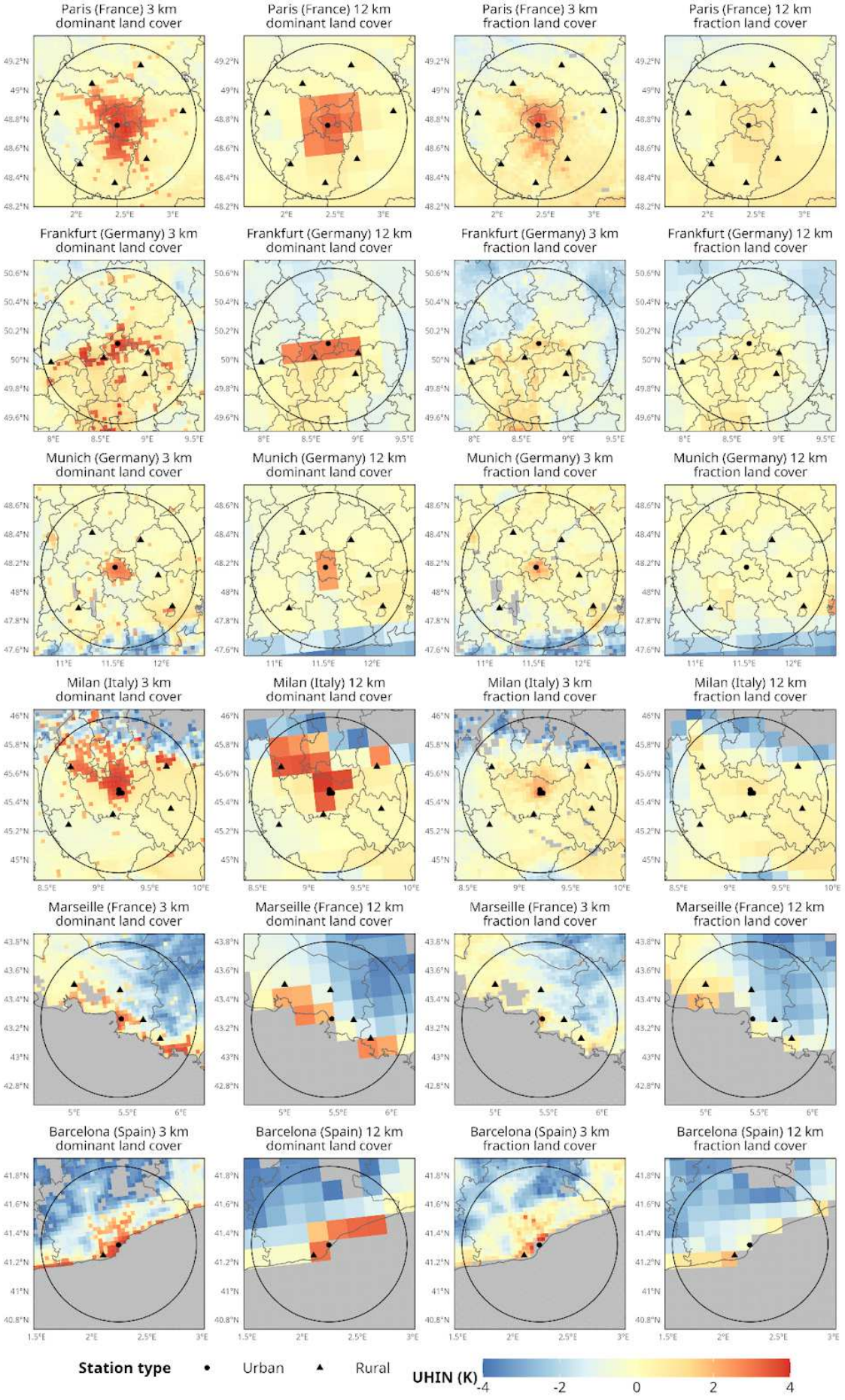

Paris (France) 3 km dominant land cover
Paris (France) 12 km dominant land cover
Paris (France) 3 km fraction land cover
Paris (France) 12 km fraction land cover
Frankfurt (Germany) 3 km dominant land cover
Frankfurt (Germany) 12 km dominant land cover
Frankfurt (Germany) 3 km fraction land cover
Frankfurt (Germany) 12 km fraction land cover
Munich (Germany) 3 km dominant land cover
Munich (Germany) 12 km dominant land cover
Munich (Germany) 3 km fraction land cover
Munich (Germany) 12 km fraction land cover
Milan (Italy) 3 km dominant land cover
Milan (Italy) 12 km dominant land cover
Milan (Italy) 3 km fraction land cover
Milan (Italy) 12 km fraction land cover
Marseille (France) 3 km dominant land cover
Marseille (France) 12 km dominant land cover
Marseille (France) 3 km fraction land cover
Marseille (France) 12 km fraction land cover
Barcelona (Spain) 3 km dominant land cover
Barcelona (Spain) 12 km dominant land cover
Barcelona (Spain) 3 km fraction land cover
Barcelona (Spain) 12 km fraction land cover
Station type
Urban
Rural
UHIN (K)
-4
-2
0
2
4

Figure 14. Representation of the combined impacts of the horizontal resolution and the approach to model gridpoint heterogeneities (fractional vs dominant cover) on average JJA UHIN. Each row represents a city. The first two columns represent one WRF simulation with a dominant cover approach at 3 km and 12 km, respectively. The two last columns represent HCLIM-AROME and HCLIM-ALADIN with a fractional approach. Model points with an elevation greater than 1000 m, or more than 5% of inland water or sea are not displayed. The black circle marks a 60km radius around the city centers.

5.3. Ongoing and future work

Some of the shortcomings highlighted in this study are already being addressed for the next generation of CORDEX CPM simulations carried out as part of the FPS URB-RCC (Langendijk et al. 2024), this time with high-resolution climate simulations dedicated to studies of urban areas. At the CPM resolution, REMO now uses a similar version of the TEB UCM (based on Hamdi and Masson, 2008) that was used by CNRM-AROME here. CNRM-AROME has switched to its next development cycle (Quenum et al. submitted), moving from SURFEX v7 to v8, and in particular to a more complex version of the UCM TEB that includes the representation of vegetation inside the urban canyon (Lemonsu et al. 2012) and a building energy module (Schoetter et al. 2017). The WRF modelling community is currently comparing all available urban parameterizations: BULK, SLUCM (Kusaka et al. 2001; Kusaka and Kimura, 2004) and MLUCM (Martilli, 2002; Chen et al. 2011), with a more careful attention on the values of urban parameters to be better suited for the studied city.

In addition, cities outside Europe will also be studied, enabling analysis of future urban conditions in different climates, where high-resolution climate information is of paramount importance. The work presented here can serve as a basis for future analyses. While this article focuses solely on temperature and, more specifically, on the UHI effect, future studies will analyse relationships between different meteorological variables and the UHI effect, as well as their representation in climate models.

Acknowledgement

The authors gratefully acknowledge the WCRP-CORDEX-FPS on Convective phenomena at high resolution over Europe and the Mediterranean (FPSCONV-ALP-3), as well as the contributing climate modeling centers for the EURO-CORDEX RCM simulations and for making them freely available. They also thank the research data exchange infrastructure and services provided by Jülich Supercomputing Centre, Germany, as part of Helmholtz Data Federation initiative.

The authors also gratefully acknowledge the different provider of the meteorological station data used: Météo-France from Paris and Marseille (France), the Deutscher Wetterdienst for Frankfurt and Munich (Germany), ARPA Lombardia for Milan (Italy) and the Servei Meteorològic de Catalunya for Barcelona (Spain).

B. L. R. works was supported as part of the CIRCE project which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 101067769.

C.C., J. F., O.M.G, and S.S., and N. Z. gratefully acknowledge the support of Horizon Europe project Impetus4Change (I4C, grant id. 101081555)

R. M. C. and P.M.M.S. acknowledge FCT, I.P./MCTES for 2021.01280.CEECIND/CP1650/CT0006 (https://doi.org/10.54499/2021.01280.CEECIND/CP1650/CT0006) and its support through national funds (PIDDAC): LA/P/0068/2020 (https://doi.org/10.54499/LA/P/0068/2020) , UID/50019/2025 (https://doi.org/10.54499/UID/50019/2025), UID/PRR/50019/2025 (https://doi.org/10.54499/UID/PRR/50019/2025), UID/PRR2/50019/2025

N. B. acknowledges the Partnership for advanced computing in Europe for awarded access to Piz Daint at ETH Zürich/Swiss National Supercomputing Centre (CSCS, Switzerland) on which ETHZ simulations have been performed.

IPSL's simulations were granted access to the HPC resources of TGCC under the allocations 2018-A0030106877, 2019-A0030106877 and 2020-A0030106877 made by GENCI.

IPSL also acknowledges funding from the HORIZON 2020 EUCP (European Climate Prediction System) project (https:// www. eucp- project. eu, grant agreement No. 776613) and the ESPRI-IPSL mesocenter for data storage and computing services. IPSL also gratefully acknowledges the support of the LEFE program from INSU/CNRS.

CNRM acknowledges funding from the HORIZON 2020 EUCP (European Climate Prediction System) project (https:// www. eucp- project. eu, grant agreement No. 776613)

Simulations by FZJ acknowledge computing time on JURECA supercomputer at Jülich Supercomputing Centre of the Forschungszentrum Jülich, Grant cjjsc39.

CICERO acknowledges the support from Norwegian National E-Infrastructure Services (Sigma2) in providing HPC resources under the project NN9188K.

E. K. and S. K. acknowledge that AUTH simulations and post-processing were supported by computational time granted by the National Infrastructures for Research and Technology S.A. (GRNET S.A.) at the National HPC facility—ARIS—under project pr009020_thin, as well as by the AUTH Compute Infrastructure and Resources. Technical support was provided by the AUTH Scientific Computing Office.

J. M. gratefully acknowledges the support from the Spanish MITECO and the European Commission NextGeneration EU (Regulation EU 2020/2094), through CSIC's PTI-Clima. The

UCAN simulations were carried out on the Meteo nodes of the supercomputing centre at the Instituto de Física de Cantabria (IFCA).

K. W.-S. acknowledges the support of the German Science Foundation (DFG) through project FOR 1695 for the UHOH simulations, which were carried out using the computational resources received from the supercomputing center HLRS in Stuttgart, Germany.

N. Z. gratefully funding from Ministero dell'Università e della Ricerca under the call Fondo Italiano per la Scienza 2022-2023 (FIS-2) for the project "Mediterranean Extreme Events and Tipping elements in a changing climate on multiple spatiotemporal scales", grant number FIS-2023-00159, CUP: D53C24005450001

We thank Dr. Peter Hoffmann for reading our paper draft and sharing his valuable comments.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest. The authors have not disclosed any competing interests.

References

Adinolfi, M., Raffa, M., Reder, A., & Mercogliano, P. (2020). Evaluation and expected changes of summer precipitation at convection permitting scale with COSMO-CLM over alpine space. Atmosphere, 12(1), 54.

Adinolfi, M., Raffa, M., Reder, A., & Mercogliano, P. (2023). Investigation on potential and limitations of ERA5 Reanalysis downscaled on Italy by a convection-permitting model. Climate dynamics, 61(9), 4319-4342.

Arnfield, A. J. (2003). Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. International Journal of Climatology: a Journal of the Royal Meteorological Society, 23(1), 1-26.

Asmus, C., Hoffmann, P., Pietikäinen, J. P., Böhner, J., & Rechid, D. (2023). Modeling and evaluating the effects of irrigation on land–atmosphere interaction in southwestern Europe with the regional climate model REMO2020–iMOVE using a newly developed parameterization. Geoscientific Model Development, 16(24), 7311-7337.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. Journal of hydrometeorology, 10(3), 623-643.

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., ... & Zander, M. J. (2021). The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation. Climate Dynamics, 57, 275-302.

Bartholomé, E., & Belward, A. S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing, 26(9), 1959-1977.

Basu, R., & Samet, J. M. (2002). Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. Epidemiologic reviews, 24(2), 190-202.

Bathiany, S., Belleflamme, A., Zohbi, J. E., Ney, P., Goergen, K., & Rechid, D. (2023). Increasing interannual climate variability during crop flowering in Europe. Environmental Research Letters, 18(4), 044037.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific data, 5(1), 1-12.

Belušić Vozila, A., Belušić, D., Telišman Prtenjak, M., Güttler, I., Bastin, S., Brisson, E., ... & Warrach-Sagi, K. (2024). Evaluation of the near-surface wind field over the Adriatic region: local wind characteristics in the convection-permitting model ensemble. Climate Dynamics, 62(6), 4617-4634.

Belušić, D., de Vries, H., Dobler, A., Landgren, O., Lind, P., Lindstedt, D., ... & Wu, M. (2020). HCLIM38: a flexible regional climate model applicable for different climate zones from coarse to convection-permitting scales. Geoscientific Model Development, 13(3), 1311-1333.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., ... & Køltzow, M. Ø.(2017). The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system. Monthly Weather Review, 145(5), 1919-1935.

Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C., & Fosser, G. (2020). Pan-European climate at convection-permitting scale: a model intercomparison study. Climate Dynamics, 55, 35-59.

Best, M. J. (2005). Representing urban areas within operational numerical weather prediction models. Boundary-Layer Meteorology, 114, 91-109.

Best, M. J., Grimmond, C. S. B., & Villani, M. G. (2006). Evaluation of the urban tile in MOSES using surface energy balance observations. Boundary-Layer Meteorology, 118(3), 503-525.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R., Ménard, C. B., ... & Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. Geoscientific Model Development, 4(3), 677-699.

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., ... & Yuan, H. (2021). Advances in land surface modelling. Current Climate Change Reports, 7(2), 45-71.

Bohnenstengel, S. I., Hamilton, I., Davies, M., & Belcher, S. E. (2014). Impact of anthropogenic heat emissions on London's temperatures. Quarterly Journal of the Royal Meteorological Society, 140(679), 687-698.

Bonan, G. B. (1996). Land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and users guide. Technical note (No. PB--97-131494/XAB; NCAR/TN--417-STR). National Center for Atmospheric Research, Boulder, CO (United States). Climate and Global Dynamics Div..

Breil, M., Rechid, D., Davin, E. L., de Noblet-Ducoudré, N., Katragkou, E., Cardoso, R. M., ... & Warrach-Sagi, K. (2020). The opposing effects of reforestation and afforestation on the diurnal temperature cycle at the surface and in the lowest atmospheric model level in the European summer. Journal of Climate, 33(21), 9159-9179.

Brousseau, P., Seity, Y., Ricard, D., & Léger, J. (2016). Improvement of the forecast of convective activity from the AROME-France system. Quarterly Journal of the Royal Meteorological Society, 142(699), 2231-2243.

Brutsaert, W. (1982), Evaporation into the Atmosphere: Theory, History and Applications, Environmental Fluid Mechanics, Springer

Bucchignani, E., Montesarchio, M., Zollo, A. L., & Mercogliano, P. (2016a). High-resolution climate simulations with COSMO-CLM over Italy: performance evaluation and climate projections for the 21st century. International Journal of Climatology, 36(2).

Bucchignani, E., Mercogliano, P., Rianna, G., & Panitz, H. J. (2016b). Analysis of ERA-Interim-driven COSMO-CLM simulations over Middle East–North Africa domain at different spatial resolutions. International Journal of Climatology, 36(9), 3346-3369.

Ca, V. T., Asaeda, T., & Ashie, Y. (1999). Development of a numerical model for the evaluation of the urban thermal environment. Journal of Wind Engineering and Industrial Aerodynamics, 81(1-3), 181-196.

Cacciamani, C., Battaglia, F., Patruno, P., Pomi, L., Selvini, A., & Tibaldi, S. (1995). A climatological study of thunderstorm activity in the Po Valley. Theoretical and applied climatology, 50(3), 185-203.

Caillaud, C., Somot, S., Alias, A., Bernard-Bouissières, I., Fumière, Q., Laurantin, O., ... & Ducrocq, V. (2021). Modelling Mediterranean heavy precipitation events at climate scale: an object-oriented evaluation of the CNRM-AROME convection-permitting regional climate model. Climate Dynamics, 56, 1717-1752.

Campanale, A., Adinolfi, M., Raffa, M., Schulz, J. P., & Mercogliano, P. (2025). Investigating urban heat islands over Rome and Milan during a summer period through the TERRA_URB parameterization in the ICON model. Urban Climate, 60, 102335.

Cardoso, R. M., & Soares, P. M. (2022). Is there added value in the EURO-CORDEX hindcast temperature simulations? Assessing the added value using climate distributions in Europe. Int. J. Climatol, 42, 4024-4039.

Champeaux, J. L., Masson, V., & Chauvin, F. (2005). ECOCLIMAP: a global database of land surface parameters at 1 km resolution. Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 12(1), 29-32.

Chen, D., M. Rojas, B.H. Samset, K. Cobb, A. Diongue Niang, P. Edwards, S. Emori, S.H. Faria, E. Hawkins, P. Hope, P. Huybrechts, M. Meinshausen, S.K. Mustafa, G.-K. Plattner, and A.-M. Tréguier, 2021: Framing, Context, and Methods. InClimate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change[Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 147–286, doi:10.1017/9781009157896.003.

Chen, F., & Dudhia, J. (2001a). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly weather review, 129(4), 569-585.

Chen, F., & Dudhia, J. (2001b). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part II: Preliminary model validation. Monthly Weather Review, 129(4), 587-604.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., ... & Zhang, C. (2011). The integrated WRF/urban modeling system: development, evaluation, and applications to urban environmental problems.

Cheng, J., Xu, Z., Zhu, R., Wang, X., Jin, L., Song, J., & Su, H. (2014). Impact of diurnal temperature range on human health: a systematic review. International journal of biometeorology, 58(9), 2011-2024.

Ciarlo, J. M., Coppola, E., Fantini, A., Giorgi, F., Gao, X., Tong, Y., ... & Jacob, D. (2021). A new spatially distributed added value index for regional climate models: the EURO-CORDEX and the CORDEX-CORE highest resolution ensembles. Climate Dynamics, 57(5), 1403-1424.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., ... & Cox, P. M. (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 2: carbon fluxes and vegetation dynamics. Geoscientific Model Development, 4(3), 701-722.

Clark, P., Roberts, N., Lean, H., Ballard, S. P., & Charlton-Perez, C. (2016). Convection-permitting models: a step-change in rainfall forecasting. Meteorological Applications, 23(2), 165-181.

Colin, J., Déqué, M., Radu, R., & Somot, S. (2010). Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique. Tellus A: Dynamic Meteorology and Oceanography, 62(5), 591-604.

Coppola, E., Sobolowski, S., Pichelli, E., Raffaele, F., Ahrens, B., Anders, I., ... & Warrach-Sagi, K. (2020). A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. Climate Dynamics, 55, 3-34.

Coppola, E., Stocchi, P., Pichelli, E., Torres Alavez, J. A., Glazer, R., Giuliani, G., ... & Giorgi, F. (2021). Non-Hydrostatic RegCM4 (RegCM4-NH): model description and case studies over multiple domains. Geoscientific Model Development, 14(12), 7705-7723.

Cornes, R. C., Van Der Schrier, G., Van Den Besselaar, E. J., & Jones, P. D. (2018). An ensemble version of the E-OBS temperature and precipitation data sets. Journal of Geophysical Research: Atmospheres, 123(17), 9391-9409.

Cox, P. M., Betts, R. A., Bunton, C. B., Essery, R. L. H., Rowntree, P. R., & Smith, J. (1999). The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. Climate Dynamics, 15, 183-203.

Daniel, M., Lemonsu, A., Déqué, M., Somot, S., Alias, A., & Masson, V. (2019). Benefits of explicit urban parameterization in regional climate modeling to study climate and city interactions. Climate Dynamics, 52(5), 2745-2764.

Davies, M., Hamilton, I., Steadman, P., Stone, A., Ridley, I., & Evans, S. (2008). London's anthropogenic heat emissions–implications for building design'. In World Renewable Energy Congress X.

Davies, M., Steadman, P., & Oreszczyn, T. (2008). Strategies for the modification of the urban climate and the consequent impact on building energy use. Energy Policy, 36(12), 4548-4551.

De Vries, H., Lenderink, G., Van Meijgaard, E., Van Ulft, B., & De Rooy, W. (2024). Western Europe's extreme July 2019 heatwave in a warmer world. Environmental Research: Climate, 3(3), 035005.

Dee, D. P., Uppala, S., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., ... & Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the royal meteorological society, 137(656), 553-597.

Demuzere, M., K. De Ridder, and N. P. M. van Lipzig (2008), Modeling the energy balance in Marseille: Sensitivity to roughness length parametrizations and thermal admittance, J. Geophys. Res., 113, 1–19, doi:10.1029/2007JD009113.

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., ... & Bechtel, B. (2022). A global map of Local Climate Zones to support earth system modelling and urban scale environmental science. Earth System Science Data Discussions, 2022, 1-57.

Déqué, M., Alias, A., Somot, S., & Nuissier, O. (2016). Climate change and extreme precipitation: the response by a convection-resolving model. Research activities in atmospheric and oceanic modelling CAS/JSC Working group on numerical experimentation Report, (46).

Diez-Sierra, J., Quintana, Y., Langendijk, G. S., Milovac, J., Demuzere, M., Nogherotto, R., ... & Fernández, J. (2025). A global CORDEX-based dataset delineating urban areas and their surroundings to assess climate change in megacities. Scientific Data, 12(1), 1961.

Doms, G., Förstner, J., Heise, E., Herzog, H.-J., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schulz J.-P., Vogel, G. (2011). A Description of the Nonhydrostatic Regional COSMO Model Part II : Physical Parameterization. Consortium for Small-Scale Modelling (COSMO).

Dosio, A. (2016). Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models. Journal of Geophysical Research: Atmospheres, 121(10), 5488-5511.

Dosio, A. (2017). Projection of temperature and heat waves for Africa with an ensemble of CORDEX Regional Climate Models. Climate Dynamics, 49(1), 493-519.

Ebi, K. L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., ... & Jay, O. (2021). Hot weather and heat extremes: health risks. The lancet, 398(10301), 698-708.

ECMWF-IFS: Part IV : PHYSICAL PROCESSES (CY33R1), Technical Report, 2008.

EEA (2018) Copernicus Land Monitoring Service – High Resolution Layer Imperviousness Product Specifications. Imperviousness Density 2006. European Union's Copernicus Land Monitoring Service information, https://doi.org/10.2909/17b26358-df10-4def-9991-fb395ce64933.

European Environmental Agency (2013). CORINE Land Cover (CLC) 2006, CRC/TR32 Database (TR32DB). Version 17. Copenhagen, Denmark: Copernicus Land Monitoring Service. Retrieved from https://www.eea.europa.eu/data-and-maps/data/clc-2006-raster-3

Eidenshink, J. C., & Faundeen, J. L. (1994). The 1 km AVHRR global land data set: first stages in implementation. International Journal of Remote Sensing, 15(17), 3443-3462.

Essery, R. L. H., Best, M. J., Betts, R. A., Cox, P. M., & Taylor, C. M. (2003). Explicit representation of subgrid heterogeneity in a GCM land surface scheme. Journal of Hydrometeorology, 4(3), 530-543.

Eunice Lo, Y. T., Mitchell, D. M., Bohnenstengel, S. I., Collins, M., Hawkins, E., Hegerl, G. C., ... & Stott, P. A. (2020). UK climate projections: Summer daytime and nighttime urban heat island changes in England's major cities. Journal of Climate, 33(20), 9015-9030.

Faroux, S., Kaptué Tchuenté, A. T., Roujean, J. L., Masson, V., Martin, E., & Le Moigne, P. (2013). ECOCLIMAP-II/Europe: A twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. Geoscientific Model Development, 6(2), 563-582.

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. Journal of Advances in Modeling Earth Systems, 12, e2018MS001453. https://doi.org/10.1029/2018MS001453

Flanner, M. G. (2009). Integrating anthropogenic heat flux with global climate models. Geophysical Research Letters, 36(2).

Fosser, G., Gaetani, M., Kendon, E. J., Adinolfi, M., Ban, N., Belušić, D., ... & Tölle, M. H. (2024). Convection-permitting climate models offer more certain extreme rainfall projections. NPJ Climate and atmospheric science, 7(1), 51.

Freedman, D., & Diaconis, P. (1981). On the histogram as a density estimator: L 2 theory. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 57(4), 453-476.

Fumière, Q., Déqué, M., Nuissier, O., Somot, S., Alias, A., Caillaud, C., ... & Seity, Y. (2020). Extreme rainfall in Mediterranean France during the fall: added value of the CNRM-AROME Convection-Permitting Regional Climate Model. Climate Dynamics, 55, 77-91.

Garbero, V., Milelli, M., Bucchignani, E., Mercogliano, P., Varentsov, M., Rozinkina, I., ... & Repola, F. (2021). Evaluating the urban canopy scheme TERRA_URB in the COSMO model for selected European cities. Atmosphere, 12(2), 237.

Geiger, Rudolf (1954). "Klassifikation der Klimate nach W. Köppen" [Classification of climates after W. Köppen]. Landolt-Börnstein – Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik, alte Serie. Vol. 3. Berlin: Springer. pp. 603–607.

Giorgi, F. (2006, December). Regional climate modeling: Status and perspectives. In Journal de Physique IV (proceedings) (Vol. 139, pp. 101-118). EDP sciences.

Giorgi, F., & Gutowski Jr, W. J. (2015). Regional dynamical downscaling and the CORDEX initiative. Annual review of environment and resources, 40(1), 467-490.

Giorgi, F., Coppola, E., Solmon, F., Mariotti, L., Sylla, M. B., Bi, X., ... & Brankovic, C. (2012). RegCM4: model description and preliminary tests over multiple CORDEX domains. Climate Research, 52, 7-29.

Giorgi, F. (2019). Thirty years of regional climate modeling: where are we and where are we going next?. *Journal of Geophysical Research: Atmospheres*, *124*(11), 5696-5723.

Grimmond, C. S. B., Blackett, M., Best, M. J., Baik, J. J., Belcher, S. E., Beringer, J., ... & Zhang, N. (2011). Initial results from Phase 2 of the international urban energy balance model comparison. International Journal of Climatology, 31(2), 244-272.

Ha, M. T., Bastin, S., Drobinski, P., Fita, L., Polcher, J., Bock, O., ... & Wulfmeyer, V. (2024). Precipitation frequency in Med-CORDEX and EURO-CORDEX ensembles from 0.44° to convection-permitting resolution: impact of model resolution and convection representation. Climate dynamics, 62(6), 4515-4540.

Hagemann, S. (2002). An improved land surface parameter dataset for global and regional climate models.

Hagemann, S., Botzet, M., Dümenil, L. Y. D. I. A., & Machenhauer, B. (1999). Derivation of global GCM boundary conditions from 1 km land use satellite data.

Hamdi, R., & Masson, V. (2008). Inclusion of a drag approach in the Town Energy Balance (TEB) scheme: Offline 1D evaluation in a street canyon. Journal of Applied Meteorology and Climatology, 47(10), 2627-2644.

Hamilton, I. G., Davies, M., Steadman, P., Stone, A., Ridley, I., & Evans, S. (2009). The significance of the anthropogenic heat emissions of London's buildings: A comparison against captured shortwave solar radiation. Building and Environment, 44(4), 807-817.

Harman, I. N., & Belcher, S. E. (2006). The surface energy balance and boundary layer over urban street canyons. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 132(621), 2749-2768.

Harman, I. N., Best, M. J., & Belcher, S. E. (2004a). Radiative exchange in an urban street canyon. Boundary-Layer Meteorology, 110, 301-316.

Harman, I. N., Barlow, J. F., & Belcher, S. E. (2004b). Scalar fluxes from urban street canyons part II: model. Boundary-Layer Meteorology, 113, 387-410.

Hong, S. Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. Monthly weather review, 134(9), 2318-2341.

Jackson, Trisha & Feddema, Johannes & Oleson, Keith & Bonan, Gordon & Bauer, John. (2010). Parameterization of Urban Characteristics for Global Climate Modeling. Annals of the Association of American Geographers. 100. 848-865. 10.1080/00045608.2010.497328.

Jacob, D., & Podzun, R. (1997). Sensitivity studies with the regional climate model REMO. Meteorology and atmospheric physics, 63, 119-129.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., ... & Yiou, P. (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. Regional environmental change, 14, 563-578.

Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., ... & Wulfmeyer, V. (2020). Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community. Regional environmental change, 20, 1-20.

Kanda, M., Kanega, M., Kawai, T., Moriwaki, R., & Sugawara, H. (2007). Roughness lengths for momentum and heat derived from outdoor urban scale models. Journal of Applied Meteorology and Climatology, 46(7), 1067-1079.

Keat, W. J., Kendon, E. J., & Bohnenstengel, S. I. (2021). Climate change over UK cities: the urban influence on extreme temperatures in the UK climate projections. Climate dynamics, 57(11), 3583-3597.

Keat, W. J., Short, C. J., & Kendon, E. J. (2024). Are convection-permitting climate projections reliable for urban planning over Africa? A case study of Johannesburg. Atmospheric Science Letters, 25(11), e1264.

Keppas, S. C., Papadogiannaki, S., Parliari, D., Kontos, S., Poupkou, A., Tzoumaka, P., ... & Melas, D. (2021). Future climate change impact on urban heat island in two mediterranean cities based on high-resolution regional climate simulations. Atmosphere, 12(7), 884.

Kim, S. W., & Brown, R. D. (2021). Urban heat island (UHI) intensity and magnitude estimations: A systematic literature review. Science of the Total Environment, 779, 146389.

Knist, S., Goergen, K., Buonomo, E., Christensen, O. B., Colette, A., Cardoso, R. M., ... & Simmer, C. (2017). Land-atmosphere coupling in EURO-CORDEX evaluation experiments. Journal of Geophysical Research: Atmospheres, 122(1), 79-103.

Kondo, H., & Liu, F. H. (1998). A study on the urban thermal environment obtained through one-dimensional urban canopy model. Journal of Japan Society for Atmospheric Environment/Taiki Kankyo Gakkaishi, 33(3), 179-192.

Kondo, H., Genchi, Y., Kikegawa, Y., Ohashi, Y., Yoshikado, H., & Komiyama, H. (2005). Development of a multi-layer urban canopy model for the analysis of energy consumption in a big city: Structure of the urban canopy model and its basic performance. Boundary-Layer Meteorology, 116(3), 395-421.

Köppen, W., (1884). Die W¨armezonen der Erde, nach der Dauer der heissen, gem¨assigten und kalten Zeit und nach der Wirkung der W¨arme auf die organische Welt betrachtet (The thermal zones of the earth according to the duration of hot, moderate and cold periods and to the impact of heat on the organic world). – Meteorol. Z. 1, 215–226. (translated and edited by Volken E. and S. Brönnimann. – Meteorol. Z. 20 (2011), 351–360)

Kotlarski, S. (2007). A subgrid glacier parameterisation for use in regional climate modelling (Doctoral dissertation, University of Hamburg Hamburg).

Kusaka, H., & Kimura, F. (2004). Coupling a single-layer urban canopy model with a simple atmospheric model: Impact on urban heat island simulation for an idealized case. Journal of the Meteorological Society of Japan. Ser. II, 82(1), 67-80.

Kusaka, H., Kondo, H., Kikegawa, Y., & Kimura, F. (2001). A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models. Boundary-layer meteorology, 101, 329-358.

Kusaka, H., Suzuki-Parker, A., Aoyagi, T., Adachi, S. A., & Yamagata, Y. (2016). Assessment of RCM and urban scenarios uncertainties in the climate projections for August in the 2050s in Tokyo. Climatic Change, 137(3), 427-438.

Langendijk, G. S., Fernandez, J., Demuzere, M., Diez-Sierra, J., Quintana, Y., Fita, L., ... & Yuan, J. (2025). Representation of global mega-cities and their urban heat island in CORDEX-CORE regional climate model simulations. npj Urban Sustainability.

Langendijk, G. S., Halenka, T., Hoffmann, P., Adinolfi, M., Campino, A. A., Asselin, O., ... & Yuan, J. (2024). Towards better understanding the urban environment and its interactions with regional climate change-The WCRP CORDEX Flagship Pilot Study URB-RCC. Urban Climate, 58, 102165.

Langendijk, G. S., Rechid, D., & Jacob, D. (2019). Urban areas and urban–rural contrasts under climate change: What does the EURO-CORDEX ensemble tell us?—Investigating near surface humidity in Berlin and its surroundings. Atmosphere, 10(12), 730.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., ... & Slater, A. G. (2011). Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. Journal of Advances in Modeling Earth Systems, 3(1).

Le Roy, B., Lemonsu, A., & Schoetter, R. (2021). A statistical–dynamical downscaling methodology for the urban heat island applied to the EURO-CORDEX ensemble. Climate Dynamics, 56(7), 2487-2508.

Lemonsu, A., Caillaud, C., Alias, A., Riette, S., Seity, Y., Le Roy, B., ... & Lucas-Picher, P. (2023). What added value of CNRM-AROME convection-permitting regional climate model compared to CNRM-ALADIN regional climate model for urban climate studies? Evaluation over Paris area (France). Climate Dynamics, 61(3), 1643-1661.

Lemonsu, A., Masson, V., Shashua-Bar, L., Erell, E., & Pearlmutter, D. (2012). Inclusion of vegetation in the Town Energy Balance model for modelling urban green areas. Geoscientific Model Development, 5(6), 1377-1393.

Lenderink, G., de Vries, H., Fowler, H. J., Barbero, R., van Ulft, B., & Van Meijgaard, E. (2021). Scaling and responses of extreme hourly precipitation in three climate experiments with a convection-permitting model. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 379(2195).

Li, C., Zhang, N., Ren, K., Zhao, W., Wu, J., & Sun, Y. (2024). Biophysical drivers of seasonal hysteresis of urban heat islands across climates and urban landscapes. Journal of Geophysical Research: Atmospheres, 129(9), e2023JD040446.

Li, D., Bou-Zeid, E., Barlage, M., Chen, F., & Smith, J. A. (2013). Development and evaluation of a mosaic approach in the WRF-Noah framework. Journal of Geophysical Research: Atmospheres, 118(21), 11-918.

Lipson, M. J., Grimmond, S., Best, M., Abramowitz, G., Coutts, A., Tapper, N., ... & Pitman, A. J. (2024). Evaluation of 30 urban land surface models in the Urban-PLUMBER project: Phase 1 results. Quarterly Journal of the Royal Meteorological Society, 150(758), 126-169.

Liu, J., Varghese, B. M., Hansen, A., Zhang, Y., Driscoll, T., Morgan, G., ... & Bi, P. (2022). Heat exposure and cardiovascular health outcomes: a systematic review and meta-analysis. The Lancet Planetary Health, 6(6), e484-e495.

Liu, Y., Chen, F., Warner, T., & Basara, J. (2006). Verification of a mesoscale data-assimilation and forecasting system for the Oklahoma City area during the Joint Urban 2003 field project. Journal of applied meteorology and climatology, 45(7), 912-929.

Loridan, T., & Grimmond, C. S. B. (2012). Multi-site evaluation of an urban land-surface model: Intra-urban heterogeneity, seasonality and parameter complexity requirements. Quarterly Journal of the Royal Meteorological Society, 138(665), 1094-1113.

Lucas-Picher, P., Argüeso, D., Brisson, E., Tramblay, Y., Berg, P., Lemonsu, A., ... & Caillaud, C. (2021). Convection-permitting modeling with regional climate models: Latest developments and next steps. Wiley Interdisciplinary Reviews: Climate Change, 12(6), e731.

Luu, L. N., van Meijgaard, E., Philip, S. Y., Kew, S. F., de Baar, J. H., & Stepek, A. (2023). Impact of surface roughness changes on surface wind speed over western Europe: A study with the regional climate model RACMO. Journal of Geophysical Research: Atmospheres, 128(12), e2022JD038426.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics, 50-60.

Manoli, G., Fatichi, S., Bou-Zeid, E., & Katul, G. G. (2020). Seasonal hysteresis of surface urban heat islands. Proceedings of the National Academy of Sciences, 117(13), 7082-7089.

Maraun, D., & Widmann, M. (2018). Statistical downscaling and bias correction for climate research. Cambridge University Press.

Marelle, L., Myhre, G., Steensen, B. M., Hodnebrog, Ø., Alterskjær, K., & Sillmann, J. (2020). Urbanization in megacities increases the frequency of extreme precipitation events far more than their intensity. Environmental Research Letters, 15(12), 124072.

Marcotullio, P. J., Keßler, C., & Fekete, B. M. (2022). Global urban exposure projections to extreme heatwaves. Frontiers in Built Environment, 8, 947496.

Martilli, A. (2007). Current research and future challenges in urban mesoscale modelling. International Journal of Climatology: A Journal of the Royal Meteorological Society, 27(14), 1909-1918.

Martilli, A., Clappier, A., & Rotach, M. W. (2002). An urban surface exchange parameterisation for mesoscale models. Boundary-layer meteorology, 104, 261-304.

Masson, V. (2000). A physically-based scheme for the urban energy budget in atmospheric models. Boundary-layer meteorology, 94, 357-397.

Masson, V. (2006). Urban surface modeling and the meso-scale impact of cities. Theoretical and applied climatology, 84, 35-45.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., ... & Voldoire, A. (2013). The SURFEXv7. 2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geoscientific Model Development, 6(4), 929-960.

McCarthy, M. P., Harpham, C., Goodess, C. M., & Jones, P. D. (2012). Simulating climate change in UK cities using a regional climate model, HadRM3. International Journal of Climatology, 32(12).

McGregor, G., Belcher, S., Hacker, J., Kovats, S., Salmond, J., Watkins, R., ... & Wooster, M. (2006). London's Urban Heat Island-A report to the Greater London Authority.

Michau, Y., Lemonsu, A., Lucas-Picher, P., & Caillaud, C. (2023). Evaluation of the Urban Heat Island of 12 cities of France in a high-resolution regional climate model simulation. Urban Climate, 47, 101386.

Michau, Y., Lemonsu, A., Lucas-Picher, P., Bastin, S., Caillaud, C., de Vries, H., ... & Coppola, E. (2025). Projected evolution of the Urban climate and heatwaves using an ensemble of convection-permitting regional climate models. Climatic Change, 178(8), 154.

Michau, Y., Lemonsu, A., Lucas-Picher, P., Schneider, M., & Caillaud, C. (2024). On the future evolution of heatwaves in French cities and associated rural areas: Insights from a convection-permitting model. Urban Climate, 55, 101920.

Mishra, V., Ganguly, A. R., Nijssen, B., & Lettenmaier, D. P. (2015). Changes in observed climate extremes in global urban areas. Environmental Research Letters, 10(2), 024005.

Moemken, J., Reyers, M., Feldmann, H., & Pinto, J. G. (2018). Future changes of wind speed and wind energy potentials in EURO-CORDEX ensemble simulations. Journal of Geophysical Research: Atmospheres, 123(12), 6373-6389.

Molina, M. O., Careto, J. M., Gutiérrez, C., Sánchez, E., Goergen, K., Sobolowski, S., ... & Soares, P. M. M. (2024). The added value of simulated near-surface wind speed over the Alps from a km-scale multimodel ensemble. Climate Dynamics, 1-19.

Molina, M. O., Sánchez, E., & Gutiérrez, C. (2020). Future heat waves over the Mediterranean from an Euro-CORDEX regional climate model ensemble. Scientific reports, 10(1), 8801.

Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. Contrib. Geophys. Inst. Acad. Sci. USSR, 151(163), e187.

Nabat, P., Somot, S., Cassou, C., Mallet, M., Michou, M., Bouniol, D., ... & Saint-Martin, D. (2020). Modulation of radiative aerosols effects by atmospheric circulation over the Euro-Mediterranean region. Atmospheric Chemistry and Physics, 20(14), 8315-8349.

Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. Journal of the Meteorological Society of Japan. Ser. II, 87(5), 895-912.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., ... & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres, 116(D12).

Noël, B., Van De Berg, W. J., Van Meijgaard, E., Kuipers Munneke, P., Van De Wal, R. S. W., & Van Den Broeke, M. R. (2015). Evaluation of the updated regional climate model RACMO2. 3: summer snowfall impact on the Greenland Ice Sheet. The Cryosphere, 9(5), 1831-1844.

Noilhan, J., & Mahfouf, J. F. (1996). The ISBA land surface parameterisation scheme. Global and planetary Change, 13(1-4), 145-159.

Noilhan, J., & Planton, S. (1989). A simple parameterization of land surface processes for meteorological models. Monthly weather review, 117(3), 536-549.

Nunez, M., & Oke, T. R. (1977). The energy balance of an urban canyon. Journal of Applied Meteorology and Climatology, 16(1), 11-19.

Oke, T. R. (1973). City size and the urban heat island. Atmospheric Environment (1967), 7(8), 769-779.

Oke, T. R. (1988). The urban energy balance. Progress in Physical geography, 12(4), 471-508.

Oke, T. R., Mills, G., Christen, A., & Voogt, J. A. (2017). Urban climates. Cambridge university press.

Oleson, K. W., Bonan, G. B., Feddema, J., & Vertenstein, M. (2008b). An urban parameterization for a global climate model. Part II: Sensitivity to input parameters and the simulated urban heat island in offline simulations. Journal of Applied Meteorology and Climatology, 47(4), 1061-1076.

Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M., & Grimmond, C. S. B. (2008a). An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. Journal of Applied Meteorology and Climatology, 47(4), 1038-1060.

Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M., & Kluzek, E. (2010). Technical description of an urban parameterization for the Community Land Model (CLMU). NCAR, Boulder, 10, D6K35RM9.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stöckli, R., Wang, A., Yang, Z.-L., Zeng, X., & Zeng, X. (2010). Technical description of version 4.0 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-478+STR, 257 pp.

Oliveira, A., Lopes, A., & Soares, A. (2022). Excess heat factor climatology, trends, and exposure across European functional urban areas. Weather and Climate Extremes, 36, 100455.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., & Mcaneney, J. (2007). Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. Journal of climate, 20(17), 4356-4376.

Pichelli, E., Coppola, E., Sobolowski, S., Ban, N., Giorgi, F., Stocchi, P., ... & Vergara-Temprado, J. (2021). The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation. Climate Dynamics, 56, 3581-3602.

Pietikäinen, J.-P., Sieck, K., Buntemeyer, L., Frisius, T., Nam, C., Hoffmann, P., Pop, C., Rechid, D., and Jacob, D.: REMO2020: a modernized modular regional climate model, EGUsphere (2025 preprint), https://doi.org/10.5194/egusphere-2025-1586, 2025.

Porson, A., Clark, P. A., Harman, I. N., Best, M. J., & Belcher, S. E. (2010a). Implementation of a new urban energy budget scheme in the MetUM. Part I: Description and idealized simulations. Quarterly Journal of the Royal Meteorological Society, 136(651), 1514-1529.

Porson, A., Clark, P. A., Harman, I. N., Best, M. J., & Belcher, S. E. (2010b). Implementation of a new urban energy budget scheme into MetUM. Part II: Validation against observations and model intercomparison. Quarterly Journal of the Royal Meteorological Society, 136(651), 1530-1542.

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., ... & Duda, M. G. (2017). The weather research and forecasting model: Overview, system efforts, and future directions. Bulletin of the American Meteorological Society, 98(8), 1717-1737.

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., ... & Jacob, D. (2016). Precipitation in the EURO-CORDEX 0.11^ ◦ 0. 11◦ and 0.44^ ◦ 0. 44◦ simulations: high resolution, high benefits?. Climate dynamics, 46, 383-412.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., ... & Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. Reviews of geophysics, 53(2), 323-361.

Quenum, G. M. L. D., Lemonsu, A., Corneille, L., Harader-Coustau, E., Caillaud, C., Somot, S., Monteiro, D., Machado, T., Li, S., Chevé, M., Richard, Y. (2026 submitted). Automatic mapping of urban heat island and population exposure based on high-resolution climate simulations: application to French cities. Climate Dynamics.

Rechid, D. (2009). On biogeophysical interactions between vegetation phenology and climate simulated over Europe (Doctoral dissertation, University of Hamburg Hamburg).

Rechid, D., Raddatz, T. J., & Jacob, D. (2009). Parameterization of snow-free land surface albedo as a function of vegetation phenology based on MODIS data and applied in climate modelling. Theoretical and applied Climatology, 95(3), 245-255.

Redon, E. C., Lemonsu, A., Masson, V., Morille, B., & Musy, M. (2017). Implementation of street trees within the solar radiative exchange parameterization of TEB in SURFEX v8. 0. Geoscientific Model Development, 10(1), 385-411.

Redon, E., Lemonsu, A., & Masson, V. (2020). An urban trees parameterization for modeling microclimatic variables and thermal comfort conditions at street level with the Town Energy Balance model (TEB-SURFEX v8. 0). Geoscientific Model Development, 13(2), 385-399.

Rockel, B., Will, A., & Hense, A. (2008). The regional climate model COSMO-CLM (CCLM). Meteorologische zeitschrift, 17(4), 347.

Romanello, M., Di Napoli, C., Green, C., Kennard, H., Lampard, P., Scamman, D., ... & Costello, A. (2023). The 2023 report of the Lancet Countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms. The Lancet, 402(10419), 2346-2394.

Rossa, A., Ferrario, M. E., Sansone, M., & Monai, M. (2012, September). Climatology of the static stability of the night-time Po Valley pbl from radio sondes and passive microwave radiometers. In Proceedings of the 9th International Symposium on Tropospheric Profiling, edited by: Cimini, D., Di Girolamo, P., Marzano, FS, and Rizi (Vol. 50).

Rummukainen, M. (2010). State-of-the-art with regional climate models. Wiley Interdisciplinary Reviews: Climate Change, 1(1), 82-96.

Sailor, D. J., & Lu, L. (2004). A top–down methodology for developing diurnal and seasonal anthropogenic heating profiles for urban areas. Atmospheric environment, 38(17), 2737-2748.

Salamanca, F., Zhang, Y., Barlage, M., Chen, F., Mahalov, A., & Miao, S. (2018). Evaluation of the WRF-urban modeling system coupled to Noah and Noah-MP land surface models over a semiarid urban environment. Journal of Geophysical Research: Atmospheres, 123(5), 2387-2408.

Sangelantoni, L., Sobolowski, S., Lorenz, T., Hodnebrog, Ø., Cardoso, R. M., Soares, P. M., ... & Bastin, S. (2024). Investigating the representation of heatwaves from an ensemble of km-scale regional climate simulations within CORDEX-FPS convection. Climate Dynamics, 62(6), 4635-4671.

Schoetter, R., Kwok, Y. T., de Munck, C., Lau, K. K. L., Wong, W. K., & Masson, V. (2020). Multi-layer coupling between SURFEX-TEB-v9. 0 and Meso-NH-v5. 3 for modelling the urban climate of high-rise cities. Geoscientific Model Development, 13(11), 5609-5643.

Schoetter, R., Masson, V., Bourgeois, A., Pellegrino, M., & Lévy, J. P. (2017). Parametrisation of the variety of human behaviour related to building energy consumption in the Town Energy Balance (SURFEX-TEB v. 8.2). Geoscientific Model Development, 10(7), 2801-2831.

Schrodin, R., & Heise, E. (2001). The multi-layer version of the DWD soil model TERRA_LM (p. 16). Offenbach, Germany: DWD.

Schulz, J. P., & Vogel, G. (2020). Improving the processes in the land surface scheme TERRA: Bare soil evaporation and skin temperature. Atmosphere, 11(5), 513.

Schulz, J.-P., and G. Vogel (2017), An improved representation of the land surface temperature including the effects of vegetation in the COSMO model, in Geophysical Research Abstracts, pp. 19, EGU2017–7896, European Geosciences Union General Assembly, Vienna.

Schulz, J.-P., Vogel, G., Becker, C., Kothe, S., Rummel, U., and Ahrens, B.: Evaluation of the ground heat flux simulated by a multi-layer land surface scheme using high-quality observations at grass land and bare soil, Meteorologische Zeitschrift, 25, 607–620, https://doi.org/10.1127/metz/2016/0537, 2016.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., ... & Masson, V. (2011). The AROME-France convective-scale operational model. Monthly Weather Review, 139(3), 976-991.

Semmler, T. (2002). Der Wasser-und Energiehaushalt der arktischen Atmosphäre (Doctoral dissertation, University of Hamburg, Hamburg).

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. D., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., Zhou, B. and Allan, R. (2021) Weather and climate extreme events in a changing climate. In: Masson-Delmotte, V. P., Zhai, A., Pirani, S. L. and Connors, C. (eds.) Climate Change 2021: The Physical Science Basis: Working Group I contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, pp. 1513-1766. doi: 10.1017/9781009157896.013

Shen, P., Zhao, S., Zhou, D., Lu, B., Han, Z., Ma, Y., ... & Liu, S. (2024). Surface and canopy urban heat island disparities across 2064 urban clusters in China. Science of the Total Environment, 955, 177035.

Sismanidis, P., Bechtel, B., Perry, M., & Ghent, D. (2022). The seasonality of surface urban heat islands across climates. Remote Sensing, 14(10), 2318.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., ... & Powers, J. G. (2008). A description of the advanced research WRF version 3. NCAR technical note, 475(125), 10-5065.

Soares, P. M., & Cardoso, R. M. (2018). A simple method to assess the added value using high-resolution climate distributions: application to the EURO-CORDEX daily precipitation. International Journal of Climatology, 38(3), 1484-1498.

Soares, P. M., Careto, J. A., Cardoso, R. M., Goergen, K., Katragkou, E., Sobolowski, S., ... & Bastin, S. (2024). The added value of km-scale simulations to describe temperature over complex orography: the CORDEX FPS-Convection multi-model ensemble runs over the Alps. Climate dynamics, 62(6), 4491-4514.

Sørland, S. L., Brogli, R., Pothapakula, P. K., Russo, E., Van de Walle, J., Ahrens, B., ... & Thiery, W. (2021). COSMO-CLM regional climate simulations in the Coordinated Regional Climate Downscaling Experiment (CORDEX) framework: a review. Geoscientific Model Development, 14(8), 5125-5154.

Sparrow, E., & Cess, R. D. (1970). Radiation heat transfer, Chap. 3–4, appendices A & B. Thermal Science Series, Brooks/Cole, 75-136.

Steensen, B. M., Marelle, L., Hodnebrog, Ø., & Myhre, G. (2022). Future urban heat island influence on precipitation. Climate Dynamics, 58(11), 3393-3403.

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., ... & Joly, A. (2018). The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. Geoscientific Model Development, 11(1), 257-281.

Tuholske, C., Caylor, K., Funk, C., Verdin, A., Sweeney, S., Grace, K., ... & Evans, T. (2021). Global urban population exposure to extreme heat. Proceedings of the National Academy of Sciences, 118(41), e2024792118.

UN-Habitat (2022). Envisaging the future of cities. World cities report, 2022, 42.

Van Meijgaard, E., Van Ulft, L. H., Lenderink, G., De Roode, S. R., Wipfler, E. L., Boers, R., & van Timmermans, R. M. A. (2012). Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe (No. KVR 054/12). KvR.

van Meijgaard, E., Van Ulft, L. H., Van de Berg, W. J., Bosveld, F. C., Van den Hurk, B. J. J. M., Lenderink, G., & Siebesma, A. P. (2008). The KNMI regional atmospheric climate model RACMO, version 2.1 (p. 43). De Bilt, The Netherlands: KNMI.

Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., ... & Yiou, P. (2013). The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Climate dynamics, 41, 2555-2575.

Viterbo, P., & Beljaars, A. C. (1995). An improved land surface parameterization scheme in the ECMWF model and its validation. Journal of climate, 8(11), 2716-2748.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., ... & Xavier, P. (2017). The Met Office unified model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. Geoscientific Model Development, 10(4), 1487-1520.

Wilby, R. L., & Wigley, T. M. (1997). Downscaling general circulation model output: a review of methods and limitations. Progress in physical geography, 21(4), 530-548.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics bulletin, 1(6), 80-83.

Wouters, H., Demuzere, M., Blahak, U., Fortuniak, K., Maiheu, B., Camps, J., ... & van Lipzig, N. P. (2016). The efficient urban canopy dependency parametrization (SURY) v1. 0 for atmospheric modelling: description and application with the COSMO-CLM model for a Belgian summer. Geoscientific Model Development, 9(9), 3027-3054.

Wouters, H., Demuzere, M., De Ridder, K., & van Lipzig, N. P. (2015). The impact of impervious water-storage parametrization on urban climate modelling. Urban Climate, 11, 24-50.

Wouters, H., Varentsov, M., Blahak, U., Schulz, J. P., Schättler, U., Bucchignani, E., & Demuzere, M. (2017, April). User guide for TERRA_URB v2. 2: The urban-canopy land-surface scheme of the COSMO model.

Yang, B., Zhang, Y., & Qian, Y. (2012). Simulation of urban climate with high-resolution WRF model: A case study in Nanjing, China. Asia-Pacific Journal of Atmospheric Sciences, 48(3), 227-241.

Yang, C., & Zhao, S. (2023). Diverse seasonal hysteresis of surface urban heat islands across Chinese cities: Patterns and drivers. Remote Sensing of Environment, 294, 113644.

Yow, D. M. (2007). Urban heat islands: Observations, impacts, and adaptation. Geography Compass, 1(6), 1227-1251.

Zerbst, L., Ban, N., Kohlhauser, I., Medvedova, A., Le Roy, B., Rotach, M., Ward, H. (2026, submitted). Analysis of urban temperatures using the first multi-model ensemble of km-scale regional climate models in the Alps. Climate Dynamics.

Zhang, Z., Paschalis, A., Mijic, A., Meili, N., Manoli, G., van Reeuwijk, M., & Fatichi, S. (2022). A mechanistic assessment of urban heat island intensities and drivers across climates. *Urban Climate*, *44*, 101215.

Zhou, B., Lauwaet, D., Hooyberghs, H., De Ridder, K., Kropp, J. P., & Rybski, D. (2016). Assessing seasonality in the surface urban heat island of London. Journal of Applied Meteorology and Climatology, 55(3), 493-505.

Zhou, B., Rybski, D., & Kropp, J. P. (2013). On the statistics of urban heat island intensity. Geophysical research letters, 40(20), 5486-5491.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- LeRoyetal2026ClimDynRCMCPMcitiesSupMat.pdf