A statistical–dynamical downscaling methodology for the urban heat island applied to the EURO-CORDEX ensemble

Benjamin Le Roy1 · Aude Lemonsu1 · Robert Schoetter1

Received: 10 July 2020 / Accepted: 21 December 2020 / Published online: 15 January 2021 © The Author(s) 2021

Abstract
Regional Climate Models (RCMs) are the primary climate information available to public stakeholders and city-planners to support local adaptation policies. However, with resolution in the order of ten kilometres, RCMs do not explicitly represent cities and their influence on local climate (e.g. Urban Heat Island; UHI). Downscaling methods are required to bridge the gap between RCMs and city scale. A statistical–dynamical downscaling methodology is developed to quantify the UHI of the city of Paris (France), based on a Local Weather Types (LWTs) classification combined with short-term high-resolution (1-km) urban climate simulations. The daily near-surface temperature amplitude, specific humidity, precipitation, wind speed and direction simulated by the RCMs are used for the LWTs attribution. The LWTs time series is associated to randomly selected days simulated with the mesoscale atmospheric model Meso-NH coupled to the urban canopy model Town Energy Balance to calculate the UHI corresponding to the successive LWTs. The downscaling methodology is applied to the EURO-CORDEX ensemble driven by the ERA-Interim reanalysis, and evaluated for the 2000–2008 period against station observations and a 2.5-km reanalysis. The short-term dynamical simulations slightly underestimate and overestimate near-surface minimum and maximum air temperature respectively, but capture the UHI intensity with biases in the order of a tenth of a degree. RCMs show significant differences in the variables used for the LWTs attribution, but the seasonal LWT frequencies are captured. Consequently, the reconstructed temperature fields maintain the small biases of the Meso-NH simulations and the statistical–dynamical downscaling greatly improves the UHI compared to the raw data of RCMs.

Keywords Urban heat island · Downscaling · EURO-CORDEX · Mesoscale modelling · Model evaluation

1 Introduction
Anthropogenic global warming (Stocker et al. 2013) is expected to cause regional climate changes (Christensen 2013), which will have in turn, several local impacts (e.g. on the urban environment; Wilby 2007). Urban areas are already the most densely populated and economically active places in the world and are expected to grow further in the future (UN DESA 2019). Moreover, due to their specific characteristics (e.g. artificialisation of surfaces) and the activities they host, they are already subject to specific hazards (e.g. air pollution; Duh et al. 2008; Landrigan et al., 2018; Salmond et al. 2018, heat waves and thermal stress; Kovats and Hajat 2008; Gosling et al. 2009; Dousset et al. 2011, and floods; Hammond et al. 2015; Mignot et al. 2019). Consequently, decision makers and stakeholders need information in order to be able to think the city of tomorrow and to develop the necessary adaptations to future changes (Revi et al. 2014; IPCC 2018). As a response, scientists conduct impact studies focused on the regional or local scale in many different fields such as air pollution (Künzli et al. 2000; Vautard et al. 2003), energy consumption (Santamouris et al. 2001; Ewing and Rong 2008; Kolokotroni et al. 2012), thermal comfort and heat stress (Patz et al. 2005; Lemonsu et al. 2013, 2014, 2015; Lauwaet et al. 2020; Viguié et al. 2020), urban ecology (Muratet et al. 2007; Hope et al 2008; Buyantuyev and Wu 2009; Wu 2014) or hydrology (Foster et al 1994; Mitchell et al 2001; Berne et al. 2004; Barles 2007; Fletcher et al. 2013). Impact models are driven by climate information either coming from Earth System Models (ESMs) or Regional Climate Models (RCMs). The RCMs horizontal spatial resolution is in the order of the
ten of kilometres which enables them to directly provide climatic forcings for some impact studies (e.g. air quality; Hogref et al. 2004; Leung 2005; Menut et al. 2013, heatwaves; Vautard et al. 2013, hydrology; Graham et al. 2007, energy production; Tobin et al. 2015; Gutiérrez et al. 2020). However, urban related studies present two main challenges. The size of the urban areas and the high heterogeneity of urban surfaces require a fine spatial resolution and a detailed representation of land covers and their physical properties. In addition, climatic conditions, especially air temperature in the atmospheric boundary layer, are locally modified by the presence of the city (i.e. the Urban Heat Island, UHI; Oke 1973; Arnfield 2003). In most cases, the spatial resolution and surface parameterisations of RCMs are not sufficient and adapted to represent those high resolution processes and feedbacks, and therefore the RCM results need to be downscaled (Rummukainen 2010). In order to derive higher spatial resolution information from RCMs, two types of downscaling exist, each with advantages and disadvantages: Statistical and Dynamical Downscaling (SD and DD, respectively; Wilby and Wigley 1997). SD approaches link large scale meteorological variables (i.e. predictors) with local ones (predictands) to find statistical relationships. These relationships are then applied to the same resolution of RCM output in which they were calibrated to obtain higher resolution meteorological variables. SD is computationally cheap as it does not require additional numerical simulations, but it can make strong assumptions regarding the universal validity of the statistical relationships and especially their stationarity in future climate conditions. On the other hand, DD approaches use the outputs from lower resolution climate models to drive finer resolution models. They assume no statistical relationships and provide physical consistency of the results, however, the finer the resolution is, the more computationally expensive the DD approaches become. These studies are often limited to a single GCM or RCM ensemble member and a single emission scenario, as well as reduced time slices that might call into question the reliability of resulting impact analysis. Statistical–dynamical downscaling methodologies (SDD; Frey-Buness et al. 1995; Boé et al. 2006) have been developed to combine the advantages of both approaches. For the SDD of the UHI, weather typing approaches have been developed with the aim to add the high-resolution UHI field to coarse resolution RCM results. The postulate that supports this approach is that the UHI phenomena is highly dependent on meteorological conditions, varying in intensity according to wind speed and cloudiness conditions, and in spatial structure according to regional flow. Using large-scale or local meteorological variables they determine specific weather situations associated with a particular response of the urban climate to the local meteorological conditions. The UHI effect is then explicitly represented by conducting high-resolution urban climate simulations, whose results are combined with the RCM output.

Duchêne et al. (2020) developed an analogs’ based SDD methodology to reconstruct Brussels’ (Belgium) UHI in regional climate projections using the ALARO model at 4-km resolution coupled with the state-of-the-art 3D Urban Canopy Model (UCM) Town Energy Balance (TEB; Masson 2000). Their approach relies on a 4-years continuous simulation of the city’s UHI whose duration necessarily constraints the spatial resolution that can be achieved. By statistical analogy, they link each day of a given regional climate projection according to its daily atmospheric variables with the most similar day of the 4-years high-resolution urban simulation. The RCM fields are then corrected by integrating the UHI signature of the city of the day considered. Hoffmann et al. (2018) reconstructed Hamburg’s (Germany) UHI in future climate projections by combining a weather pattern classification and 1-km resolution urban climate simulations with the mesoscale model METRAS. Compared to the previous method, they free themselves from the resolution limitations related to the computation time since they only simulate selected meteorological situations discriminating the UHI’s intensity and spatial pattern. However, METRAS does not include an UCM. Instead, the specifics of the urban surfaces are taken into account by the modification of the surface parameters related to radiative, energy, and water exchanges.

The method presented here follows a weather typing approach similar to that of Hoffmann et al. (2018) rather than the analogue approach of Duchêne et al. (2020) with two main improvements: (1) the high resolution urban climate simulations are done for a limited number of days that are representative of the weather types, and not in a continuous way, thus limiting the computational cost; (2) these urban climate simulations are performed with the same UCM TEB used by Duchêne et al. (2020) but with a higher horizontal resolution of 1 km and a more sophisticated version of the model including in-canyon urban vegetation (Lemonsu et al. 2012) and a building energy model (Bueno et al. 2012; Pigeon et al. 2014).

The present study describes the SDD method which consists in processing time series of meteorological variables simulated by RCMs and representative of the meteorological conditions in the rural areas surrounding the city (Fig. 1a). Local day-to-day weather types discriminating the UHI’s intensity and spatial pattern are determined based on the RCM output (Fig. 1b). High resolution urban climate simulations for these weather types are performed using a mesoscale atmospheric model and the results are combined with the rural background RCM data to obtain a climatological time series of high-resolution temperature maps (Fig. 1c). For the weather types classification, a database developed by Hidalgo et al. (2018)
for 51 French cities and a set of high resolution urban simulations already performed and presented in detail by Schoetter et al. (2020) and Gardes et al. (2020) are available. The method is here applied and evaluated for the city of Paris (France) with the aim of correcting daily temperature values from RCM projections for the urban influence on air temperature and for producing daytime and nighttime UHI diagnostics. The SDD methodology aims to bring high resolution information in climate projections to addresses UHI issues and their impacts. Accordingly, it is used to the spatial downscaling of air temperature and related variables (humidity and atmospheric radiation), but is not suitable for the reconstruction of the city’s effect on wind speed, precipitation, and other indicators that are not strongly related to air temperature.

The aim of the study is to evaluate the capacity of the SDD methodology to reconstruct the seasonality of UHIs both in terms of intensities and spatial extent. The weather type climatology, the high resolution urban climate simulations, and the SDD methodology are presented in Sect. 2. Then, the different steps of the approach are evaluated in Sect. 3. The final results are presented in Sect. 4 and completed by a discussion on the possible limitations of the approach.

2 Data and method

This section presents the successive steps of the SDD methodology and the different databases it relies on.

2.1 Local weather type climatology

Weather type classifications used for statistical downscaling are developed to link large scale predictors such as atmospheric circulation regimes with local scale predictands (Cassou et al. 2005). For the present case study over the Paris region, we use the existing Local Weather Types (LWTs) classification developed by Hidalgo and Jougla (2018) for 51 French cities, because it meets the specific needs of urban climate studies. Unlike traditional classifications related to the synoptic-scale meteorological situation (Philipp et al. 2010; Beck and Philipp 2010), the LWTs are related to the different states of the planetary boundary layer in the rural areas adjacent to the city and are attributed based on near-surface meteorological variables (Hidalgo and Jougla 2018). Applied in the vicinity of cities, the LWTs make it possible to link the evolution...
of meteorological situations to the day-to-day variability of the UHI phenomenon.

Based on the previous work of Hidalgo et al. (2014), Hidalgo and Jougla (2018) developed a statistical methodology to produce a LWTs classification. They used the “Partition Around Medoids” clustering algorithm (PAM; Kaufman and Rousseeuw, 1990) in combination with Gower’s distance (Gower 1971) as a measure of dissimilarity for a set of five daily variables: the daily near-surface (2 m above the ground) thermal amplitude (dT) computed from the minimum and maximum air temperatures (TN and TX), specific humidity (Q), precipitation rate (RR), wind speed (FF) and direction (DD) computed from the zonal and meridional wind components (U and V). The meteorological data required for the LWTs attribution were taken from a reanalysis produced with the French numerical weather prediction model AROME (Seity et al. 2011) with a 2.5-km resolution over the 10-year period 2000–2009 (Pouponneau et al. 2017). This classification method was evaluated for Toulouse (Hidalgo and Jougla 2018), and Orléans, Paris, and Tours (Jougla et al. 2019). For each city, the meteorological variables used for the LWTs classification were taken from the AROME reanalysis for a reference rural grid point located near the city to represent the local climate but far enough to avoid urban influences.

The LWTs classification for Paris consists of 12 LWTs. Figure 2 presents the seasonal occurrence frequencies of the LWTs for the 2000–2008 period. Every LWT can be found in every season with the exception of LWT09 and 12 for winter (DJF). Seasonal predominances are found: LWT01, 02, 03, 04 and 06 in DJF, and LWT05, 07, 08, 09, 10, 11 and 12 in summer (JJA). Except for LWT08, no LWT is specific to the transition seasons, i.e. spring (MAM) or autumn (SON). These two seasons are represented by a combination of all possible LWTs. The different steps of the classification are presented in “Appendix 1” and more details on the Paris LWTs classification are given by Jougla et al (2019).

2.2 Short-term high-resolution urban climate simulations

2.2.1 Model configuration

Mesoscale atmospheric models incorporating an UCM can simulate the urban climate at horizontal resolutions as fine as 100 m × 100 m. These models not only have higher resolution than RCMs but also better represent mesoscale meteorological phenomena (e.g. an urban breeze) and near-surface processes (i.e. in the urban canopy layer). They make use of finer land cover databases with more details regarding the geometry (e.g. building height, plan area building density, aspect ratio of the street canyons) and the physical parameters (e.g. albedo, emissivity, thermal conductivity and capacity) of the city.

Schoetter et al. (2020) and Gardes et al. (2020) conducted high-resolution urban climate simulations for 42 French cities including Paris based on the LWTs approach. They used the non-hydrostatic limited-area research atmospheric model Meso-NH (Lac et al. 2018) driven by lateral boundary conditions provided by the European Centre for Medium-Range Weather Forecast (ECMWF) Integrated Forecasting System (IFS) high-resolution operational forecast analysis (ECMWF 2019). Two intermediary domains centred over the city of Paris were used to downscale the ECMWF-IFS analysis up to a horizontal resolution of 1 km (Table 1). Surface-atmosphere exchanges in Meso-NH are calculated using the SURFEX platform (Masson et al. 2013), which includes the UCM TEB (Masson 2000; Lemonsu et al. 2012; Bueno et al.
A statistical–dynamical downscaling methodology for the urban heat island applied to the…

2012; Pigeon et al. 2014). TEB is a state-of-the-art UCM which represents the complex urban geometry as an urban street canyon of infinite length (Oke 1981). More details regarding the physical parameterisations are given in Kwok et al. (2019). The present study uses the exact same model configuration as in Gardes et al. (2020) for the city of Paris over the Meso-NH model domain at 1-km horizontal resolution (Fig. 3).

Schoetter et al. (2020) simulated an ensemble of past meteorological situations representative of each LWT for each season. They found that for each combination of LWT and season, ideally 6 days should be simulated to obtain a robust average UHI pattern during these specific meteorological conditions. With 12 LWTs and 4 seasons, this would result in 288 days to be simulated. However, not all LWTs occur in each season and some LWTs occur for less than 6 days in the 10-year LWTs attribution period, which reduces the number of days to be simulated slightly. For LWTs which occur less than 6 days in one season during the 10-year attribution period, all days belonging to this combination of LWT and season are simulated. For LWTs which occur for more than 6 days during one season, the days are selected to prioritize more stationary meteorological situations (e.g. 3 consecutive days with the same LWT), since we assume that the UHI pattern for these more stationary situations is better representative of the specific influence of this LWT on the local meteorological conditions. As a result, a total of 255 days were simulated (not counting two days of spinup before each simulation).

For each day, two simulations were conducted: (1) SIMURB for which the city of Paris was represented using a detailed urban database for French cities (Schoetter et al. 2019) containing information on urban morphology (Bocher et al. 2018), building construction materials (Tornay et al. 2017), and human behaviour related to building energy consumption (Schoetter et al. 2017); and

| Domain | Extent (km) | Horizontal resolution (km) | Time step (s) | Parametrisation of deep convection | Parametrisation of shallow convection and dry thermals | Mixing length calculation |
|--------|-------------|---------------------------|--------------|----------------------------------|-----------------------------------------------------|--------------------------|
| D1     | 1600×1600   | 8                         | 20           | Kain and Fritsch (1990)          | Pergaud et al. (2009)                                | Bougeault and Lacarrère (1989) |
| D2     | 400×400     | 2                         | 20           | None                             | Pergaud et al. (2009)                                | Bougeault and Lacarrère (1989) |
| D3     | 200×200     | 1                         | 20           | None                             | Pergaud et al. (2009)                                | Bougeault and Lacarrère (1989) |

Fig. 3 Map of the city of Paris and the area studied for the SDD methodology (a). EURO-CORDEX domain (b). Comparison of the EURO-CORDEX domain resolution and the Meso-NH resolution (c)
(2) \( \text{SIM}_{\text{RUR}} \) for which the urban land cover was replaced with the predominant natural land cover in the surroundings of Paris from the ECOsystem CLImate MAP (ECOCLIMAP)-I database (Masson et al. 2003; Chameaux et al. 2005), which is “warm temperate crops”. The two simulations being otherwise identical, the difference in the meteorological fields allows to isolate the effect of the city of Paris on the local climate.

### 2.2.2 Computation of the urban influence on daily air temperature

The Meso-NH urban climate simulations are used here to derive the urban influence on air temperature, hereafter called the urban increment (Böhnenstengel 2011). While the UHI denotes the difference in air temperature between an urban and a rural location, and can be calculated based on observations or numerical model results; the urban increment is the air temperature difference between two numerical simulations at the same location, one with urbanisation and one without. The local effect of the city is calculated in the same way as what Duchêne et al. (2020) called the “urban signature”, according to:

\[
\Delta T_{\text{URB}}(x, y, z, t) = T_{\text{URB}}(x, y, z, t) - T_{\text{RUR}}(x, y, z, t)
\]

with \( T_{\text{URB}} \) and \( T_{\text{RUR}} \) the air temperature simulated with \( \text{SIM}_{\text{URB}} \) and \( \text{SIM}_{\text{RUR}} \), respectively. \( x, y, z \) the longitude, latitude, vertical height above ground level of the Meso-NH grid point, and \( t \) the time.

The high-resolution simulations not only provide information about the city’s effect on the local climate but can also be used to refine the regional and topographic effects, which are not included in \( \Delta T_{\text{URB}} \). Such regional gradients are obtained by subtracting the average background temperature from the 2D air temperature fields of \( \text{SIM}_{\text{RUR}} \):

\[
\Delta T_{\text{REG}}(x, y, z, t) = T_{\text{URB}}(x, y, z, t) - T_{\text{RUR}}(x, y, z, t)
\]

with \( T_{\text{RUR}} \) the spatial average of the horizontal temperature field (i.e. the spatial average of all grid points in \( \text{SIM}_{\text{RUR}} \) at height \( z \) and time \( t \)). The combination of the urban effect (\( \Delta T_{\text{URB}} \)) and the regional refinement (\( \Delta T_{\text{REG}} \)) yields the temperature correction (\( \Delta T_{\text{COR}} \)) that will be added to the coarse resolution RCM air temperature time series:

\[
\Delta T_{\text{COR}}(x, y, z, t) = \Delta T_{\text{URB}}(x, y, z, t) + \Delta T_{\text{REG}}(x, y, z, t)
\]

This correction (thereafter referred to as urban increment for simplicity) can be computed for any height above ground \( z \) or model output frequency available in the high-resolution urban climate simulations. Model results from \( \text{SIM}_{\text{URB}} \) are evaluated in Sect. 3.1.

### 2.3 Regional climate model ensemble

#### 2.3.1 EURO-CORDEX

The Coordinated Regional Downscaling Experiment (CORDEX; Giorgi et al. 2009) is an internationally coordinated framework initiated by the World Climate Research Program to define a common guideline regarding regional climate modelling. RCMs are either forced by General Circulation Models (GCMs) from the CMIP5 global climate projections (Taylor et al. 2012) or by the European Centre for Medium-Range Weather Forecasts Reanalysis (ECMWF ERA-Interim; Dee et al. 2011) for the RCMs evaluation.

To produce LWT time series from RCM data following the Hidalgo and Jouglé (2018) method, the average climatic conditions representative for the rural areas surrounding Paris need to be calculated. To do so, the daily variables (TN, TX, RR, Q, U, V) are extracted for grid points surrounding the city, by considering that urban areas extend on a domain of \( 3 \times 3 \) RCM grid points (Fig. 3c). The meteorological variables are then averaged over those grid points to obtain daily time series. These are used to perform a day-to-day LWT attribution (“Appendix 1”) to build a time series of LWTs for the 12 RCM evaluation experiments available (Table 2) over the EUR-11 domain (Fig. 3). RCMs cover different time periods of more than 20 years and are evaluated against a reanalysis with the AROME model for the common time period 2000–2008 (Sect. 3.2).

#### 2.3.2 Correction of RCM temperature with the urban influence on air temperature

RCM LWT time series are used to determine the urban increment to be combined with the daily temperature time series simulated by each RCM. The Meso-NH simulation results provide around six days (i.e. six dates) for every LWT and season which allows to obtain robust UHI patterns free of small-scale meteorological variability (Schoetter et al. 2020) and offers the opportunity to describe the intra-cluster variability for each LWT. For each successive day of the LWTs time series constructed from the RCM data, a Meso-NH day (date) is randomly selected from the associated LWT/season pool. The random selection is done equiprobably such as to not bias the mean for each LWT and season while providing the desired day-to-day variability. By concatenating the selected days from the Meso-NH simulations, time series of 2D 1-km temperature fields are constructed and combined with the RCM daily time series:

\[
T_{\text{SDD}}(x, y, z, t) = T_{\text{RCM}}(z, t) + \Delta T_{\text{COR}}(x, y, z, t)
\]
A statistical–dynamical downscaling methodology for the urban heat island applied to the…

with $T_{SDD}$ the downscaled temperature field combining the horizontally homogeneous rural climate simulated by the RCM ($T_{RCM}$) and the high-resolution urban increment ($\Delta T_{COR}$). Figure 4 illustrates the method for an example of TN/TX RCM time series. Based on RCM daily meteorological variables, a LWT is assigned to each day (and indicated at the bottom in Fig. 4). In accordance with the LWT and the season, a Meso-NH day is randomly selected from the pool of up to 6 days available for each season and LWT; the successive dates are listed on the right side of Fig. 4. Finally, urban increments of TN and TX corresponding to this day are added to the RCM TN and TX. The figure highlights that these 2D fields of urban increment result in different air temperature corrections for urban ($\Delta T_{COR,URB}$) and rural ($\Delta T_{COR,RUR}$) grid points. In addition, the methodology, here focused on daily temperatures, can also be extended to produce sub-daily atmospheric temperature fields, which can then be used to drive impact models (such as energy building

### Table 2: The EURO-CORDEX EUR-11 RCM evaluation ensemble used for the present study

| Institute | RCM             | Start           | End             | References                                |
|-----------|-----------------|-----------------|-----------------|-------------------------------------------|
| CNRM      | ALADIN53        | 1979-01-01      | 2008-12-31      | Colin et al. (2010)                       |
| CNRM      | ALADIN63        | 1979-01-01      | 2018-12-31      | Daniel et al. (2019)                      |
| CLMcom    | CCLM4-8-17      | 1989-01-01      | 2008-12-31      | Doms and Baldauf (2011); Doms et al. (2011) |
| CLMcom-ETH| COSMO-crCLIM-v1-1| 1979-01-01      | 2010-12-31      | Leutwyler et al. (2017)                    |
| MOHC      | HadREM3-GA7-05  | 1982-12-01      | 2012-12-01      | In preparation, see Walters et al. (2019) for HadGEM3-GC3.1 |
| DMI       | HIRHAM5         | 1989-01-01      | 2011-12-31      | Christensen et al. (2007)                 |
| KNMI      | RACMO22E        | 1979-01-01      | 2012-12-31      | van Meijgaard et al. (2008; 2012)         |
| SMHI      | RCA4            | 1980-01-01      | 2010-12-31      | Samuelsson et al. (2015)                  |
| MPI-CSC   | REMO2009        | 1989-01-01      | 2008-12-31      | Jacob et al. (2012)                       |
| GERICs    | REMO2015        | 1979-01-02      | 2012-12-31      |                                            |
| ICTP      | RegCM4-6        | 1979-01-01      | 2016-11-30      | Giorgi et al. (2012)                      |
| IPSL      | WRF381P         | 1989-01-01      | 2008-12-31      | Skamarock et al. (2005)                   |
models) at the city or neighbourhood scale. This step is presented in “Appendix 2”.

3 Evaluation

3.1 Evaluation of high-resolution urban climate simulations

The quality of the reconstructed temperature fields is directly related to the ability of the high-resolution UCM to correctly simulate UHIs in term of intensity, spatial pattern and seasonal evolution. A systematic evaluation of the 1-km horizontal resolution Meso-NH simulations (coupled to the TEB model) is done for the 255 days by comparison to 124 meteorological station observations available from the Météo-France observation network (Fig. 3). Observed TN and TX are compared station by station to the model output using the closest Meso-NH grid point. The five stations located in a 10 km × 10 km sub-domain centred on Paris are classified as urban, whereas the 119 stations located further from the centre of Paris than a box of 50 km × 50 km and in non-urban areas (according to the ECOCLIMAP-II land cover map; Faroux et al. 2013) are considered as rural. The average TN and TX are computed from urban and rural stations separately, and the corresponding UHIs (referred to as UHIN and UHIX, respectively) are computed from their difference. Quantitative evaluation is made using the mean seasonal bias (Δ), root mean squared error (RMSE), and hit-rate (HR in %) defined as:

\[
HR = \frac{1}{N} \sum_{d=1}^{N} n_d \times 100 \text{ with } n_d = \left\{ \begin{array}{ll}
1, & \text{if } |T_{M,d} - T_{O,d}| \leq \Delta T \\
0, & \text{else}
\end{array} \right.
\]

with \(d\) the index for the days and \(N\) the total number of days, \(T_M\) the modelled air temperature, \(T_O\) the observed air temperature, and \(\Delta T\) the desired accuracy threshold. It is here prescribed to 2 K for TN and TX and 1 K for the UHI. HR values range between 0 and 100%, with 0% (100%) meaning that no (every) day is in the desired accuracy range.

Table 3 lists the average seasonal scores weighted according to the 2000–2009 AROME LWT frequencies (Jougla et al. 2019) for urban and rural TN and TX, and associated UHIs. Meso-NH slightly overestimates TN, with a systematic positive bias whatever the season and the urban or rural environment. The positive bias is maximal during the winter season (\(\Delta T_{\text{N URB}} = 1.5\) K and \(\Delta T_{\text{N RUR}} = 1.4\) K in DJF) in accordance with higher rmse, lower hit-rates, and minimal in the spring season (0.5 and 0.6 K for the urban and rural environment, respectively). Due to the comparable and small \(\Delta T_{\text{N}}\) in urban and rural areas, UHIN is well simulated. Absolute bias (rmse) does not exceed 0.2 K (1.0 K). Hit-rates vary between 68 and 92% depending on the season, with lowest values in JJA and SON because the LWTs associated with the highest UHIN intensities are the most frequent in these seasons. Opposite results are found for TX. The biases are systematically negative and their absolute values are larger than for TN (\(\Delta T_{\text{X URB,ALL}} = -1.6\) K and \(\Delta T_{\text{X RUR,ALL}} = -2.3\) K). The scores vary with the season, poorer model performance is found in JJA than in DJF. Model performance also differs between urban and rural
areas. On average, TX_RUR is less well simulated than TX_URB with a bias (hit-rate) of −2.3 K (40%), compared with −1.6 K (58%) for TX_URB. These urban/rural discrepancies of the bias might be due to a shortcoming in soil moisture initialisation, as well as an insufficient number of spin-up days (2) used. This could explain why the worst simulated season is JJA. The combination of too humid soil, strong solar radiation, and developed vegetation leads to an overestimation of evapotranspiration thus leading to an underestimation of daytime air temperature especially in the rural areas. From these biases in TX results an overestimation of UHIX (ΔUHIX = 0.7 K).

In summary, the Meso-NH model coupled with TEB at 1 km horizontal resolution is able to simulate realistic UHIs in the nighttime. It slightly overestimates UHIs in the daytime, especially in JJA, but the biases remain acceptable.

### 3.2 Evaluation of EURO-CORDEX RCMs

Intrinsic errors in RCM simulations can impact the time series of downscaled temperature fields through the frequencies of LWTs that can be biased and that determine the corresponding UHI patterns. To understand the possible LWTs attribution differences between RCMs and reanalysis, the EURO-CORDEX daily variables are first compared to the corresponding AROME reanalysis variables rather than to the meteorological station observations. While Pouponneau et al. (2017) and Hidalgo and Jougla (2018) highlighted the high quality of the AROME reanalysis both in terms of the individual meteorological variable’s biases and the LWTs attribution, it is not expected to perfectly match station observations. Therefore, the RCMs evaluation presented below focuses more on the agreement between RCMs and AROME than on a general assessment of the quality of the RCMs.

#### 3.2.1 Seasonal evaluation of air temperature and precipitation

The results of the EURO-CORDEX RCMs, providing the rural background climatic conditions, are compared to the AROME reanalysis. TN, TX, the resulting daily temperature range dT, and RR are evaluated for the 2000–2008 period.

Figure 5 compares the RCM quantile distributions against those of AROME for DJF and JJA. In DJF, TN is mostly overestimated by the RCMs, except for ALADIN53 and
RACMO22E. Almost all quantiles simulated by the RCMs are within a 2 K accuracy range (grey area in Fig. 5), with the exception of the lowest quantiles. Biases are accentuated for negative temperatures, resulting in an even larger spread between RCMs. For TX, an overestimation is found for mid to high quantiles, with the exception of CCLM4-8-17 and HIRHAM5. Again, a greater ensemble spread is observed for the lowest quantiles. This time, it does not result from the exacerbation of biases but rather from the fact that some RCMs underestimate and others overestimate low values of $dT_{JJA}$. Larger biases and/or larger ensemble spread for the lower quantiles of DJF temperature can be due to the large sensitivity of DJF temperatures on differences in the regional temperature field since there is a strong west–east air temperature gradient in DJF in Europe and might also be due to the snow-albedo feedback. The biases of $dT$ are very variable because they can be the result of a compensation or an accumulation of bias in $TN_{DJF}$ and $TX_{DJF}$. Eight RCMs mostly underestimate it, two slightly overestimate it, and two strongly (outside of the 1 K range). The largest deviations are not found for the first quantiles, where the greatest spread and biases of $TN_{DJF}$ and $TX_{DJF}$ are found, but rather for the medium to high $dT_{DJF}$. The four RCMs aforementioned, for which important biases for TN and TX are found are outliers for $dT$: ALADIN53 and RACMO22E on the upper part with high $dT_{DJF}$ strongly overestimated mainly because of their underestimation of $TN_{DJF}$ and to a lesser extent CCLM4-8-17 and HIRHAM5 with lowest values of $dT$.

In JJA, TN are systematically overestimated, with the exception of ALADIN53(RCA4) lowest(highest) quantiles. The overall ensemble spread is larger than for $TN_{DJF}$ and relatively constant across the entire distribution except for the highest quantiles. Conversely to $TN_{DJF}$ where RCM biases are accentuated towards the distribution tails, for $TN_{JJA}$, some RCMs show a relatively constant bias throughout their distribution and better results for high quantiles (e.g. REMO2009 and REMO2015). For $TX_{JJA}$, a lot of variability is noted between RCMs (i.e. some overestimate the whole distribution like RegCM4-6, some others only the tails like REMO2015). Nevertheless, $TX_{JJA}$ are on average mostly overestimated, with the exception of HIRHAM5, RACMO22E, RCA4 that underestimate it on their whole distribution. The largest deviations are observed for the largest $TX_{JJA}$ values with half of the models underestimating and the other half overestimating up to 4 K (ALADIN63).

The large ensemble spread for $TX_{JJA}$ might be due to the differences in the soil moisture availability simulated by the different RCMs. As for $dT_{DJF}$, $dT_{JJA}$ is almost systematically underestimated (except for ALADIN53 on the whole distribution, and ALADIN63 and WRF381P for the largest and lowest quantiles, respectively) and these deviations are more important for larger $dT$. This might be due to the fact that the 2.5-km resolution of the reanalysis is able to capture

more extreme TN and TX than the 12-km resolution RCMs. The RCMs with the worst performance in terms of $dT_{JJA}$ relative to the AROME reanalysis are not the ones with the worst performances in terms of TN and TX (e.g. RCA4 and RegCM4-6) but rather those with large differences between their TN and TX biases (e.g. HIRHAM5).

For precipitation, $RR_{DJF}$ lower quantiles are systematically overestimated. This might be due to an overestimation of drizzle in RCMs (Piani et al. 2010; Maraun 2013). While the drizzle bias is also present in the AROME reanalysis (not shown here), the biases are smaller, which might be due to the finer resolution of the reanalysis. RCMs systematically underestimate the number of dry days (i.e. $RR < 0.1$ mm d$^{-1}$; with the exception of CCLM4-8–17 and COSMO-crCLIM-v1-1 and HadREM3-GA7-05). Every RCM overestimates rainy days ($1 < RR < 10$ mm d$^{-1}$) up to 12% for RCA4. In JJA, RCMs show a similar overestimation of lower quantiles, but with smaller biases. This is both due to an almost systematic overestimation of lower quantiles in RCMs with the exception of ALADIN63, CCLM4-8-17, COSMO-crCLIM-v1-1 and HIRHAM5) and an underestimation in AROME. The frequencies of days with more than 10 mm precipitation (which in turn would affect the Meso-NH days selected to represent the urban influence on air temperature. Even if the RCM climatologies matched the AROME reanalysis perfectly, differences in LWTs attribution on a day-to-day basis would still be expected due to the local scale meteorological variability (e.g. different wind directions or precipitating events), but on a seasonal average, the RCM LWT frequencies should be very similar to those obtained for the AROME reanalysis. These differences are objectively quantified using the Perkins et al. (2013) skill score (SSC, in %):

$$SSC = \sum_{LWT=1}^{12} \min(f_{M,LWT}, f_{R,LWT})$$

with $f$ the frequency of the LWT in the model $M$ or the reference $R$. The SSC’s theoretical range lies between 0 and 100%, with 0% meaning that the RCM only simulates LWTs that do not occur in the AROME reanalysis, and 100%
A statistical–dynamical downscaling methodology for the urban heat island applied to the…

meaning a perfect match between RCM and AROME reanalysis LWT frequencies.

Figure 6 illustrates the frequencies of every LWT for DJF and each EURO-CORDEX RCM as well as the associated SSCs. As discussed in Sect. 2.1, the only LWTs typically observed in DJF are LWT01, 02, 03, 04 and 06. This is found for all RCMs except for ALADIN53 (and RACMO22E to a lesser extent) which simulates some meteorological situations classified as LWT05, 07, 08. These LWTs are characterised by high dT, which are strongly overestimated by ALADIN53. Consequently, it has the lowest SSC (79%) while all other RCMs present an SSC of at least 85%. For LWT01 and LWT02, two groups of RCMs stand out: ALADIN63, RACMO22E, RCA4, RegCM4-6 and WRF381P simulate LWT01(02) frequencies well with a slight under(over) estimation whereas other RCMs show the
opposite with larger deviations. However, these two LWTs mainly differ in wind speed (FF) and RR, and exhibit a quite similar UHIN, therefore no large bias of the downscaled UHIN can be expected due to this shortcoming. LWT04 that results in the strongest UHIN (UHIN,DJF,LWT4 = 3 K), is the only LWT with winds coming from the North-East and its frequency is underestimated in every RCM and especially in ALADIN53 and WRF381P. The frequency of LWT03 is, with the exception of ALADIN53 for which it is underestimated, either well simulated (in ALADIN63, CCLM4-8-17, COSMO-crCLIM-v1-1, HADREM3-GA7-05 and RACMO22E), or strongly overestimated. The frequency of LWT06 is overestimated in every RCM and especially in CCLM4-8-17 and COSMO-crCLIM-v1-1. This is explained by the overestimation of days with winds coming from the South-East which are always associated with either LWT06 or LWT10.

In JJA, the LWTs occurring are LWT05, 07, 08, 09, 10, 11 and 12 (Fig. 7). There are more frequency differences between RCMs for a given LWT than in DJF [i.e. there is no general trend besides LWT05(07) whose frequency is systematically under(over)estimated]. This might be due to the

Fig. 7  Same as Fig. 6, but for JJA
fact that during JJA the synoptic scale circulation regimes are less relevant for the local scale meteorological conditions than for DJF, it is therefore more likely that the different RCMs simulate different meteorological conditions in the Paris region for the same lateral forcing at continental scale. As a result, the majority of RCMs has lower (or equal) values for SSC_{JJA} than SSC_{DJF}, with the exception of ALADIN53 and REMO2015. The overall attribution is still satisfactory with an average SCC_{JJA} of 85%. The LWTs in JJA discriminate the UHI_N more than in DJF, therefore LWT attribution errors could impact more significantly the quality of the UHI_N reconstruction. Nonetheless, RCMs all yield slightly overestimated but quite realistic frequencies for LWT10, which is associated with the strongest UHI_N (average UHI_{N,JJA,LWT10} = 4.3 K). The positive bias found in dT_{JJA} for ALADIN53 might explain its overestimation of the LWT09 frequency (+7 points) and underestimation of the LWT07 frequency (−6 points), which are relatively similar LWTs mainly differing in dT. HIRHAM5 which strongly underestimates dT_{JJA} overestimates the LWT03 frequency, which is low in JJA and high in DJF, at the expense of the frequency of LWT09. For every RCM except ALADIN53, LWTs typically more specific to other seasons and associated to small UHI_N intensities are more frequent than in the AROME reanalysis.

In summary, the EURO-CORDEX RCMs tend to well reproduce the LWT frequencies found in the AROME reanalysis for DJF and JJA. Specific exceptions for some RCMs/LWTs are found; they are related to the biases of the simulated meteorological predictors of the LWT. For some RCMs, these deviations are related to LWTs with distinct UHI_N, and thus might impact the statistically–dynamically downscaled UHIs.

3.3 Evaluation of the statistically–dynamically downscaled UHIs

To evaluate the ability of the methodology to reconstruct robust UHI patterns, seasonal maps of downscaled TN and TX are produced for DJF and JJA. Meso-NH days are randomly selected from the appropriate pool of LWT and season for each RCM and the corresponding masks of correction $\Delta T_{\text{COR}}$ (Eq. 3 in Sect. 2.2.2) are temporally averaged. These masks are not combined with the TN and TX time series of the RCM and directly analysed in order to assess the impact of errors in LWT frequencies (while leaving aside the RCM biases). Hereafter, the seasonal averages of $\Delta T_{\text{N,COR}}$ and $\Delta T_{\text{X,COR}}$ are called $\Delta T_{\text{N}}$ and $\Delta T_{\text{X}}$, respectively. The same approach is also applied to AROME’s TN and TX time series to reconstruct reference $\Delta T_{\text{N}}$ and $\Delta T_{\text{X}}$ seasonal maps.

![Fig. 8 Seasonal maps (DJF, JJA) of $\Delta T_{\text{N,COR}}$ obtained using the LWTs attributed for the AROME reanalysis and spatial bias (compared to AROME) when using the LWTs attributed from the EURO-CORDEX RCM ensemble](image)
Figure 8 presents the RCMs ΔTN difference maps for DJF (top) and JJA (bottom) compared to AROME’s. All RCMs are close to the AROME reference with differences never exceeding 0.4 K for an UHI of around 2–3 K, depending on the season. For DJF, the bias maps mostly highlight regional patterns with low bias over the city. The most important regional biases are noted in the north-west of the domain for ALADIN53, RACMO22E, and WRF381P. These RCMs have in common that they either overestimate the frequency of LWTs associated with above-average DJF temperature in this area (e.g. LWT05 and 07 for ALADIN53 and RACMO22E) or they underestimate the frequency of LWTs with lower than average DJF temperature (e.g. LWT04 for ALADIN53 and WRF381P). The overestimation of LWT05 and 07 frequencies in ALADIN53 (and to a lesser extent RACMO22E) also results in a small positive bias over the city. Other regional patterns are found, such as the positive bias in the North–East of the domain shared by multiple RCMs (HIRHAM5, REMO2009, and REMO2015), which might be due to an overestimation (underestimation) of the LWT01 (LWT02) frequency associated to higher (lower) temperature in this area.

In JJA, more variability is found between the statistically–dynamically downscaled ΔTN for the different RCMs. Regional patterns similar to those found in DJF are also present for some RCMs (e.g. the positive bias in the North-West of the domain for CCLM4-8-17 and HIRHAM5), but the results for JJA are characterised by larger biases in the urban area. This can be explained by the overall larger deviations found in the LWT frequencies in JJA, especially since the Meso-NH UHI cover a larger range of intensities in JJA than in DJF. Nonetheless, a comparable maximal absolute bias of 0.4 K is found for both seasons, that means the relative bias is lower in JJA than in DJF. With the exception of COSMO-crCLIM-v1 and HIRHAM5, most RCMs lead to a positive bias over the city. Especially, the UHI statistically–dynamically downscaled with ALADIN63 and WRF381P (and to a lesser extent ALADIN53) are too strong and have a too high spatial extent. Each one of those three RCMs overestimate the frequency of one of the four LWTs in JJA associated with strong UHI (LWT07 for WRF381P, LWT08 for ALADIN63 and LWT09 for ALADIN53) at the expense of LWTs with low UHI (e.g. LWT11).

For ΔTX, no bias is observed over the city (Fig. 9). Instead, a regional gradient of the bias common to most RCMs is noted in the main wind direction (south-west to north-east). As a result, we hypothesise that differences in ΔTX are more driven by the air circulation rather than the surface, which is consistent with daytime UHI characteristics. In JJA, as for ΔTN, more variability is found between RCMs, but this time without urban bias.

![Fig. 9](image_url)
In summary, some differences of the statistically–dynamically downscaled ΔTN and ΔTX are found between RCMs and the AROME reanalysis. Biases are more important in JJA for ΔTN both in rural areas in relation with topographic and climatological gradients, and over the city for some RCMs. We hypothesise that the strongest observed biases are linked to an overestimation of the frequency of the LWTs with large UHI intensities in the mesoscale simulations. Nevertheless, biases remain small (absolute values below 0.4 K) compared to UHI intensities, and the results for the RCMs agree well with those of the AROME reanalysis.

4 Discussion

The SDD methodology developed in the present study to produce high-resolution climatological average UHI fields assumes that the RCM’s horizontal resolution is not sufficient to adequately represent the urban climate (e.g. an urban breeze). Furthermore, with resolutions in the order of ten kilometres, RCMs can not simulate cities with a specific urban parameterisation. Based on a community survey, Langendijk et al. (2019) have gathered different information about the surface parameters used in several EURO-CORDEX RCMs and especially regarding the urban representation. While sometimes available, UCMs are not activated for EURO-CORDEX simulations, therefore urban land cover is either treated as rock surfaces, with adjusted parameters (e.g. roughness length for RACMO22E, as well as roughness length and albedo for REMO2009 and WRF381P, Langendijk et al. 2019), or disregarded. To our knowledge, only CCLM4-8-17 and COSMO-crCLIM-v1-1 replaced Paris’ city by natural land covers to completely eliminate the UHI effect in their simulation. RCMs also use different land cover databases, different land surface schemes, and different methodologies to deal with subgrid heterogeneity of land covers (e.g. REMO2009 uses a fractional approach, RACMO22E a dominant tile approach and COSMO-CLM a tiling approach, i.e. treats each sub-grid land cover type separately). All these disparities can impact simulation results,

![Fig. 10](image-url)  
**Fig. 10** Average TN and TX simulated by the EURO-CORDEX RCM ensemble for DJF and JJA during 2000–2008. Below and above season-average temperatures are represented in blue and red, respectively.
e.g. the near-surface air temperature simulated in the urban environment; Jänicke et al. (2017).

An in-depth analysis of these different approaches could help to understand the differences observed between RCMs but is beyond the scope of the present study. Figure 10 presents the average maps of TN and TX using results of all EURO-CORDEX RCMs for DJF and JJA. Even without explicit urban representations, an influence of the city on local air temperature is found, especially for TN_{JJA}. UHIs directly simulated by the RCMs can be compared to the ones reconstructed by the SDD methodology.

Table 4 lists the mean UHI intensities for TN and TX in DJF and JJA computed in the same way as for the Meso-NH evaluation in Sect. 3.1. UHI intensities are also computed directly in each EURO-CORDEX RCM (UHI_{RCM}) as well as in the statistically–dynamically downscaled maps of ΔTN and ΔTX (UHI_{SDD}), and the observation (UHI_{OBS}). Night-time UHIs directly simulated by RCMs (UHI_{N,RCM}) are almost all positive but systematically underestimated. The small positive UHI_{N} found in the RCMs for which the urban land cover has been replaced by rural land covers might be only due to topographic and regional temperature gradients (e.g. for CCLM4-8–17 and COSMO-crCLIM-v1-1 UHI_{N,RCM,JJA} = 0.2 and 0.4 K, respectively). UHI_{N,RCM} is only overestimated by WRF381P in JJA. The underestimation of UHI_{N} intensities is more important in DJF than in JJA. This could be explained by not taking into account the anthropogenic heat fluxes that are the main driver of the UHI phenomenon in DJF, since heating of buildings is a large contribution to anthropogenic heat flux in DJF, whereas air conditioning of buildings during JJA is not commonly applied in Paris during the selected time period. While RCMs using rock covers are able to simulate positive UHI_{N,RCM}, they also systematically overestimate UHI_{X,RCM}. This might be due to the thermal properties as well as the lack of shading inherent to the approach of representing cities using rock covers. Because rock surfaces are subject to a high amount of incoming solar radiation, they heat up strongly during the day, resulting in an overestimation of UHI_{X,RCM} that is more important in JJA than DJF because of the greater amount of incoming solar radiation. Conversely, at night, they do not have the heat storage capacity provided by the 3D geometry of the city and cool down more quickly, resulting in lower TN_{URB} and an underestimation of UHI_{N,RCM,JJA}.

Table 4 lists the averaged seasonal UHI_{N} and UHI_{X} for DJF and JJA during 2000–2008. UHIs are calculated in the same way as for the Meso-NH evaluation (Sect. 3.1).

| Observation          | UHI_{N} (K) |   | UHI_{X} (K) |   |
|----------------------|-------------|---|-------------|---|
|                      | DJF         | JJA| DJF         | JJA|
|                      | RCM SDD     | RCM SDD | RCM SDD | RCM SDD |
| AROME reanalysis     | – 2.3       | – 3.0       | – 1.6       | – 2.0       |
| ALADIN53             | 0.4 2.4     | 2.1 3.2     | 0.0 1.6     | 0.8 2.0     |
| ALADIN63             | 0.5 2.3     | 1.3 3.3     | 0.2 1.6     | 0.9 2.0     |
| CCLM4-8–17           | 0.2 2.3     | 0.2 3.2     | 0.3 1.6     | 1.0 1.9     |
| COSMO-crCLIM-v1-1    | 0.3 2.3     | 0.4 3.0     | 0.4 1.6     | 0.6 1.9     |
| HadREM3-GA7-05       | 1.1 2.2     | 2.4 3.1     | 0.7 1.6     | 0.9 1.9     |
| HIRHAM5              | – 0.2 2.3   | 1.0 3.0     | – 0.2 1.6   | 0.8 1.9     |
| RACMO22E             | 0.1 2.3     | 1.0 3.1     | 0.1 1.6     | 0.6 1.9     |
| RCA4                 | 0.4 2.2     | 1.0 3.1     | 0.5 1.6     | 1.6 1.9     |
| REMO2009             | 1.2 2.2     | 2.7 3.1     | 0.5 1.6     | 1.0 1.9     |
| REMO2015             | 0.0 2.2     | 0.9 3.1     | – 0.1 1.6   | 2.0 1.9     |
| RegCM4-6             | 0.1 2.3     | 1.0 3.1     | 0.1 1.6     | 1.9 1.9     |
| WRF381P              | 2.0 2.2     | 4.5 3.3     | 0.7 1.6     | 1.8 1.9     |
| ENSEMBLE mean        | 0.5 2.3     | 1.5 3.1     | 0.3 1.6     | 1.2 1.1     |
UHI\textsubscript{SDD} shows overall relatively small deviations from the AROME reanalysis. The largest deviation of UHI\textsubscript{N,SDD,JJA} is found for WRF381P, which is 0.3 K too large, which corresponds only to an overestimation of 10% of UHI\textsubscript{OBS}. Biases are even smaller during the day because UHI\textsubscript{X} are quite similar between LWTs both in Meso-NH and in the station observations. Overall, the end results of the statistical–dynamical downscaling are highly satisfactory. The UHI\textsubscript{N} intensities are clearly improved by the downscaling compared to the RCM results with average biases of the ensemble in the order of a tenth of a degree. For UHI\textsubscript{X}, DJF intensities are also improved due to the downscaling even though they are slightly overestimated. However, because of the Meso-NH TX biases noted in JJA, UHI\textsubscript{X,SDD,JJA} intensities are overestimated. The dynamical uncertainties coming from the Meso-NH biases are much greater than the statistical uncertainties associated to the LWT attribution errors. Consequently, to improve the overall results of the methodology, efforts should be focused on improving the Meso-NH simulations.

5 Conclusion

A statistical–dynamical downscaling methodology has been developed to produce high-resolution climatological daily UHI fields combining the climatology of the rural areas surrounding the city of Paris (France) from regional climate projections and the urban influence on the local climate simulated with a mesoscale atmospheric model. The methodology has been applied to 12 RCMs from the EURO-CORDEX ensemble for the 2000–2008 period. The different steps of the downscaling approach are evaluated against urban and rural weather station observations. The short-term high-resolution urban climate simulations slightly underestimate TN and overestimate TX. UHIs are overall well simulated, especially in the nighttime with average errors in the order of one tenth of a degree. Important disparities between RCMs are observed on the daily variables used for the LWT attribution. Nevertheless these differences have on average relatively small impacts on the LWT frequencies simulated by the RCMs. Consequently the reconstructed temperature fields maintain the small Meso-NH biases and allow to strongly improve the UHI intensities compared to those directly simulated by the RCMs.

Limitations common to other SDD methodologies are also present here, such as the stationarity hypothesis of LWTs in a future climate, as well as the stationarity of simulated UHI patterns. Nevertheless the satisfactory results obtained here, combined with the easy applicability of the method encourage its application to other cities where high-resolution urban climate simulations are available. In addition, most of the biases and uncertainties observed during the evaluation are related to the dynamical part of the methodology and could be improved without calling into question the statistical assumptions made. Other steps, specific to the reconstruction of atmospheric forcings have yet to be evaluated, as well as the quantification of the improvement of such a method compared to more simplistic (i.e. other SD) or complex (DD) ones.

By taking into account the UHI signal in climate projections, the SDD method should make it possible to study the evolution of temperature in the city under climate change conditions, and to derive standard meteorological indicators related to thermal effects, e.g. summer days, hot days or tropical nights (Karl et al. 1999; Peterson et al. 2001). Besides the impact diagnostics that can be directly deduced from the method, the data produced could also be used to feed more specific and high-resolution impact models, such as building energy models to evaluate energy consumption (Schoetter et al. 2020), or the indoor conditions of exposure of inhabitants to heat. Other environmental issues linked to urban climate but not directly related to air temperature, e.g. air quality or hydrology (Gidhagen 2020), imply different methodologies for impacts studies. Especially, the impact assessment of climate change on water resources in cities, require specific downscaling methods adapted to rainfall. A reflection is underway on these questions, in order to combine the two approaches and produce complete climatic forcing datasets.

Given the available databases, the methodology could be applied to all major French cities, as well as to other regional climate simulations (e.g. the next generation of EURO-CORDEX based on CMIP6; Eyring et al. 2016). But, it will first be applied on the EURO-CORDEX climate projections to study the effects of climate change on the UHI.

Appendix 1: Local weather type classification and attribution

The Local Weather Types (LWTs) classification developed by Hidalgo and Jouglà (2018) uses the “Partition Around Medoids” clustering algorithm (PAM; Kaufman and Rousseeuw 1990) with Gower’s distance as a measure of dissimilarity (Gower 1971).

The classification uses five daily variables: the daily thermal amplitude (dT = TX – TN in °C), the daily mean precipitation rate (RR in mm h\textsuperscript{-1}), the daily mean specific humidity (Q in kg kg\textsuperscript{-1}), the daily mean wind speed (FF in m s\textsuperscript{-1}) and the daily mean wind direction (obtained from the daily average of the hourly wind components and expressed in 4 wind quadrants of 90° from the north).

The PAM algorithm uses a matrix of dissimilarity to represent the distance between each pair of days in the AROME re-analysis time series. To produce this matrix, the partial
similarity ($S_v$) is computed for each variable $V$ and each pair of days ($D_1, D_2$) by normalising the absolute difference between the two days by the maximum amplitude of $V$ over the complete period ($\Delta V = V_{\text{MAX}} - V_{\text{MIN}}$)

$$s_{V(D_1,D_2)} = 1 - \frac{|V_{D_1} - V_{D_2}|}{\Delta V}$$

In the case of the wind direction, which is considered here as a categorical variable (in sectors from 1 to 4), the partial similarity is expressed as:

$$s_{V(D_1,D_2)} = \begin{cases} 1, & \text{if } V_{D_1} = V_{D_2} \\ 0, & \text{else} \end{cases}$$

Gower’s index of similarity ($S_G$) is computed for each pair of days as the average of the partial similarities of the five variables taken into account for the classification:

$$S_{G(D_1,D_2)} = \frac{1}{5} \sum_{j=1}^{5} s_{V(D_1,D_2)}$$

And Gower’s distance ($D_G$) which is used to produce the matrix is expressed as:

$$D_{G(D_1,D_2)} = 1 - S_{G(D_1,D_2)}$$

From the dissimilarity matrix, the PAM algorithm chooses a first set of $k$ medoids, and assigns each day to the nearest medoid by minimising the Gower distance, thus forming $k$ clusters. To find a satisfactory non-random first set of clusters, PAM first selects the day with the lowest sum of the distances to all others, and then selects other days (until $k$ is reached) always minimising the total sum of the distances (Schubert and Rousseeuw 2019).

From this initial set of medoids and clusters, PAM swaps all the medoids with all the available days and keeps the day associated with the smallest sum of the distances as a new medoid. The days are then again associated with the closest medoid. These two steps (inversion then reallocation) are repeated until the sum of the distances no longer decreases (Schubert and Rousseeuw 2019).

Hidalgo and Jougla (2018) conducted a sensitivity study to determine the optimal number of clusters. They applied several iterations of PAM by varying the number of prescribed clusters from 4 to 15 and then compared, for the different LWT classifications, the quality of the seasonal diurnal cycles reconstructed from the daily variables (see their temporal downscaling method in Hidalgo and Jougla 2018). For the Paris region, the classification into 12 LWTs (or clusters) gave the best results.

The attribution of LWTs on a new dataset on the basis of an existing classification follows the previous steps with some modifications. A new dissimilarity matrix is constructed for the new data, following the same principle, but retaining for normalisation the maximum amplitude $\Delta V$ of the reference classification. The PAM algorithm uses the 12 medoids of the reference classification to assign each day of the new period to one of the clusters by minimising the distance within a cluster while maximising the distance between clusters.

### Appendix 2: Sub-daily atmospheric forcing computation

The downscaling methodology presented and evaluated here is restricted to daily minimum and maximum temperatures but can be extended to sub-daily time-steps.

While the LWTs attribution step remains the same, the combination of RCM sub-daily outputs with Meso-NH is more challenging. The finest sub-daily temporal resolution of the RCMs (e.g. three hours for EURO-CORDEX) is to be combined with Meso-NH. The LWTs being defined from daily variables, the simplest approach would be to combine with the three-hourly urban increment of the same Meso-NH day (from 0 to 21 UTC). However, the thermal signature of the city (i.e. the UHI) does not result from the same physical processes during the day and night. Daytime UHI depends on the current atmospheric conditions whereas the nocturnal UHI is strongly governed by the conditions during the previous day, especially the incoming solar radiation that drives the process of heat storage in the construction materials. Therefore, rather than concatenating the daily cycles of the urban increment simulated by Meso-NH at 0 UTC, the transition time between two daily cycles is shifted to 9 UTC when temperature differences between urban and rural areas are expected to be at minimum, ensuring a smoother transition between two cycles. This involves starting the Meso-NH daily cycle at 12 UTC for the selected date and then extending it to 9 UTC using the following day. This date can correspond to a different LWT but it is assumed that the urban increments between 0 and 9 UTC are influenced by the previous day’s conditions. For this purpose, the next day of every simulation is also simulated (i.e. 104 days) resulting in a total of 359 simulated days. Following this approach, time series of sub-daily climate projections maps of 200x200 km taking into account the effect of the city on the atmosphere are created and can, in turn, be used to drive impact models.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated
otherwise in a credit line to the material. If material is not included in
the article’s Creative Commons licence and your intended use is not
permitted by statutory regulation or exceeds the permitted use, you will
need to obtain permission directly from the copyright holder. To view a
copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Arnfield AJ (2003) Two decades of urban climate research: a review
of turbulence, exchanges of energy and water, and the urban heat
island. Int J Climatol J R Meteorol Soc 23(1):1–26. https://doi.
org/10.1002/joc.859

Barles S (2007) Urban metabolism and river systems: an historical
perspective–Paris and the Seine, 1790–1970

Beck C, Philipp A (2010) Evaluation and comparison of circula-
tion type classifications for the European domain. Phys Chem
Earth Parts A/B/C 35(9–12):374–387. https://doi.org/10.1016/j.
pce.2010.01.001

Berne A, Delrieu G, Creutin JD, Obled C (2004) Temporal and spatial
resolution of rainfall measurements required for urban hydrol-
ogy. J Hydrol 299(3–4):166–179. https://doi.org/10.1016/j.jhydr.
el.2004.08.002

Bocher E, Petit G, Bernard J, Palominos S (2018) A geoprocessing
framework to compute urban indicators: the MApUCE tools
chain. Urban Clim 24:153–174. https://doi.org/10.1016/j.uclim.
2018.01.008

Boé J, Terray L, Habets F, Martin E (2006) A simple statistical–
dynamical downscaling scheme based on weather types and condi-
tional resampling. J Geophys Res Atmos 111:D23. https://
doi.org/10.1029/2005JD006889

Bohnenstengel SI, Evans S, Clark PA, Belcher S (2011) Simula-
tions of the London urban heat island. Q J R Meteorol Soc
137(659):1625–1640. https://doi.org/10.1002/qj.855

Bougeault P, Lacarrere P (1989) Parameterization of oroography-
induced turbulence in a mesoscale–scale model. Mon Weather
Rev 117(8):1872–1890

Bueno B, Pigeon G, Norford LK, Zibouche K, Marchadier C
(2013) Climate phenomena and their relevance to urban areas:
proceedings of the international conference organized
by the institution of civil engineers and held in London, 2–3
June 1993. Thomas Telford Publishing, pp 43–63. https://doi.
org/10.1175/CL13506.1

Champeaux JL, Masson M, Chauvin F (2005) ECOCLIMAP: a global
database of land surface parameters at 1 km resolution. Meteorol
Appl 12:29–32. https://doi.org/10.1017/S1350482705001519

Christensen OB, Drews M, Christensen JH, Dethloff K, Ketelsen K,
Hebestadt I, Rinke A (2007) The HIRHAM regional climate
model. Version 5 (beta)

Christensen JH, Kanikicharla KK, Aldrian E, An SL, Cavalcanti IFA, de
Castro M, et al. (2013) Climate phenomena and their relevance for
future regional climate change. In: Climate change 2013 the
physical science basis: working group I contribution to the fifth
assessment report of the intergovernmental panel on climate
change. Cambridge University Press, pp 1217–1308. https://doi.
org/10.1017/CBO9781107415324.028

Colin J, Déqué M, Radu R, Somot S (2010) Sensitivity study of heavy
precipitation in Limited Area Model climate simulations: influ-
ence of the size of the domain and the use of the spectral nudging

technique. Tellus A Dyn Meteorol Oceanogr 62(5):591–604.
https://doi.org/10.1111/j.1600-0870.2010.00467.x

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. Clim Dyn
52(5–6):2745–2764. https://doi.org/10.1007/s00382-018-4289-x

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi
S et al (2011) The ERA-Interim reanalysis: configuration and
performance of the data assimilation system. Q J R Meteorol Soc
137(656):553–597. https://doi.org/10.1002/qj.828

Doms G, Baldauf M (2011) A description of the nonhydrostatic
regional COSMO model. Part I. Dynamics and numerics.
Deutscher Wetterdienst, Offenbach

Doms G, Förstner J, Heise E, Herzog HJ, Mironov D, Raschendorf-
er M et al (2011) A description of the nonhydrostatic regional
COSMO model. Part II. Physical parameterization. Deutscher
Wetterdienst, Offenbach

Dousset B, Gourmelen F, Laaïdi K, Zeghnoun A, Giraudet E, Bretin
P, Mauri E, Vautard R (2011) Satellite monitoring of sum-
mer heat waves in the Paris metropolitan area. Int J Climatol
31(2):313–323. https://doi.org/10.1002/joc.2222

Duchêne F, Van Schaeybroeck B, Caluwaerts S, De Troch R, Hamdi
R, Termonia P (2020) A Statistical–dynamical methodology to
downslope regional climate projections to urban scale. J Appl
Meteorol Climatol 59(6):1109–1123. https://doi.org/10.1175/
JAMC-D-19-0104.1

Duh JD, Sandhas V, Chang H, George LA (2008) Rates of urbani-
sation and the resiliency of air and water quality. Sci Total
Environ 400(1–3):238–256. https://doi.org/10.1016/j.scitot.
env.2008.05.002

ECMW (2019) Changes in ECMWF model-evolution of the IFS.
https://www.ecmwf.int/en/forecasts/documentation-and-support/
changes-ecmwf-model

Ewing R, Rong F (2008) The impact of urban form on US residen-
tial energy use. Housing Policy Debate 19(1):1–30. https://doi.
or/10.1080/10511482.2008.9521624

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Tay-
lor KE (2016) Overview of the Coupled Model Intercomparison
Project Phase 6 (CMIP6) experimental design and organization.
Geoscience Model Dev 9(5):1937–1958. https://doi.org/10.5194/
gmd-8-10539-2015

Faroux S, Kaptué Tchuenté AT, Roujean JL, Masson V, Martin E,
Moigne PL (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. Geosci Model Dev 6(2):563–582. https://
doi.org/10.5194/gmd-6-563-2013

Fletcher TD, Andrieu H, Hamel P (2013) Understanding, manage-
ment and modelling of urban hydrology and its consequences
for receiving waters: a state of the art. Adv Water Resour
51:261–279. https://doi.org/10.1016/j.adwatre.2012.09.001

Foster SSD, Morris BL, Lawrence AR (1994). Effects of urbanization
on groundwater recharge. In: Groundwater problems in urban
areas: proceedings of the international conference organized
by the institution of civil engineers and held in London, 2–3
June 1993. Thomas Telford Publishing, pp 43–63. https://doi.
or/10.1680/gpua.19744.0005

Frey-Buness F, Heimann D, Sausen R (1995) A statistical-dynamical
downscaling procedure for global climate simulations. Theoret
Appl Climatol 50(3–4):117–131. https://doi.org/10.1007/BF008
66111

Gardes T, Schoetter R, Hidalgo J, Long N, Marquès E, Masson V
(2020) Statistical prediction of the nocturnal urban heat island
intensity based on urban morphology and geographical fac-
tors: An investigation based on numerical model results for a

 Springer
large ensemble of French cities. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2020.139253

Ghidagen L, Olsson J, Amorim JH, Asker C, Belusic D, Carvalho AC, Lindstedt D (2020) Towards climate services for European countries: lessons learnt from the Copernicus project Urban SIS. Urban Clim 31:100549

Giorgi F, Jones C, Asrar GR (2009) Addressing climate information needs at the regional level: the CORDEX framework. World Meteorol Organ Bull 58(3):175

Giorgi F, Coppola E, Solmon F, Mariotti L, Sylva MB, Bi X et al (2012) RegCM4: model description and preliminary tests over multiple CORDEX domains. Clim Res 52:7–29. https://doi.org/10.3354/cr01018

Gosling SN, Lowe JA, McGregor GR, Pelling M, Malamud BD (2009) Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. Clim Change 92(3–4):299–341. https://doi.org/10.1007/s10584-008-9441-x

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics. https://doi.org/10.2307/2528823

Graham LP, Andréasson J, Carlsson B (2007) Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods—a case study on the Lule River basin. Clim Change 81(1):293–307. https://doi.org/10.1007/s10584-006-9215-2

Gutiérrez C, Somot S, Nabat P, Corre L, van Meijgaard E et al (2012) Future evolution of surface solar radiation and photovoltaic potential in Europe: investigating the role of aerosols. Environ Res Lett 15(3):034035. https://doi.org/10.1088/1748-9326/ab6666

Hammond MJ, Chen AS, Jordjievski S, Butler D, Mark O (2015) Urban flood impact assessment: a state-of-the-art review. Urban Water J 12(1):14–29. https://doi.org/10.1080/1573062X.2013.857421

Hidalgo J, Jogla R (2018) On the use of local weather types classification to improve climate understanding: an application on the urban climate of Toulouse. PLoS One 13(12):e0208138. https://doi.org/10.1371/journal.pone.0208138

Hidalgo J, Masson V, Baehr C (2014) From daily climatic scenarios to atmospheric forcing fields to force soil–vegetation–atmosphere transfer models. Front Environ Sci 2:40. https://doi.org/10.3389/fenvs.2014.00040

Hoffmann P, Schoetter R, Schlinzien H (2018) Statistical–dynamical downscaling of the urban heat island in Hamburg, Germany. Meteorol Z 27:89–109. https://doi.org/10.1127/metzw/2016/0773

Hogrefe C, Lynn B, Civerolo K, Ku JY, Rosenthal J, Rosenzweig C et al (2014) Simulating changes in regional air pollution over the eastern United States due to changes in global and regional climate and emissions. J Geophys Res Atmos 109:D22. https://doi.org/10.1029/2009JD004690

Hope D, Gries C, Zhu W, Fagan WF, Redman CL, Grimm NB et al (2008) Socioeconomics drive urban plant diversity. Urban Ecol. Springer, Boston, pp 339–347. https://doi.org/10.1007/978-0-387-73412-5_21

IPCC (2018) Global Research and Action Agenda on Cities and Climate Change Science. Cities IPCC, Geneve

Jacob D, Elizalde A, Haensler A, Hagemann S, Kumar P, Podzun R et al (2012) Assessing the transferability of the regional climate model REMO to different coordinated regional climate downscaling experiment (CORDEX) regions. Atmospheric 3(1):181–190. https://doi.org/10.3390/atmos3010181

Jänicke B, Meier F, Fenners D, Fehrenbach U, Holtmann A, Scherer D (2017) Urban–rural differences in near-surface air temperature as resolved by the Central Europe Refined analysis (CER): sensitivity to planetary boundary layer schemes and urban canopy models. Int J Climatol 37(4):2063–2079. https://doi.org/10.1002/joc.4835

Jougla R, Hidalgo J, Pouponneau B (2019) Identification des situations météorologiques locales pour une cinquantaine de villes françaises. La Météorol. https://doi.org/10.4267/2042/70370

Kain JS, Fritsch JM (1990) A one-dimensional entraining/detraining plume model and its application in convective parameterization. J Atmos Sci 47(23):2784–2802

Karl TR, Nicholls N, Ghazi A (1999) Clivar/GCOS/WMO workshop on indices and indicators for climate extremes workshop summary. Weather and climate extremes. Springer, Dordrecht, pp 3–7

Kaufman L, Rousseauw P (1990) Partitioning around medoids (program pam). Find Groups Data Introd Cluster Anal 344:68–125

Kolokotroni M, Ren X, Davies M, Mavrogiani A (2012) London’s urban heat island: impact on current and future energy consumption in office buildings. Energy Build 47:302–311. https://doi.org/10.1016/j.enbuild.2011.12.019

Kovats RS, Hajat S (2008) Heat stress and public health: a critical review. Annu Rev Public Health 29:41–55. https://doi.org/10.1146/annurev.publ health.29.020907.090843

Künzli N, Kaiser R, Medina S, Studnicka M, Chanel O, Filliger P et al (2000) Public-health impact of outdoor and traffic-related air pollution: a European assessment. Lancet 356(9232):795–801. https://doi.org/10.1016/S0140-6736(00)02652-3

Kwok YT, Schoetter R, Lau KKL, Hidalgo J, Ren C, Pigeon G, Masson V (2019) How well does the local climate zone scheme discern the thermal environment of Toulouse (France)? An analysis using numerical simulation data. Int J Climatol 39(14):5292–5315. https://doi.org/10.1002/joc.6140

Lac C, Chaboureau P, Masson V, Pinty P, Tulet P, Escobar J et al (2018) Overview of the Meso-NH model version 5.4 and its applications. Geosci Model Dev 11:1929–1969. https://doi.org/10.5194/gmd-11-1929-2018

Landrigan PJ, Fuller R, Acosta NJ, Adey O, Arnold R, Baldé AB et al (2018) The Lancet Commission on pollution and health. Lancet 391(10119):462–512. https://doi.org/10.1016/S0140-6736(17)32345-0

Langendijk GS, 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. https://doi.org/10.3390/atmos10120730

Lauwaet D, Maiheu B, De Ridder K, Boënne W, Hooyberghs H, Demuzere M, Verdonck ML (2020) A new method to assess fine-scale contrasts under climate change: what does the EURO-CORDEX model ensemble tell us?—Investigating near surface humidity in Berlin and its surroundings. Atmosphere 10(12):730. https://doi.org/10.3390/atmos10120730

Leung LR, Gustafson WI Jr (2005) Potential regional climate change sensitivity to planetary boundary layer schemes and urban canopy models. Int J Climatol 37(4):2063–2079. https://doi.org/10.1002/joc.4835

Leutwyler D, Lüthi D, Ban N, Fuhrer O, Schär C (2017) Evaluation of the convection-resolving climate modeling approach on
simulations. Urban Clim 20:75–93. https://doi.org/10.1016/j.uclim.2017.03.002

United Nations, Department of Economic and Social Affairs, Population Division (2019) World Urbanization Prospects: the 2018 Revision (ST/ESA/SER.A/420). New York: United Nations

van Meijgaard E, van Ulft LH, van de Berg WJ, Bosveld FC, van den Hurk BJJM, Lenderink G, Siebesma AP (2008) The KNMI regional atmospheric climate model RACMO, version 2.1. KNMI, De Bilt, p 43

van Meijgaard E, van Ulft LH, Lenderink G, De Roode SR, Wipfler EL, Boers R, van Timmermans RMA (2012) Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe (No. KVR 054/12). KVR.

Vautard R, Menut L, Beekmann M, Chazette P, Flamant PH, Gombert D et al (2003) A synthesis of the Air Pollution Over the Paris Region (ESQUIF) field campaign. J Geophys Res Atmos 108:D17. https://doi.org/10.1029/2003JD003380

Vautard R, Gobiet A, Jacob D, Belda M, Colette A, Déqué M et al (2013) The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Clim Dyn 41(9–10):2555–2575. https://doi.org/10.1007/s00382-013-1714-z

Viguie V, Lemonsu A, Hallegatte S, Beaulant AL, Marchadier C, Mason V et al (2020) Early adaptation to heat waves and future reduction of air-conditioning energy use in Paris. Environ Res Lett. https://doi.org/10.1088/1748-9326/ab6a24

Walters D, Baran AJ, Boutle I, Brooks M, Earnshaw P, Edwards J et al (2019) The Met Office Unified Model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. Geosci Model Dev 12(5):1909–1963. https://doi.org/10.5194/gmd-12-1909-2019

Wilby RL (2007) A review of climate change impacts on the built environment. Built Environ 33(1):31–45. https://doi.org/10.2148/benv.33.1.31

Wilby RL, Wigley TM (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr 21(4):530–548. https://doi.org/10.1177/030913339702100403

Wu J (2014) Urban ecology and sustainability: the state-of-the-science and future directions. Landsc Urban Plann 125:209–221. https://doi.org/10.1016/j.landurbplan.2014.01.018

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.