A Statistical–Dynamical Methodology to Downscale Regional Climate Projections to Urban Scale

FRANÇOIS DUCHÊNE AND BERT VAN SCHAEYBROECK
Royal Meteorological Institute of Belgium, Brussels, Belgium

STEVEN CALUWAERTS
University of Ghent, Ghent, Belgium

ROZEMIEN DE TROCH AND RAFIQ HAMDI
Royal Meteorological Institute of Belgium, Brussels, Belgium

PIET TERMONIA
Royal Meteorological Institute of Belgium, Brussels, and University of Ghent, Ghent, Belgium

(Manuscript received 6 May 2019, in final form 17 April 2020)

ABSTRACT

The demand of city planners for quantitative information on the impact of climate change on the urban environment is increasing. However, such information is usually extracted from decadelong climate projections generated with global or regional climate models (RCMs). Because of their coarse resolution and unsuitable physical parameterization, however, their model output is not adequate to be used at city scale. A full dynamical downscaling to city level, on the other hand, is computationally too expensive for climatological time scales. A statistical–dynamical computationally inexpensive method is therefore proposed that approximates well the behavior of the full dynamical downscaling approach. The approach downscales RCM simulations using the combination of an RCM at high resolution (H-RES) and a land surface model (LSM). The method involves the setup of a database of *urban signatures* by running an H-RES RCM with and without urban parameterization for a relatively short period. Using an analog approach, these signatures are first selectively added to the long-term RCM data, which are then used as forcing for an LSM using an urban parameterization in a stand-alone mode. A comparison with a full dynamical downscaling approach is presented for the city of Brussels, Belgium, for 30 summers with the combined ALADIN–AROME model (ALARO-0) coupled to the Surface Externalisée model (SURFEX) as H-RES RCM and SURFEX as LSM. The average bias of the nocturnal urban heat island during heat waves is vanishingly small, and the RMSE is strongly reduced. Not only is the statistical–dynamical approach able to correct the heat-wave number and intensities, it can also improve intervariable correlations and multivariate and temporally correlated indices, such as Humidex.

1. Introduction

The global mean surface air temperature has been increasing since the preindustrial period, and this increase is projected to continue throughout the twenty-first century (Masson-Delmotte et al. 2018). People living in cities are more exposed to heat as cities feature climate conditions with temperatures that are typically higher than the surrounding rural areas, a phenomenon called the urban heat island (UHI; Oke et al. 2017; Bader et al. 2018). Relative to rural inhabitants, city residents are therefore subject to stronger heat stress during hot events (Li and Bou-Zeid 2013; Lemonsu et al. 2015; Hamdi et al. 2016; Founda and Santamouris 2017; Chapman et al. 2017; Wouters et al. 2017). With almost

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAMC-D-19-0104.s1.

Corresponding author: François Duchêne, francois.duchene@meteo.be

DOI: 10.1175/JAMC-D-19-0104.1

50% of the global population currently living in an urban environment and as this fraction is projected to increase [United Nations Human Settlements Programme (UN-Habitat) 2010], it is relevant to provide improved climate projections at the city scale (Rosenzweig et al. 2015; Berckmans et al. 2019).

Several exceptionally strong heat-wave (HW) events occurred over Europe in the most recent decades (Gabriel and Endlicher 2011). For instance, more than 70 000 deaths were attributed to the European heat wave of summer 2003 (Robine et al. 2008). HWs are projected to become longer, more intense, and more likely to occur with global warming (Jacob et al. 2014; Guerreiro et al. 2018) and are expected to have a huge impact on human lives (e.g., Forzieri et al. 2017). In Antwerp (Belgium), Martinez et al. (2018) showed an expected increase in mortality by a factor of 1.7 and 4.5 attributable to heat in the near and far future, respectively. Lauwaet et al. (2016) found that the number of future HW days in Belgian cities will increase 2 times faster than in the rural surroundings, even if the UHI is expected to decrease slightly in Brussels upon increase of incoming longwave radiation since the larger storage capacity of buildings in urban areas buffer the increase of longwave radiation in contrast to rural areas that will warm more (Oleson et al. 2011; Hamdi et al. 2014a).

The common approach to estimate the uncertainty of long-term (i.e., more than 30 years) climate projections is by use of an ensemble of climate simulations. The global circulation models participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) provide such an ensemble at a horizontal resolution of 100–200 km, and therefore they do not capture regional, let alone local, effects. Regional ensemble information is mainly generated by the approach of dynamic downscaling (see Fig. 1) using regional climate models (RCMs). For instance, for Europe, the Coordinated Regional Climate Downscaling Experiment (CORDEX; Giorgi et al. 2009) provides an ensemble of climate projections with 12.5-km resolution. However, even at that extent over Europe, interactions between atmosphere and city (Oke et al. 2017) are not simulated, and, moreover, the RCM models typically lack an urban scheme (Hamdi et al. 2014a; Daniel et al. 2019). The urban–atmosphere interactions can be captured by including an urban parameterization (UP) into an RCM at high resolution (H-RES), that is, below 5 km (Masson 2000; Martilli et al. 2002; Hamdi and Masson 2008). Further downscaling to urban ranges can then be done using a land surface model (LSM) with an urban parameterization in a stand-alone mode or using a microscale model (MM) that can simulate processes up to meter-scale resolutions (Moonen et al. 2012).

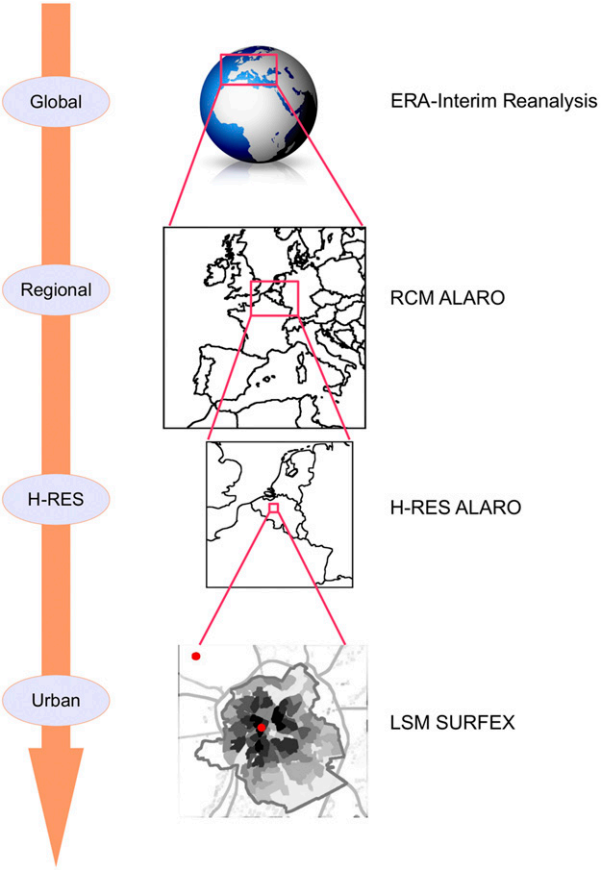

FIG. 1. Schematic representation of the dynamical downscaling approach that allows one to go from the global over the regional and the high-resolution (H-RES) down to the urban scale. The approach shown is the one used for Brussels in this work. The global reanalysis ERA-Interim dataset was downscaled over western Europe using the ALARO RCM at 20-km resolution. Further downscaling over Belgium was done with ALARO coupled with SURFEX (4-km resolution). In this step, three types of simulations are conducted: the urban parameterization has or has not been activated for two simulations, respectively, and our method has been applied on a third one. Last, a domain covering Brussels is simulated with SURFEX offline (1-km resolution). Dynamic downscaling for the three first steps involved sequential use of the meteorological conditions as lateral boundary conditions, and the last step used the H-RES output as upper-air forcing. The two red dots over Brussels indicate the locations of the urban and rural stations (Molenbeek and Brussegem).

Three types of methodologies are typically used to downscale to the urban level: dynamic downscaling (DD), statistical downscaling, and statistical–dynamical downscaling (SDD). These approaches are tabulated in Table 1 along with their fulfillment of three key criteria. These criteria describe whether the downscaling approach 1) is able to adequately simulate physical interactions between the city and the atmosphere, 2) provides time series of different variables that are

TABLE 1. Different downscaling approaches/methods and their fulfillment of three criteria. The checkmark indicates that a feature is satisfied, and an X indicates that it is not. The rightmost approach (new SDD) refers to the method presented in the current work. This approach aims to approximate the full dynamical downscaling approach.

Criterion	Full DD: RCM → H-RES UP → LSM/MM	DD: RCM → LSM	Statistical downscaling	SDD	New SDD: RCM → H-RES calibration → LSM
Adequate city–atmosphere interactions	✓	X	X	✓	✓
Output physically and spatiotemporally consistent	✓	✓	X	X	✓
Computationally manageable	X	✓	✓	✓	✓

physically consistent both in space and time and among variables, and 3) is currently computationally manageable for climatological ensemble purposes.

In principle, the best downscaling approach from regional to city scale is the full dynamical downscaling (*full DD: RCM → H-RES UP → LSM/MM*) that satisfies criteria 1 and 2, but it is computationally too expensive and therefore fails criterion 3. Although coordinated efforts (e.g., at the European level; Hewitt and Lowe 2018) are currently exploring the H-RES scales, such simulations are exceedingly computationally expensive and largely lack urban parameterizations (Hamdi et al. 2014a; Termonia et al. 2018). An inexpensive alternative dynamic downscaling approach (denoted as *RCM → LSM*) involves the direct use of RCM output as input forcing for a stand-alone and computationally inexpensive LSM at city scale. Yet, the lack of well-resolved atmosphere–city feedback processes in the forcing was shown to lead to unrealistic simulation of the UHI (Hamdi et al. 2014a). An intermediate downscaling step to H-RES is therefore necessary that adequately reproduces the urban–atmosphere interactions below the range of the RCM model (Hamdi et al. 2014a; Tsiringakis et al. 2019). Therefore, the RCM → LSM approach fails to satisfy criterion 1 (see Table 1).

Statistical downscaling methods use statistical relationships between observed large-scale weather variables and local-scale variables. They are computationally very inexpensive but are typically tuned for a few city locations and a single variable (Kershaw et al. 2010; Hoffmann et al. 2012; Gutiérrez et al. 2013; Le Bras and Masson 2015; Liu et al. 2018; Wilby 2008). Therefore, they generally fail to capture complex interactions (criterion 1; Table 1) and give physically coherent results for different meteorological variables (criterion 2).

Recently, SDD techniques have emerged that aim to combine the benefits of statistical and dynamical downscaling (Hoffmann et al. 2018). Typically, these approaches first produce detailed simulations at city level for a relatively short period. Then, using relations between large-scale RCM and local-scale information, these detailed results are statistically recombined in order to extract the information required. For instance, Hoffmann et al. (2018) used a weather-type classification based on cluster analysis to study the evolution of UHI with climate change. The UHI associated with each weather type was obtained by dynamically H-RES downscaling to 1-km resolution for a few years. The authors then performed a cluster-weighted average over the 1-km UHI pattern associated with a long-term RCM simulation where the weights were taken proportional to the weather-type frequencies of the RCM. Another urban SDD method is the cuboid method (Früh et al. 2011; Žuvela-Aloise et al. 2014; Bokwa et al. 2018). This method considers a few covariates (typically temperature, humidity, and wind) to fully characterize a particular variable of interest such as daily maximum or minimum temperature. For each prevailing wind direction, a limited set of (typically eight) idealized computationally demanding microscale simulations is performed by combining all combinations of the maximum and minimum boundaries of all covariates. The downscaling of RCM variables is then done by multilinear interpolation of the microscale simulation results, weighted by RCM values of the covariates.

The SDD methods described so far are relatively computationally inexpensive (criterion 3; Table 1), and their city-scale simulations capture city–atmosphere interactions (criterion 1). Although the approaches are useful for extracting climate averages, important issues exist when they are used to reconstruct long climate time series. Indeed, upon recombining the limited set of city-scale simulations, correlations of climate variables will generally not be physically consistent. Such correlations are, however, very important to characterize heat waves. Moreover, the linear interpolation used by the cuboid method may lead to physically unrealistic results because of the nonlinearities. In sum, existing SDD methods fail to satisfy criterion 2 (see Table 1).

Rather than the calculation of a sole variable such as the UHI, the aim here is to propose a downscaling approach that yields physically consistent time series on climate time scales (i.e., 30 years) for multiple variables. Such time series could then be used as input for further

downscaling, impact modeling, or the calculation of indices that require adequate correlations in time (e.g., HWs, which are by definition very sensitive to the crossing of the daily temperature thresholds) or intervariable correlations [e.g., human comfort indices such as "Humidex" or the universal thermal climate index (UTCI), the definitions of which are based on several variables, which should then be calibrated in the same way]. The main motivation for this work is therefore to propose an SDD method that reproduces the results of the full downscaling approach but using limited computational resources. This would allow the translation of an ensemble of regional climate projections (e.g., from CORDEX) from its regional down to city-level scale, thereby propagating the model and scenario uncertainties.

Statistical downscaling approaches (Maraun and Widmann 2018; Wilby 2008; Hoffmann et al. 2012) and SDD methods (Hoffmann et al. 2018; Früh et al. 2011; Žuvela-Aloise et al. 2014; Bokwa et al. 2018) act on the final model output of RCMs. The calibration is done on the basis of a limited set of H-RES RCM experiments with urban parameterization that allows one to construct a database of so-called *urban signatures*, as shown in Fig. 2. Once this database is built one can skip the computationally demanding H-RES downscaling step as in Fig. 1 by applying the signature on the RCMs. The additional downscaling step with the LSM then guarantees the physical and spatiotemporal coherence of the output (criterion 2). A proof of concept is provided here by comparing to what extent our method is able to reproduce the results of a full downscaling approach over Brussels, Belgium. An evaluation of the full downscaling approach against observations was presented in Hamdi et al. (2015, 2016).

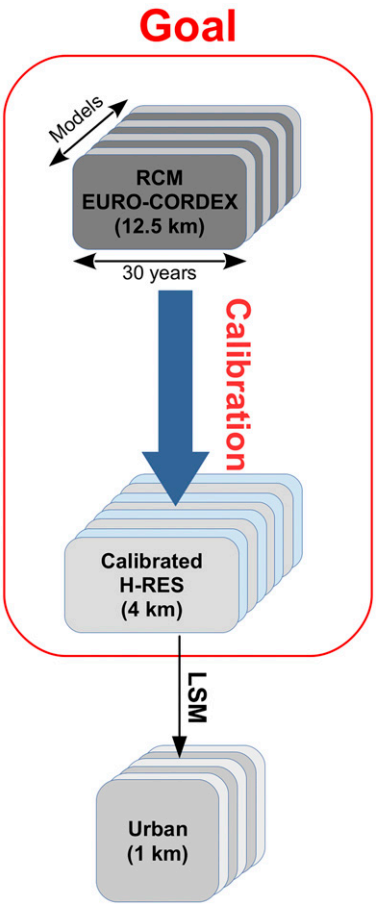

FIG. 2. The overall goal of the statistical–dynamical downscaling method is to translate an ensemble of RCM results (e.g., from EURO-CORDEX at 12.5 km) to city-level scale. Instead of using a computationally expensive dynamic downscaling step using an H-RES RCM (see Fig. 1), a statistical calibration step is used for which the method is described in Fig. 3, below. Next, the calibrated H-RES output is used as forcing for an LSM in stand-alone mode. For our evaluation approach, the indicated resolutions are different from the ones indicated here (see Table 2).

2. Data and methods

a. Calibration method

The ultimate goal of this method is to be able to downscale a large set of RCM climate projections (e.g., EURO-CORDEX at 12.5 km) to city level. This is done by first calibrating RCM output to H-RES scale by adding the spatial signature of the urban–atmosphere interactions for all meteorological variables (Fig. 2). The calibrated data are then used as forcing for an LSM in stand-alone mode at 1-km resolution. This paper focuses on the evaluation of the proposed method, and the application of the method using the EURO-CORDEX ensemble RCM of future climate projection (up to 2100) will be presented in a future work.

Figure 3 outlines the different steps of the proposed calibration method. The first or *training* step of the calibration consists of creating an H-RES training database containing a few years (four in our proof-of-concept study) of H-RES climate simulations with no urban parameterization (H-RES NO-UP) and a training base with urban parameterization (H-RES UP), both covering the same period. For each day d of this period, the signature $\Delta(d)$ is defined as the difference between the corresponding variables with activated (UP) and deactivated (NO-UP) urban parameterizations. More specifically, for variable i, location r, and hour of the day t:

$$\text{urban signature} \equiv \Delta_{i,r,t}(d) = X_{i,r,t}^{\text{H-RES,UP}}(d) - X_{i,r,t}^{\text{H-RES,NO-UP}}(d). \quad (1)$$

Here $X^{\text{H-RES,UP}}$ and $X^{\text{H-RES,NO-UP}}$ represent the variables from the H-RES RCM training dataset with and

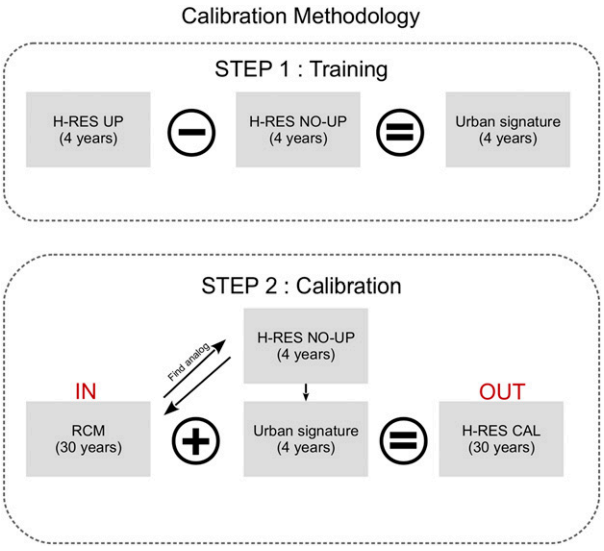

FIG. 3. Outline of the two main steps of the calibration method. The first or training step consists of creating the urban signature database on the basis of a limited number of years simulated using an H-RES RCM with activated urban parameterization (H-RES UP) and deactivated urban parameterization (H-RES NO-UP). The urban signature is obtained by subtracting H-RES NO-UP from H-RES UP [see Eq. (1)]. For each day of the RCM time series, the second or calibration step consists of finding the most analogous day within the H-RES NO-UP dataset [minimizing Eq. (2)] and adding the corresponding urban signature to the RCM results [see Eq. (3)]. The final output consists of a climatological H-RES time series that can then be used for further downscaling to urban scales (see Fig. 2).

without urban parameterization, respectively. The term urban signature is reminiscent of the urban fingerprint used in Giovannini et al. (2011), although, in our quantitative approach it only captures urban and not orographic effects.

Once constructed, the urban signature database serves to correct the RCM outputs in the second step of the calibration (see Fig. 3). The correction of the RCM result of a given day d_{RCM} starts by finding the most similar day in the training set (from the H-RES NO-UP). To select this day, all variables are first standardized by subtracting their mean and dividing by their standard deviation (for each location separately). The day d_{min} that is most analogous to a day d_{RCM} is obtained by minimizing the following cost function J with respect to d:

$$J(d) = \sum_{r}\sum_{t}\sum_{i}\left[\overline{X}^{\mathrm{RCM}}_{i,r,t}(d_{\mathrm{RCM}}) - \overline{X}^{\mathrm{H-RES,NO-UP}}_{i,r,t}(d)\right]^2. \quad (2)$$

Here $\overline{X}^{\mathrm{RCM}}_{i,r,t}$ and $\overline{X}^{\mathrm{H-RES,NO-UP}}_{i,r,t}$ are the standardized RCM and H-RES NO-UP variables, respectively. Both generally feature distinct spatial resolutions (RCM and H-RES). Simple interpolation schemes (bilinear or nearest neighbor as used in this study), potentially including simple orographic corrections, are therefore suggested to downscale the RCM contribution to the H-RES resolution prior to subtraction.

TABLE 2. Model resolutions used in the downscaling approaches for the regional climate model, high-resolution RCM, and land surface model. For the evaluation, both RCM and H-RES have a resolution of 4 km because the RCM is in fact the H-RES without urban parameterization in this step, which is different from the RCM in Fig. 1. The RCM/H-RES model used in the current work is ALARO, and the LSM is SURFEX. The goal is to use the presented method in a future work to downscale the EURO-CORDEX simulations (different models) that have a resolution of 12.5 km and to use ALARO at 4 km and SURFEX at 1 km to downscale.

	RCM	H-RES calibration	LSM
Evaluation	4 km	4 km	1 km
EURO-CORDEX downscaling	12.5 km	4 km	1 km

Last, after finding day d_{min} of the training dataset that is the best analog day of d_{RCM}, the latter is calibrated by adding the urban signature to all forcing variables i, all spatial points r, and all times of the day t:

$$\begin{aligned}\mathrm{Calibration} &\equiv X^{\mathrm{Cor-RCM}}_{i,r,t}(d_{\mathrm{RCM}})\\ &= X^{\mathrm{RCM}}_{i,r,t}(d_{\mathrm{RCM}}) + \Delta_{i,r,t}(d_{\mathrm{min}}). \quad (3)\end{aligned}$$

Again, simple interpolation schemes are recommended to convert the RCM to the H-RES resolution. The calibrated result $X^{\mathrm{Cor-RCM}}(d_{\mathrm{RCM}})$ is used as the forcing of the LSM (see Fig. 2).

b. Model evaluation setup

The proposed SDD method aims at reproducing the results of a full dynamical downscaling effort (full DD: RCM → H-RES UP → LSM) but at a limited computational cost. The evaluation effort therefore compares the proposed downscaling scheme with a full dynamic downscaling. Dynamic downscaling is done for the city of Brussels using the reanalysis ERA-Interim as boundary forcing of the combined ALADIN–AROME RCM (ALARO-0; De Troch et al. 2013), coupled with the surface scheme Surface Externalisée (SURFEX; Masson et al. 2013) as described further in section 2c. To investigate the physical and spatiotemporal consistency of the output (criterion 2 in Table 1), we focus on HWs, the UHI, and their relation but also consider other variables. The period considered is 1981–2010 and covers June, July, August, and September, which is henceforth simply denoted as summer.

As tabulated in Table 2, for the urban, H-RES, and RCM resolutions we use 1, 4, and 4 km, respectively.

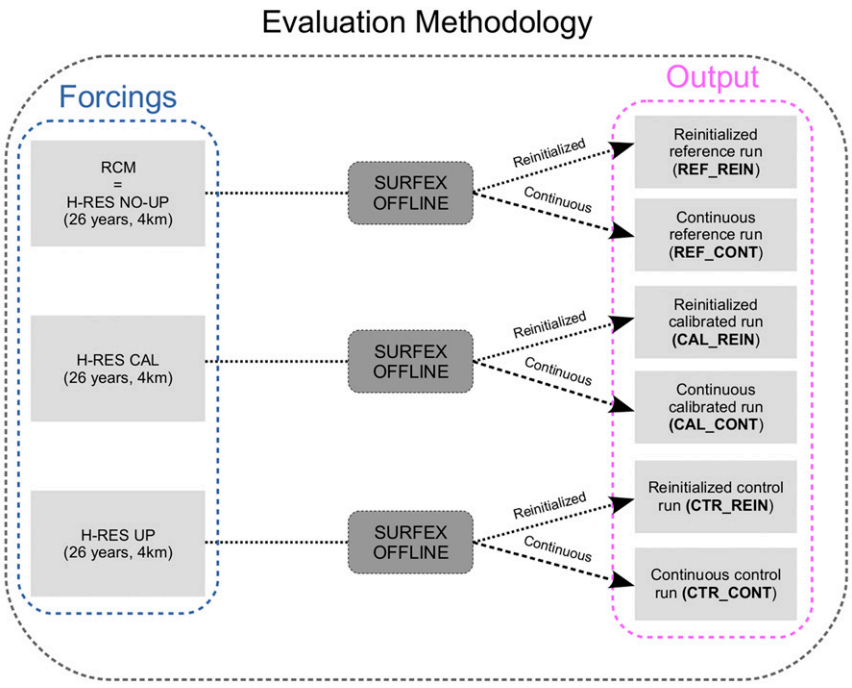

FIG. 4. Schematic representation of the different datasets used for our evaluation method. The leftmost column corresponds to the forcing input for the stand-alone or offline simulations with the LSM SURFEX to downscale to 1-km resolution over Brussels. The rightmost column corresponds to the results used in our evaluation approach. The forcing input includes the data from the RCM, the high-resolution calibrated (H-RES CAL) model generated using the method of Fig. 3, and a high-resolution one generated with an RCM that includes activated urban parameterization (H-RES UP). The output includes a reference run that is equivalent to the RCM → LSM approach from Table 1 and a control one that is equivalent to the full dynamical downscaling approach in Table 1. Both daily reinitialized soil variables (REIN) and once-initialized continuous (CONT) SURFEX simulations are generated. Note that in our evaluation method the RCM is the same as the H-RES NO-UP.

Here, the evaluation of the method is applied on a 4-km resolution simulation without urban parameterization. The rationale behind this is that it allows establishing a competitive reference method (REF) against which our SDD method can be tested in a perfect setup configuration. A comparison against lower-resolution model outputs (e.g., 12.5 km from CORDEX) would be unfair as added value might simply be caused by their lack of spatial detail. The control (CTR) dataset is taken here as the 1-km LSM downscaling forced by the 4-km run with activated urban parameterization (H-RES UP) and validated against observations in Hamdi et al. (2015, 2016).

The training years consist of four summers (2003, 2005, 2006, and 2010) taken out of the last decade, all considered as the warmest years (Masson-Delmotte et al. 2018) during which different HWs occurred over Brussels (Hamdi et al. 2016). Four years allow us to have sufficient climatological representation while at the same time keeping the computing time low. The training years were taken out from the evaluation set.

Figure 4 summarizes the 26-yr datasets that are used for evaluation, while Table S1 in the online supplemental material summarizes the naming of the experiments. The REF data, equivalent to the RCM → LSM approach in Table 1, are obtained by direct input of the RCM outputs (here also equivalent to H-RES NO-UP) as forcing for the LSM. The calibrated (CAL) one uses the H-RES calibrated results as LSM input, and, finally, the CTR run or full dynamic downscaling approach is equivalent to the RCM → H-RES UP → LSM approach in Table 1. For each forcing, two LSM simulations are performed with different initialization methods. More specifically, the soil parameters are either reinitialized each day (REIN) (Best and Grimmond 2014; Berckmans et al. 2017) or initialized once and evolved continuously thereafter (CONT). Soil reinitializing of the H-RES CAL is done by using the soil parameters of the H-RES UP setup. By comparing a run with daily reinitialized soil (prognostic) variables to the one with continuous soil variables, one can estimate the impact of reinitialization

on the urban signature. Note that a full daily reinitialization of the reinitialized reference run (REF_REIN) is not possible for all variables. This is due to absence of the urban-related surface variables (temperature of the road, building, etc.) in the forcing dataset. Therefore, these variables are initialized here with their summer-averaged value.

c. Dynamic downscaling approach

Here the utilized model setup is described, and we refer the reader to Hamdi et al. (2015) for more details. The RCM used is the numerical forecast model ALARO-0 coupled to the LSM SURFEX, version 5 (v5; Hamdi et al. 2014a,b). The ALARO model is designed to run at convection-permitting resolutions (Gerard et al. 2009; De Troch et al. 2013). The externalized land and ocean surface platform SURFEX (Masson et al. 2013) can be used either in a coupled mode with ALARO, by exchanging surface fluxes and atmospheric forcing at every time step, or in offline mode where the atmospheric drivers are derived from model output and fed to the LSM. In such a case, there is no feedback between the surface and the upper part of the atmosphere. However, SURFEX is coupled to a surface boundary layer (described in Hamdi and Masson 2008) and has several air layers between the soil and the forcing level. Large-scale forcing, turbulence, drag, and canopy forces are therefore taken into account. Note that a higher degree of surface–atmosphere coupling is established by use of 3D microscale models with levels beyond the surface boundary layer. Each grid box is subdivided into four tiles: nature, urban, sea (or ocean), and lakes. Each of these tiles has its own parameterization but lacks horizontal detail. In SURFEX, the Town Energy Balance (TEB) single-layer urban canopy model (Masson 2000) is used for the urban tile. TEB is based on the canyon approach where a building block is represented as a roof, two facing walls, and a street separating them. This approach simplifies the radiation exchanges to keep the computing time low. The land coverage is based on the global ECOCLIMAP database (Masson et al. 2003). Energy, water, and fluxes are computed with each parameterization scheme and aggregated at the grid scale weighted by the tile fractions following the global database. There are 16 land-cover classes in ECOCLIMAP among which 8 are used to describe urban areas. For Brussels, ''dense urban'' and ''temperate suburban'' are the dominant classes. Around Brussels, the vegetation tiles are mainly crops and temperate forest. Different properties (including radiative and thermal properties, albedo, emissivity, leaf area index, and geometry of buildings) are kept constant throughout the simulation and can be found in Masson et al. (2003).

The dynamic downscaling is composed of three steps (see Fig. 1 for the exact domains): (i) A 20-km ALARO simulation over western Europe forced at the boundaries by ERA-Interim (Dee et al. 2011), (ii) another 4-km ALARO one over Belgium forced at the lateral boundaries by the 20-km simulation, (iii) a 1-km SURFEX offline run over a 30-km by 30-km domain covering Brussels, forced at 17 m above ground using the 4-km output. As in the experimental setup of Hamdi et al. (2015), SURFEX offline is coupled to a surface boundary layer scheme following the method in Hamdi and Masson (2008) and Masson and Seity (2009). The 1-km CAL run is forced by calibrated [see Eq. (3)] 4-km data for all SURFEX hourly forcing variables (temperature, humidity, wind, precipitation, snow, pressure, and the downward radiative fluxes).

d. Definition of evaluation indices

Our model comparison focuses on multiple variables relevant for most urban impact studies. The UHI for Brussels is defined here as the temperature difference between Molenbeek in the city center and the rural Brussegem, situated 13 km northwest from the city center (see red dots in Fig. 1). The nocturnal and daytime UHI, denoted UHI_N and UHI_D, are the differences of daily minimum and maximum temperatures, respectively.

Multiple HW definitions exist in the literature, but here the definition of the Belgian public health authorities is used. According to this definition, an HW is a period of at least three consecutive days with average daily minimum and maximum temperatures that exceed 18° and 30°C, respectively. This evaluation of HW days is performed for each location of interest and for all presented experiments. The nocturnal HW intensity (MIN_I) is defined as the sum of the $T_{\min}$ exceedances above 18°C during an HW:

$$\mathrm{MIN\_I} = \sum_d (T_{\min,d} - 18°\mathrm{C})\theta(T_{\min,d} - 18°\mathrm{C}). \quad (4)$$

Here θ is the Heaviside or unit step function (Abramowitz and Stegun 1972). The daytime HW intensity is defined similarly with the threshold set at 30°C:

$$\mathrm{MAX\_I} = \sum_d (T_{\max,d} - 30°\mathrm{C})\theta(T_{\max,d} - 30°\mathrm{C}). \quad (5)$$

There are several commonly used heat-stress indices that account for the effect of temperature as well as additional environmental, physiological, and behavioral variables such as humidity, air movement, solar radiation,

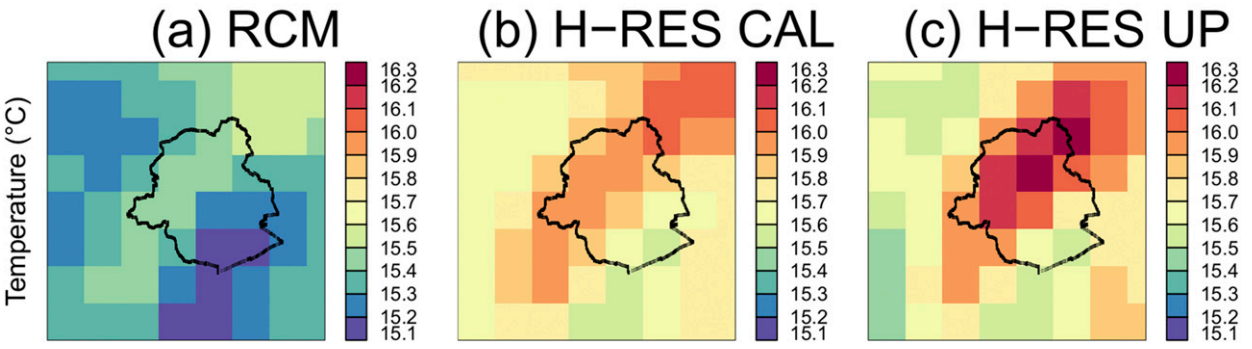

FIG. 5. The average temperature used as forcing for the LSM as tabulated in the leftmost column in Fig. 4. This includes the temperature from (a) the RCM, (b) the H-RES CAL model generated using the method of Fig. 3, and (c) the high-resolution one generated with H-RES UP. Averages are taken over 26 summers of the evaluation period for the three 4-km experiments. The H-RES calibration data approximate well those of H-RES UP.

metabolic rate, and age (Budd 2008). Here we use Humidex, which combines both temperature and vapor pressure in a formula developed by Masterson and Richardson (1979):

$$\text{Humidex} = T + \frac{5°\text{C}}{9\,\text{hPa}}(e - 10\,\text{hPa}), \qquad (6)$$

where T is the air temperature (°C) and e is the vapor pressure (hPa) calculated from T and relative humidity. The relation between Humidex and different categories of human comfort, ranging from comfortable to very dangerous, is tabulated in Table S2 in the online supplemental material.

e. *Review of model evaluation*

An extensive validation of our CTR run against observations was performed in Hamdi et al. (2015, 2016), the most relevant results of which can be summarized as follows:

- For both minimum T_{min} and maximum T_{max} temperature, the model compares well with the observations. The mean summer bias for the rural reference station (Brussegem) is 0.5° and −0.6°C for T_{min} and T_{max}, respectively. For the urban station (Molenbeek), the mean summer bias is 0.0° and −0.8°C for T_{min} and T_{max}, respectively.
- During summer, the model overestimates by 0.3 and 0.2 K the nocturnal and daytime UHI, respectively. During HW events, there is an underestimation by 0.9 K for both the nocturnal and daytime UHI. However, the higher UHI phenomenon during HW events is still simulated by the model. The root-mean-square error (RMSE) for the nocturnal UHI for the summer is 1.7 K, whereas it is negligible for the daytime. A PDF of the observed and simulated temperature and UHI is also shown on Fig S4 of the online supplemental material and will be discussed in more detail in the results section. The agreement between the observations and the model is good, with the maximum of the probability density that peaks at 29% and 56%, respectively.
- The mean observed summer precipitation is well simulated with only a slight underestimation of 0.5 mm day^{-1} in August.
- The model reproduces well the difference in relative humidity (RH) between the urban and rural environments but underestimates RH during the warm season by approximately 10% at a suburban station in Uccle.
- For wind speed and cloud cover, the model satisfactorily reproduces the observations during the summer.

The comparison of the results of our calibration method and the reference simulation with the observations will be performed in the next section.

3. Results

a. *Univariate evaluation of the calibration methodology*

The results presented henceforth concern the REIN simulations, and the ones of the continuous simulations will be presented in section 3b. The results of this section assess to what extent the calibration is able to reproduce the results of the full dynamic downscaling approach (CTR_REIN). The results focus on air temperature and UHI, but results for wind and relative humidity are also provided.

Prior to the comparison of the 1-km results, the (H-RES) urban signature is briefly considered. Figure 5 shows the average summer temperature of the three 4-km forcing [H-RES UP, RCM (=H-RES NO-UP), and H-RES CAL; see Fig. 4]. Although of coarser resolution than the LSM output of 1 km, it is clear that the calibrated H-RES result (H-RES CAL) approximates well the one

TABLE 3. Mean, bias, and RMSE of the nocturnal UHI [UHI_N (K)] during summer and heat-wave events for the reference (REF), calibrated (CAL), reinitialized (REIN), and continuous (CONT) runs. The difference is calculated between Molenbeek in the city center and Brussegem, a rural location outside Brussels (see Fig. 1). Note that the bias and RMSE are evaluated with respect to the control run (CTR_REIN or CTR_CONT), and all are averages over the 26-yr evaluation period.

	REF_REIN UHI_N		CAL_REIN UHI_N		REF_CONT UHI_N		CAL_CONT UHI_N	
	Summer	HW	Summer	HW	Summer	HW	Summer	HW
Mean (K)	1.0	0.6	1.6	2.2	0.7	0.6	0.9	0.9
Bias (K)	−0.5	−1.5	0.0	0.0	0.4	0.6	0.2	0.3
RMSE (K)	0.5	2.2	0.1	0.5	0.4	1.0	0.3	0.9

of the H-RES run with urban parameterization (H-RES UP). The forcing temperature at 17 m in the city center (Molenbeek) is 15.4°, 16.0°, and 16.2°C for the RCM (=H-RES NO-UP), H-RES CAL, and H-RES UP, respectively. The 17-m forcing corresponds to the lowest level of the ALARO model.

Table 3 tabulates the average nocturnal UHI during summer and during HW events for the reference (REF_REIN) and calibrated (CAL_REIN) runs and their bias and RMSE with respect to the control run (CTR_REIN). The calibration method is able to eliminate the large negative bias of nocturnal UHI (−0.5 K) of the summer UHI of the reference run (REF_REIN). The observed phenomenon of an augmented UHI during HWs (Zhou and Shepherd 2010; for Brussels, see Hamdi et al. 2016) is not reproduced by the reference run, thereby giving rise to a negative nocturnal UHI bias of −1.5 K (third column of Table 3). This failure of the reference run can be traced back to the initialization at constant value of the prognostic urban variables (see end of section 2b). However, it confirms the fact that the REF_REIN run, obtained by directly coupling RCM model results to an LSM, amounts to unrealistic results and requires calibration. Again, our method provides an almost perfect elimination of bias during HWs. To the same degree, the daytime UHI (based on daily maximum temperatures) is fully calibrated, both during the full summer period and during HWs events (see Table S3 in the online supplemental material).

The evaluation results presented so far are obtained by comparing with the control run. A full spatial and subdaily evaluation against observations is out of the scope of this study. However, the observed summer-averaged nocturnal UHI of 1.8 K compares well with the one of the calibrated run, which is 1.6 K (Hamdi et al. 2016). During HWs, there is an underestimation of around 1 K. The observed values of the daytime UHI (0.8 K) approximate well the ones of the control and calibrated model (0.9 K).[1] The RMSE for the reference run (2.1 K) is reduced by calibration to the level of the control run (RMSE = 1.7 K), implying a near-perfect RMSE correction. For the daytime UHI, existing differences between control and reference runs are minor (0.1 K). Supplemental Fig. S4 also compares the observed and modeled probability densities of the minimal temperatures at the rural and urban stations and their associated urban heat islands. It is shown that there is a strong agreement of the probability densities between the observations and the full dynamical simulation but a weaker agreement with the reference run. Moreover, the calibration methodology improves the probability distribution of the reference run. The peak goes from 72% for the REF_REIN down to 52% for the CAL_REIN, very close to 56% of the CTR_REIN compared to the observations' peak at 29%. The calibration does improve the temperature at the urban site but has limited effect at the rural station, as expected.

Apart from the overall elimination of the bias, the calibration scheme also strongly reduces the RMSE of the reference run (REF_REIN; see Table 3), both for the entire summer (80% reductions) and during the HW events (77% reduction). The UHI used so far is derived from daily minimal or maximal temperatures and at one rural and one urban station only. However, our calibration method also improves the UHI at subdaily time scales as clearly shown in Fig. 6 and for the entire domain over Brussels as shown in Fig. 7. The southwest–northeast orientation of the UHI pattern seen in Fig. 7 can be attributed to a comparable pattern of urbanization (see, e.g., Fig. 1 in Hamdi et al. 2014a) and to heat advection (Hamdi et al. 2015).

The spatial patterns of daily minimum and maximum wind speed and relative humidity are shown in Figs. S2 and S3 of the online supplemental material. Calibration provides clear improvements for each variable considered. For minimum wind speed, the results for the reference, control, and calibration runs are all similar. Note that our method does not aim to add value for precipitation at the urban resolution because the final step of our downscaling method involves the use of an LSM and therefore does not resolve upper-air processes.

[1] The observed values reported here are averages over the extended-summer periods of the 26-yr evaluation period and therefore slightly differ from those in Hamdi et al. (2016).

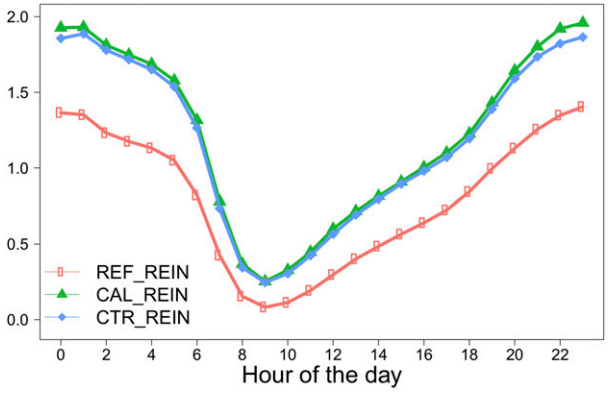

FIG. 6. Diurnal cycle of the UHI (K) averaged over the 26 summers of the evaluation period for the three reinitialized 1-km experiments. These include the reference run (REF_REIN; red), the calibration run (CAL_REIN; green), and the control one (CTR_REIN; blue) as described in Fig. 4. The calibration method clearly improves the diurnal cycle of UHI with respect to the CTR as compared with the REF.

b. *Calibration of heat-wave features*

Apart from an overall improvement, the calibration method is also able to improve the results during HW days (see Table 3). This suggests that the calibration induces more than a mere shift of the average UHI by reestablishing the temporal correlations. This is indeed confirmed in Fig. 8, which shows the average nocturnal UHI per summer month. There is a clear increase of the UHI of the control run during HWs (Fig. 8b) as compared with the summer average (Fig. 8a). As opposed to the reference run that underestimates the HW intensity by a factor of 4 in July and August, the calibration realistically reproduces this effect. The clear improvements from calibration of the monthly averaged are caused by the reproduction of the individual HWs (not shown). The HW characteristics in the urban location are summarized in Fig. 9. The nocturnal HW intensity [MIN_I; see Eq. (4)] is strongly improved to a value that differs only 3% from the control result (CTR_REIN).

c. *Calibration of multivariate variables of heat comfort*

Apart from reestablishing climatological averaged values, one may ask whether the calibration method is capable of reestablishing correlations among different variables. These are relevant, for instance, when using the results for calculating heat-comfort indices, for use in impact models or for further downscaling.

Figure S1 in the online supplemental material shows the diurnal cycle of the correlation between temperature and relative humidity in the city center. The anticorrelations in

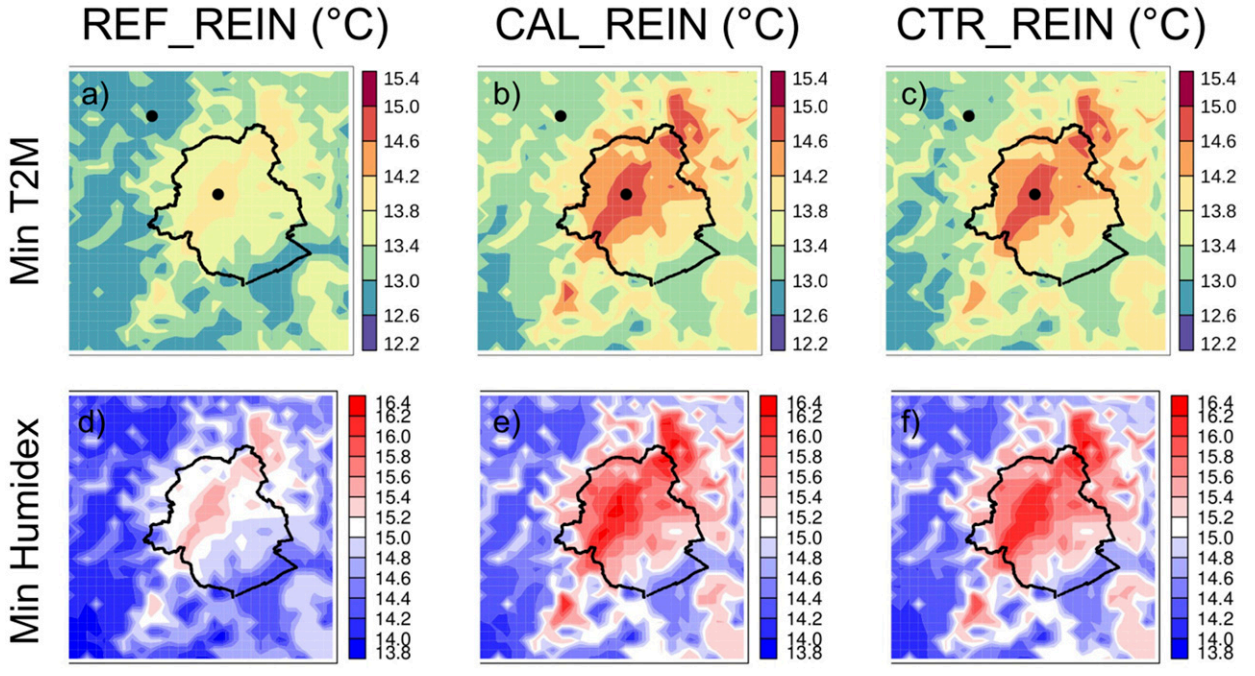

FIG. 7. Spatial pattern of average (a)–(c) daily minimum temperature and (d)–(f) daily minimum Humidex [see Eq. (6)] over Brussels. Both averages are taken over 26 summers of the evaluation period for the three 1-km experiments. These include (left) REF_REIN, (center) CAL_REIN, and (right) CTR_REIN as described in Fig. 4. The two black dots over Brussels indicate the locations of the urban and rural stations (Molenbeek and Brussegem).

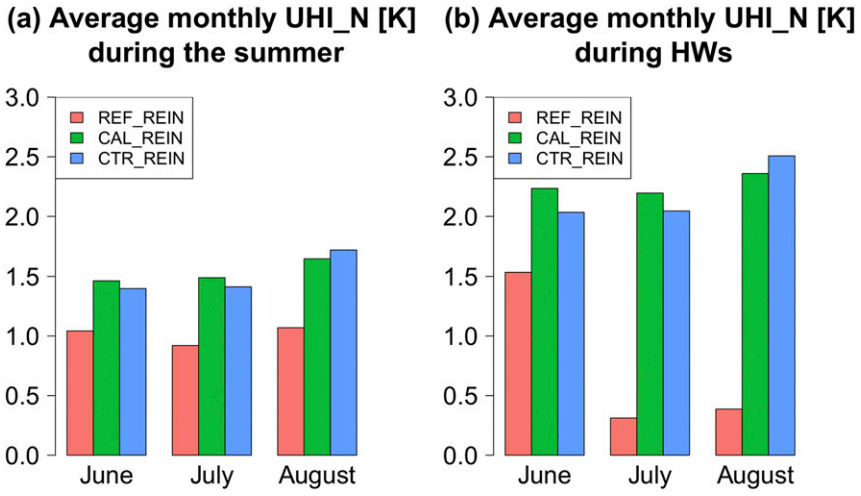

FIG. 8. The average monthly UHI_N (K) by month. Averages are taken over the 26 summers of the evaluation period from the three 1-km experiments. These include REF_REIN (red), CAL_REIN (green), and CTR_REIN (blue) as described in Fig. 4. Shown are (a) the average on the complete summer data and (b) averages taken over all HW days. There were no heat waves in September.

the reference and the control runs only differ during the night, and the calibration is clearly capable of reestablishing the correlations. This leads to an improved representation of Humidex, an index for heat comfort that combines temperature and relative humidity [see Eq. (6)]. Figures 7d–f show the average (minimal) Humidex. The calibration method strongly improves the spatial pattern, especially during the night, although there is a slight overestimation of the calibrated run as compared to the control one.

Figure 10 shows the daily minimum Humidex averaged over the HW days (at the urban location for each experiment separately). In the rural and urban areas the reference run overestimates and underestimates, respectively. Again, the calibration is able to strongly alleviate these differences.

d. Calibration results of the continuous setup

Here we discuss the results of the 1-km CONT simulations, which are obtained by initializing the soil variables once and evolve continuously thereafter. Although the calibration scheme does not fully eliminate the bias in the nocturnal UHI (see Table 3) with respect to the control run (CTR_CONT), there is a clear improvement. More specifically, the nocturnal UHI bias is reduced from 0.4 to 0.2 K, and the RMSE is reduced from 0.4 to 0.3 K during summertime. During HWs the bias is lower by a factor of 2, but the RMSE remains at the same level.

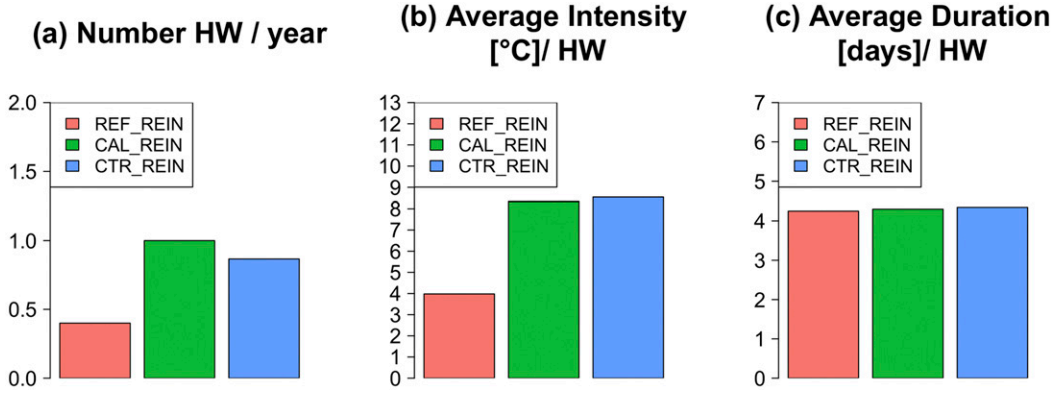

FIG. 9. Heat-wave characteristics, averaged over the 26 summers of the evaluation period: (a) average number of HWs per year, (b) average nocturnal intensity during an HW [see Eq. (4)], and, (c) average HW duration. The three 1-km experiments include REF_REIN (red), CAL_REIN (green), and CTR_REIN (blue) as described in Fig. 4.

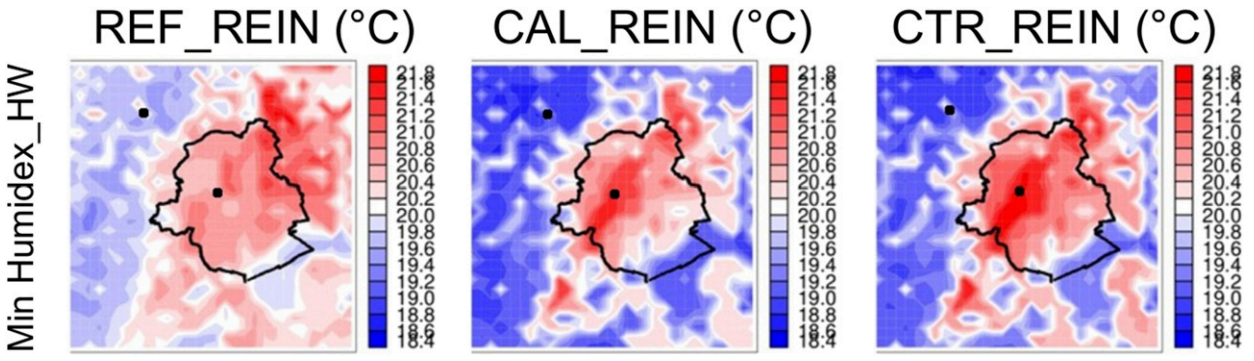

FIG. 10. As in Figs. 7d–f, but for Humidex [see Eq.(6)] only during heat waves. Averaged are taken over all HW days detected in the city center.

Figure 11 shows the spatial pattern of the average minimum and maximum summer temperature over Brussels. There is a clear spatial improvement relative to the reference when using the calibration method. Although the difference in temperature is small, this has already had an impact on extreme events such as the number of HW. The total number of HWs in the city center is 17, 26, and 39 for the REF_CONT, CAL_CONT, and CTR_CONT, respectively. Therefore, the calibration amounts to 53% more HWs as compared with the reference. The nocturnal HW intensity [see Eq. (4)] is only slightly increased (by 16%) by calibration, while the daytime HW intensity [see Eq. (5)] and the duration of HW is not drastically impacted. The impact of calibration on Humidex is comparable to the one on temperature (see Fig. 11). More specifically, the average daily minimum Humidex in the city center is 20.2°, 20.5°, and 21.0°C for the reference, calibrated, and control runs, respectively.

4. Conclusions

An ensemble of decadelong climate projections (i.e., more than 30 years) is necessary to estimate uncertainties and, in principle, an ensemble of RCM projections could be used for that purpose. However, RCMs fail to resolve city-scale processes due their coarse spatial detail and since they typically lack an urban scheme. Since a full dynamic downscaling approach from RCM

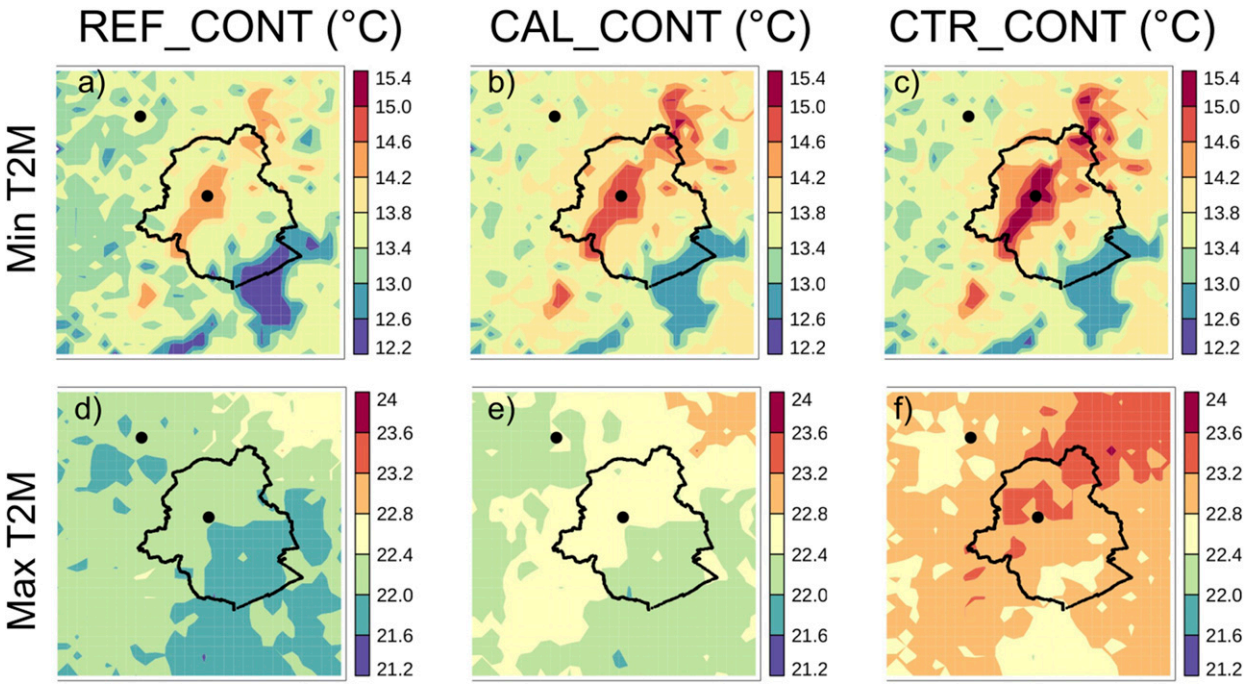

FIG. 11. As in Figs. 7a–c, but for the continuous (CONT) simulation instead of the reinitialized simulations and showing the daily maximum temperature as well as the daily minimum temperature.

to the urban range is computationally too expensive, a statistical–dynamical method is proposed that approximates this full dynamic downscaling. Moreover, as opposed to previously proposed downscaling methods, ours is computationally manageable, adequately reproduces city–atmosphere interactions, and amounts to time series with adequate intervariable and spatiotemporal correlations.

The method consists of selectively adding an urban signature to the RCM outputs and then using this as input of the LSM in a stand-alone mode. A proof of concept is presented for Brussels for 26 summer seasons using the RCM ALARO-0 and the LSM SURFEX.

The evaluation is done against a full dynamic downscaling approach. The calibration method is able to reduce the overall RCM bias in nocturnal UHI from −0.5 K to below 0.1 K, while during heat waves the bias is reduced from 1.5 K to below 0.1 K. Similarly, the RMSE is strongly reduced. The calibration method goes beyond a simple overall adjustment as it also reestablishes temporal correlations such as the number and intensity of heat waves, spatial patterns, and intervariable correlations such as the one of Humidex. Moreover, the spatial pattern of Humidex is also well calibrated during heat waves. Apart from the aforementioned simulations that use daily reinitialized soil variables, the calibration method was also applied using continuously evolved soil variables. Although strongly reduced, the improvements from calibration were substantial. This highlights the need for adequate soil initialization.

The proposed approach will be used to downscale regional climate model ensembles such as the one from the EURO-CORDEX archive (Jacob et al. 2014; Giot et al. 2016) to city level in order to propagate the model and scenario uncertainties from regional to city scale.

Acknowledgments. This work is supported by URCLIM, which has received funding from EU's H2020 Research and Innovation Program under Grant Agreement 690462. The authors are very grateful to the anonymous reviewers for their constructive comments. The data for Molenbeek station were obtained from IRCEL-CELINE (the Belgian Interregional Environment Agency).

REFERENCES

Abramowitz, M., and I. A. Stegun, 1972: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 10th ed. U.S. Department of Commerce, 1046 pp.

Bader, D. A., and Coauthors, 2018: Urban climate science. *Climate Change and Cities: Second Assessment Report of the Urban Climate Change Research Network (ARC3.2)*, C. Rosenweig et al., Eds., Cambridge University Press, 27–60.

Berckmans, J., O. Giot, R. De Troch, R. Hamdi, R. Ceulemans, and P. Termonia, 2017: Reinitialised versus continuous regional climate simulations using ALARO-0 coupled to the land surface model SURFEXv5. *Geosci. Model Dev.*, **10**, 223–238, https://doi.org/10.5194/gmd-10-223-2017.

——, R. Hamdi, and N. Dendoncker, 2019: Bridging the gap between policy-driven land use changes and regional climate projections. *J. Geophys. Res. Atmos.*, **124**, 5934–5950, https://doi.org/10.1029/2018JD029207.

Best, M., and C. S. B. Grimmond, 2014: Importance of initial state and atmospheric conditions for urban land surface models' performance. *Urban Climate*, **10**, 387–406, https://doi.org/10.1016/j.uclim.2013.10.006.

Bokwa, A., and Coauthors, 2018: Urban climate in central European cities and global climate change. *Acta Climatol.*, **51–52**, 7–35, https://doi.org/10.14232/acta.clim.2018.52.1.

Budd, G. M., 2008: Wet-bulb globe temperature (WBGT)—Its history and its limitations. *J. Sci. Med. Sport*, **11**, 20–32, https://doi.org/10.1016/j.jsams.2007.07.003.

Chapman, S., J. E. M. Watson, A. Salazar, M. Thatcher, and C. A. McAlpine, 2017: The impact of urbanization and climate change on urban temperatures: A systematic review. *Landscapes Ecol.*, **32**, 1921–1935, https://doi.org/10.1007/s10980-017-0561-4.

Daniel, M., A. Lemonsu, M. Déqué, S. Somot, A. Alias, and V. Masson, 2019: Benefits of explicit urban parametrization in regional climate modeling to study climate and city interactions. *Climate Dyn.*, **52**, 2745–2764, https://doi.org/10.1007/s00382-018-4289-x.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

De Troch, R., R. Hamdi, H. Van de Vyver, J.-F. Geleyn, and P. Termonia, 2013: Multiscale performance of the ALARO-0 model for simulating extreme summer precipitation climatology in Belgium. *J. Climate*, **26**, 8895–8915, https://doi.org/10.1175/JCLI-D-12-00844.1.

Forzieri, G., A. Cescatti, F. B. Silva, and L. Feyen, 2017: Increasing risk over time of weather-related hazards to the European population: A data-driven prognostic study. *Lancet Planet. Health*, **1**, e200–e208, https://doi.org/10.1016/S2542-5196(17)30082-7.

Founda, D., and M. Santamouris, 2017: Synergies between urban heat island and heat waves in Athens (Greece), during an extremely hot summer (2012). *Nat. Sci. Rep.*, **7**, 10973, https://doi.org/10.1038/s41598-017-11407-6.

Früh, B., and Coauthors, 2011: Estimation of climate-change impacts on the urban heat load using an urban climate model and regional climate projections. *J. Appl. Meteor. Climatol.*, **50**, 167–184, https://doi.org/10.1175/2010JAMC2377.1.

Gabriel, K. M. A., and W. R. Endlicher, 2011: Urban and rural mortality rates during heat waves in Berlin and Brandenburg, Germany. *Environ. Pollut.*, **159**, 2044–2050, https://doi.org/10.1016/j.envpol.2011.01.016.

Gerard, L., J.-M. Piriou, R. Brožková, J.-F. Geleyn, and D. Banciu, 2009: Cloud and precipitation parameterization in a meso-gammascale operational weather prediction model. *Mon. Wea. Rev.*, **137**, 3960–3977, https://doi.org/10.1175/2009MWR2750.1.

Giorgi, F., C. Jones, and G. R. Asrar, 2009: Addressing climate information needs at the regional level: The CORDEX framework. *WMO Bull.*, **58**, 175–183.

Giot, O., and Coauthors, 2016: Validation of the ALARO-0 model within the EURO-CORDEX framework. *Geosci. Model Dev.*, **9**, 1143–1152, https://doi.org/10.5194/gmd-9-1143-2016.

Giovannini, L., D. Zardi, and M. Franceschi, 2011: Analysis of the urban thermal fingerprint of the city of Trento in the Alps. *J. Appl. Meteor. Climatol.*, **50**, 1145–1162, https://doi.org/10.1175/2010JAMC2613.1.

Guerreiro, S. B., R. J. Dawson, C. Kilsby, E. Lewis, and A. Ford, 2018: Future heat-waves, droughts and floods in 571 European cities. *Environ. Res. Lett.*, **13**, 034009, https://doi.org/10.1088/1748-9326/aaaad3.

Gutiérrez, J. M., D. San-Marin, S. Brands, R. Manzanas, and S. Herrera, 2013: Reassessing statistical downscaling techniques for their robust application under climate change conditions. *J. Climate*, **26**, 171–188, https://doi.org/10.1175/JCLI-D-11-00687.1.

Hamdi, R., and V. Masson, 2008: Inclusion of a drag approach in the town energy balance (TEB) scheme: Offline 1-D evaluation in a street canyon. *J. Appl. Meteor. Climatol.*, **47**, 2627–2644, https://doi.org/10.1175/2008JAMC1865.1.

——, H. Van de Vyver, R. De Troch, and P. Termonia, 2014a: Assessment of three dynamical urban climate downscaling methods: Brussels's future urban heat island under an A1B emission scenario. *Int. J. Climatol.*, **34**, 978–999, https://doi.org/10.1002/joc.3734.

——, and Coauthors, 2014b: Evaluating the performance of SURFEXv5 as a new land surface scheme for the ALADINcy36 and ALARO-0 models. *Geosci. Model Dev.*, **7**, 23–39, https://doi.org/10.5194/gmd-7-23-2014.

——, O. Giot, R. De Troch, A. Deckmyn, and P. Termonia, 2015: Future climate of Brussels and Paris for the 2050s under the A1B scenario. *Urban Climate*, **12**, 160–182, https://doi.org/10.1016/j.uclim.2015.03.003.

——, F. Duchêne, J. Berckmans, A. Delcloo, C. Vanpoucke, and P. Termonia, 2016: Evolution of urban heat wave intensity for the Brussels Capital Region in the ARPEGE-Climat A1B scenario. *Urban Climate*, **17**, 176–195, https://doi.org/10.1016/j.uclim.2016.08.001.

Hewitt, C. D., and J. A. Lowe, 2018: Toward a European climate prediction system. *Bull. Amer. Meteor. Soc.*, **99**, 1997–2001, https://doi.org/10.1175/BAMS-D-18-0022.1.

Hoffmann, P., O. Krueger, and K. H. Schlünzen, 2012: A statistical model for the urban heat island and its application to a climate change scenario. *Int. J. Climatol.*, **32**, 1238–1248, https://doi.org/10.1002/joc.2348.

——, R. Schoetter, and K. Heinke Schlünzen, 2018: Statistical-dynamical downscaling of the urban heat island in Hamburg, Germany. *Meteor. Z.*, **27**, 89–109, https://doi.org/10.1127/metz/2016/0773.

Jacob, D., and Coauthors, 2014: EURO-CORDEX: New high-resolution climate change projections for European impact research. *Reg. Environ. Change*, **14**, 563–578, http://doi.org/10.1007/s10113-013-0499-2.

Kershaw, T., M. Sanderson, D. Coley, and M. Eames, 2010: Estimation of the urban heat island for UK climate change projections. *Build. Serv. Eng. Res. Tech.*, **31**, 251–263, https://doi.org/10.1177/0143624410365033.

Lauwaet, D., K. De Ridder, S. Saeed, E. Brisson, F. Chatterjee, N. P. M. Lipzig, B. Maiheu, and H. Hooyberghs, 2016: Assessing the current and future urban heat island of Brussels. *Urban Climate*, **15**, 1–15, https://doi.org/10.1016/j.uclim.2015.11.008.

Le Bras, J., and V. Masson, 2015: A fast and spatialized urban weather generator for long-term urban studies at the city-scale. *Front. Earth Sci.*, **3**, 27, https://doi.org/10.3389/feart.2015.00027.

Lemonsu, A., V. Viguié, M. Daniel, and V. Masson, 2015: Vulnerability to heat waves: Impact of urban expansion scenarios on urban heat island and heat stress in Paris (France). *Urban Climate*, **14**, 586–605, https://doi.org/10.1016/j.uclim.2015.10.007.

Li, D., and E. Bou-Zeid, 2013: Synergistic interactions between urban heat islands and heat waves: The impact in cities is larger than the sum of its parts. *J. Appl. Meteor. Climatol.*, **52**, 2051–2064, https://doi.org/10.1175/JAMC-D-13-02.1.

Liu, D., and Coauthors, 2018: A new model to downscale urban and rural surface and air temperatures evaluated in Shanghai, China. *J. Appl. Meteor. Climatol.*, **57**, 2267–2283, https://doi.org/10.1175/JAMC-D-17-0255.1.

Maraun, D., and M. Widmann, 2018: *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press, 347 pp.

Martilli, A., A. Clappier, and M. W. Rotach, 2002: An urban surface exchange parameterisation for mesoscale models. *Bound.-Layer Meteor.*, **104**, 261–304, https://doi.org/10.1023/A:1016099921195.

Martinez, G. S., and Coauthors, 2018: Heat and health in Antwerp under climate change: Projected impacts and implications for prevention. *Environ. Int.*, **111**, 135–143, https://doi.org/10.1016/j.envint.2017.11.012.

Masson, V., 2000: A physically-based scheme for the urban energy budget in atmospheric models. *Bound.-Layer Meteor.*, **94**, 357–397, https://doi.org/10.1023/A:1002463829265.

——, and Y. Seity, 2009: Including atmospheric layers in vegetation and urban offline surface schemes. *J. Appl. Meteor. Climatol.*, **48**, 1377–1397, https://doi.org/10.1175/2009JAMC1866.1.

——, J. L. Champeaux, F. Chauvin, C. Meriguet, and R. Lacaze, 2003: A global database of land surface parameters at 1 km resolution in meteorological and climate models. *J. Climate*, **16**, 1261–1282, https://doi.org/10.1175/1520-0442-16.9.1261.

——, and Coauthors, 2013: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of Earth surface variables and fluxes. *Geosci. Model Dev.*, **6**, 929–960, https://doi.org/10.5194/gmd-6-929-2013.

Masson-Delmotte, V., and Coauthors, Eds., 2018: Global warming of 1.5°C. World Meteorological Organization, 32 pp.

Masterson, M. J., and F. A. Richardson, 1979: *Humidex: A Method of Quantifying Human Discomfort due to Excessive Heat and Humidity*. Environment Canada, Atmospheric Environment Service, 45 pp.

Moonen, P., T. Defraeye, V. Dorer, B. Blocken, and J. Carmeliet, 2012: Urban physics: Effect of the micro-climate on comfort, health and energy demand. *Front. Archit. Res.*, **1**, 197–228, https://doi.org/10.1016/j.foar.2012.05.002.

Oke, T., G. Mills, A. Christen, and J. Voogt, 2017: *Urban Climates*. Cambridge University Press, 546 pp., https://doi.org/10.1017/9781139016476.

Oleson, K. W., G. B. Bonan, J. Feddema, and T. Jackson, 2011: An examination of urban heat island characteristics in a global climate model. *Int. J. Climatol.*, **31**, 1848–1865, https://doi.org/10.1002/joc.2201.

Robine, J. M., S. L. K. Cheung, S. Le Roy, H. Van Oyen, C. Griffiths, J. P. Michel, and F. R. Herrmann, 2008: Death toll exceeded 70,000 in Europe during the summer of 2003. *C. R. Biol.*, **331**, 171–178, https://doi.org/10.1016/j.crvi.2007.12.001.

Rosenzweig, C., W. Solecki, P. Romero-Lankao, S. Mehrotra, S. Dhakal, T. Bowman, S. Ali Ibrahim, 2015: Climate change and cities: Second assessment report of the Urban Climate Change Research Network (ARC3.2)—Summary for city leaders. Urban Climate Change Research Network Rep., 25 pp., http://uccrn.org/files/2015/12/ARC3-2-web.pdf.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Termonia, P., and Coauthors, 2018: The CORDEX.be initiative as a foundation for climate services in Belgium. *Climate Serv.*, **11**, 49–61, https://doi.org/10.1016/j.cliser.2018.05.001.

Tsiringakis, A., G. J. Steeneveld, A. A. Holtslag, S. Kotthaus, and S. Grimmond, 2019: On- and off-line evaluation of the single-layer urban canopy model in London summertime conditions. *Quart. J. Roy. Meteor. Soc.*, **145**, 1474–1489, https://doi.org/10.1002/qj.3505.

UN-Habitat, 2010: State of the world's cities 2010/2011: Bridging the urban divide—Overview and key findings. UN-Habitat Rep., 244 pp., https://sustainabledevelopment.un.org/content/documents/11143016_alt.pdf.

Wilby, R. L., 2008: Constructing climate change scenarios of urban heat island intensity and air quality. *Environ. Plann.*, **35B**, 902–919, https://doi.org/10.1068/b33066t.

Wouters, H., and Coauthors, 2017: Heat stress increase under climate change twice as large in cities as in rural areas: A study for a densely populated midlatitude maritime region. *Geophys. Res. Lett.*, **44**, 8997–9007, https://doi.org/10.1002/2017GL074889.

Zhou, Y., and J. M. Shepherd, 2010: Atlanta's urban heat island under extreme heat conditions and potential mitigation strategies. *Nat. Hazards*, **52**, 639–668, https://doi.org/10.1007/s11069-009-9406-z.

Žuvela-Aloise, M., R. Koch, A. Neureiter, R. Böhm, and S. Buchholz, 2014: Reconstructing urban climate of Vienna based on historical maps dating to the early instrumental period. *Urban Climate*, **10**, 490–508, https://doi.org/10.1016/j.uclim.2014.04.002.