Assessment of the Paris urban heat island in ERA5 and offline SURFEX-TEB (v8.1) simulations using METEOSAT land surface temperature product

Miguel Nogueira[1], Alexandra Hurduc[1], Sofia Ermida[1,2], Daniela C. A. Lima[1], Pedro M. M. Soares[1], Frederico Johannsen[1], Emanuel Dutra[1,2]

[1]Instituto Dom Luiz, IDL, Faculty of Sciences, University of Lisbon, 1749-016 Lisbon, Portugal
[2]Instituto Português do Mar e da Atmosfera, IPMA, 1749-077 Lisbon, Portugal

Correspondence to: Frederico Johannsen (jfjohannsen@fc.ul.pt)

Abstract. Cities concentrate people, wealth, emissions, and infrastructures, thus representing a challenge and an opportunity for climate change mitigation and adaptation. This places an urgent demand for accurate urban climate projections to help organizations and individuals making climate smart-decisions. However, most of the large ensembles of global and regional climate model simulations do not include sophisticated urban parameterizations. Here, we use the city of Paris as a case study to show that this is the case for the fifth (and latest) generation reanalysis from the European Centre for Medium-Range Weather Forecasts (ERA5) and for simulations employing the widely used bulk bare rock approach to urban climate parameterization. Subsequently, we leveraged on the hourly resolution of ERA5 and the Satellite Application Facility Land Surface Analysis (LSA-SAF) land surface temperature product to demonstrate the significant added value of employing the SURFEX land-surface model coupled to Town Energy Balance (TEB) urban canopy model in simulating the Parisian Surface Urban Heat Island (SUHI) during daytime and the urban heat island during both daytime and nighttime. Our results showed the significant added value of SURFEX-TEB in reproducing the observed daytime and nighttime Parisian urban heat island effect. An annual average bias magnitude reduction of 0.5ºC was observed for daytime and around 1.5ºC for nighttime when compared to ERA5 and bare rock approach. Also, SURFEX-TEB revealed an overall better performance in reproducing the observed daytime SUHI, whilst the added value of SURFEX-TEB was lower during nighttime (but still slightly better than ERA5 and the bare rock approach), due to the lack of land-atmosphere feedbacks in the proposed offline framework. Finally, the offline SURFEX-TEB framework applied here demonstrates the added value of using more comprehensive urban parameterizations to simulate the urban climate, therefore, improving urban climate projections.

1 Introduction

Urban areas accommodate nearly half of the global population, and this fraction is projected to increase to 68% by 2050 according to the World Health Organization (WHO, 2018). Moreover, cities concentrate wealth, infrastructures, and emissions - being responsible for about 75% of the global greenhouse gas emissions from energy consumption (IPCC, 2014).

30   Consequently, understanding and simulating the urban climate evolution is a key task for climate change assessments and for designing climate change adaptation and mitigation strategies.

The urban areas are characterized by drastic land-use changes which are responsible for increased trapping and absorption of solar radiation, reduced evapotranspiration, and decreased nighttime cooling in built-up areas. As a result, cities typically have warmer air and surface temperatures compared to nearby rural environments. This is the well-known urban heat island

35   (UHI) effect, which has been found over multiple cities across the globe (see e.g., Deilami et al., 2018 for a recent review).

The identification and quantification of the UHI dynamics have proven to be challenging. Despite its widespread emergence in urban environments, the UHI is sensitive to the specific land surface characteristics and meteorological conditions, hence displaying significant variability between different locations and periods. Indeed, previous investigations reported several different relevant UHI dependencies, including city size and population density (Oke, 1982; Clinton & Gong, 2013; Manoli

40   et al., 2019), urban vegetation coverage (Kaloustian & Diab, 2015; Peng et al., 2012; Zhou et al., 2014; Nogueira & Soares, 2019), background climate conditions (namely precipitation and wind, Zhou et al., 2013; Lemonsu et al., 2013; Zhao et al., 2014; Manoli et al., 2019) and urban morphology (e.g., city geometry, building height, construction materials, etc., Oke, 1973; 1982; Zhou et al., 2017; Krayenhoff et al., 2018; Nogueira & Soares, 2019; Masson et al., 2020). Heat release resulting from human activities has also been shown to modulate the UHI (De Munck et al., 2013; Schoetter et al., 2020).

45   Moreover, surface and near-surface air temperature over "natural" regions also display large sensitivity to the complex land use and land cover patterns (e.g, Beljaars et al., 1996; Koster et al., 2004; Johannsen et al., 2019; Nogueira et al., 2020a, 2021), which represents an additional layer of complexity to the UHI.

Investigations of the UHI based on contrasting in situ temperature and surface fluxes observations from cities and neighboring rural locations are generally unable to capture its complex spatial heterogeneity, particularly for large cities,

50   resulting in large uncertainties in the UHI characterization (Stewart, 2011; Schwarz et al., 2011; Stewart and Oke, 2012). The development of dense urban meteorological station networks allowed to partially overcome these limitations, but the temporal and spatial coverage of such networks remains too narrow to fully characterize the urban induced climate modulation (Muller et al., 2013; Konstantinov et al., 2018). Remote sensing techniques provide a widely used alternative for comprehensive characterization of the UHI and its variability, providing reliable estimates for numerous land surface

55   properties with wide spatial coverage and adequate spatial and temporal sampling, including land surface temperature (LST), land use and land cover (LULC) maps, soil moisture, rainfall, and snow, amongst others (see Balsamo et al., 2018 for a recent review).

Numerous works revealed the existence of a surface urban heat island (SUHI), referring to warmer LST in urban areas compared to its rural environment (e.g., Roth et al., 1989; Imhoff et al., 2010; Schwarz et al., 2011; Peng et al., 2012; Zhao

60   et al., 2014; Zhou et al., 2017). Yet, these studies identified significant differences between the UHI and SUHI, including the maximum UHI hour and seasonality, and the relationship between thermal contrast magnitude and land use. Moreover, LST estimates are often restricted to clear-sky conditions since, typically, the available all-sky estimates are restricted by very

coarse spatial resolution which is inappropriate to characterize the urban environments (Masson et al., 2020). The LST estimates are also often constrained by the time of satellite overpass, which limits the temporal resolution.

65 Urban climate simulations generated by physically-based numerical models can potentially circumvent some of the limitations of in situ and remote sensing observational products. Specifically, coherent information for multiple relevant variables with high spatial and temporal coverage and resolutions may be obtained. Additionally, due to the complexity and diversity of cities around the world, the city scale climate properties are specific and often limited to a particular location. Moreover, while observations cover the past, numerical simulations can be extended to the future and, therefore, consider

70 different scenarios of future socio-economic evolution, urban development, and adaptation strategies, as shown, for example, in Georgescu et al. (2014) where it was shown how urban planning could help offset the global warming effect in U.S. cities in the future.

Most large ensembles of global and regional climate model simulations have simplified representations of the urban environment (Garuma, 2018; Zhao et al., 2021). Furthermore, the available large ensembles of Earth System Models (ESMs)

75 and Global Climate Models (GCMs) typically have coarse spatial resolutions (~100 km), which are inadequate for representing most of the city-scale processes. Typically, state-of-the-art large multi-model ensembles of Regional Climate Models (RCMs) have grid-resolutions on the order of tens of kilometers which is still inappropriate to simulate many aspects of the urban climate system (e.g., Langendijk et al., 2019; Nogueira et al., 2020b; McNorton et al., 2021). The next generation of RCM ensembles will have a resolution of a few kilometers, allowing a better simulation of the local climate

80 variability (Jacob et al., 2020). Indeed, several pilot studies have suggested significant added value in including urban canopy models (UCMs) to parameterize interactions between the urban surface and the atmosphere in RCMs with resolutions of a few kilometers (Chen et al., 2011; Kusaka et al., 2012; Hamdi et al., 2012; Lemonsu et al., 2014; Daniel et al., 2019; Garuma, 2018; Schoetter et al., 2020). However, the use of UCM coupled to RCMs is not a standard procedure for long-time/century climate simulations (and is not projected to be in the next generation of multi-model RCM ensembles) due

85 to its very high computational costs, resulting in a poor representation of many aspects of urban climate in those RCM ensemble datasets (Langendijk et al., 2019; Nogueira et al., 2020b).

The use of land-surface models (LSM) coupled to a UCM, forced offline by atmospheric data, provides a computationally efficient option for urban climate simulation. This approach overcomes the computational and resolution limitations of ESMs, GCMs and RCMs, but comes at the cost of neglecting the urban land-atmosphere feedbacks, providing only

90 diagnostics for the surface and near-surface variables. One may estimate the urban impact on surface and near-surface air temperature and humidity, near-surface wind, latent and sensible heat fluxes, but not on clouds, precipitation, or local circulations. Despite those limitations, recent studies have demonstrated the added value of this approach in reproducing key features of observed urban climate compared to traditional climate simulations (without representation of urban processes), including the UHI and the frequency, intensity and duration of urban extreme temperature events (Broadbent et al., 2018;

95 Conlon et al., 2016; Daniel et al., 2018; Kaloustian and Diab, 2015; Lemonsu et al., 2013, 2015; Nogueira and Soares, 2019; Hamdi et al., 2020; Viguié et al., 2020; Nogueira et al., 2020b). Leveraging the competitive computational cost of offline

LSM-UCM simulations, these studies explored the local climate response to multiple different urbanization patterns and emission scenarios over relatively long periods and at high spatial resolution. Additionally, Nogueira and Soares (2019) demonstrated how this type of framework may be used to disentangle the impact of land-use change, from large-scale warming induced by greenhouse gas emissions, and from natural climate variability. Other approaches to tackle this problem have been suggested in the past (e.g. Georgescu et al., 2014; Krayenhoff et al., 2018; Broadbent et al., 2020). This represents a critical task for anthropogenic climate change attribution and for designing effective mitigation strategies. The added value of the offline framework has also been demonstrated in simulating the impact of changes in vegetation cover patterns over non-urban regions (Johannsen et al., 2019; Nogueira et al., 2020a, 2021).

The present study assesses the ability of the LSM-UCM approach to downscale ERA5 reanalysis, the fifth, and latest, generation reanalysis from the European Centre for Medium-Range Weather Forecasts to resolutions of a few kilometers over dense urban areas. Specifically, we analyze the added value of the Météo-France SURFEX (Surface Externalisée) surface modelling platform (Le Moigne, 2018) in improving the simulation of the UHI and SUHI over Paris, a European mega-city characterized by a well-known strong urban heat island effect (e.g., Sarkar & De Ridder, 2011; De Munck et al., 2013; Hamdi et al., 2015; Lemonsu et al., 2015; Daniel et al., 2019). SURFEX is particularly relevant in this context since it has shown to perform particularly well in offline urban simulations (e.g., Hamdi et al., 2015; Lemonsu et al., 2015; Nogueira & Soares, 2019; Nogueira et al., 2020b). Previously, Nogueira et al. (2020b) used the offline LSM approach to perform downscaling of the EURO-CORDEX simulation ensemble for the historical and future periods over an urban grid-point inside the city of Lisbon and a neighboring rural grid-point. The results highlighted the poor representation of the UHI effect in the EURO-CORDEX RCMs and suggested the added value of the online approach for simulating the UHI effect. However, this study was limited to two single-column simulations and focused on the daily maximum and daily minimum 2-meter air temperature, which are diagnostic variables in the offline LSM approach. Here, we use high-resolution LST satellite data to investigate the spatial structure of the Parisian SUHI and its diurnal cycle in ERA5 and to assess the added value of the offline LSM downscaling approach, contrasting against the often-used bulk bare rock urban parameterization approach in large ensembles of GCMs/RCMs.

2. Methods

2.1 Observations and Reanalysis

Parisian daily maximum and minimum temperatures (respectively T_{max} and T_{min}) for the 2004-2018 period were obtained from two weather stations (Fig. 1), retrieved from the Global Summary of the Day (GSOD), produced by the National Climatic Data Center (NCDC), which includes quality control checks and random error removal. The first station located in the Montsouris public park (48.82N, 2.33E), in Paris city center, was used as reference to characterize the Parisian urban temperature. The second station located in Melun (southeast of Paris - 48.61N, 2.67E), in a natural environment, was used as

reference to characterize the Parisian surroundings temperature. These two stations were also previously employed by Hamdi et al. (2015) and Daniel et al. (2019) to characterize the Parisian UHI.

130   The LSA-SAF LST estimates are derived from the outgoing thermal infrared radiation (TIR) measured at top-of-atmosphere by the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard the Meteosat Second Generation (MSG) series by employing a generalized "split-window" technique (Freitas et al., 2010). The TIR spectral band (8–13 µm) is particularly appropriate as it presents relatively weak atmospheric attenuation under clear-sky conditions and includes the peak of the Earth's spectral radiance (Li et al., 2013; Ermida et al., 2019). The LSA-SAF LST estimates were available every 15 minutes

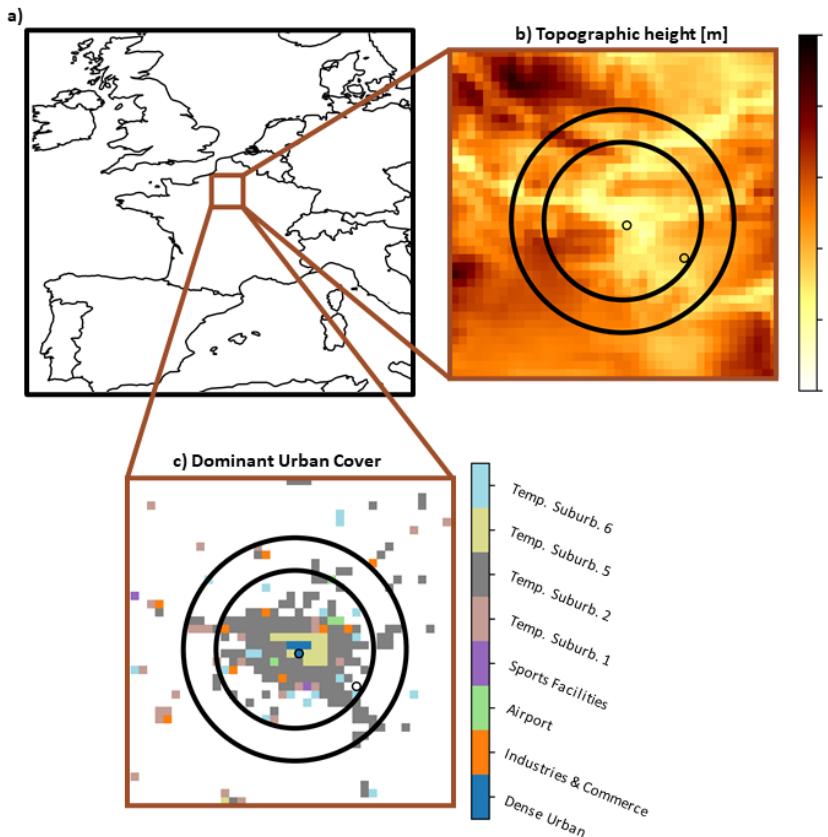
135   from 2004 to present-day over land pixels within the MSG disk, comprising satellite zenith view angles between 0º and 80º, with a 3 km resolution at the nadir. The LSA-SAF LST estimates were aggregated as the average at 00, 15, 30, and 45 minutes for each hour. Then, the hourly mean for the period ranging from 2004 through 2018 was computed. LST obtained through remote sensing is intrinsically directional due to the heterogeneity of the land surface. Still, given the model resolution considered in this study, LST's ability to evaluate model data should not be affected by its directional property.

140

Figure 1: Study domain identified by the brown square in panel a). Panel b) shows a zoom on the simulation domain with the color shading representing topographic height. Panel c) shows a zoom on the simulation domain with the color scheme representing the dominant urban classes for grid boxes where urban fraction exceeds 0.1 (grid boxes where urban fraction is below 0.1 are painted

ERA5 is the latest-generation global atmospheric reanalysis produced by the ECMWF, extending from 1979 to the present (although a preliminary version of an extension to 1950 is already available). ERA5 is based on a recent version of the ECMWF Integrated Forecast System (IFS cycle 41r2), including several improvements compared to the version used in ERA-Interim (the ECMWF's previous generation reanalysis, Dee et al., 2011). Namely, ERA5 features increased temporal, horizontal and vertical resolutions (respectively 1 hour, ~31 km and 137 vertical levels extending from surface to 0.01 hPa) (see Hersbach et al., 2020 for a detailed description of ERA5), and an increased number and more recent versions of a wide variety of observational datasets are assimilated. Additionally, ERA5 benefits from improvements to several model parameterizations (e.g. convection and microphysics) and to the four-dimensional variational data assimilation scheme. Furthermore, it also presents an overall better accuracy in representing several climate variables compared to ERA-Interim, including LST, near-surface air temperature, wind, radiation, and rainfall (e.g., Urraca, 2018; Beck et al., 2019; Johannsen et al., 2019; Rivas and Stoffelen, 2019; Nogueira, 2020).

2.2 SURFEX simulations

We performed a set of high-resolution (0.05°×0.05°) simulations of Paris and its surroundings (cf. Fig. 1) using the Météo-France SURFEX (Surface Externalisée) version 8.1 (Le Moigne, 2018) modelling platform. SURFEX couples multiple physical-based models over all types of natural surfaces, including the Interaction between Soil Biosphere and Atmosphere (ISBA) land-surface scheme over natural land surfaces (Calvet et al., 1998; Gibelin et al., 2006) and the Town Energy Balance (TEB) Urban Canopy Model (UCM) over urban surfaces. TEB uses the urban canyon approach (Oke, 1987) to simulate key urban physical processes on the local climate, including the possibility to account for the effects of vegetation and water bodies (see Masson et al., 2000 and Masson et al., 2013 for a detailed description of TEB). The SURFEX simulations were performed in an offline setup forced by ERA5 fields – namely surface pressure, precipitation, short- and long-wave radiative fluxes, and air temperature, humidity, and wind speed at 40 m height above sea level (above the Parisian urban canopy height). The simulations started in January 2003 and extended until the end of 2018, with a 15-minute time-step.

Two different SURFEX experiments were carried here. In the first one, denoted SFX-ROCK, the city grid boxes were described as rock covers. This bulk urban parameterization is often employed in large ensembles of regional climate simulations (e.g., Daniel et al., 2019; Langendijk et al., 2019; Davin et al. 2019; Nogueira et al., 2020b). The second experiment, denoted SFX-TEB, employed the TEB UCM for urban grid boxes. Both experiments considered a multilayer soil diffusion scheme with 14 soil layers and a single-level canopy layer, which has previously been demonstrated to be adequate over European mid-rise cities such as Paris (Schoetter et al., 2020; Trusilova et al. (2016)). It is worth pointing out, however, that several studies presenting RCMs combined with more complex urban schemes in short-term case studies have previously shown added value in simulating urban climate (e.g. Salamanca et al., 2010; Salamanca et al., 2011).

2.3 Assessment of the simulated UHI and SUHI

The UHI was defined here as the 2-meter air temperature (T_{2m}) difference between the urban (Montsouris) and the rural (Melun) station locations. For ERA5 and SURFEX simulations this was estimated using the respective nearest-neighbor grid boxes. Although the two-point difference approach cannot account for the complex spatial heterogeneity of urban environments and their surroundings, the limited number of observations available to this study defined this particular choice. Nonetheless, the complex spatial heterogeneity of the Paris area was accounted by the SUHI definition considered here. Following the methodology employed in previous works (Peng et al., 2012; and Zhou et al., 2013, 2017) the SUHI was defined as the difference between the average temperature within the considered urban cluster and the average temperature within an equal area belt around it. This approach combined land cover data (from ECOCLIMAP-II) with LSA-SAF LST. The urban cluster was defined as the grid boxes with urban fraction greater than 66% within the inner circle shown in Fig. 1. The surrounding belt is also shown in Fig. 1. The SUHI was computed for ERA5, SFX-ROCK and SFX-TEB using the same approach, considering the same urban fractions from the ECOCLIMAP-II land cover data for all datasets.

The daytime and nighttime UHI and SUHI were evaluated separately. The daytime maximum T_{2m} and LST (denoted T_{max} and LST_{max} respectively) were computed as the maximum temperature within the 11 to 18 UTC interval. The nighttime minimum T_{2m} and LST (denoted T_{min} and LST_{min} respectively) were computed as the minimum temperature over the 00 to 07 UTC interval. We computed two error metrics for evaluating the UHI and SUHI in the different model-based datasets. The first was the mean bias calculated following Eq. (1):

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(m_k - o_k), \tag{1}$$

where m_k and o_k are respectively model simulated and observed values and N is the total number of days in the historical time-series. The mean bias measures the models' systematic errors. The second was the Perkins skill score (Perkins et al., 2007), henceforth denoted S, which measures the models' ability to reproduce the observed probability distribution functions (PDFs):

$$S = 100 \times \sum_{i=1}^{B}\left[Z_{m,i}, Z_{o,i}\right], \tag{2}$$

where $min[x, y]$ represents the minimum between two values, Z_m and Z_o are the modeled and observed empirical PDFs, respectively, and B is the total number of bins used to compute the empirical PDFs (here we used steps of 1ºC). S provides a measure of similarity between modeled and observed empirical PDFs, with S=100% if the model reproduces the empirical PDF perfectly and decreasing towards zero as the similarity between the PDFs decreases. Both error metrics were also applied to evaluate the daytime and nighttime LST over the study domain grid boxes and the daytime and nighttime T_{2m} at the two station locations.

3. Results

3.1 Intercomparison of the simulated SUHI over Paris

The observed LST averaged over the 2004-2018 period showed a clear signature of the Parisian SUHI effect during daytime (Fig. 2a) and nighttime (Fig. 2b). The SUHI was not reproduced by ERA5 during daytime (Fig. 2c) nor nighttime (Fig. 2d). The results also highlighted that ERA5 0.25º resolution is too coarse to reproduce the complex urban climate patterns, even for a relatively large city as Paris. The simulation SFX-ROCK also failed to reproduce the Paris SUHI during daytime (Fig. 2e) and nighttime (Fig. 2f). As expected, the higher resolution did improve the simulation of topographic effects on surface temperature south and northwestern of Paris (cf. Fig. 1b), although, the bare rock approach misrepresented the city LST modulation. In contrast, the LST patterns resulting from the SFX-TEB simulation showed a clear signature of the Paris SUHI during daytime (Fig. 2g), which was closer to the observed pattern. During nighttime, SFX-TEB shows some signature of the SUHI (Fig. 2h) but underestimates its magnitude.

The improved ability of SFX-TEB in reproducing the urban LST over Paris was evidenced by the large reduction of the median Bias over urban grid boxes (i.e., where the urban fraction was above 2/3), from -7.0ºC in ERA5 and -6.7ºC in SFX-ROCK to 1.5ºC in SFX-TEB (Fig. 3a). Over natural surfaces, this reduction was lower, from -2.5ºC in ERA5 to -1.5ºC in SFX-ROCK and -1.3ºC in SFX-TEB. The 0.2ºC difference between SFX-ROCK and SFX-TEB was due to the considered definition of natural surfaces, encompassing all grid-boxes with urban fractions below 1/3. Indeed, in cases where the urban fraction was zero, the SFX-ROCK and SFX-TEB simulations were identical. Notice, however, that the differences amongst different datasets over natural surfaces were within the typical uncertainty associated with LSA-SAF LST estimates, which is of the order of 2ºC (Trigo et al., 2015). Over mixed surfaces (i.e., urban fractions between 1/3 and 2/3) the median Bias for annual averaged LST was -3.8ºC for ERA5, -3.3ºC for SFX-ROCK and 0.5ºC for SFX-TEB. These differences were also within the typical observational uncertainty. Yet, the large improvements over urban surfaces and identical performances over natural surfaces suggest that SFX-TEB represented an improvement over mixed surfaces too, particularly when the grid-box urban fraction approaches 2/3.

The largest reductions to daytime LST over urban areas' systematic errors occurred during MAM, where the median Bias was -8.8ºC for ERA5, -7.6ºC for SFX-ROCK, and 2.1ºC for SFX-TEB (Fig. 3e). A large reduction of the daytime LST systematic error was also found during SON, where the median Bias was -6.5ºC for ERA5 and SFX-ROCK, and -0.5ºC for SFX-TEB (Fig. 3i). During JJA, the median Bias was -8.2ºC for ERA5, -5.2ºC for SFX-ROCK, and 5.1ºC for SFX-TEB (Fig. 3g). Finally, during DJF the median Bias was -4.1ºC for ERA5, -6.3ºC for SFX-ROC, and -2.8ºC for SFX-TEB (Fig. 3c). Notice that, on the seasonal scale, SFX-TEB was not always the best performing model during daytime (for example, ERA5 was the best performing model during DJF over natural surfaces, and SFX-ROCK was the best performing during JJA over mixed surfaces).

During nighttime, the differences in annual average LST systematic errors amongst different simulations (Fig. 3b) were lower than during daytime. Specifically, over urban surfaces, the median Bias was -1.8ºC in ERA5, -2.0ºC in SFX-ROCK,

and -1.7ºC in SFX-TEB. Over mixed surfaces, the median Bias was 0.1ºC in ERA5 and -0.1ºC in SFX-ROCK, and near null

240   in SFX-TEB. Finally, over natural surfaces, the nighttime median Bias was 1.0ºC in ERA5 and 0.7ºC in SFX-ROCK and

SFX-TEB. Notice that these differences amongst simulations were within the observational uncertainty. This was also true

for all systematic differences in nighttime LST amongst different datasets on the seasonal scale (Fig. 3d, f, h, j).

Figure 4 evidences the clear reduction of the simulated LST systematic errors during daytime hours in SFX-TEB compared

to ERA5 and SFX-ROCK over urban (Fig. 4c) and mixed (Fig. 4b) grid boxes. Over natural surfaces, the daytime

245   performance was similar for SFX-ROCK and SFX-TEB, both slightly outperforming ERA5 (Fig. 4a). Finally, during the

night hours, Figure 4 showed a similar performance in reproducing the average LST over all surface types.

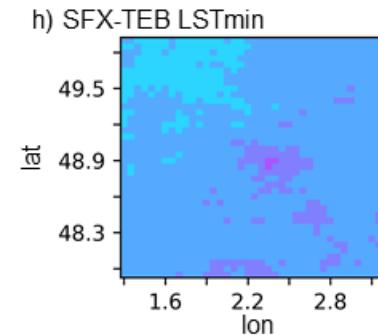
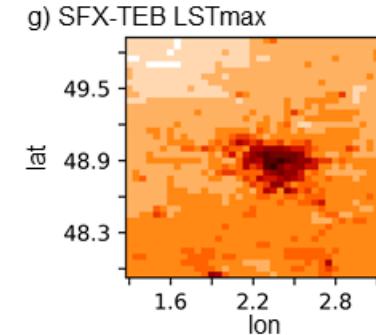
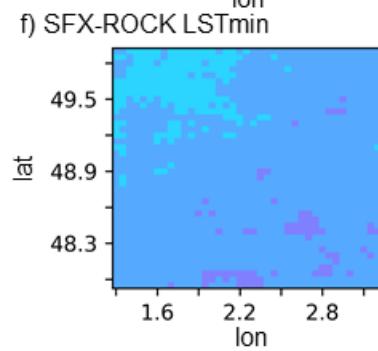
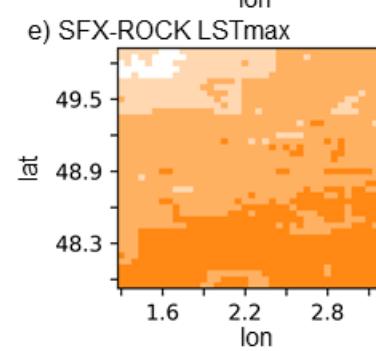
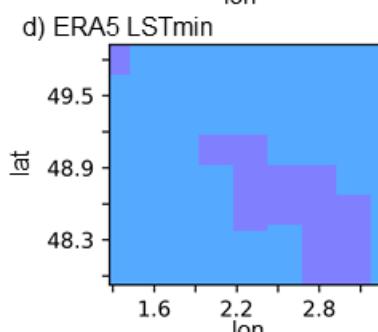
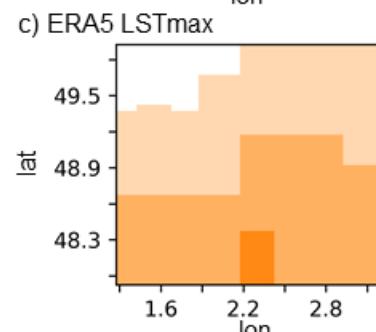
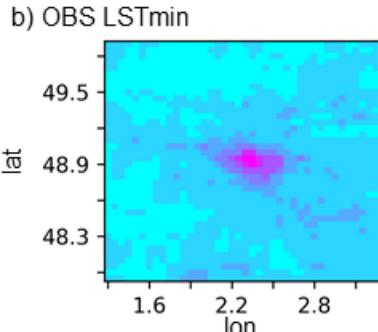
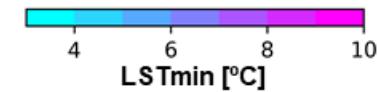
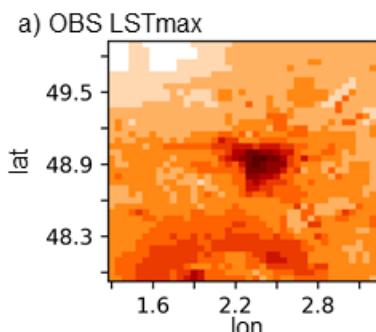
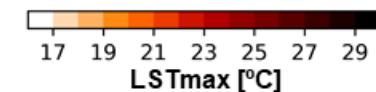



Figure 2: Maps of LST averaged over the 2004-2018 period during daytime (left column) and nighttime (right column) over the study domain computed from LSA-SAF (a) and b)), ERA5 (c) and d)), SFX-ROCK (e) and f)), and SFX-TEB (g) and h)).

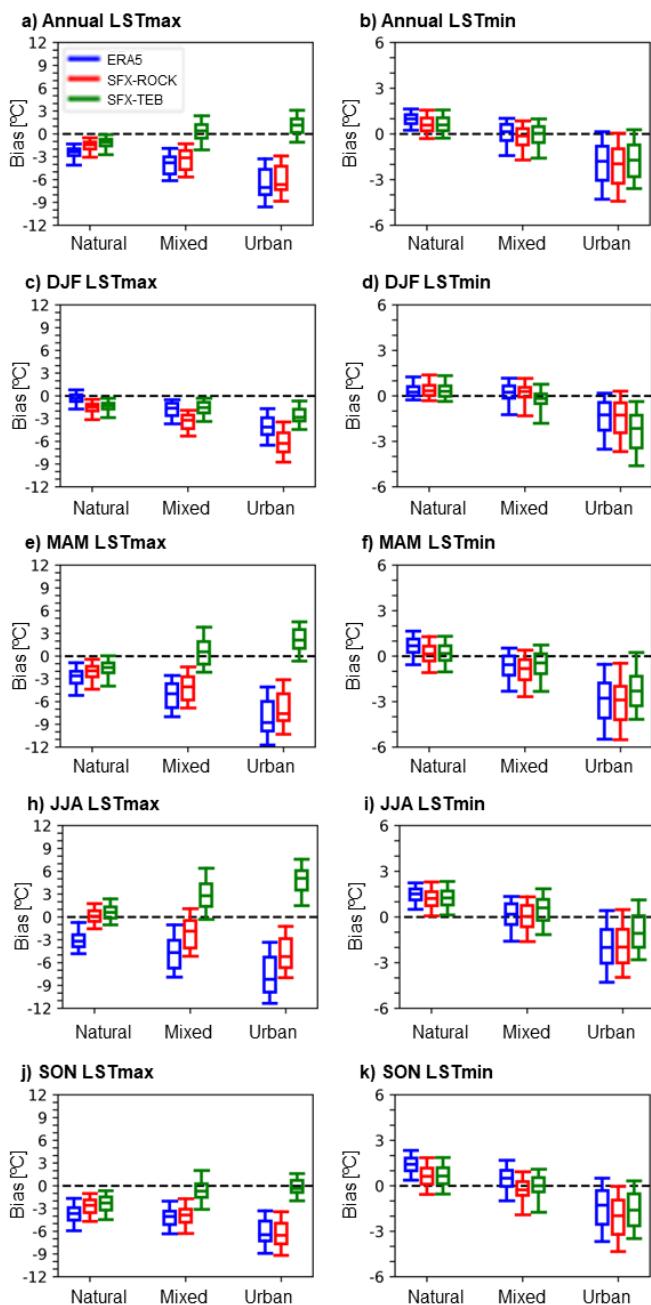



Figure 3: Boxplots of LST bias computed over the 2004-2018 period for daytime (left column) and nighttime (right column) for ERA5 (blue), SFX-ROCK (red), and SFX-TEB (green). The boxplots represent the bias spread for grid boxes classified as natural surfaces (grid boxes with urban fraction below 0.33), mixed surfaces (grid boxes with urban fraction between 0.33 and 0.66), and urban surfaces (grid boxes with urban fractions above 0.66). From top to bottom, the rows represent the bias computed for the full annual cycle, DJF, MAM, JJA and SON.

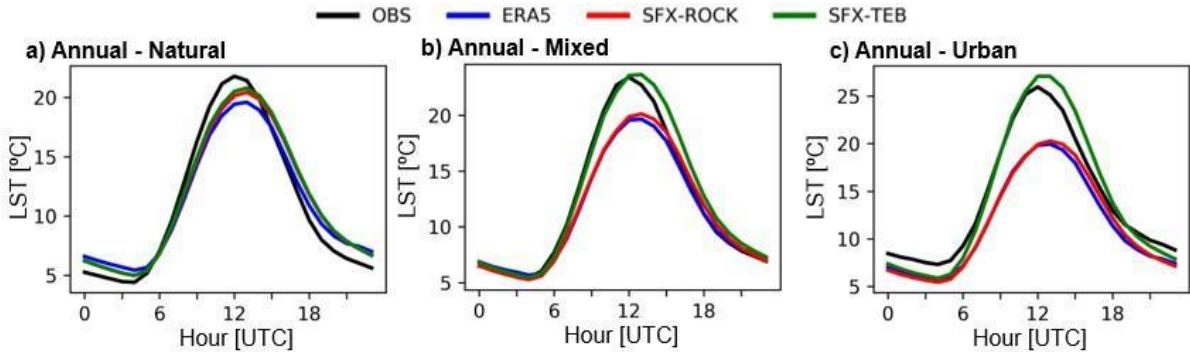



Figure 4: Average diurnal cycle of LST computed over the 2004-2018 period from LSA-SAF (black), ERA5 (blue), SFX-ROCK (red), and SFX-TEB (green) for grid boxes classified as a) natural surfaces (grid boxes with urban fraction below 0.33), b) mixed surfaces (grid boxes with urban fraction between 0.33 and 0.66), and c) urban surfaces (grid boxes with urban fractions above 0.66).

The SFX-TEB ability to reproduce the observed annual daytime LST PDF was better than ERA5 and SFX-ROCK over all surface types (Fig. 5a). Over urban surfaces, the median S score for daily maximum was 42% in ERA5, 49% in SFX-ROCK, and 62% in SFX-TEB. Over mixed surfaces, the median S score for daily maximum LST was 58% in ERA5, 64% in SFX-ROCK, and 67% in SFX-TEB. Over natural surfaces, the median S score for daily maximum LST was 66% in ERA5, 69% in SFX-ROCK, and 70% in SFX-TEB. The largest improvements in daily maximum LST S score associated with SFX-TEB emerged during DJF (Fig. 5c), MAM (Fig. 5e), and SON (Fig. 5i), and lowest during JJA (Fig. g, in fact, SFX-ROCK outperformed SFX-TEB over mixed surfaces during summer). The small differences in nighttime LST amongst simulations were also reflected in the S score, both on the annual (Fig. 5b) and seasonal scales (Figs. 5d, f, h, j).

At a seasonal scale the observed SUHI reveals a clear seasonal cycle at daytime and a rather constant value at nighttime (Figs. 6a, b). At daytime (nighttime), the winter and summer SUHI effect amounts to around 3 and 6ºC (2 and 2.5ºC), respectively. UHI presents a less pronounced seasonal cycle at daytime and similar values at nighttime in relation to SUHI (Figs. 9a, b). At daytime (nighttime), the winter and summer UHI effect amounts to around 0.6 and 0.25ºC (2 and 2.75ºC), respectively. These results are similar to what Roy et al. (2020) found when studied the intensity and spatial extent of the UHI and SUHI over Paris using 1-km resolution observational datasets. Both studies agree in the SUHI and UHI annual cycle, although displaying some differences in their intensities, namely in daytime SUHI (4ºC in summer, 2ºC in winter) and nighttime UHI (2ºC in summer, 1ºC in winter) (Fig. 9b), both more intense in our study. These differences may arise from a number of reasons: the temporal ranges considered were different (2004-2018 in our study vs 2000-2016); Roy et al. (2020) considered a much larger rural area and the LST satellite data was retrieved from MODIS (which has higher spatial resolution than SEVIRI but at the cost of lower temporal resolution, with only two daily observations); finally, the T2M observations were generated from a gridded dataset developed at Météo-France while ours were obtained directly from two in-situ weather stations.

SFX-TEB overestimated the observed daytime SUHI effect during MAM, JJA, and SON, while underestimating this effect during DJF (Fig. 6a). This resulted in an annual average overestimation of the SUHI intensity for this simulation. In contrast,

285 ERA5 and SFX-ROCK largely underestimated the daytime SUHI effect over Paris (Fig. 6a). Indeed, misrepresentation of the urban radiative budget resulted in a nearly null SUHI effect in these simulations, as also illustrated by Fig. 2. One important result is that the magnitude of the daytime SUHI overestimation in SFX-TEB was smaller than the underestimation in ERA5 and SFX-ROCK, meaning that SFX-TEB improved the representation of the SUHI effect over all seasons (Fig. 6c). The statistical distribution of the daytime SUHI effect intensity was greatly improved in SFX-TEB

290 compared to ERA5 and SFX-TEB throughout all seasons (Fig. 6e) - the annual average S score for the daytime SUHI PDF was 14% in ERA5, 27% in SFX-ROCK, and 81% in SFX-TEB.

During nighttime, the generally similar performance of all simulations over all types of surfaces resulted in a similar performance in reproducing the average nighttime UHI effect throughout all seasons (Fig. 6b). All simulations - ERA5, SFX-ROCK, and SFX-TEB – underestimated the observed nighttime Parisian SUHI, with the differences amongst

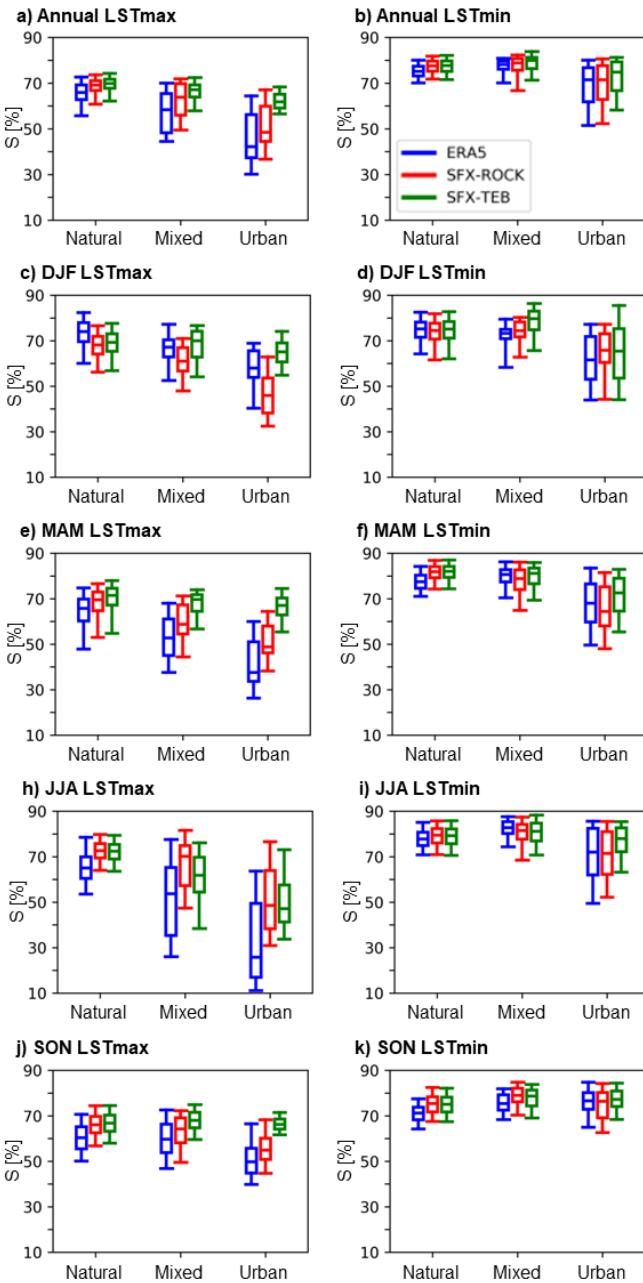
295 simulations being within observational uncertainty (Fig. 6d). Nonetheless, the results revealed a better ability of the SFX-TEB in reproducing the nighttime SUHI statistics, which showed an annual average S score of 39%, clearly above the 11% and 18% found for ERA5 and SFX-ROCK respectively (Fig. 6f). This improved representation of the nighttime SUHI statistics was found during all seasons, being largest during JJA and lowest during DJF (Fig. 6f).

300

Figure 5: Boxplots of S score computed over the 2004-2018 period from ERA5 (blue), SFX-ROCK (red), and SFX-TEB (green) for natural surfaces (grid boxes with urban fraction below 0.33), mixed surfaces (grid boxes with urban fraction between 0.33 and 0.66), and urban surfaces (grid boxes with urban fractions above 0.66). From top to bottom, the rows represent the S score computed for the full annual cycle, DJF, MAM, JJA, and SON.

305

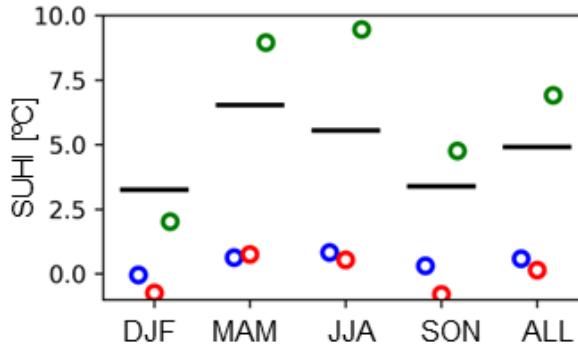
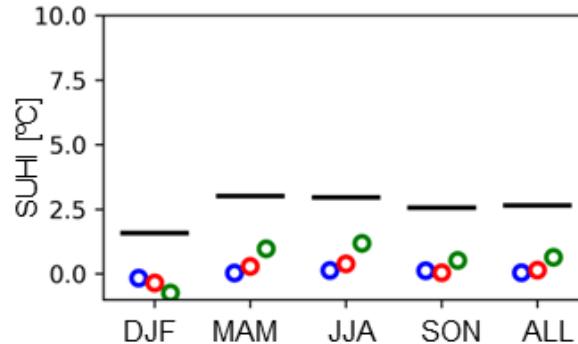
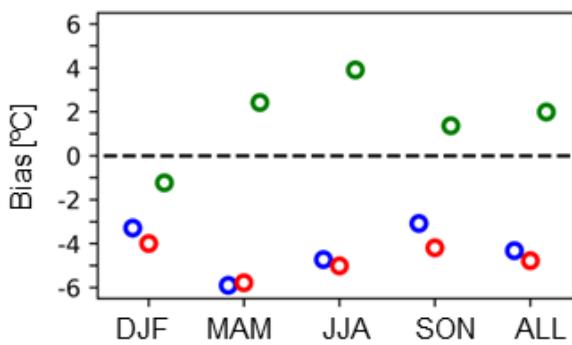
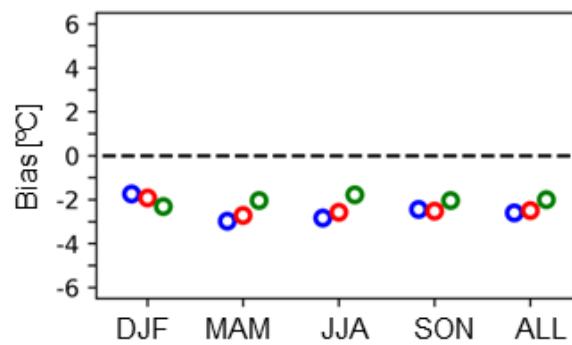
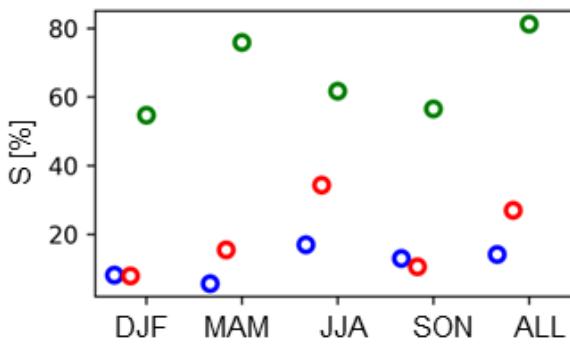
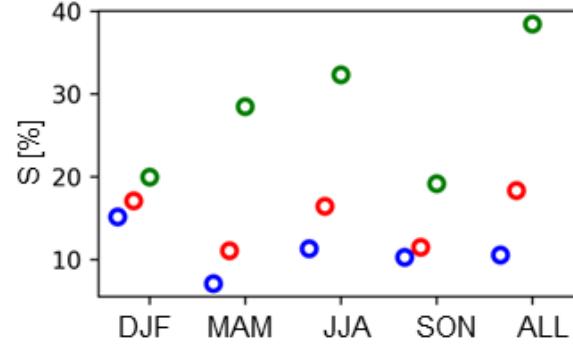



Figure 6: Annual and seasonal average Parisian SUHI magnitude computed over the 2004-2018 period during a) daytime and b) nighttime. The corresponding annual and seasonal average Bias for daytime is represented in c) and for nighttime in d), and the corresponding annual and seasonal S score is represented in e) for daytime and f) for nighttime. The LSA-SAF is represented by black lines, ERA5 by blue markers, SFX-ROCK by red markers, and SFX-TEB by green markers.

The observations show a relatively small range of the diurnal amplitude of the Parisian SUHI, which is largely overestimated by the SFX-TEB simulation (Fig. 7a). This reflects the overestimation of daytime LST and underestimation of nighttime LST over the urban grid boxes, as discussed above, in all seasons except winter (Fig. 7c). This issue is particularly

pronounced during summer (Fig. 7g). Nonetheless, SFX-TEB still represented an improvement compared to ERA5 and SFX-ROCK. On the one hand, ERA5 and SFX-ROCK failed to simulate any urban to rural contrast throughout the full annual cycle (Figs. 7a, c, e, g, and i). On the other hand, the annual average systematic errors in the simulated LST were lower in SFX-TEB compared to ERA5 and SFX-ROCK (Fig. 7b), although the differences only exceeded the observational uncertainty during daytime. The daytime reductions occurred mostly during DJF (Fig. 7d), MAM (Fig. 7f), and SON (Fig. 7j), while the daytime |Bias| values during JJA (Fig. 7h) and the nighttime |Bias| values during all seasons were within observational uncertainty.

3.2 Intercomparison of the simulated UHI over Paris

The T_{2m} in ERA5 did not show any evidence of the UHI over Paris during daytime (Fig. 8a) nor during nighttime (Fig. 8b). The UHI effect did not emerge in SFX-ROCK during day- nor nighttime (respectively Fig. 8c and Fig. 8d). In contrast, the SFX-TEB T_{2m} showed a signature of the UHI over Paris during daytime (Fig. 8e) and nighttime (Fig. 8f), in agreement with previous studies reporting the existence of the UHI effect over Paris (Lemonsu et al., 2014; Daniel et al., 2019). The enhanced performance in simulating the Paris UHI for SFX-TEB compared to ERA5 and SFX-ROCK was confirmed by the comparison against the UHI estimated from station observations (Fig. 9). During daytime, ERA5 and SFX-ROCK underestimated the observed Parisian UHI effect throughout all seasons, while SFX-TEB slightly overestimated the UHI (Fig. 9a), resulting in an overall reduction of the Bias in all seasons while also changing its signal (Fig. 9c). On the annual average, the systematic error magnitude reduced from 0.7ºC in ERA5 and SFX-ROCK to 0.2ºC in SFX-TEB. Moreover, SFX-TEB also improved the statistics of the daily UHI magnitude throughout all seasons (Fig. 9e), resulting in an overall S score value of 87%, well above the 59% and 56% respectively, found for ERA5 and SFX-ROCK.

The results also showed that SFX-TEB improved the simulation of the UHI during nighttime when compared to ERA5 and SFX-ROCK (Fig. 9b), reducing the Bias error over all seasons by more than 1ºC, resulting in an overall |Bias| value of 0.6ºC in SFX-TEB, 2.5ºC in ERA5, and 2.2ºC in SFX-ROCK (Fig. 9d). SFX-TEB largely improved the representation of nighttime UHI statistical distribution in ERA5 and SFX-ROCK throughout all seasons (Fig. 9f). The overall S score was 25% for ERA5, 31% for SFX-ROCK, and 79% for SFX-TEB.

The annual average diurnal cycle of the Paris UHI clearly illustrated the contrasting results from ERA5 and SFX-ROCK to SFX-TEB. SFX-TEB showed an annual averaged UHI intensity varying between +1.9ºC during the afternoon and night, and +0.1ºC during the morning (Fig. 10a), whilst ERA5 and SFX-ROCK showed a nearly zero UHI effect throughout the entire diurnal cycle. The afternoon and nighttime UHI effect in SFX-TEB was highest during MAM (Fig. 10c) but remained above +1.5ºC throughout all seasons (Fig. 10b-e), while the morning UHI was strongest during DJF (Fig. 10b) and its minimum reached close to zero values during MAM (Fig. 10c), JJA (Fig. 10d), and SON (Fig. 10e).

The relatively strong magnitude of the nighttime UHI in SFX-TEB corresponds to a significant improvement compared to ERA5 and SFX-ROCK. This result contrasts with the relatively small magnitude of the nighttime SUHI in SFX-TEB, which largely underestimates observations. The better performance of SFX-TEB in simulating the UHI compared to the SUHI was

likely related to the ability of the model to represent part of the nighttime urban canopy layer heating associated with prescribed anthropogenic heat fluxes, while the lack of land-atmosphere feedbacks inhibits this warmer canopy layer from affecting the LST.

350 The annual averaged SUHI measured as a difference between the urban and rural station locations under all-sky conditions computed from SFX-TEB peaked around 12h local time with a value of +5.0ºC, and reduced throughout the afternoon stabilizing at a value around +1.5ºC during the night and morning (Fig. 10f). The inability of ERA5 and SFX-ROCK to simulate the SUHI throughout the diurnal cycle was also clearly evidenced in Fig. 10f. The SUHI was strongest for all hours of the day during MAM (Fig. 10h) and JJA (Fig. 10i) and weakest during DJF (Fig. 10g) and SON (Fig. 10j). We highlight

355 the clear contrast between the UHI and SUHI diurnal cycles in Fig. 10. While the former peaked during late afternoon and night and decreased sharply during the morning, the latter showed opposite behavior peaking around midday and reaching the minimum during nighttime and late afternoon. This result was related to the thermal inertia of the canopy layer resulting in the well-known lag between surface and near-surface air warming.

However, the analysis of the annual averaged SUHI measured as a difference between the urban and rural station locations

360 under clear-sky conditions revealed significant differences in the average diurnal cycle simulated by SFX-TEB and observations (Fig. 10k). Indeed, SFX-TEB clearly overestimated the SUHI diurnal amplitude: the maximum amplitude of the SUHI average diurnal cycle was 2.9ºC in observations and 6.8ºC in SFX-TEB. This large overestimation resulted from SFX-TEB underestimation of the observed nighttime SUHI and overestimation during daytime. This effect was strongest during MAM (Fig. 10m) and JJA (Fig. 10n) and weakest during SON (Fig. 10o). During DJF (Fig. 10l), SFX-TEB underestimated

365 the observed clear-sky SUHI intensity throughout the entire diurnal cycle. Despite its limitations, SFX-TEB represents an improvement compared to ERA5 and SFX-ROCK which fail to simulate the Parisian SUHI effect throughout the entire diurnal cycle (Figs. 10 l-o).

The strong overestimation of the annual averaged daytime SUHI in SFX-TEB was associated with a sharp difference in the surface turbulent heat fluxes between the urban and rural station locations. The urban site showed lower average latent heat

370 flux (LH) values than the rural site by -75 Wm^{-2} around noon (Fig. 10p) and higher sensible heat flux (SH) values by around +75 Wm^{-2} (Fig. 10u). ERA5 and SFX-ROCK showed significantly smaller differences in LH and SH between urban and rural locations in agreement with their inability to simulate the UHI and SUHI. The differences in SH and LH between urban and rural locations in SFX-TEB were strongest during MAM (respectively Fig. 10r and 10w) and JJA (respectively Fig. 10s and 10x), and weaker during SON (respectively Fig. 10t and 10y) and DJF (respectively Fig. 10q and 10v). These results

375 suggest a direct link between the daytime overestimation of daytime SUHI in SFX-TEB with the lack of surface-atmosphere feedback in agreement with the recent findings of McNorton et al. (2021),

Finally, we noticed that the underestimation of the observed SUHI estimated as a point difference between the two station locations during night hours and the overestimation during morning and afternoon shown in Figs. 10k to 10o was coherent with the results presented in Fig. 6 for the daily minimum and daily maximum SUHI using the area-averaged SUHI

380 definition. Moreover, these results were also coherent with the overestimation of the daily maximum UHI and

underestimation of the daily minimum UHI presented in Fig. 9, including the daytime underestimation of both UHI and SUHI during DJF.

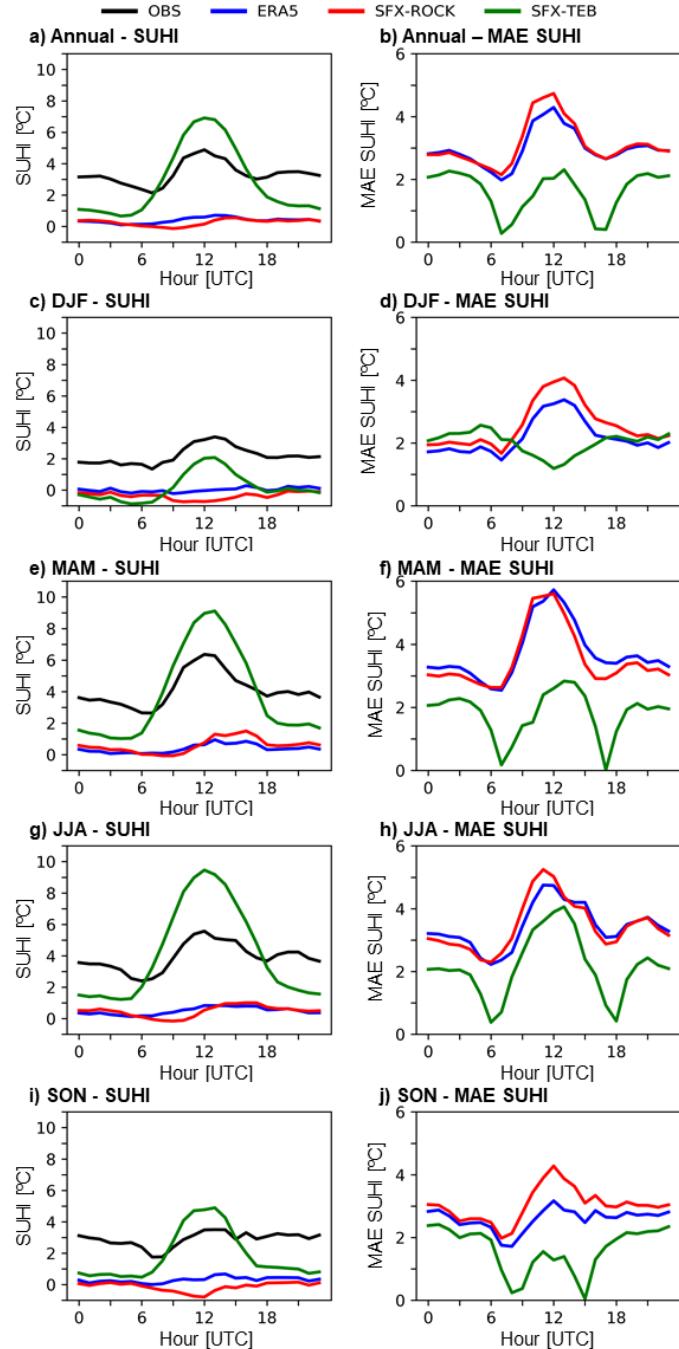



Figure 7: Average SUHI diurnal cycle computed over the 2004-2018 period for a) full year, c) DJF, e) MAM, g) JJA, and i) SON. The different colors represent LSA-SAF (black), ERA5 (blue), SFX-ROCK (red), and SFX-TEB (green). The respective |Bias| diurnal cycles are represented in b) full year, d) DJF, f) MAM, h) JJA, and j) SON using LSA-SAF as reference.

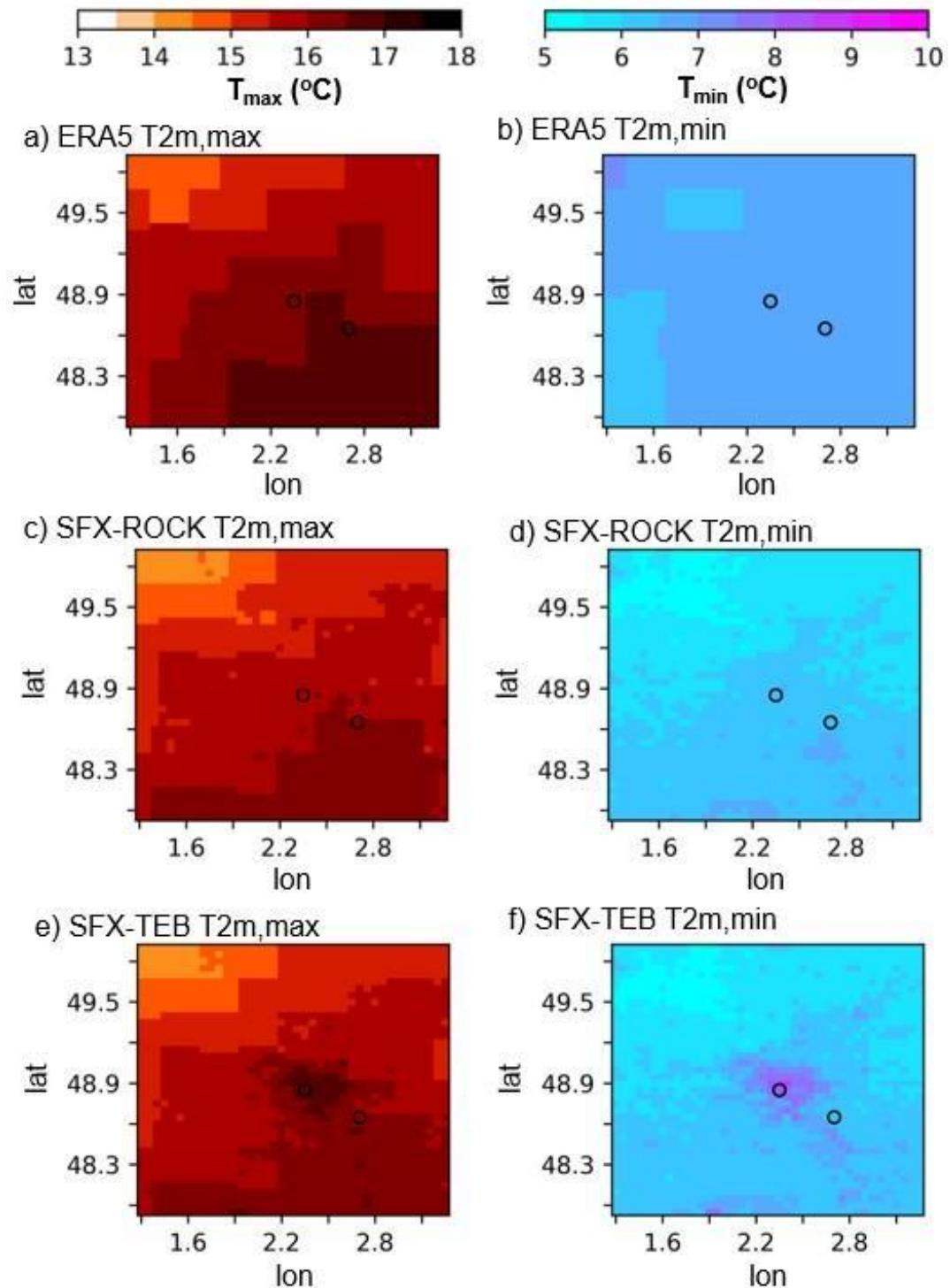



Figure 8: T$_{2m}$ averaged over the 2004-2018 period during daytime (left column) and nighttime (right column) over the study domain computed from ERA5 (a) and b)), SFX-ROCK (c) and d)), and SFX-TEB (e) and f)).

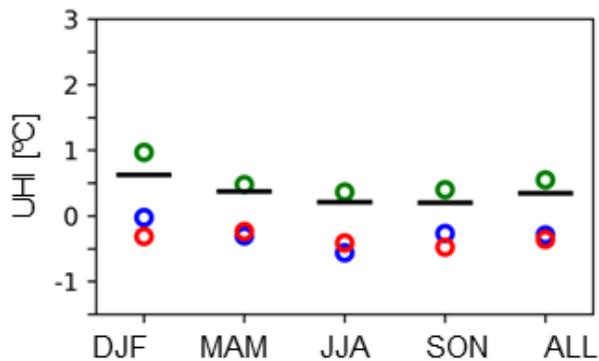
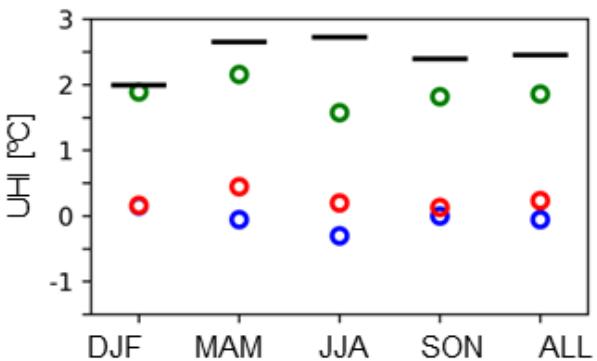
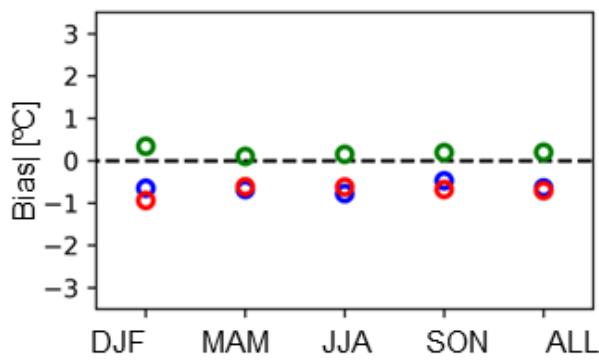
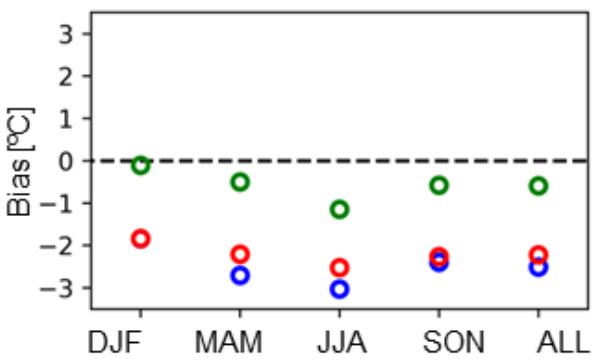
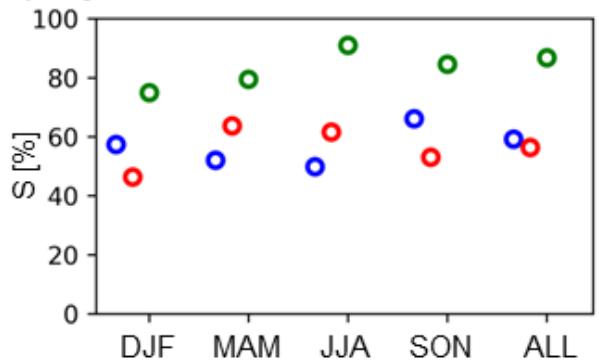
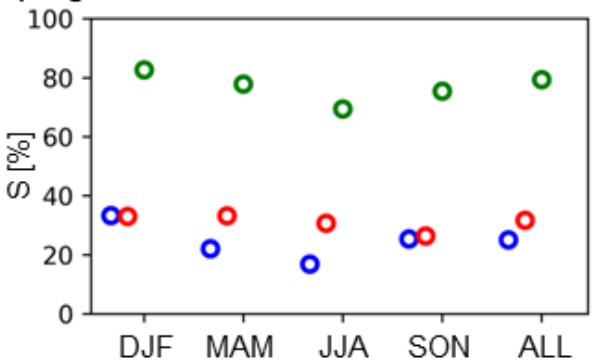



Figure 9: Annual and seasonal average Parisian UHI magnitude computed over the 2004-2018 period during a) daytime and b) nighttime. The corresponding annual and seasonal average Bias for daytime is represented in c) and for nighttime in d), and the corresponding annual and seasonal S score is represented in e) for daytime and f) for nighttime. The observed UHI is represented by black lines, ERA5 by blue markers, SFX-ROCK by red markers, and SFX-TEB by green markers.

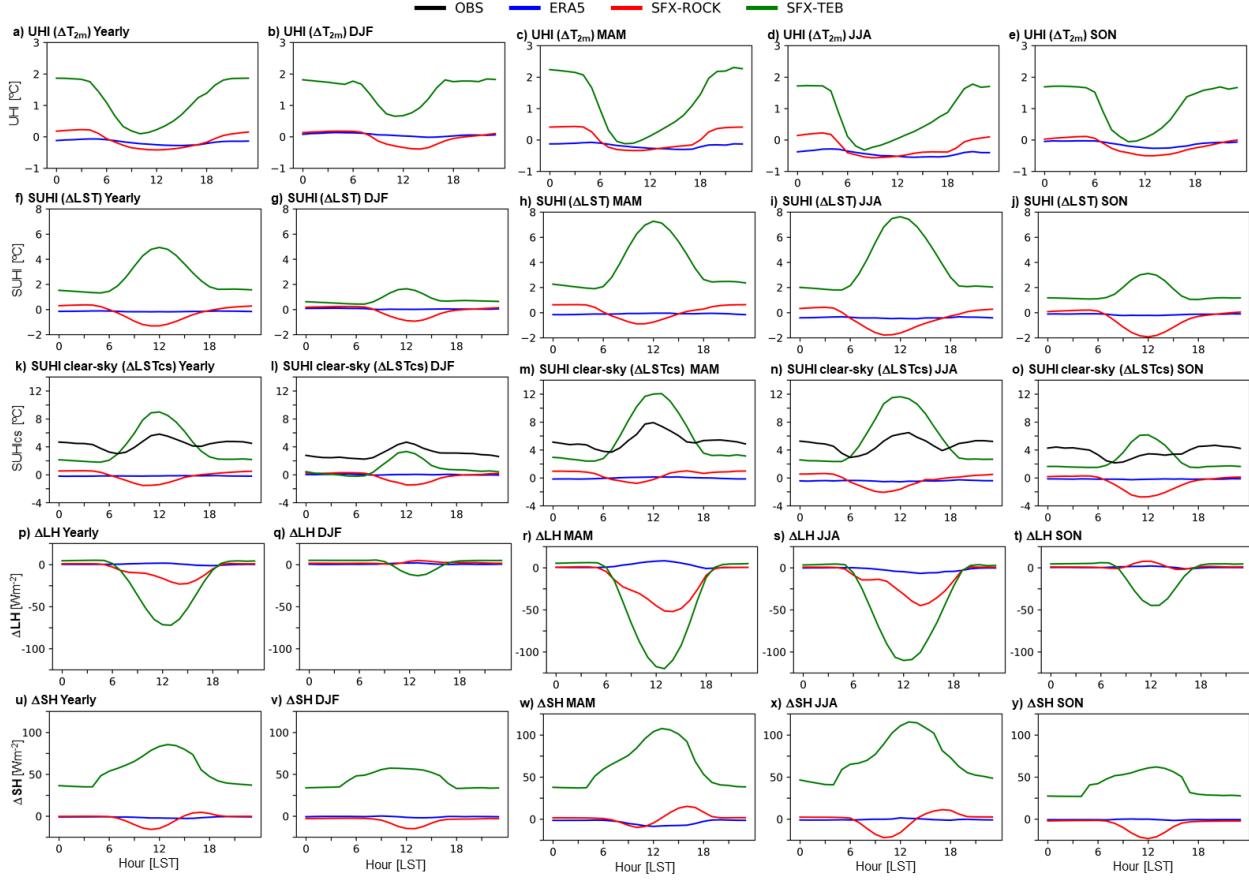



Figure 10: Average diurnal cycle of UHI (top row), SUHI under all-sky conditions (second row), SUHI under clear-sky conditions (third row), ΔLH (difference between surface latent heat flux between urban and rural areas, fourth row), ΔSH (difference between surface sensible heat flux between urban and rural areas, last row). The columns from left-to-right represent averages taken over full year, DJF, MAM, JJA, and SON. The average diurnal cycles were computed over the 2004-2018 period. The different colors represent LSA-SAF (black), ERA5 (blue), SFX-ROCK (red), and SFX-TEB (green).

4. Conclusions

We assessed the added value of the SURFEX offline simulations to downscale ERA5 reanalysis over urban areas, focusing on the urban heat island effect over the city of Paris. The relevance of this study is threefold. First, ERA5 is a widely used product for climate research and applications but an evaluation of its ability to represent urban climate is lacking. Second, it adds to recent works demonstrating the added value of the SURFEX offline downscaling framework for high-resolution computationally efficient urban climate simulation. Specifically, we leveraged on the hourly resolution of ERA5, the SURFEX simulations, and the LSA-SAF LST product to perform an unprecedented evaluation of the Parisian SUHI diurnal cycle, its key governing processes, and the ability of the different model setups to simulate them. Finally, the

intercomparison of long-duration high-spatial and temporal resolution datasets allow to explore the key sensitivities of the UHI and SUHI and their implications for the development of urban climate modelling.

Our results showed that ERA5 does not reproduce the observed UHI nor SUHI effects over Paris due to lack of representation of urban processes in the ECMWF IFS and, also, the relatively coarse grid resolution which is inappropriate for representing the highly heterogeneous urban environments. Increasing the grid resolution to the order of a few kilometers or less was found to be an important but not sufficient condition for simulating the urban heat island. Indeed, the high-resolution simulation employing the bulk bare rock approach to urban climate parameterization commonly used in large ensembles of ESMs, GCMs and RCMs (see, e.g., Daniel et al., 2019; Langendijk et al., 2019; Zhao et al., 2021) was unable to reproduce the observed UHI and SUHI. In contrast, the high-resolution simulation employing an UCM for urban parameterization showed significant added value in reproducing the Parisian SUHI during daytime when compared to ERA5 and SFX-ROCK.

Specifically, SFX-TEB reduced the systematic errors of daily maximum LST over urban pixels for all seasons (yearly average reduction of 4.2ºC and 3.9ºC respectively compared to ERA5 and SFX-ROCK). Moreover, SFX-TEB also improved the simulation of the observed daytime LST PDF over urban areas during all seasons compared to ERA5 and SFX-ROCK (with S values up to 20% higher). On the other hand, the differences in LST_{min} over urban surfaces and in LST_{max} and LST_{min} over natural surfaces were within observational uncertainty. In other words, SFX-TEB significantly improved the simulation of urban daytime LST, while displaying similar skill to ERA5 and SFX-ROCK during nighttime over urban areas and, also, over natural surfaces throughout the entire diurnal cycle.

SFX-TEB displayed the overall better performance in reproducing the observed daytime SUHI over Paris and its magnitude variability (with the respective S score increasing by 67% compared to ERA5). During nighttime, we found similar performance amongst ERA5, SFX-ROCK and SFX-TEB in reproducing the nighttime SUHI effect throughout all seasons, with a generalized underestimation of the observed SUHI magnitude. Notice, however, that SFX-TEB showed a slight improvement in simulating the nighttime SUHI spatial pattern and temporal variability compared to ERA5 and SFX-ROCK.

ERA5 and SFX-ROCK did not reproduce the Parisian UHI effect throughout the entire diurnal cycle. In contrast, SFX-TEB displayed significant added value in simulating the observed UHI during daytime and nighttime, resulting in an annual average bias magnitude reduction of 0.5ºC for daytime and more than 1.5ºC for nighttime compared to ERA5 and SFX-ROCK. The distribution of daily UHI variability was also improved in SFX-TEB, with the S score increased by roughly 30% for daytime and 50% for nighttime. The improved nighttime performance of SFX-TEB in simulating UHI relative to the SUHI may be explained by the ability of SFX-TEB to warm the urban canopy layer by anthropogenic heat releases combined with the lack of land-atmosphere feedbacks, which inhibits an LST response to the nighttime UHI. Finally, an analysis of the diurnal cycle of the simulated surface turbulent heat fluxes suggests that SFX-TEB overestimates the urban/rural contrasts in SH and LH during daytime (particularly during warmer months). However, ERA5 and SFX-ROCK consistently underestimate the T_{2m}, LST and turbulent flux contrast throughout the entire diurnal cycle. It is important to note

that uncertainties in remotely sensed LST due to its directional property should be negligible in this study due to the coarse spatial resolution of the SURFEX model.

In summary, we highlight the large potential of the offline SURFEX-TEB framework for urban climate projections given its ability to produce computationally efficient high-resolution climate projections with increased accuracy compared to ESMs, GCMs, and RCMs used in large ensemble datasets. Moreover, its relatively small computational cost allows to perform a large number of climate experiments to investigate the impact of city-scale climate adaptation and mitigation strategies under different future emission scenarios. This framework may be improved in the future by including simplified representations of key land-atmospheric feedbacks. Specifically, the possibility to improve the ability of this framework in reproducing the nighttime SUHI by including coherent dynamical corrections to the forcing temperature and downwelling longwave radiation fields based on the T_{2m} diagnostic in the previous timestep will be investigated in a subsequent work.

Code availability

The SURFEX modeling platform of Météo-France is open source and can be downloaded freely at http://www.umr-cnrm.fr/surfex/ (CNRM, 2016). It uses the CECILL-C license, a French equivalent to the L-GPL license (http://cecill.info/licences/Licence_CeCILL_V1.1-US.html; CEA-CNRS-Inria, 2013). It is updated at a relatively low frequency (every 3 to 6 months). If more frequent updates are needed – or if what is required is not in Open-SURFEX (DrHOOK, FA/LFI formats or GAUSSIAN grid) – you are invited to follow the procedure to get an SVN account and to access real-time modifications of the code (see the instructions in the first link). In this study, SURFEX's version 8.1 was used.

Data availability

ERA5 data can be obtained freely from the Copernicus Climate Change Service Information website (https://climate.copernicus.eu/, Copernicus Climate Change Service (C3S), 2019).

The LSA-SAF LST can be obtained freely from their website (https://landsaf.ipma.pt/, last access: 13 December 2021).

The NCDC CSOD weather station data can be obtained from their website (https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516, last access: 14 December 2021).

The considered fields (LST, T2m) from the simulations (SFX-TEB, SFX-ROCK) considered in the present study are freely available at https://doi.org/10.5281/zenodo.5780448 (Nogueira et al., 2021).

Author contributions

Conceptualization was done by MN, PMS and ED. LST data was acquired and processed by AH and SE. ERA5 processing was carried out by DL. Formal analysis and the original draft writing were done by MN. All authors contributed to writing, reviewing, and editing the paper.

475 **Competing interests**

The authors declare that they have no conflict of interest.

Acknowledgments

The authors wish to acknowledge the CONTROL and LEADING projects, both funded by FCT. Daniela C.A. Lima was funded by FCT under project LEADING (PTDC/CTA-MET/28914/2017). Pedro M.M. Soares would like to acknowledge
480 the financial support of FCT through project UIDB/50019/2020 – IDL and EEA-Financial Mechanism 2014-2021 and the Portuguese Environment Agency through Pre-defined Project-2 National Roadmap for Adaptation XXI (PDP-2). Frederico Johannsen was funded by FCT under the research grant UI/BD/151498/2021. The authors would also like to acknowledge the financial support by FCT through project UIDB/50019/2020 – Instituto Dom Luiz. Finally, all authors acknowledge the ECMWF for producing ERA5.

485 **References**

Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Blyth, E., Bousserez, N., Boussetta, S., Brown, A., Buizza, R., Buontempo, C., Chevallier, F., Choulga, M., Cloke, H., Cronin, M. F., Dahoui, M., De Rosnay, P., Dirmeyer, P. A., Drusch, M., Dutra, E., Ek, M. B., Gentine, P., Hewitt, H., Keeley, S. P. E., Kerr, Y., Kumar, S., Lupu, C., Mahfouf, J.-F., McNorton, J., Mecklenburg, S., Mogensen, K., Muñoz-Sabater, J., Orth, R., Rabier, F., Reichle, R., Ruston,
490 B., Pappenberger, F., Sandu, I., Seneviratne, S. I., Tietsche, S., Trigo, I. F., Uijlenhoet, R., Wedi, N., Woolway, R. I. and Zeng, X.: Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review, Remote Sensing, 10(12), 2038, doi:10.3390/rs10122038, 2018.

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J., Adler, R. F., and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, Hydrol. Earth
495 Syst. Sci., 23, 207–224, https://doi.org/10.5194/hess-23-207-2019, 2019.

Beljaars, A. C. M., Viterbo, P., Miller, M. J., and Betts, A. K.: The Anomalous Rainfall over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies, Mon. Weather Rev., 124, 362–383, https://doi.org/10.1175/1520-0493(1996)124<0362:TAROTU>2.0.CO;2, 1996.

Broadbent, A. M., Coutts, A. M., Tapper, N. J., and Demuzere, M.: The cooling effect of irrigation on urban microclimate during heatwave conditions, Urban Clim., 23, 309–329, https://doi.org/10.1016/j.uclim.2017.05.002, 2018.

Broadbent, M. A., Krayenhoff, S. E., and Georgescu, M.: The motley drivers of heat and cold exposure in 21st century US cities, Proceedings of the National Academy of Sciences, 117, 21108–21117, https://doi.org/10.1073/pnas.2005492117, 2020.

Calvet, J.-C., Noilhan, J., Roujean, J.-L., Bessemoulin, P., Cabelguenne, M., Olioso, A., and Wigneron, J.-P.: An interactive vegetation SVAT model tested against data from six contrasting sites, Agric. For. Meteorol., 92, 73–95, https://doi.org/10.1016/S0168-1923(98)00091-4, 1998.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., and Zhang, C.: The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems, Int. J. Climatol., 31, 273–288, https://doi.org/10.1002/joc.2158, 2011.

Clinton, N. and Gong, P.: MODIS detected surface urban heat islands and sinks: Global locations and controls, Remote Sensing of Environment, 134, 294–304, https://doi.org/10.1016/j.rse.2013.03.008, 2013.

Conlon, K., Monaghan, A., Hayden, M., and Wilhelmi, O.: Potential Impacts of Future Warming and Land Use Changes on Intra-Urban Heat Exposure in Houston, Texas, PLoS One, 11, 1–19, https://doi.org/10.1371/journal.pone.0148890, 2016.

Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS), available at: https://cds.climate.copernicus.eu/ (last access: 13 December 2021), 2019.

Daniel, M., Lemonsu, A., and Viguié, V.: Role of watering practices in large-scale urban planning strategies to face the heat-wave risk in future climate, Urban Clim., 23, 287–308, https://doi.org/10.1016/j.uclim.2016.11.001, 2018.

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

Davin, E. L., Rechid, D., Breil, M., Cardoso, R. M., Coppola, E., Hoffmann, P., Jach, L. L., Katragkou, E., de Noblet-Ducoudré, N., Radtke, K., Raffa, M., Soares, P. M. M., Sofiadis, G., Strada, S., Strandberg, G., Tölle, M. H., Warrach-Sagi, K., and Wulfmeyer, V.: Biogeophysical impacts of forestation in Europe: first results from the LUCAS (Land Use and Climate Across Scales) regional climate model intercomparison, Earth Syst. Dynam., 11, 183–200, https://doi.org/10.5194/esd-11-183-2020, 2020.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N.,

and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc.,137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Deilami, K., Kamruzzaman, M., and Liu, Y.: Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures, International Journal of Applied Earth Observation and Geoinformation, 67, 30–42, https://doi.org/10.1016/j.jag.2017.12.009, 2018.

de Munck, C., Pigeon, G., Masson, V., Meunier, F., Bousquet, P., Tréméac, B., Merchat, M., Poeuf, P., and Marchadier, C.: How much can air conditioning increase air temperatures for a city like Paris, France?, International Journal of Climatology, 33, 210–227, https://doi.org/10.1002/joc.3415, 2013.

Ermida, S. L., Trigo, I. F., DaCamara, C. C., Jiménez, C., and Prigent, C.: Quantifying the Clear-Sky Bias of Satellite Land Surface Temperature Using Microwave-Based Estimates, Journal of Geophysical Research: Atmospheres, 124, 844–857, https://doi.org/10.1029/2018JD029354, 2019.

Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., and Gottsche, F.-M.: Quantifying the Uncertainty of Land Surface Temperature Retrievals From SEVIRI/Meteosat, IEEE Transactions on Geoscience and Remote Sensing, 48, 523–534, https://doi.org/10.1109/TGRS.2009.2027697, 2010.

Garuma, G. F.: Review of urban surface parameterizations for numerical climate models, Urban Climate, 24, 830–851, https://doi.org/10.1016/j.uclim.2017.10.006, 2018.

Georgescu, M., Morefield, E. P., Bierwagen, G. B., and Weaver, P. C.: Urban adaptation can roll back warming of emerging megapolitan regions, Proceedings of the National Academy of Sciences, 111, 2909–2914, https://doi.org/10.1073/pnas.1322280111, 2014.

Gibelin, A.-L., Calvet, J.-C., Roujean, J.-L., Jarlan, L., and Los, S. O.: Ability of the land surface model ISBA-A-gs to simulate leaf area index at the global scale: Comparison with satellites products, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006691, 2006.

Hamdi, R., Degrauwe, D., and Termonia, P.: Coupling the Town Energy Balance (TEB) Scheme to an Operational Limited-Area NWP Model: Evaluation for a Highly Urbanized Area in Belgium, Weather and Forecasting, 27, 323–344, https://doi.org/10.1175/WAF-D-11-00064.1, 2012.

Hamdi, R., Giot, O., Troch, R. D., Deckmyn, A., and Termonia, P.: 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, 2015.

Hamdi, R., Kusaka, H., Doan, Q.-V., Cai, P., He, H., Luo, G., Kuang, W., Caluwaerts, S., Duchêne, F., Van Schaeybroek, B., and Termonia, P.: The State-of-the-Art of Urban Climate Change Modeling and Observations, Earth Systems and Environment, 4, 631–646, https://doi.org/10.1007/s41748-020-00193-3, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G.,

de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Imhoff, M. L., Zhang, P., Wolfe, R. E., and Bounoua, L.: Remote sensing of the urban heat island effect across biomes in the continental USA, Remote Sensing of Environment, 114, 504–513, https://doi.org/10.1016/j.rse.2009.10.008, 2010.

570 IPCC: 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, https://doi.org/10.1017/CBO9781107415324, 2014.

Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., Benestad, R., Boberg, F., Buonomo, E., Cardoso, R. M., Casanueva, A., Christensen, O. B., Christensen, J. H., Coppola, E., De Cruz, L., Davin, E. L., Dobler, A.,
575 Domínguez, M., Fealy, R., Fernandez, J., Gaertner, M. A., García-Díez, M., Giorgi, F., Gobiet, A., Goergen, K., Gómez-Navarro, J. J., Alemán, J. J. G., Gutiérrez, C., Gutiérrez, J. M., Güttler, I., Haensler, A., Halenka, T., Jerez, S., Jiménez-Guerrero, P., Jones, R. G., Keuler, K., Kjellström, E., Knist, S., Kotlarski, S., Maraun, D., van Meijgaard, E., Mercogliano, P., Montávez, J. P., Navarra, A., Nikulin, G., de Noblet-Ducoudré, N., Panitz, H.-J., Pfeifer, S., Piazza, M., Pichelli, E., Pietikäinen, J.-P., Prein, A. F., Preuschmann, S., Rechid, D., Rockel, B., Romera, R., Sánchez, E., Sieck, K., Soares, P. M.
580 M., Somot, S., Srnec, L., Sørland, S. L., Termonia, P., Truhetz, H., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community, Regional Environmental Change, 20, 51, https://doi.org/10.1007/s10113-020-01606-9, 2020.

Johannsen, F., Ermida, S., Martins, J. P. A., Trigo, I. F., Nogueira, M., and Dutra, E.: Cold Bias of ERA5 Summertime Daily Maximum Land Surface Temperature over Iberian Peninsula, Remote Sensing, 11(21), https://doi.org/10.3390/rs11212570,
585 2019.

Kaloustian, N. and Diab, Y.: Effects of urbanization on the urban heat island in Beirut, Urban Climate, 14, 154–165, https://doi.org/10.1016/j.uclim.2015.06.004, 2015.

Konstantinov, P., Varentsov, M. and Esau, I.: A high density urban temperature network deployed in several cities of Eurasian Artic, Environmental Research Letters, 13(7), https://doi.org/10.1088/1748-9326/aacb84, 2018.

590 Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y. C., Taylor, C. M., Verseghy, D., Vasic, R., Xue, Y., and Yamada, T.: Regions of strong coupling between soil moisture and precipitation., Science, 305, 1138–1140, https://doi.org/10.1126/science.1100217, 2004.

Krayenhoff, E. S., Moustaoui, M., Broadbent, A. M., Gupta, V., and Georgescu, M.: Diurnal interaction between urban
595 expansion, climate change and adaptation in US cities, Nature Climate Change, 8, 1097–1103, https://doi.org/10.1038/s41558-018-0320-9, 2018.

Kusaka, H., Hara, M., and Takane, Y.: Urban Climate Projection by the WRF Model at 3-km Horizontal Grid Increment: Dynamical Downscaling and Predicting Heat Stress in the 2070's August for Tokyo, Osaka, and Nagoya Metropolises, Journal of the Meteorological Society of Japan. Ser. II, 90B, 47–63, https://doi.org/10.2151/jmsj.2012-B04, 2012.

600 Langendijk, G. S., Rechid, D. and Jacob, D.: 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, doi:10.3390/atmos10120730, 2019.

Le Moigne, P. (Ed.): SURFEX Scientific Documentation, V8.1, available at: http://www.umr-cnrm.fr/surfex/IMG/pdf/surfex_scidoc_v8.1.pdf (last access: 21 December 2021), 2018.

605 Lemonsu, A., Kounkou-Arnaud, R., Desplat, J., Salagnac, J.-L., and Masson, V.: Evolution of the Parisian urban climate under a global changing climate, Climatic Change, 116, 679–692, https://doi.org/10.1007/s10584-012-0521-6, 2013.

Lemonsu A., Beaulant A. L., Somot S., and Masson V.: Evolution of heat wave occurrence over the Paris basin (France) in the 21st century, Clim Res, 61, 75–91, https://doi.org/10.3354/cr01235, 2014.

Lemonsu, A., Viguié, V., Daniel, M., Masson, V.: Vulnerability to heat waves: Impact of urban expansion scenarios on 610 urban heat island and heat stress in Paris (France), Urban Climate, 14, 586–605, https://doi.org/10.1016/j.uclim.2015.10.007, 2015.

Le Roy, B., Lemonsu, A., Kounkou-Arnaud, R., Brion, D., and Masson, V.: Long time series spatialized data for urban climatological studies: A case study of Paris, France, International Journal of Climatology, 40, 3567–3584, https://doi.org/10.1002/joc.6414, 2020.

615 Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sensing of Environment, 131, 14–37, https://doi.org/10.1016/j.rse.2012.12.008, 2013.

Manoli, G., Fatichi, S., Schläpfer, M., Yu, K., Crowther, T. W., Meili, N., Burlando, P., Katul, G. G., and Bou-Zeid, E.: Magnitude of urban heat islands largely explained by climate and population, Nature, 573, 55–60, 620 https://doi.org/10.1038/s41586-019-1512-9, 2019.

Masson, V.: A Physically-Based Scheme For The Urban Energy Budget In Atmospheric Models, Boundary-Layer Meteorology, 94, 357–397, https://doi.org/10.1023/A:1002463829265, 2000.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., 625 Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: 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, 2013.

630 Masson, V., Lemonsu, A., Hidalgo, J., and Voogt, J.: Urban Climates and Climate Change, Annu. Rev. Environ. Resour., 45, 411–444, https://doi.org/10.1146/annurev-environ-012320-083623, 2020.

McNorton, J. R., Arduini, G., Bousserez, N., Agustí-Panareda, A., Balsamo, G., Boussetta, S., Choulga, M., Hadade, I., and Hogan, R. J.: An Urban Scheme for the ECMWF Integrated Forecasting System: Single-Column and Global Offline

Application, Journal of Advances in Modeling Earth Systems, 13, e2020MS002375, https://doi.org/10.1029/2020MS002375, 2021.

Muller, C. L., Chapman, L., Grimmond, C. S. B., Young, D. T., and Cai, X.: Sensors and the city: a review of urban meteorological networks, International Journal of Climatology, 33, 1585–1600, https://doi.org/10.1002/joc.3678, 2013.

Nogueira, M.: Inter-comparison of ERA-5, ERA-interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences, Journal of Hydrology, 583, 124632, https://doi.org/10.1016/j.jhydrol.2020.124632, 2020.

Nogueira, M. and Soares, P. M. M.: A surface modelling approach for attribution and disentanglement of the effects of global warming from urbanization in temperature extremes: application to Lisbon, 14, 114023, https://doi.org/10.1088/1748-9326/ab465f, 2019.

Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I. F., Ermida, S. L., Martins, J. P. A., and Dutra, E.: Role of vegetation in representing land surface temperature in the CHTESSEL (CY45R1) and SURFEX-ISBA (v8.1) land surface models: a case study over Iberia, Geosci. Model Dev., 13, 3975–3993, https://doi.org/10.5194/gmd-13-3975-2020, 2020a.

Nogueira, M., Lima, D. C. A., and Soares, P. M. M.: An integrated approach to project the future urban climate response: Changes to Lisbon's urban heat island and temperature extremes, Urban Climate, 34, 100683, https://doi.org/10.1016/j.uclim.2020.100683, 2020b.

Nogueira, M., Boussetta, S., Balsamo, G., Albergel, C., Trigo, I. F., Johannsen, F., Miralles, D. G., and Dutra, E.: Upgrading Land-Cover and Vegetation Seasonality in the ECMWF Coupled System: Verification With FLUXNET Sites, METEOSAT Satellite Land Surface Temperatures, and ERA5 Atmospheric Reanalysis, Journal of Geophysical Research: Atmospheres, 126, e2020JD034163, https://doi.org/10.1029/2020JD034163, 2021.

Nogueira, M., Hurduc, A., Ermida, S., Lima, D.C., Soares, P.M., Johannsen, F., and Dutra, E.: Assessment of the Paris urban heat island in ERA5 and offline SURFEX-TEB (v8.1) simulations using METEOSAT land surface temperature product (Version 1), Zenodo, https://doi.org/10.5281/zenodo.5780448, 2021.

Oke, T. R.: City size and the urban heat island, Atmospheric Environment (1967), 7, 769–779, https://doi.org/10.1016/0004-6981(73)90140-6, 1973.

Oke, T. R.: The energetic basis of the urban heat island, Quarterly Journal of the Royal Meteorological Society, 108, 1–24, https://doi.org/10.1002/qj.49710845502, 1982.

Oke, T.R.: Boundary layer climates, 2nd ed, Quarterly Journal of the Royal Meteorological Society, Wiley, London, https://doi.org/10.1002/qj.49711448412, 1987.

Peng, S., Piao, S., Ciais, P., Friedlingstein, P., Ottle, C., Bréon, F.-M., Nan, H., Zhou, L., and Myneni, R. B.: Surface Urban Heat Island Across 419 Global Big Cities, Environ. Sci. Technol., 46, 696–703, https://doi.org/10.1021/es2030438, 2012.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., and McAneney, J.: Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions, Journal of Climate, 20, 4356–4376, https://doi.org/10.1175/JCLI4253.1, 2007.

Belmonte Rivas, M. and Stoffelen, A.: Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT, Ocean Sci., 15, 831–852, https://doi.org/10.5194/os-15-831-2019, 2019.

670 Roth, M., Oke, T. R., and Emery, W. J.: Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology, International Journal of Remote Sensing, 10, 1699–1720, https://doi.org/10.1080/01431168908904002, 1989.

Salamanca, F., Krpo, A., Martilli, A., and Clappier, A.: A new building energy model coupled with an urban canopy parameterization for urban climate simulations—part I. formulation, verification, and sensitivity analysis of the model, Theoretical and Applied Climatology, 99, 331, https://doi.org/10.1007/s00704-009-0142-9, 2009.

675 Salamanca, F., Martilli, A., Tewari, M., and Chen, F.: A Study of the Urban Boundary Layer Using Different Urban Parameterizations and High-Resolution Urban Canopy Parameters with WRF, Journal of Applied Meteorology and Climatology, 50, 1107–1128, https://doi.org/10.1175/2010JAMC2538.1, 2011.

Sarkar, A. and De Ridder, K.: The Urban Heat Island Intensity of Paris: A Case Study Based on a Simple Urban Surface Parametrization, Boundary-Layer Meteorology, 138, 511–520, https://doi.org/10.1007/s10546-010-9568-y, 2011.

680 Schoetter, R., Kwok, Y. T., de Munck, C., Lau, K. K. L., Wong, W. K., and Masson, V.: Multi-layer coupling between SURFEX-TEB-v9.0 and Meso-NH-v5.3 for modelling the urban climate of high-rise cities, Geosci. Model Dev., 13, 5609–5643, https://doi.org/10.5194/gmd-13-5609-2020, 2020.

Schwarz, N., Lautenbach, S., and Seppelt, R.: Exploring indicators for quantifying surface urban heat islands of European cities with MODIS land surface temperatures, Remote Sensing of Environment, 115, 3175–3186, 685 https://doi.org/10.1016/j.rse.2011.07.003, 2011.

Stewart, I. D.: A systematic review and scientific critique of methodology in modern urban heat island literature, International Journal of Climatology, 31, 200–217, https://doi.org/10.1002/joc.2141, 2011.

Stewart, I. D. and Oke, T. R.: Local Climate Zones for Urban Temperature Studies, Bulletin of the American Meteorological Society, 93, 1879–1900, https://doi.org/10.1175/BAMS-D-11-00019.1, 2012.

690 Trigo, I. F., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., and Sandu, I.: Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling, 120, 12,096-12,111, https://doi.org/10.1002/2015JD023812, 2015.

Trusilova, K., Schubert, S., Wouters, H., Früh, B., Grossman-Clarke, S., Demuzere, M., and Becker, P.: The urban land use in the COSMO-CLM model: a comparison of three parameterizations for Berlin, 25, 231–244, 695 https://doi.org/10.1127/metz/2015/0587, 2016.

Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de-Pison, F. J., Kaspar, F., and Sanz-Garcia, A.: Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data, Solar Energy, 164, 339–354, https://doi.org/10.1016/j.solener.2018.02.059, 2018.

Viguié, V., Lemonsu, A., Hallegatte, S., Beaulant, A.-L., Marchadier, C., Masson, V., Pigeon, G., and Salagnac, J.-L.: Early

700   adaptation to heat waves and future reduction of air-conditioning energy use in Paris, 15, 075006, https://doi.org/10.1088/1748-9326/ab6a24, 2020.

World Health Organization: Global status report on alcohol and health 2018, World Health Organization, Geneva, 2018.

Zhao, L., Lee, X., Smith, R. B., and Oleson, K.: Strong contributions of local background climate to urban heat islands, Nature, 511, 216–219, https://doi.org/10.1038/nature13462, 2014.

705   Zhao, L., Oleson, K., Bou-Zeid, E., Krayenhoff, E. S., Bray, A., Zhu, Q., Zheng, Z., Chen, C., and Oppenheimer, M.: Global multi-model projections of local urban climates, Nature Climate Change, 11, 152–157, https://doi.org/10.1038/s41558-020-00958-8, 2021.

Zhou, B., Rybski, D., and Kropp, J. P.: On the statistics of urban heat island intensity, Geophysical Research Letters, 40, 5486–5491, https://doi.org/10.1002/2013GL057320, 2013.

710   Zhou, J., Yu, B. and Qin, J.: Multi-Level Spatial Analysis for Change Detection of Urban Vegetation at Individual Tree Scale, Remote Sensing, 6(9), 9086–9103, doi:10.3390/rs6099086, 2014.

Zhou, B., Rybski, D., and Kropp, J. P.: The role of city size and urban form in the surface urban heat island, Scientific Reports, 7, 4791, https://doi.org/10.1038/s41598-017-04242-2, 2017.