Calibrating and validating the InVEST urban cooling model: Case studies in France and the United States

Perrine Hamel[1,2], Martí Bosch[3], Léa Tardieu[4,5], Aude Lemonsu[6], Cécile de Munck[6], Chris Nootenboom[7], Vincent Viguié[8], Eric Lonsdorf [8], James A. Douglass[9], Richard S. Sharp[10]

[1]Asian School of the Environment, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
[2]Earth Observatory of Singapore, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
[3]Urban and Regional Planning Community, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[4]TETIS, INRAE, AgroParisTech, CIRAD, CNRS, Univ Montpellier, Montpellier, France
[5]CIRED, Ecole des Ponts ParisTech, AgroParisTech, Cirad, CNRS, EHESS, Université Paris-Saclay, F-94736, Nogent-sur-Marne, France
[6]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[7]Ecognosis, USA
[8]The Department of Environmental Sciences, Emory College, GA, 30322 GA, USA
[9]Natural Capital Project, Dept. of Biology and Woods Institute for the Environment, Stanford University, CA 94305 USA
[10]Spring, 5455 Shafter Avenue, Oakland, CA 94618USA

Correspondence to: Perrine Hamel (Perrine.hamel@ntu.edu.sg)

Abstract. Understanding the cooling service provided by vegetation in cities is important to inform urban policy and planning. However, the performance of decision-support tools estimating heat mitigation for urban greening strategies is not systematically evaluated. Here, we further develop a calibration algorithm and evaluate the performance of the Urban Cooling model developed within the open-source InVEST (Integrated Valuation of Ecosystem Services and Tradeoffs) software. The Urban Cooling model estimates air temperature reduction due to vegetation based on four predictors: shade provision, evapotranspiration, albedo, and building density and was designed for data-rich and data-scarce situations. We apply the calibration algorithm and evaluate the model in two case studies (Paris, France, and Minneapolis-St Paul, USA) by examining the spatial correlation between InVEST predictions and reference temperature data at 1-km horizontal resolution. In both case studies, model performance was high for nighttime air temperatures, which is an important indicator of human wellbeing. After calibration, we found a medium performance for surface temperatures during daytime but a low performance for daytime air temperatures in both case studies, which may be due to model and data limitations. We illustrate the model adequacy for urban planning by testing its ability to simulate a green infrastructure scenario in the Paris case study. The predicted air temperature change compared well with that of an alternative physics-based model ($r^2=0.55$ and $r^2=0.85$, for air daytime and nighttime temperatures, respectively). Finally, we discuss opportunities and challenges for the use of such parsimonious decision-support tools, highlighting their importance to mainstream ecosystem services information in urban planning.

1 Introduction

35 The urban heat island is increasingly documented, varying in European cities from 1 to 10°C with an average of 6°C for a sample of 110 European cities (Santamouris, 2016). The phenomenon involves an increase in air and surface temperatures in urban areas due to the modification of the energy budget (Oke, 1982). This has direct health, economic and energy consumption implications (Lehmann, 2014; Santamouris, 2020) as excessive heat has been associated with increases in energy consumption for cooling purposes, increases in ground level ozone and particulate matter concentrations, and hospital admissions due to

40 cardiovascular conditions (Gosling et al., 2009; Hémon & Jougla, 2004; Lai & Cheng, 2009; Reid et al., 2012; Santamouris, 2015; Viguié et al., 2020; Wang et al., 2017). To reduce these effects, policy-makers and urban planners are increasingly turning to blue-green infrastructure (e.g., street trees, green roofs, urban parks), a cost-effective option for urban cooling that also produces multiple co-benefits (Bolund & Hunhammar, 1999; Corburn, 2009; Cortinovis & Geneletti, 2019; Rosenzweig et al., 2006; Villanueva-Solis, 2017).

45 Blue-green infrastructure influences air temperatures and thermal comfort at several scales. At a local scale – a tree or building, shading can reduce air temperature under the canopy (Kroeger et al., 2018; McDonald et al., 2016; Shashua-Bar & Hoffman, 2000). Street trees can also indirectly improve pedestrian comfort and reduce the use of air conditioning in neighboring buildings, thus avoiding additional heat generation (Viguié et al., 2020). Green roofs and walls change the heat and energy balance of buildings: by absorbing incident solar radiation to support biological functions, vegetation acts as a screen and

50 reduces seasonal temperature variations – although to a limited extent compared to other insulating materials (Eumorfopoulou & Kontoleon, 2009). At a larger scale, urban parks provide an "oasis effect" reducing air temperatures by up to 6°C (Eliasson, 1996; Jauregui, 1990; Kroeger et al., 2018; Potchter et al., 2006; Spronken-Smith & Oke, 1999; Yu et al., 2020; Ziter et al., 2019). The effect is influenced by park size (Cao et al., 2010; Yu et al., 2020), composition (Potchter et al., 2006), and local climatic conditions (Shashua-Bar & Hoffman, 2000; Yu et al., 2020). In a systematic literature review, Bowler et al. (2010)

55 showed that parks larger than 2-3 ha are systematically cooler than the rest of the city. Recent work by Wong et al. (2021) suggests a lower threshold of 1 ha. For such large parks, the cooling capacity depends mainly on evapotranspiration, and can extend to 800 m, although more frequently closer to 100 m (Kroeger et al., 2018; Wong et al., 2021).

Quantitative estimates of the cooling service provided by vegetation come from two main areas. Research in urban climatology examines the physical processes contributing to air cooling (Phelan et al., 2015). These studies often involve complex

60 numerical models –meso-scale or micro-scale (Lemonsu et al., 2012; Meili et al., 2020), which focus on one or several processes explaining microclimate and increasingly incorporating vegetation effects (Bartesaghi Koc et al., 2018). Such models require long development and calibration processes and therefore significant human and time resources. The same applies to geostatistical models that require large geospatial datasets (e.g., Cheung et al., 2021). On the other hand, an increasing body of literature in the fields of urban ecology and ecosystem services science examines the cooling service through

65 indicators associated with land use (Derkzen et al., 2015; Farrugia et al., 2013; Larondelle & Haase, 2013; Nedkov et al., 2017; Zardo et al., 2017), as reviewed in previous work (Hamel et al., 2021). These models further simplify the physical processes

underlying the cooling dynamics, the effect of the morphological parameters of the buildings, and the daily or seasonal temperature variations, often by attributing a "cooling capacity" to a broadly defined land use type.

Cooling service models routinely measure the cooling effect of vegetation with land surface temperature, which is readily accessible from satellite data (Manoli et al., 2019; Zhao et al., 2014). Surface temperature influences thermal comfort and is often correlated with air temperature, making it an interesting proxy for microclimate studies. However, this temperature alone does not suffice to estimate economic or health implications of the urban heat island (Martilli et al., 2020; Venter et al., 2021), reducing the relevance of models estimating land surface temperature in policy making.

These model limitations lead to a lack of decision-support tools that accurately quantify the cooling effect of vegetation on air temperatures and therefore its impact in socio-economic terms. A recent review of open-source tools found that only a few ecosystem services software tools were adequate for general urban planning purposes—i.e. applicable in any geography, and flexible enough to represent different types of blue-green infrastructure and decision contexts (Hamel et al., 2021). InVEST (Integrated Valuation of ecosystem services and tradeoffs) is one of the most popular tools developed to support ecosystem services assessments by quantifying and mapping the benefits provided by blue-green infrastructure in urban or non-urban environments (Natural Capital Project, 2022). The Urban Cooling model within InVEST, developed by the author team, addresses some of the limitations stated above by estimating air temperature (instead of land surface temperature only), being applicable in cities all around the world by using readily available data.

The InVEST Urban Cooling model has recently been applied to a European urban dataset (Cortinovis et al., 2022), and to individual cities in Europe, and India (Bosch et al., 2021; Kadaverugu et al., 2021; Zawadzka et al., 2021). However, attempts to validate the model against observed or modelled data are much rarer. In previous work, Bosch et al. (2021) found good performance of the model in Lausanne, although the observation data was sparse with 11 weather stations over 112 km^2. Zawadzka et al. (2021) found that the model explained between 48% and 60% of the variability in land surface temperature for a Summer day in three towns in the United Kingdom. Despite these notable efforts, the model has not been extensively tested and validated against spatialized air temperature data.

The goal of this paper is to evaluate the performance of the InVEST Urban Cooling model in different contexts: daytime and nighttime, and with different levels of data availability. In doing so, we further develop an open-source calibration algorithm to facilitate future applications of the model (Bosch et al., 2021). Given the intended use of InVEST as a decision-support tool, we focus model performance assessment on two aspects: the ability to represent spatial variations in temperatures, and the ability to represent temporal changes in temperature due to landscape changes. We test the model through two case studies with contrasted climates: the Paris metropolitan area, France, and the Twin Cities, Minnesota, USA. In the following, we present the model performance assessment and discuss the strengths and limitations of the tool and its recommended use to inform urban planning policies.

2 Methods

We assess the performance of the InVEST model for the current land uses land covers (LULCs) by comparing its outputs with available ("best estimates") air temperature and land surface temperature. Since daytime temperatures are highly influenced by convection and atmospheric turbulence, confounding the effect of land use (Le Roy et al., 2020), we hypothesized that the InVEST daytime outputs would better capture variations in land surface temperatures, rather than air temperatures. On the contrary, nighttime temperatures are strongly influenced by land cover (in particular built infrastructure) and thermal processes, so we expected to find a stronger correlation between InVEST and nighttime air temperature. Given the purpose of the model to support decision-making, we also assess the model's ability to capture the effect of a change in LULC on urban cooling.

2.1 Case studies

2.1.1 Paris metropolitan area

The Paris metropolitan area, situated in the Île-de-France administrative region, is our main case study. Île-de-France spans over 12,000 km^2, with land use being predominantly agriculture (50%), forests (24%), and artificial areas (22.5%) (see Figure A1 in Appendix). The climate in Paris is oceanic (humid temperate, no dry season and warm summers). Average rainfall over the period 1981-2010 was 637 mm and summer (June to August) temperatures averaged 19.7°C, with increasingly frequent temperature peaks in the middle of summer (Météo-France). The study period is the heat wave of 8-13 August 2003, described in previous work (de Munck et al., 2018), which led to maximum daytime temperatures of over 39°C in the region. A greening scenario for the Paris metropolitan area was developed by de Munck et al. (2018) to simulate the implementation of low and high vegetation (60% grass and small shrubs 40% of deciduous trees) for 50% of available ground surface (except roads and areas already covered by vegetation). This represents the greening of pavements, squares, carparks, and some roofs over a surface of 199 km^2, or a 23% increase in green areas, resulting in a reduction of up to 2°C in maximum daily mean temperatures over the study area (de Munck et al., 2018).

2.1.2 Twin Cities

The second case study, to validate the model in a different climate zone, is in the Twin Cities metropolitan area surrounding the cities of Minneapolis and St Paul in Minnesota, USA (see Figure A2 in Appendix). The climate of the Twin Cities is classified as hot-summer humid continental. We studied the regional heat wave event of July 22, 2016, and the average air temperatures over the period 2011-2014, building on previous studies of urban heat island in the region (Smoliak et al., 2015).

2.2 Model description

The InVEST Urban Cooling model, hereafter "InVEST model", is fully described in the InVEST software User's manual (Natural Capital Project, 2022). We provide a summary of key equations here to orient readers. The model computes air temperature prior to air mixing (T_{nomix}, in degree Celsius) on each pixel as a function of a background rural reference

4

temperature (T_{ref}) modified by a local heat factor. The latter is expressed as the maximum urban heat island intensity for the city (UHI_{max}) modulated by the local heat mitigation (HM), such that, for a given pixel i:

130
$$T_{nomix}(i) = T_{ref} + (1 - HM(i)) \times UHI_{max}$$

[1]

While this temperature does not account for mixing due to atmospheric turbulence, actual air temperature, T_{air}, is estimated from T_{nomix} using a spatial moving average algorithm with search radius r_{mix}. This radius varies with time, depending on the lateral mixing due to atmospheric turbulence, and can be estimated through calibration (Section 2.4).

135 The proportion of heat mitigation relative to UHI_{max}, with values ranging between 0 and 1, is derived from the cooling capacity (CC) of the LULC type on a given pixel and the proximity to large parks. Following an approach proposed by others (Kunapo et al., 2018; Zardo et al., 2017), the cooling capacity during daytime is expressed as a function of shade, evapotranspiration, and albedo:

$$CC(i) = W_S.S(i) + W_A.A(i) + W_E.E(i),$$

140
[2]

where S(i), A(i), and E(i) are unitless indices ranging from 0 to 1 that characterize shade, albedo, and evapotranspiration on the pixel i, respectively. They are each weighted by a coefficient (W_S, W_A, W_E) constant across the study area. S represents the proportion of shade for a given LULC type, e.g.,1 for a land cover type completely covered by canopy or high buildings, and 0 for bare lands. A is the albedo value of the land cover. E is calculated from reference evapotranspiration (a raster termed

145 ET_0, in mm) and the crop coefficient (K_c, no unit) for the LULC type, and then normalized by the maximum value of reference evapotranspiration in the area of interest according to:

$$E(i) = \frac{K_c(i).ET_0(i)}{max(ET_0)}$$

[3]

For nighttime temperatures, the model calculates cooling capacity as the complement of building density (B), an indicator
150 representing urban compactness and highly correlated to heat storage capacity (Wong et al., 2021). Building capacity is normalized across the area and can be obtained from building density data: a value of 1 implies that the pixel is covered by the buildings with the highest energy retention, often the tallest building, corresponding to a maximum nighttime urban heat island effect:

$$CC(i) = 1 - B(i),$$

155
[4]

To account for the effect of large parks (>2 ha), both in nighttime and daytime, the heat mitigation factor HM equals a distance-weighted average of the CC values from surrounding areas (CC_{park}). The algorithm for this distance-weighted average, for a pixel i, is as follows:

$$GA(i) = area \cdot \sum_{j \in c(d^{\circ}_{cool})} g(j)$$

5

160 [5a]

$$CC_{park}(i) = \sum_{j \in \hat{(i)}} g(j) \cdot CC_j \cdot exp(-d\,(i,j)/d_{cool})$$

[5b]

$$HM_i = \begin{cases} CC_i & if\ \ CC_i \geq CC_{park_i} \vee GA_i < 2ha \\ CC_{park_i} otherwise \end{cases}$$

[5c]

165   where GA(i) is the total area of green spaces in a buffer of radius d_{cool} (the distance over which a green space has a cooling
effect); $area$ is the area of a pixel in ha; c(d_{cool}) is the buffer area of radius d_{cool}; $g(j)$ is 1 if pixel j is green space, 0 otherwise;
CC_{park}(i) is the cooling capacity including the influence of parks; and $d(i,j)$ is the distance between pixel i and j.

In plain words, if the amount of green spaces surrounding a pixel (GA) is less than 2 ha, the value of HM on the pixel equals
CC (Eq. 5c), assuming little cooling effect outside the park other than through air mixing due to atmospheric turbulence
170   (defined by r_{mix}). The threshold size of 2 ha is obtained from the literature (Bowler et al., 2010; also see Discussion).

2.3 Input and calibration data

2.3.1 Paris metropolitan area

Input data. LULC data for the Île-de-France region were obtained from the regional urban planning agency (Institut Paris
Region, 2019) for the year 2003 (Figure S1). Reference evapotranspiration (ET_0) for August was obtained from monthly
175   modeled climatological data for the region, averaged for the 1985-2005 period (ALADIN model, Stéfanon et al., 2015). Using
long-term average reference evapotranspiration instead of 2003 reference evapotranspiration has a limited effect on outputs
given that the InVEST model only uses relative values (see Eq. 3 above), which have lower temporal variability than absolute
values. LULC parameter values for shade, crop evapotranspiration coefficient, albedo, and building intensity were assigned
based on a combination of expert opinion and literature review (see Appendix B).
180   *Reference temperature data.* In the Paris case study, we compared the InVEST model outputs with two datasets of land surface
temperature and air temperature. Land surface temperature maps with a 1-km horizontal resolution were retrieved from MODIS
satellite products (Wan, 2013). They were obtained for August 13th 2003 at daytime and nighttime since data were missing at
the two dates closer to the peak of the heat wave (August 11th and 12th). We also considered higher resolution data from Landsat
but due to its lower frequency (16 days) there was no data around the study period. MODIS land surface temperature data have
185   been used in many surface urban heat island studies globally and regionally (Chakraborty & Lee, 2019; Li et al., 2017).
Predictions of air temperature maps with a 1-km horizontal resolution were obtained from the simulations performed by de
Munck et al. (2018) with the physical land-surface model TEB/Surfex (Lemonsu et al., 2012; Masson, 2000; Masson et al.,
2013). The model computes the energy and water budgets on a geographic domain of 100 km x 100 km centered on Paris with
a user-defined regular grid, and for natural and urban areas, taking land cover and building characteristics as inputs (Masson,
190   2000). The outputs used in this study were the 6-day average air temperatures from 8-13 August 2003. The 1st and 99th

percentiles were extracted from these air temperature data to define the extreme temperature conditions over the domain during the heat wave (to exclude visible outliers). The background reference rural temperature (T_{ref}) was set to the 1st percentile, while the maximum urban heat island intensity (UHI_{max}) was the difference between both percentiles. Because the TEB/Surfex data were available for a 100 km x 100 km window, the rest of the analyses are presented for this domain, rather than the entire Ile-de-France region. The three datasets InVEST, TEB/Surfex and MODIS were resampled to a common 1-km grid for comparison, using bilinear interpolation.

2.3.2 Twin Cities

Input data. We used 2016 LULC data from the National Land Cover Dataset (NLCD; Homer et al., 2020). Details on LULC parameters values are provided in Appendix B and in previous work (Hamel et al., 2021). Reference evapotranspiration for July 2016 was obtained from globally available data from the Consultative Group on International Agricultural Research (Trabucco & Zomer, 2018).

Reference temperature data. We compared air temperature as modeled by InVEST to two data sources: the midday land surface temperature maps of the study area during a regional heatwave on July 22, 2016, retrieved from Landsat data at a 30-m horizontal resolution for daytime and MODIS for nighttime (Wan, 2013); and rasters of average summertime (June, July, August, 2011-2014) daytime and nighttime air temperatures across the study area, interpolated from a dense network of temperature sensors (~170 stations over 5000 km^2, interpolated with by cokriging using impervious surfaces) (Smoliak et al., 2015). We projected the reference temperature data to match the InVEST output coordinate system and resampled all datasets to match the resolution (~1 km) of Smoliak et al.'s data using bilinear interpolation. The background reference rural temperature (T_{ref}) was defined using the average air temperature in July for the metropolitan area obtained from the US National Weather Service (23.2 °C; NOAA 2020) and the maximum urban heat island intensity (UHI_{max}) was taken from a global assessment of urban heat islands (2.05°C; Chakraborty & Lee, 2019).

2.4 Model calibration and performance assessment

The model outputs were first assessed without calibration, using default parameter values (500 m, 100 m, 0.2, 0.2, 0.6, for r_{mix}, d_{cool}, W_a, W_e, W_s, respectively). For calibration, we further developed an optimization algorithm used in previous work (Bosch et al., 2021). The algorithm starts with default parameter values and implements a simulated annealing optimization to derive the values of the five parameters for daytime (r_{mix}, d_{cool}, W_a, W_e, W_s) or two parameters for nighttime (r_{mix}, d_{cool},) until the algorithm converges to a solution that minimizes r^2 (or a limit number of 100 iterations is reached). The main improvement on the calibration tool developed for this study is the ability to use as reference temperatures either point data (e.g., a network of stations) or raster data (as is the case for the temperature data in this study). Other improvements are minor and of technical nature. (They include testing for the compatibility of user inputs, updating deprecated packages, and improving the code efficiency and readability, as documented in the source code.). The tool reports several performance metrics: mean absolute error (MAE), root mean square error (RMSE), and r^2. We Following previous work (Bosch et al., 2021), we selected these

7

metrics since MAE and RMSE are useful quantification of the uncertainty in model outputs with physical quantities (expressed in °C), which is important ~~from a user perspective~~for users to understand the impact of errors. However, MAE and RMSE also depend on UHI_{max}, which means that performance might be artificially good for areas with small urban heat island magnitudes. For this reason, we also report r^2 (the default performance criterion for the optimization). The source code for the calibration tool can be found at: https://github.com/martibosch/invest-ucm-calibration/tree/v0.6.0 (and on Zenodo: Please refer to the Code Availability section) and a user guide available at: https://invest-ucm-calibration.readthedocs.io/en/latest/usage.html. In addition, we performed a one-at-a-time sensitivity analyses to further understand how the calibration parameters influence model outputs. Ranges of variation for each parameter are provided in Appendix C.

To demonstrate the application of the InVEST model in practice and assess the model's ability to represent the effect of a change in LULC, we also examined the effect of a greening strategy on urban cooling in the first case study, where comparison data were available from an alternative model. We used the greening scenario LHV50 (50% of impervious areas covered by low and high vegetation) developed and simulated by de Munck et al. (2018) and described in Section 2.1. In InVEST, we represented this scenario by changing the LULC properties (shade, albedo, crop coefficient) for urban categories. Specifically, we estimated the proportion of available ground, as defined above, for each category of the LULC map, and computed the weighted average of initial parameter values and values for an urban forest. For example, for the "Parks or Gardens" LULC category, it was estimated that 15% of the ground was available, so a weighted average of the shade value for Forest (1) and original park value (0.5) was computed. The resulting parameter values are shown in Supplementary Material.

3 Results

3.1 Model performance prior to calibration

3.1.1 Daytime temperatures

For the Paris case study, prior to calibration, the InVEST daytime temperatures were moderately correlated with daytime land surface temperature (r^2=0.60), with a MAE of 3°C, but showed no correlation with air temperatures (Table 1). In the Twin Cities, the InVEST results had low correlation with LST (r^2=0.20) but slightly higher for air temperatures (r^2=0.29). Of note, comparisons between InVEST and land surface temperature in the Twin Cities revealed a large MAE (11.90°C) as surface temperatures can be much higher (up to 52°C) than air temperatures (up to 26°C) based on the reference temperature data.

Visual observations of the difference between the two maps suggest that the model overestimates the temperatures in forested areas (e.g., to the west and southern areas, see Figure S1 for LULC map), as well as in the dense urban areas in the city of Paris (centre, Figure 1). In the Twin Cities, the model overestimates temperatures in agricultural areas, while underestimating temperatures around wetland areas (Figure 2 and Figure S1 for detailed LULC map). In both case studies, the model exhibited a higher variability than reference data.

3.1.2 Nighttime temperatures

As expected, the InVEST nighttime temperatures were correlated with both air and surface temperature, although more strongly with air temperature in Paris (r^2=0.84 for Paris, and r^2=0.70 for the Twin Cities, Table 1). The MAE is also significantly lower when calculated with respect to air temperatures than land surface temperatures for both cases (0.5 against 3.0°C for Paris and 0.5 against 2.5°C for Twin Cities).

Nevertheless, the maps of the differences between InVEST and nighttime air temperature show important spatial heterogeneities. For Paris, this suggests that the model severely underestimates temperatures in the core area (by about 3°C) and overestimates temperatures in the new urban developments (by about 1°C, Figure 1 and Figure S1). In the Twin Cities, InVEST systematically underestimates temperatures in the city center, outlying developed suburbs, and around bodies of water (e.g. in the west and the confluence of the Minnesota and Mississippi rivers south of the city center), while overestimating temperatures in the surrounding agricultural hinterlands (Figure 2 and Figures S2).

3.2 Model performance after calibration

The model performance remained relatively stable after calibration for nighttime air temperatures in both Paris and the Twin Cities, reaching r^2 values of 0.84 and 0.73, respectively. Daytime temperatures, however, showed very low correlations with best estimates in Paris and medium correlation (r^2=0.33) in the Twin Cities. The sensitivity analysis conducted for the Paris case study confirmed that most model parameters had only a limited effect on correlations with land surface temperature data, which could explain the limited effect of calibration (Figures C1 and C2 in Appendix C). Only r_{mix} had a significant effect on r^2 (with a range of r^2 values ranging from 0.30 to 0.62 for r_{mix} ranging from 100 to 5000 m). Because of the very low r^2 values for daytime temperatures in Paris, we considered calibration unsuccessful and we do not report the calibrated values.

The Twin Cities daytime calibrated values of the parameters are reported in Table 2 (with UHI_{max}=11.7°C). The shade weight remains the highest after calibration, and values only changed by <15%. Nighttime calibrated parameter values suggest that there is less air mixing, with both r_{mix} and d_{cool} having lower values than during daytime in the Twin Cities (Table 2).

Table 1: InVEST model performance for daytime and nighttime air (Tair) and surface (LST) temperatures (MAE: mean absolute error) before and after calibration. Post-calibration values are shown in parentheses.

		Paris		Twin Cities	
		r^2	MAE (°C)	r^2	MAE (°C)
T_{air}	Day	<0.01 (0.01)	1.40 (1.30)	0.29 (0.33)	0.48 (0.43)
	Night	0.84 (0.84)	0.52 (0.52)	0.70 (0.73)	0.48 (0.48)
LST	Day	0.60 (0.62)	3.06 (2.80)	0.20 (0.43)	11.9 (8.70)
	Night	0.45 (0.47)	2.99 (6.10)	0.59 (0.76)	2.54 (2.48)

Table 2: Calibrated coefficient values for air temperature in the two case studies.

	Paris	Twin Cities

Daytime	r_{mix} (m)	n.a.	771
	d_{cool} (m)	n.a.	109
	W_a ; W_e ; W_s	n.a.	0.21; 0.17; 0.62
Nighttime	r_{mix} (m)	500	630
	d_{cool} (m)	100	66

280
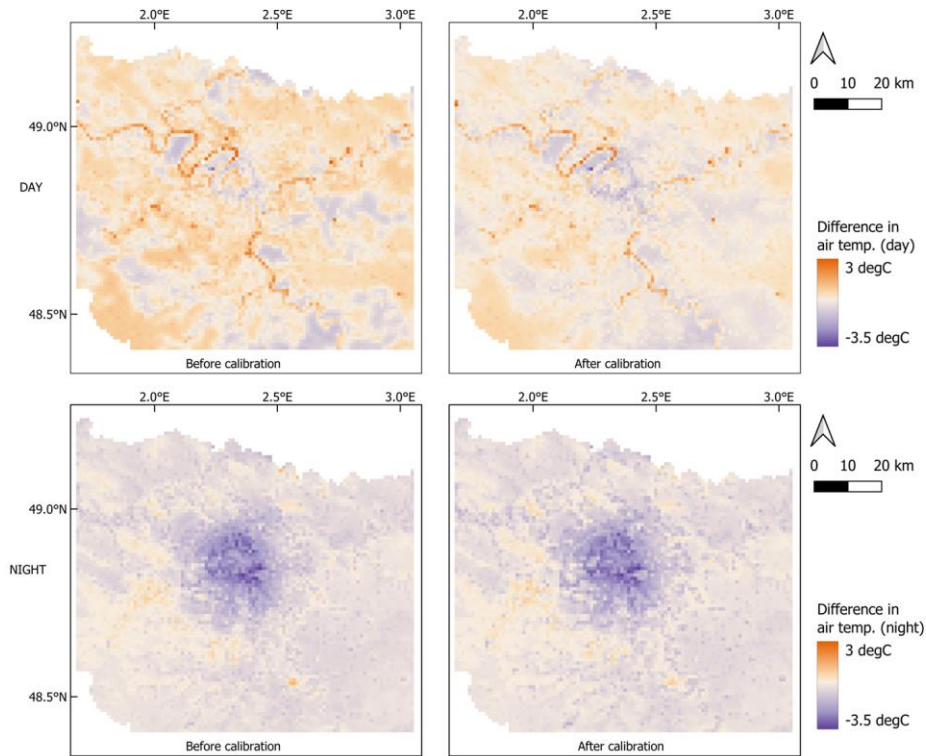
Figure 1: Difference between modelled and reference air temperatures (°C) for the Paris region for daytime and nighttime simulations, pre- and post-calibration
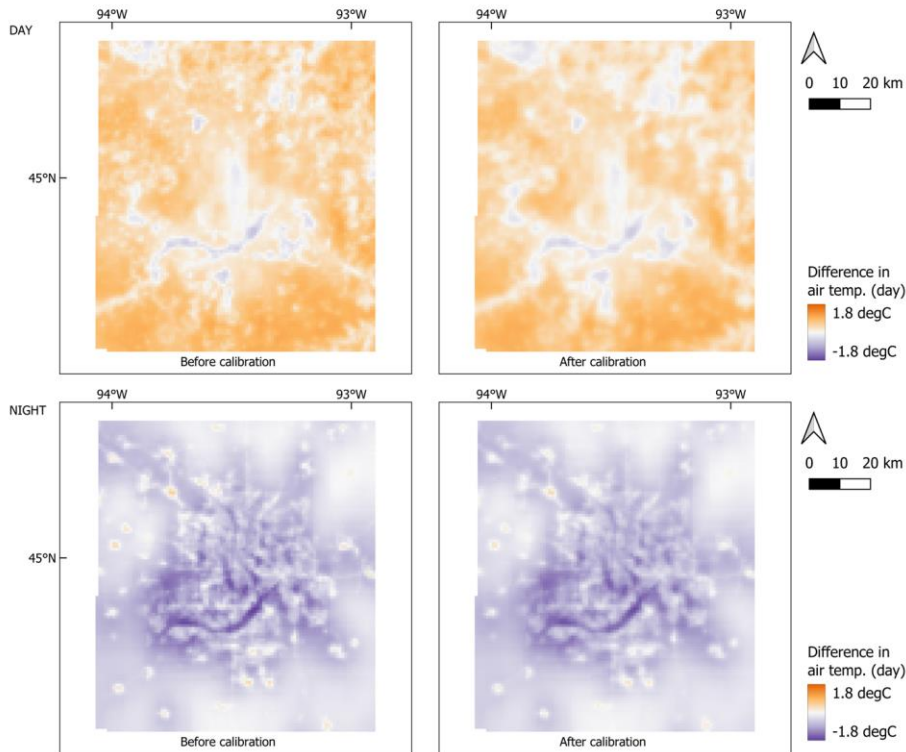
Figure 2: Difference between modelled and reference air temperatures (°C) for the Twin Cities region for daytime and nighttime simulations, pre- and post-calibration

3.3 Effect of the greening scenario

Given the limited effect of calibration for daytime and nighttime air temperature data (see Section 3.2), we tested the correlation obtained with temperature estimates from the uncalibrated model. The results were satisfactory, with a medium correlation strength between InVEST and the TEB/Surfex model (r^2= 0.55 and mean absolute error of 0.07 °C) for daytime data (Figure 3). For nighttime data, correlation between the two models was stronger, with r^2= 0.85.

11
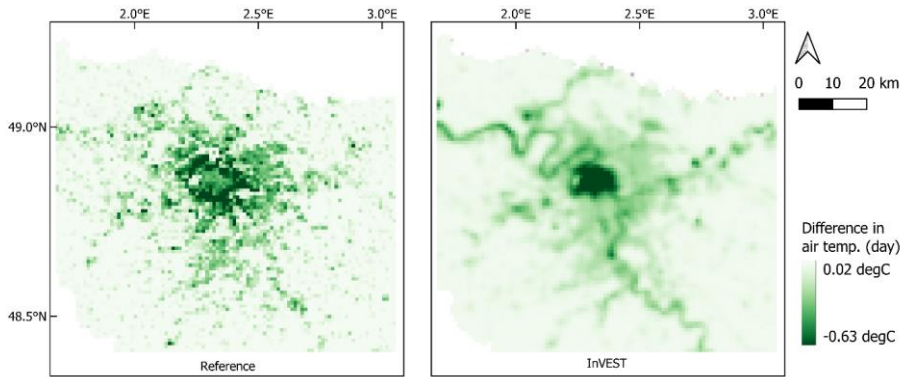
Figure 3: Cooling service (°C) during daytime resulting from the greening scenario as simulated by InVEST and the TEB/Surfex model (reference data).

4 Discussion

4.1 Model calibration and performance

Based on the two case studies, the InVEST model performance was best (r^2>0.73) for nighttime air temperature after calibration. The performance improvement due to calibration was modest in most cases (day or night, air or surface temperature), especially for nighttime air temperature in Paris where the performance was already high prior to calibration in Paris (r^2=0.84). Only in the Twin Cities for surface temperature did the calibration significantly improve model performance, possibly due to the LULC configuration in this landscape, although our analyses do not allow us to confirm this hypothesis. The modest effect due to calibration can be explained by a relatively low sensitivity to model parameters, which is suggested by the sensitivity analyses presented in Appendix (Figures C3 and C4).

For the Twin Cities, nighttime calibrated parameter values (r_{mix} and d_{cool}) are lower than for daytime, which supports the lower convection and air mixing during the night. Calibrated values of r_{mix} are 771 m and 660 m for daytime and nighttime, respectively, which compares well with the estimate of 600 m obtained from previous studies (Lonsdorf et al., 2021; Schatz & Kucharik, 2014). Similar reasoning to verify the physical interpretation of model parameters could be done for the shade, albedo, and evapotranspiration weights if such data are available in a case study: for example, calibrated values could be compared with relative proportion of shading to evapotranspiration that are studied in some cities (e.g., in Singapore, Tan et al., 2018). In both case studies, model performance for surface temperatures (r^2 ranging from 0.43 to 0.76 post-calibration)

12

compares with the study by Zawadzka et al. (2021), who found that the InVEST model explained between 48% and 60% of the variability in land surface temperature for a Summer day in three towns in the United Kingdom. We highlight, however, that the model was initially not developed for land surface temperatures. The fair performance for land surface temperatures is an artifact of the model's simplified representation of air temperatures, which imperfectly represents the local energy balance, and the strong correlation of both air and surface temperatures with LULC.

Regarding scenario assessment, the model in the Paris case study performed reasonably well (r^2= 0.55 and r^2= 0.85 for air daytime and nighttime temperatures, respectively). This confirms the potential of the tool to support urban planning studies where different scenarios might need to be compared based on their cooling potential, similar to earlier work with InVEST (Bosch et al., 2022). The effect from the greening scenario, resulting in less than 1°C cooling, is modest but in line with the literature on such large-scale implementation of green infrastructure, not only from the reference study by de Munck et al. (2018), but also for other temperate climate studies.

4.2 Limitations of the study and future works

In both our case studies, we note that the performance was poor for daytime air temperature, which does not support the use of the model for absolute temperature estimates. This poor performance may be due to ~~i or~~ oversimplifications of the physical processes involved in urban cooling. In particular, the simplification of the flow dynamics means that a model such as InVEST cannot represent urban canyon and wind effects in the city. The effect of parks is also simplified with a threshold of 2 ha for park size. While this has the advantage of reducing the number of model parameters, it also ignores the effect of smaller greenspaces might have on their surroundings (Wong et al., 2021; Yu et al., 2020). The poor performance of the model for daytime air temperature may also be attributed to errors in parameterizations, in particular the use of climate data for short periods (e.g., Aug 6-13th 2003 for Paris) vs. average values over several months, as was the case for some inputs (e.g., reference evapotranspiration in the Paris case study, or T_{ref} and UHI_{max} in the Twin Cities, see Section 2.3). Further investigation of these temporal dynamics should be explored in future work, although we highlight that they did not seem to impact the fair performance of the model for nighttime air temperatures or land surface temperatures.

An important limitation of any urban climate study examining fine-resolution spatial variations in temperature is the availability of robust reference data. Because air temperature cannot be readily derived from remote sensing data, models are routinely compared to networks of weather stations (Bosch et al., 2021; de Munck et al., 2018; Smoliak et al., 2015). This means that model performance is only assessed for a limited number of points, which are not typically representative of the diversity of LULC in a region. When data are collected specifically for model validation purposes (e.g., transect data), they are also limited by practical factors such as timing considerations (e.g., mismatch in time between the beginning and end of the transect) or lack of reference data (Stewart, 2011; Velasco, 2018).

In our study, we have used alternative models as reference data for air temperatures, either a physics-based model or a statistical interpolation model. Both of these models have limitations and uncertainties in themselves, making a fine scale understanding of the limitations of the InVEST model challenging. In other words, differences in models observed in Figures 1 and 2 might

13

also be due to errors in the reference data and future studies could examine the effect of uncertainty in reference datasets on calibration and model performance. For example, the effect of shade during daytime is poorly represented in the TEB/Surfex model used in Paris (de Munck et al., 2013). Land surface temperature datasets, on the other hand, are less prone to such limitations and are more robust when it comes to spatial distributions. Absolute values are nonetheless challenging to ascertain

350 due to temperature calculations requiring correction algorithms.

Finally, an important study limitation that could be explored in future work is the spatial resolution and scale of case studies. In both cases, the InVEST model was evaluated at 1-km horizontal resolution over an area of about 100 km by 100 km. Because the model produces outputs at the same resolution as LULC inputs, it would be interesting to evaluate its performance at a different scale and resolution, where such data are available. Relatedly, the use of local climate zones, which is common in

355 urban climate studies (Aslam & Rana, 2022), could be examined as an alternative parameterization of the InVEST model. This work would also examine the influence of reference temperature (T_{ref}) and maximum intensity (UHI_{max}), as well as the biophysical parameter values (in particular crop coefficients – notably difficult to ascertain for urban land uses), which have not been explored in the present study.

Overall, this discussion of data quality and calibrated parameter values highlights an important contribution of this study, in

360 the customized calibration tool available on Github (https://github.com/martibosch/invest-ucm-calibration and on Zenodo, please refer to the Code Availability section). Such a tool can be applied to any other city where the InVEST Urban Cooling model is applied with the only data requirement, beyond the InVEST model input data, being a reference temperature dataset (either point data or raster). The calibration tool allows for systematic calibration and model testing, which paves the way for better understanding of model limitations and strengths. Although in our case studies the performance improvement

365 from calibration was modest, future work could assess the performance of the model over multiple cities with comparable datasets, examine the potential of local climate zones for improved parameterization, or explore finer temporal resolution by linking the night and day model outputs.

5 Conclusion

In this study, we have developed a custom calibration tool to assess the performance of the InVEST Urban cooling model in

370 two case studies: in Paris, France, and the Twin Cities, MN, USA. Our analyses expand on past model testing studies by providing a much more extensive validation dataset of air temperatures (continuous data based on a reference urban climate model in Paris, and on data interpolated from a dense weather station network in the Twin Cities). The model showed good performance, assessed through mean absolute error (0.52°C) and r^2 (0.84) for nighttime air temperatures. Calibration only slightly improved model performance in the Twin Cities. For the case study of Paris, the use of the tool for scenario assessment

375 was supported by moderate (daytime) and high (nighttime) correlation with change predicted by an alternative, physics-based model (r^2=0.55 and r^2=0.85, for air daytime and nighttime temperatures, respectively). With respect to the study objectives, we conclude that the open-source model can be used to support decisions related to land use and land cover change in cities,

with greater reliability for nighttime UHI applications and for relative change (i.e., comparing scenarios to one another as opposed to using absolute values of model predictions. As these results were obtained for the case studies of Paris and the Twin Cities, the InVEST model and calibration tools should be tested in other geographies to assess model performance for urban planning applications. For research applications, future studies in other geographies will help further understand the effect of data resolution and data quality on model performance.
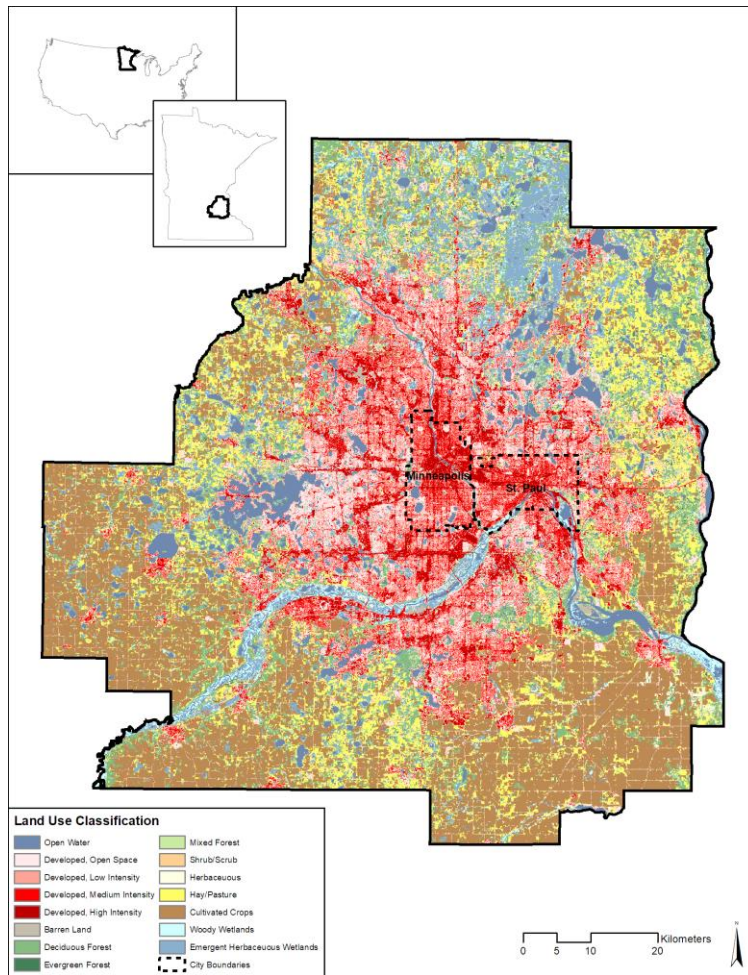
Appendices

A Input LULC data
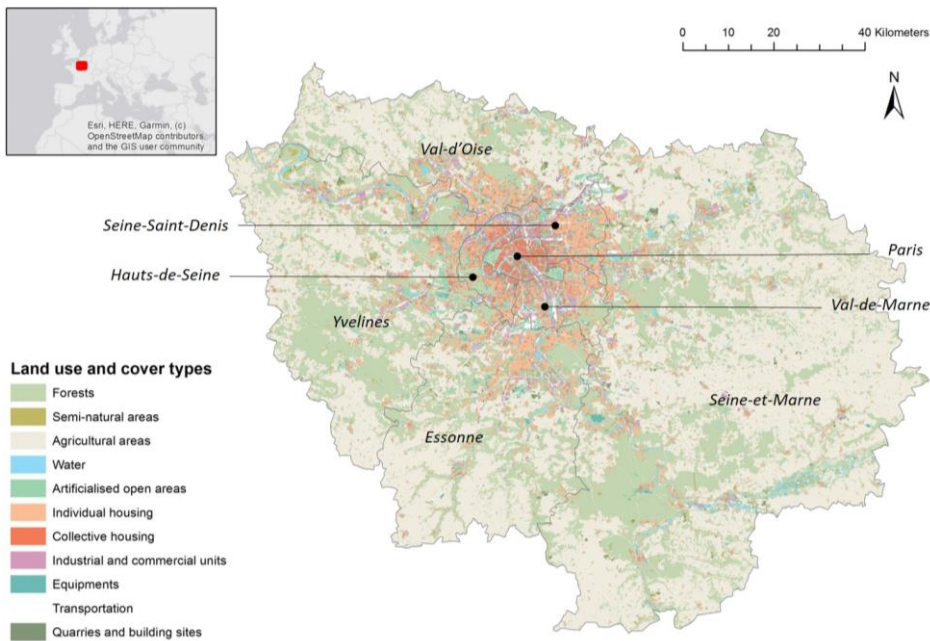
Figure A1: Land use and cover in Ile-de-France region in 2017, France (based on data developed by the Institut Paris Region)

Land Use Classification

Color	Class	Color	Class
	Open Water		Mixed Forest
	Developed, Open Space		Shrub/Scrub
	Developed, Low Intensity		Herbaceuous
	Developed, Medium Intensity		Hay/Pasture
	Developed, High Intensity		Cultivated Crops
	Barren Land		Woody Wetlands
	Deciduous Forest		Emergent Herbaceuous Wetlands
	Evergreen Forest		City Boundaries

390 **Figure A2: Location and land cover of the Twin Cities metropolitan area in Minnesota, USA, with the cities of Minneapolis and St. Paul outlined.**

B InVEST biophysical tables

InVEST biophysical tables can be found in Supplementary Material.

395

Paris case study

Albedo values for the Paris biophysical table were derived from the following three sources:

- *APUR. (2020). Atténuer les îlots de chaleur urbains - Cahier n°5 : méthodes et outils de conception des projets. https://www.apur.org/fr/nos-travaux/attenuer-ilots-chaleur-urbains-cahier-5-methodes-outils-conception-projets*

400   *(Accessed 30 April 2023)*

- *Lavigne, P., P. Brejon, and Fernandez. 1994. Architecture Climatique: Une Contribution Au Développement Durable. Edisud.*

- *Stewart, I.D., Oke, T.R., (2012). Local Climate Zones for Urban Temperature Studies. Bulletin of the American Meteorological Society 93, 1879-1900*

405 Specifically, most albedo values were taken from Stewart and Oke (2012). Exceptions are for parks, cemeteries, all industrial land uses and all facilities but hippodromes and car tracks (derived from Lavigne et al, 1994), while outdoor courts, transportation land use, quarries, dumpsites, construction sites, hippodromes and car tracks were derived from APUR (2020). Crop evapotranspiration values were derived from the FAO Irrigation paper 56 (Allen et al., 1998), as described in previous work (Tardieu et al., 2020):

410   - *Allen, R.G., Pereira, L.S., Raes, D., Smith, M. (1998) FAO Irrigation and Drainage Paper No. 56. Irrigation and Drainage, 56, No. 97. Food and Agriculture Organization of the United Nations, Rome.*

- *Tardieu L, Hamel P, Viguié V, Coste L, Levrel H (2021). Are soil sealing indicators sufficient to guide urban planning? Insights from an ecosystem services assessment in the Paris metropolitan area. Environmental Research Letters 16(10). https://doi.org/10.1088/1748-9326/ac24d0*

415 Shade values were assumed to be 1 for forested or tree-planted land uses, 0.5 for mixed land use with trees, 0.25 for commercial land use with planted trees, and 0 for all remaining land uses.

Twin Cities case study

Details on the Twin Cities biophysical tables are provided in Supplementary Information of this source:

420   - *Hamel, P., Guerry, A. D., Polasky, S., Han, B., Douglass, J. A., Hamann, M., Janke, B., Kuiper, J. J., Levrel, H., Liu, H., Lonsdorf, E., McDonald, R. I., Nootenboom, C., Ouyang, Z., Remme, R. P., Sharp, R. P., Tardieu, L., Viguié, V., Xu, D., ... Daily, G. C. (2021). Mapping the benefits of nature in cities with the InVEST software. Npj Urban Sustainability, 1(1), 25. https://doi.org/10.1038/s42949-021-00027-9*

C Sensitivity analyses

425 For the Paris case study, we performed a local sensitivity analysis for the five following parameters (ranges for each parameter are in parentheses):

- Air mixing distance r_{mix} (50 to 5000 m)
- Green area maximum cooling distance d_{cool} (50 to 1000 m)
- Cooling capacity factors W_s, W_a, and W_e (0 to 1)

430 We assessed results based on the correlation with daytime and nighttime temperatures and land surface temperatures. We found that the model was highly sensitive to r_{mix} (Figure S1), with a local maximum for a value of 500 m for daytime temperatures and 1000 m for nighttime temperatures (Figure S3). The model was less sensitive to the values of maximum cooling distance and the cooling capacity factors: r^2 varied between 0.61 and 0.63 for d_{cool}, and 0.59 et 0.64 for the weight factors.

435 These results held for nighttime temperature, where InVEST was also most sensitive to the air mixing parameter, r_{mix} (Figure S4). Sensitivity to the building density parameter B and the maximum cooling distance d_{cool} was much lower.
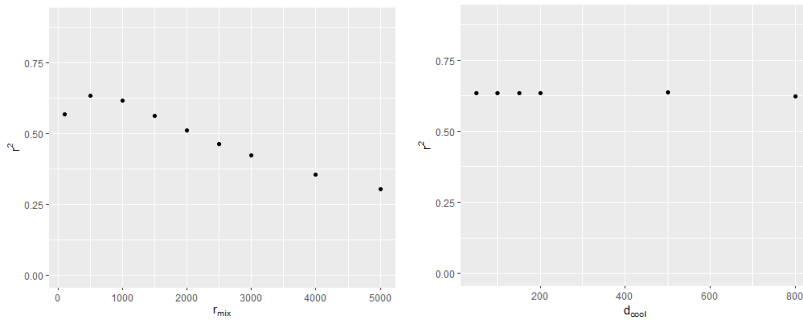
440 **Figure C1: Sensitivity of the InVEST model to the r_{mix} and d_{cool} parameters and W_e and W_s factors. Results are plotted against the correlation coefficient (r^2) with daytime land surface temperature (MODIS).**
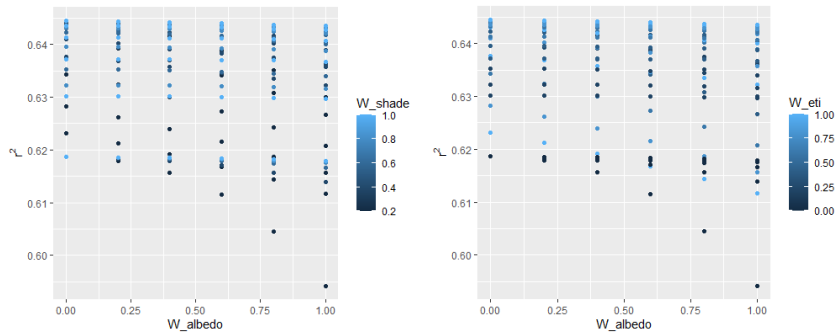
Figure C2: Sensitivity of the InVEST model to daytime weight factors W_e (W_eti), W_a (W_albedo) and W_s (W_shade). Results are plotted against the correlation coefficient (r^2) with daytime land surface temperature (MODIS).

Code availability

The code materials used in this article are available at: https://zenodo.org/record/8081822 (last access: 30 April 2023).

Data availability

The source files for the biophysical tables to run the InVEST models can be found in Supplementary Information.

Other input data such as the LULC map for the Paris case study are not publicly available.

Author contributions

PH designed the study and conducted the analysis with MB and CN, with technical input from LT, AL, CdM, and VV. JAD and RPS led the software development for the InVEST Urban Cooling model and MB developed the code for the calibration algorithm. PH wrote the first draft of the paper with inputs from LT and CN. All authors provided comments and contributed to the final version of the paper.

Competing interests

The authors declare that they have no conflict of interest.

470 **References**

Aslam, A., & Rana, I. A. (2022). The use of local climate zones in the urban environment: A systematic review of data sources, methods, and themes. *Urban Climate*, *42*, 101120. https://doi.org/10.1016/J.UCLIM.2022.101120

Bartesaghi Koc, C., Osmond, P., & Peters, A. (2018). Evaluating the cooling effects of green infrastructure: A systematic review of methods, indicators and data sources. *Solar Energy*, *166*, 486–508.
475   https://doi.org/10.1016/J.SOLENER.2018.03.008

Bolund, P., & Hunhammar, S. (1999). Ecosystem services in urban areas. *Ecological Economics*, *29*(2), 293–301. https://doi.org/10.1016/S0921-8009(99)00013-0

Bosch, M., Locatelli, M., Hamel, P., Remme, R. P., Chenal, J., & Joost, S. (2021). A spatially explicit approach to simulate urban heat mitigation with InVEST (v3.8.0). *Geoscientific Model Development*, *14*(6), 3521–3537.
480   https://doi.org/10.5194/GMD-14-3521-2021

Bosch, M., Locatelli, M., Hamel, P., Remme, R. P., Jaligot, R., Chenal, J., & Joost, S. (2022). Evaluating urban greening scenarios for urban heat mitigation: a spatially explicit approach. *Royal Society Open Science*, *8*(12), 202174. https://doi.org/10.1098/rsos.202174

Bowler, D. E., Buyung-Ali, L. M., Knight, T. M., & Pullin, A. S. (2010). A systematic review of evidence for the added
485   benefits to health of exposure to natural environments. *BMC Public Health*, *10*(1), 1–10. https://doi.org/10.1186/1471-2458-10-456/TABLES/1

Cao, X., Onishi, A., Chen, J., & Imura, H. (2010). Quantifying the cool island intensity of urban parks using ASTER and IKONOS data. *Landscape and Urban Planning*, *96*(4), 224–231. https://doi.org/10.1016/J.LANDURBPLAN.2010.03.008

Chakraborty, T., & Lee, X. (2019). A simplified urban-extent algorithm to characterize surface urban heat islands on a global
490   scale and examine vegetation control on their spatiotemporal variability. *International Journal of Applied Earth Observation and Geoinformation*, *74*, 269–280. https://doi.org/10.1016/J.JAG.2018.09.015

Cheung, P. K., Livesley, S. J., & Nice, K. A. (2021). Estimating the cooling potential of irrigating green spaces in 100 global cities with arid, temperate or continental climates. *Sustainable Cities and Society*, *71*, 102974. https://doi.org/10.1016/J.SCS.2021.102974

495 Corburn, J. (2009). Cities, Climate Change and Urban Heat Island Mitigation: Localising Global Environmental Science. *Http://Dx.Doi.Org/10.1177/0042098008099361*, *46*(2), 413–427. https://doi.org/10.1177/0042098008099361

Cortinovis, C., & Geneletti, D. (2019). A framework to explore the effects of urban planning decisions on regulating ecosystem services in cities. *Ecosystem Services*, *38*, 100946. https://doi.org/10.1016/J.ECOSER.2019.100946

20

Cortinovis, C., Olsson, P., Boke-Olén, N., & Hedlund, K. (2022). Scaling up nature-based solutions for climate-change
adaptation: Potential and benefits in three European cities. *Urban Forestry & Urban Greening*, *67*, 127450.
https://doi.org/10.1016/J.UFUG.2021.127450

de Munck, C., Lemonsu, A., Masson, V., Le Bras, J., & Bonhomme, M. (2018). Evaluating the impacts of greening scenarios
on thermal comfort and energy and water consumptions for adapting Paris city to climate change. *Urban Climate*, *23*, 260–
286. https://doi.org/10.1016/J.UCLIM.2017.01.003

Derkzen, M. L., van Teeffelen, A. J. A., & Verburg, P. H. (2015). REVIEW: Quantifying urban ecosystem services based on
high-resolution data of urban green space: an assessment for Rotterdam, the Netherlands. *Journal of Applied Ecology*,
52(4), 1020–1032. https://doi.org/10.1111/1365-2664.12469

Eliasson, I. (1996). Urban nocturnal temperatures, street geometry and land use. *Atmospheric Environment*, *30*(3), 379–392.
https://doi.org/10.1016/1352-2310(95)00033-X

Eumorfopoulou, E. A., & Kontoleon, K. J. (2009). Experimental approach to the contribution of plant-covered walls to the
thermal behaviour of building envelopes. *Building and Environment*, *44*(5), 1024–1038.
https://doi.org/10.1016/J.BUILDENV.2008.07.004

Farrugia, S., Hudson, M. D., & McCulloch, L. (2013). An evaluation of flood control and urban cooling ecosystem services
delivered by urban green infrastructure. *Https://Doi.Org/10.1080/21513732.2013.782342*, *9*(2), 136–145.
https://doi.org/10.1080/21513732.2013.782342

Gosling, S. N., Lowe, J. A., McGregor, G. R., Pelling, M., & Malamud, B. D. (2009). Associations between elevated
atmospheric temperature and human mortality: A critical review of the literature. *Climatic Change*, *92*(3–4), 299–341.
https://doi.org/10.1007/S10584-008-9441-X/METRICS

Hamel, P., Guerry, A. D., Polasky, S., Han, B., Douglass, J. A., Hamann, M., Janke, B., Kuiper, J. J., Levrel, H., Liu, H.,
Lonsdorf, E., McDonald, R. I., Nootenboom, C., Ouyang, Z., Remme, R. P., Sharp, R. P., Tardieu, L., Viguié, V., Xu, D.,
… Daily, G. C. (2021). Mapping the benefits of nature in cities with the InVEST software. *Npj Urban Sustainability*, *1*(1),
25. https://doi.org/10.1038/s42949-021-00027-9

Hémon, D., & Jougla, E. (2004). *Surmortalité liée à la canicule d'août 2003. Rapport remis au Ministre de la Santé et de la
protection sociale.*

Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., Gass, L., Funk, M., Wickham, J., Stehman, S., Auch, R.,
& Riitters, K. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land
Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 184–199.
https://doi.org/10.1016/j.isprsjprs.2020.02.019

Institut Paris Region. (2019). *Mode d'occupation du sol (MOS)*. https://www.institutparisregion.fr/mode-doccupation-du-sol-
mos.html

Jauregui, E. (1990). Influence of a large urban park on temperature and convective precipitation in a tropical city. *Energy and
Buildings*, *15*(3–4), 457–463. https://doi.org/10.1016/0378-7788(90)90021-A

Kadaverugu, R., Gurav, C., Rai, A., Sharma, A., Matli, C., & Biniwale, R. (2021). Quantification of heat mitigation by urban green spaces using InVEST model—a scenario analysis of Nagpur City, India. *Arabian Journal of Geosciences*, *14*(2), 1–13. https://doi.org/10.1007/S12517-020-06380-W/METRICS

Kroeger, T., McDonald, R. I., Boucher, T., Zhang, P., & Wang, L. (2018). Where the people are: Current trends and future potential targeted investments in urban trees for PM10 and temperature mitigation in 27 U.S. Cities. *Landscape and Urban Planning*, *177*, 227–240. https://doi.org/10.1016/j.landurbplan.2018.05.014

Kunapo, J., Fletcher, T. D., Ladson, A. R., Cunningham, L., & Burns, M. J. (2018). A spatially explicit framework for climate adaptation. *Https://Doi.Org/10.1080/1573062X.2018.1424216*, *15*(2), 159–166. https://doi.org/10.1080/1573062X.2018.1424216

Lai, L. W., & Cheng, W. L. (2009). Air quality influenced by urban heat island coupled with synoptic weather patterns. *Science of the Total Environment*, *407*(8), 2724–2733. https://doi.org/10.1016/j.scitotenv.2008.12.002

Larondelle, N., & Haase, D. (2013). Urban ecosystem services assessment along a rural–urban gradient: A cross-analysis of European cities. *Ecological Indicators*, *29*, 179–190. https://doi.org/10.1016/J.ECOLIND.2012.12.022

Le Roy, B., Lemonsu, A., Kounkou-Arnaud, R., Brion, D., & Masson, V. (2020). Long time series spatialized data for urban climatological studies: A case study of Paris, France. *International Journal of Climatology*, *40*(7), 3567–3584. https://doi.org/10.1002/JOC.6414

Lehmann, S. (2014). Low carbon districts: Mitigating the urban heat island with green roof infrastructure. *City, Culture and Society*, *5*(1), 1–18. https://doi.org/10.1016/j.ccs.2014.02.002

Lemonsu, A., Masson, V., Shashua-Bar, L., Erell, E., & Pearlmutter, D. (2012). Inclusion of vegetation in the Town Energy Balance model for modelling urban green areas. *Geoscientific Model Development*, *5*(6), 1377–1393. https://doi.org/10.5194/GMD-5-1377-2012

Lonsdorf, E. V., Nootenboom, C., Janke, B., & Horgan, B. P. (2021). Assessing urban ecosystem services provided by green infrastructure: Golf courses in the Minneapolis-St. Paul metro area. *Landscape and Urban Planning*, *208*, 104022. https://doi.org/10.1016/J.LANDURBPLAN.2020.104022

Manoli, G., Fatichi, S., Schläpfer, M., Yu, K., Crowther, T. W., Meili, N., Burlando, P., Katul, G. G., & Bou-Zeid, E. (2019). Magnitude of urban heat islands largely explained by climate and population. *Nature 2019 573:7772*, *573*(7772), 55–60. https://doi.org/10.1038/s41586-019-1512-9

Martilli, A., Roth, M., Chow, W. T. L., Demuzere, M., Lipson, M., Krayenhoff, E. S., Sailor, D., Nazarian, N., Voogt, J., Wouters, H., Middel, A., Stewart, I. D., Bechtel, B., Christen, A., & Hart, M. A. (2020). Summer average urban-rural surface temperature differences do not indicate the need for urban heat reduction. *OSF Preprints*. https://doi.org/10.31219/osf.io/8gnbf

Masson, V. (2000). A physically-based scheme for the urban energy budget in atmospheric models. *Boundary-Layer Meteorology*, *94*(3), 357–397. https://doi.org/10.1023/A:1002463829265/METRICS

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., … Voldoire, A. (2013). The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geoscientific Model Development*, *6*(4), 929–960. https://doi.org/10.5194/gmd-6-929-2013

570 McDonald, R., Kroeger, T., Boucher, T., LongZhu, W., & Salem, R. (2016). Planting healthy air: a global analysis of the role of urban trees in addressing particulate matter pollution and extreme heat. *Planting Healthy Air: A Global Analysis of the Role of Urban Trees in Addressing Particulate Matter Pollution and Extreme Heat.*

Meili, N., Manoli, G., Burlando, P., Bou-Zeid, E., Chow, W. T. L., Coutts, A. M., Daly, E., Nice, K. A., Roth, M., Tapper, N. J., Velasco, E., Vivoni, E. R., & Fatichi, S. (2020). An urban ecohydrological model to quantify the effect of vegetation on

575 urban climate and hydrology (UT&C v1.0). *Geoscientific Model Development*, *13*(1), 335–362. https://doi.org/10.5194/gmd-13-335-2020

Natural Capital Project. (2022). InVEST 3.13.0 User's Guide. Stanford University, University of Minnesota, Chinese Academy of Sciences, The Nature Conservancy, World Wildlife Fund, and Stockholm Resilience Centre. In *http://releases.naturalcapitalproject.org/invest-userguide/latest/en/index.html*.

580 Nedkov, S., Zhiyanski, M., Dimitrov, S., Borisova, B., Popov, A., Ihtimanski, I., Yaneva, R., Nikolov, P., & Bratanova-Doncheva, S. (2017). Mapping and assessment of urban ecosystem condition and services using integrated index of spatial structure. *One Ecosystem*, *2*, e14499. https://doi.org/10.3897/oneeco.2.e14499

NOAA. (2020). *NOWData: Minneapolis-St Paul Area monthly summarized data, 1981-2010, mean of monthly average temperatures*. National Weather Service Forecast Office, Twin Cities.

585 Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, *108*(455), 1–24. https://doi.org/10.1002/qj.49710845502

Phelan, P. E., Kaloush, K., Miner, M., Golden, J., Phelan, B., Silva, H., & Taylor, R. A. (2015). Urban Heat Island: Mechanisms, Implications, and Possible Remedies. *Annual Review of Environment and Resources*, *40*, 285–307. https://doi.org/10.1146/annurev-environ-102014-021155

590 Potchter, O., Cohen, P., & Bitan, A. (2006). Climatic behavior of various urban parks during hot and humid summer in the mediterranean city of Tel Aviv, Israel. *International Journal of Climatology*, *26*(12), 1695–1711. https://doi.org/https://doi.org/10.1002/joc.1330

Reid, C. E., Snowden, J. M., Kontgis, C., & Tager, I. B. (2012). The Role of Ambient Ozone in Epidemiologic Studies of Heat-Related Mortality. *Environmental Health Perspectives*, *120*(12), 1627. https://doi.org/10.1289/EHP.1205251

595 Rosenzweig, C., Solecki, W., & Slosberg, R. (2006). *Mitigating New York City's heat island with urban forestry, living roofs, and light surfaces*. https://www.google.com/url?sa=t&rct=j&esrc=s&source=appssearch&uact=8&cd=0&cad=rja&q&sig2=-StYgC-Y9r4zWxjUdVyL6w&ved=0ahUKEwiE4PSmnuL9AhUPqzcKHbMoCRY4ABABKAAwAA&url=https://www.nyserda

.ny.gov/-/media/Project/Nyserda/files/Publications/Research/Environmental/EMEP/NYC-Heat-Island-Mitigation.pdf&usg=AOvVaw3OCBlUuCArLzFf6uyEc9Zp

Santamouris, M. (2015). Regulating the damaged thermostat of the cities—Status, impacts and mitigation challenges. *Energy and Buildings*, *91*, 43–56. https://doi.org/10.1016/J.ENBUILD.2015.01.027

Santamouris, M. (2016). Innovating to zero the building sector in Europe: Minimising the energy consumption, eradication of the energy poverty and mitigating the local climate change. *Solar Energy*, *128*, 61–94. https://doi.org/10.1016/J.SOLENER.2016.01.021

Santamouris, M. (2020). Recent progress on urban overheating and heat island research. Integrated assessment of the energy, environmental, vulnerability and health impact. Synergies with the global climate change. *Energy and Buildings*, *207*, 109482. https://doi.org/10.1016/J.ENBUILD.2019.109482

Schatz, J., & Kucharik, C. J. (2014). Seasonality of the Urban Heat Island Effect in Madison, Wisconsin. *Journal of Applied Meteorology and Climatology*, *53*(10), 2371–2386. https://doi.org/https://doi.org/10.1175/JAMC-D-14-0107.1

Shashua-Bar, L., & Hoffman, M. E. (2000). Vegetation as a climatic component in the design of an urban street: An empirical model for predicting the cooling effect of urban green areas with trees. *Energy and Buildings*, *31*(3), 221–235. https://doi.org/10.1016/S0378-7788(99)00018-3

Smoliak, B. V., Snyder, P. K., Twine, T. E., Mykleby, P. M., & Hertel, W. F. (2015). Dense network observations of the Twin Cities Canopy-Layer urban heat island. *Journal of Applied Meteorology and Climatology*, *54*(9). https://doi.org/10.1175/JAMC-D-14-0239.1

Spronken-Smith, R. A., & Oke, T. R. (1999). Scale Modelling of Nocturnal Cooling in Urban Parks. *Boundary-Layer Meteorology*, *93*(2), 287–312. https://doi.org/10.1023/A:1002001408973

Stéfanon, M., Martin-StPaul, N. K., Leadley, P., Bastin, S., Dell'Aquila, A., Drobinski, P., & Gallardo, C. (2015). Testing climate models using an impact model: what are the advantages? *Climatic Change*, *131*(4), 649–661. https://doi.org/10.1007/s10584-015-1412-4

Stewart, I. D. (2011). A systematic review and scientific critique of methodology in modern urban heat island literature. *International Journal of Climatology*, *31*(2), 200–217. https://doi.org/10.1002/JOC.2141

Tan, P. Y., Wong, N. H., Tan, C. L., Jusuf, S. K., Chang, M. F., & Chiam, Z. Q. (2018). A method to partition the relative effects of evaporative cooling and shading on air temperature within vegetation canopy. *Journal of Urban Ecology*, *4*(1). https://doi.org/10.1093/JUE/JUY012

Trabucco, A., & Zomer, R. J. (2018). *Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2*. CGIAR Consortium for Spatial Information (CGIAR-CSI).

Velasco, E. (2018). Go to field, look around, measure and then run models. *Urban Climate*, *24*, 231–236. https://doi.org/10.1016/J.UCLIM.2018.04.001

Venter, Z. S., Chakraborty, T., & Lee, X. (2021). Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms. *Science Advances*, *7*(22). https://doi.org/10.1126/SCIADV.ABB9569/SUPPL_FILE/SCIADV.ABB9569_SM.PDF

Viguié, V., Lemonsu, A., Hallegatte, S., Beaulant, A. L., Marchadier, C., Masson, V., Pigeon, G., & Salagnac, J. L. (2020). Early adaptation to heat waves and future reduction of air-conditioning energy use in Paris. *Environmental Research Letters*, *15*(7). https://doi.org/10.1088/1748-9326/ab6a24

Villanueva-Solis, J. (2017). Urban Heat Island Mitigation and Urban Planning: The Case of the Mexicali, B. C. Mexico. *American Journal of Climate Change*, *6*(1), 22–39. https://doi.org/10.4236/ajcc.2017.61002

Wan, Z. (2013). *Collection-6 MODIS Land Surface Temperature Products Users' Guide*.

Wang, T., Xue, L., Brimblecombe, P., Lam, Y. F., Li, L., & Zhang, L. (2017). Ozone pollution in China: A review of concentrations, meteorological influences, chemical precursors, and effects. *Science of the Total Environment*, *575*, 1582–1596. https://doi.org/10.1016/j.scitotenv.2016.10.081

Wong, N. H., Tan, C. L., Kolokotsa, D. D., & Takebayashi, H. (2021). Greenery as a mitigation and adaptation strategy to urban heat. *Nature Reviews Earth & Environment 2021 2:3*, *2*(3), 166–181. https://doi.org/10.1038/s43017-020-00129-5

Yu, Z., Yang, G., Zuo, S., Jørgensen, G., Koga, M., & Vejre, H. (2020). Critical review on the cooling effect of urban blue-green space: A threshold-size perspective. *Urban Forestry & Urban Greening*, *49*, 126630. https://doi.org/10.1016/J.UFUG.2020.126630

Zardo, L., Geneletti, D., Pérez-Soba, M., & Van Eupen, M. (2017). Estimating the cooling capacity of green infrastructures to support urban planning. *Ecosystem Services*, *26*, 225–235. https://doi.org/10.1016/J.ECOSER.2017.06.016

Zawadzka, J. E., Harris, J. A., & Corstanje, R. (2021). Assessment of heat mitigation capacity of urban greenspaces with the use of InVEST urban cooling model, verified with day-time land surface temperature data. *Landscape and Urban Planning*, *214*, 104163. https://doi.org/10.1016/J.LANDURBPLAN.2021.104163

Zhao, L., Lee, X., Smith, R. B., & Oleson, K. (2014). Strong contributions of local background climate to urban heat islands. *Nature*, *511*(7508), 216–219. https://doi.org/10.1038/nature13462

Ziter, C. D., Pedersen, E. J., Kucharik, C. J., & Turner, M. G. (2019). Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer. *Proceedings of the National Academy of Sciences*, *116*(15), 7575 LP – 7580. https://doi.org/10.1073/pnas.1817561116