



**HAL**  
open science

## Urban heat island temporal and spatial variations: Empirical modeling from geographical and meteorological data

Jérémy Bernard, Marjorie Musy, Isabelle Calmet, Erwan Bocher, Pascal  
Kéravec

► **To cite this version:**

Jérémy Bernard, Marjorie Musy, Isabelle Calmet, Erwan Bocher, Pascal Kéravec. Urban heat island temporal and spatial variations: Empirical modeling from geographical and meteorological data. *Building and Environment*, 2017, 125, pp.423-438. 10.1016/j.buildenv.2017.08.009 . hal-01630175v2

**HAL Id: hal-01630175**

**<https://hal.science/hal-01630175v2>**

Submitted on 1 Jun 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Urban heat island temporal and spatial variations : empirical modelling from geographical and meteorological data

BERNARD Jérémy<sup>a,b,1</sup>, MUSY Marjorie<sup>a,c</sup>, CALMET Isabelle<sup>a,d</sup>, BOCHER  
Erwan<sup>e</sup>, KERAVEC Pascal<sup>a,d</sup>

<sup>a</sup> CNRS, Institut de Recherche en Sciences et Techniques de la Ville, FR 2488, École  
Centrale de Nantes - Nantes France

<sup>b</sup> CNRS, Centre de Recherche Nantais Architectures Urbanités, UMR 1563, École Nationale  
Supérieure d'Architecture de Nantes - Nantes, France

<sup>c</sup> Cerema Ouest - Nantes, France

<sup>d</sup> CNRS, Laboratoire de Recherche en Hydrodynamique, Énergétique et Environnement  
Atmosphérique, UMR 6598, École Centrale de Nantes - Nantes, France

<sup>e</sup> CNRS, Laboratoire Lab-STICC, UMR 6285 - Vannes, France

---

## Abstract

Urban air temperature varies along time and space. This contribution proposes a methodology to model these variations from empirical models. Four air temperature networks are implemented in three west France cities. The period [1.4 - 1.8] of a normalized night is investigated. Three empirical models are established, in order to predict temporal and spatial air temperature variation. They are gathered under the name of *model group*. Urban heat island (UHI) intensity is explained according to one temporal model using meteorological variables. Temperature spatial variations are explained from two models using geographical informations averaged in a 500 m radius buffer circle. The first one represents the mean UHI value for a given location, the second the variability of the UHI around its mean value. 32 model groups are calibrated from data sampled by one network. The accuracy of their estimations is tested comparing estimated to observed values from the three other networks. The most accurate model is identified and its performances are analyzed. Night-time UHI intensity variations are explained all along the year by wind speed and nebulosity values calculated after sunset. During spring and summer season UHI variations are mainly driven by Normalized Difference Vegetation Index (NDVI) value whereas building density or surface density are predominant explicative variables during colder seasons. The model is finally applied to the city of Nantes during a summer night to illustrate the interest of this work to urban planning applications. Several warm areas located outside the city center are identified according to this method. Their high UHI value is mainly caused by their low NDVI value.

---

\*Corresponding author. E-mail adress: jeremy.bernard.jb@gmail.com Address : IRSTV, École centrale de Nantes, 1 rue de la Noë, 44300 Nantes, France

*Keywords:* urban climate, air temperature observation, empirical model, urban heat island, meteorological variable

---

## 1. Introduction

Air temperature is higher in urban areas than in their surrounding (Oke, 1987). This phenomenon, called Urban Heat Island (UHI), may become a serious issue in the next decades due to global warming. According to the Intergovernmental Panel on Climate Change (IPCC) results, one of the consequences of climate change is an increase of heat waves intensity and frequency (Revel et al., 2012). The combination of UHI and climate change will impact urban comfort (Clarke and Bach, 1971), human and fauna heat stress (Lowe, 2016; Clarke, 1972), building cooling demand (Malys et al., 2012), vegetation behavior such as flowering phenology (Neil and Wu, 2006), etc.

However, urban heat island causes are known and may be classified in three categories :

- urban morphology causes : high facade density leads to short and long wave radiation trapping in the canopy (Bernabé et al., 2015), and high roughness is also responsible for wind speed reduction, preventing heat evacuation (Britter and Hanna, 2003).
- urban surfaces causes : high fraction of impervious soils reduces water storage, thus being unable to cool by evapotranspiration as vegetated soils do (Asaeda and Ca, 1993). Moreover, low building materials albedo (e.g. asphalt) increases the absorption of solar radiation which contributes to ground and air overheating (Santamouris, 2014).
- human activity causes : antropogenic heat sources (road traffic and buildings) are higher in cities than in rural areas (Sailor, 2011; de Munck et al., 2013).

Urban climate is not the only issue that decision makers have to deal with during both the design process of new neighborhoods and the definition of an urban planning strategy. They also have to take into account many other issues such as public transportation, education and health access, water, waste and energy management, functional and social mixity etc. For urban designing (construction or renovation), the choice of the urban form, of the soil type, and of the function of each building may affect each of the authority of the decision makers. To make a decision, there is a need of simple indicators to evaluate the impact of a new urban project on each of them. The objective of this article is to propose indicators dedicated to the urban climate field. Indeed, as cities become denser, there is a need to know which neighborhoods are the warmest and to evaluate the potential impact of a new urban project on air temperature.

Tools developed to tackle this issue may use either knowledge based model or empirical model. Knowledge based modeling consists in splitting the entire city into several volumes and to apply on each of them an energy balance

(Grimmond et al., 2009). One of the major shortcomings of these models is the complexity of both the method (numerical simulation) and the physical concept (fluid dynamics, radiation, convection and conduction principles) used. Empirical models are constructed from field observation where measurements are explained from the geographical context. For example, Tokyo urban climate was investigated measuring the mean night-time air temperature of 11 locations during 30 days (Yokobori and Ohta, 2009). The vegetation density of each site was calculated to explain the temperature spatial differences. Finally, the empirical model they proposed consists in a linear relationship between air temperature and vegetation density. This method has several shortcomings. We identified the four that we consider of major importance :

1. *Relevant period of interest* : within a day, air temperature is rarely homogeneous at city scale. Heterogeneity, defined as the standard deviation of all measurement at a given time, is maximum at certain periods of the day. However, the literature is conflicting concerning the best period to calculate air temperature differences between each part of the city and a reference station, and the choice of a period rather than another is rarely justified. Yan et al. (2014) compared the air temperature at 11 p.m., Oke (1976) averaged the values between one and three hours after sunset, Petralli et al. (2014) calculated the differences when air temperature was maximum, and Krüger and Givoni (2007) averaged the differences during the whole day.
2. *Cross modeling of spatial and temporal variations* : within an urban area, air temperature heterogeneity may be explained by the heterogeneity of geographical contexts (morphology, surface types, human activities). Empirical models such as the one used by Yokobori and Ohta (2009) have been proposed to model the spatial variations of the mean air temperature from geographical indicators. This air temperature heterogeneity may be exacerbated by the meteorological conditions. It is higher under clear (few clouds) and calm (low wind speed) conditions than under cloudy and windy conditions (Svensson et al., 2002). Empirical models have been proposed to model the temporal variations of the air temperature difference between a given point of the city and a reference station from meteorological indicators (Svensson et al., 2002; Krüger and Givoni, 2007). However, these relationships are only valid for the given point and thus cannot be used to estimate the temperature of an other site. There is a need to establish an overall relationship to link spatial and temporal temperature variations to both geographical and meteorological variations.
3. *Temporal variations modeling* : while some meteorological variables have an immediate impact on air temperature, some others can be delayed due to inertial effects. For example, solar radiation occurs during the daytime but still has an impact on night-time temperature differences. The relationships between observed temperature differences and their potential causes (wind speed, solar radiation, etc) must not be studied synchronically. However, few studies have investigated this point : most of the time,

the average meteorological condition (the causes) and the average temperature differences (the consequences) are calculated at the same time.

4. *Spatial variations modeling* : air temperature differences are due to energy flux differences occurring within the city. The value recorded for each flux is strongly correlated to the geographical context of the site of interest (for example, the amount of vegetation impacts latent heat flux value). However, indicators are rarely linked to an energy flux occurring within an urban volume. Establishing a clear link between a heat flux and a geographical indicator used in an empirical model would allow a more physical analyze of the model equation.

Our main goal is to propose a new methodology based on measurement analysis to estimate the spatial and temporal variations of the air temperature within an urban area, providing improvements to overcome the four main shortcomings identified in the previous models. This methodology distinguishes spatial variations which are calibrated according to geographical indicators from temporal variations which are calibrated according to meteorological conditions. The model relies on four air temperature measurement networks, which are presented in section 2. One of them is used for model calibration whereas the three others are used for model verification. Meteorological and geographical data, respectively responsible for temporal and spatial variations of the UHI, are also presented section 2. Section 3 is dedicated to the methodology description (i.e. model calibration). The three first subsections consist in the general method description. We propose solutions to solve each of the four issues described above. Two spatial and one temporal models are then obtained. They are gathered under the name of model group. In the last subsection, five variations of this general methodology are proposed, leading to the obtention of 32 model groups (32 temporal models and 16 spatial models). A method to choose the most performing model group is also proposed. The section 4 is dedicated to the analysis of this model group and its performances by comparing estimated and observed temperature values.

## 2. Materials

### 2.1. Study areas

The study areas are three west France conurbations : Nantes, Angers and La Roche-sur-Yon. According to a census performed in 2012, they have respectively 606'640, 218'657 and 52'808 inhabitants. They are located in a temperate climatic zone with a warm summer but without dry season (Cfb climatic zone according to *Köppen climate classification* - (Peel et al., 2007)). They are all further than 20 km away from any major climatic disturbing elements such as oceans or mountains. Although sea breezes may affect rural air temperature much beyond 20 km from the coast (Hu and Xue, 2016), the temperature field during night-time (where the UHI intensity is the highest) is dominated by the influence of the urban land-use (Pigeon et al., 2006). Because our contribution consists in setting the foundation of a method to solve the issues mentioned in

section 1, we make the choice not taking into account the potential sea breeze influence for this preliminary work.

## 2.2. Meteorological data

Meteorological data are splitted in two categories :

- urban air temperature networks : they are used to observe temperature heteogeneity within each urban area;
- reference stations for meteorological data : they are used to quantify regional meteorological conditions and also as references for air temperature differences between the city measurements and the surrounding rural areas.

### 2.2.1. Urban air temperature networks

Four measurement networks are implemented within our three study areas and are presented figure 1 : two are located in Nantes, one in Angers and one in La Roche-sur-Yon (LRY).

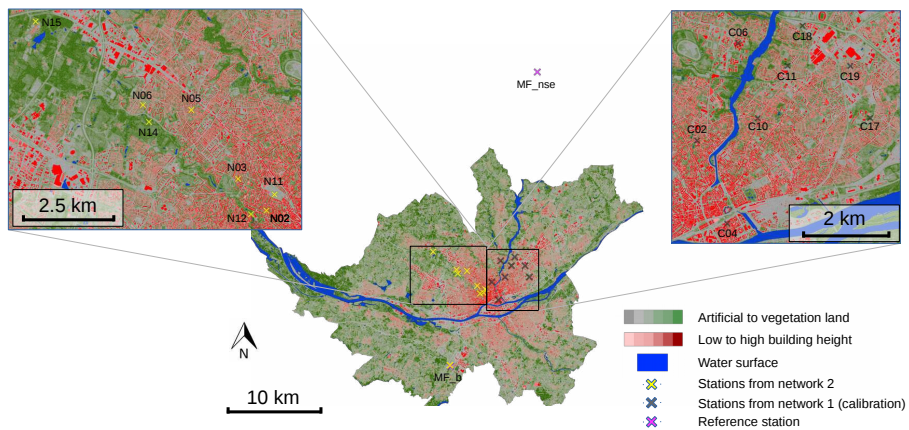
One of the two Nantes networks is used for calibration step, whereas the three other networks are used for verification purpose. The quantity of available data for each network is given in table 1. A day is considered as available only if all network stations have no missing data. The network chosen for calibration is the one having the highest number of measurement stations as well as the highest number of measurement days per season. All measurement stations sample a temperature value every 15 minutes using positive temperature coefficient (PTC) thermistor. Those sensors have an accuracy lower than 0.3 K and they are renowned for their negligible temperature time-drift. For this reason, they are supposed to be calibrated all along the experiment.

Table 1: Description of the data sets utilized for model calibration and model verification purpose (Source : Bernard (2017))

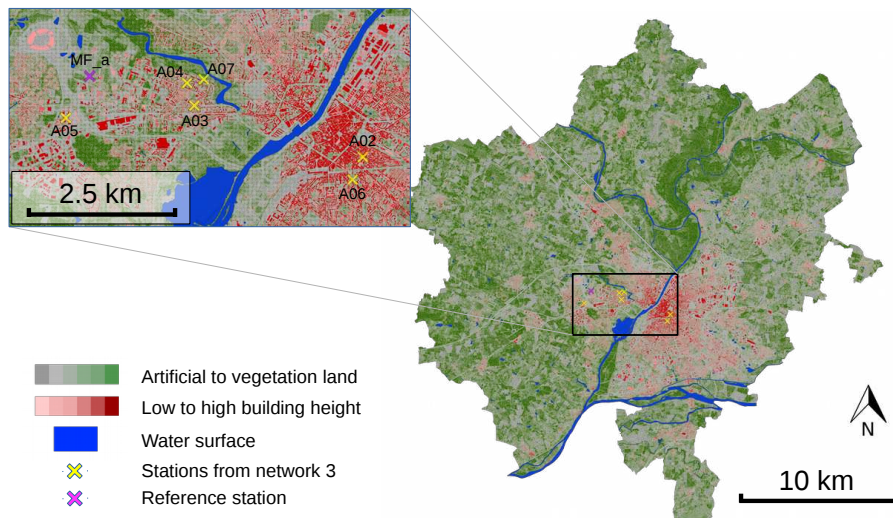
ID	City	Model purpose	stations number	Measurement period	Availability of measurement per season (days)*			
					winter	spring	summer	autumn
1	Nantes	Calibration	9	2010-2015 (4 years)	141	208	127	116
2	Nantes	Verification	8	2015 (1 year)	0	81	39	0
3	Angers	Verification	6	2015 (1 year)	5	92	90	78
4	LRY	Verification	5	2015 (1 year)	3	92	92	72

\* A day is considered as available only if all network stations have no missing data.

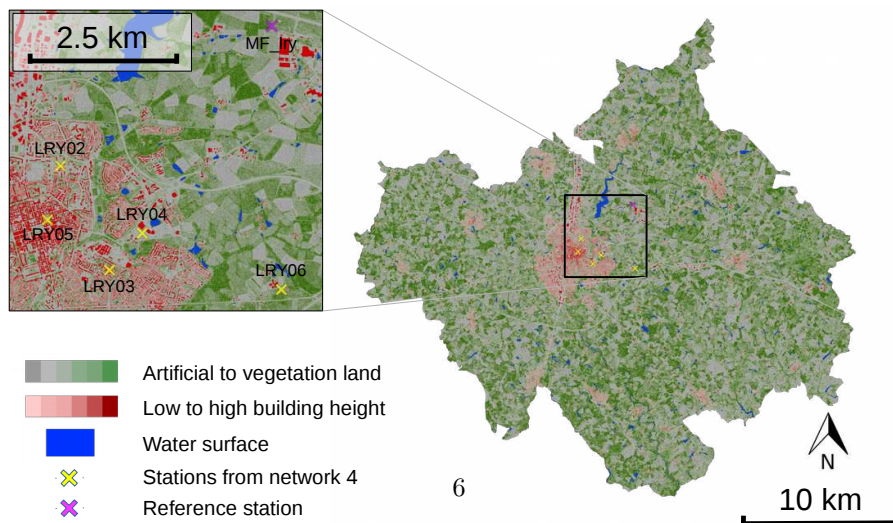
Experimental conditions are given in table 2. Calibration network (ID 1) conditions differ from the verification network ones (ID 2, 3 and 4). The measurement distance from the ground is acceptable for both verification and calibration networks since temperature gradient is negligible within the urban canopy above 2 m (Armstrong, 1974; Nakamura and Oke, 1988; Taesler, 1980). However, the radiation shield used for the calibration network is very sensitive to solar radiation whereas the one used for verification is almost not sensitive when used under high solar radiation intensity (Lacombe et al., 2011; Lacombe, 2010). Moreover, the location of a measurement station at less than 1 m from



(a)



(b)



(c)

Figure 1: Location of the measurement stations of the conurbation of Nantes (a), Angers (b) and LRY (c)

Table 2: Experimental conditions for each network

Network ID	Measurement height from the ground	Measurement distance from a facade	Natural ventilation shield used (shape and volume)
1	2 - 5 m	$\geq 0.3$ m	cylinder - 0.1 dm <sup>3</sup>
2, 3 and 4	3 m	$\geq 1$ m	cuboid - 1 dm <sup>3</sup>

a vertical surface may lead to considerable errors during day-time (Oke, 2004). For those reasons and also because temperature differences are often lower during day-time than during night-time, only night-time temperature differences will be investigated in this study.

### 2.2.2. Reference stations for meteorological data

Because the empirical models developed in this study must be applied to any other french city, data used both to characterize meteorological conditions and as temperature reference must be standardized. Meteo-France stations are chosen for three reasons :

- they respect measurement recommendations of the World Meteorological Organization (local vegetation height lower than 10 cm, no heat source or water surfaces in a 100 m radius circle, etc.),
- they are located in each of the major urban areas of France,
- they have long-term records that cover our measurement period (2010 to 2016).

The UHI empirical model is calibrated with Nantes data and should be applied to the other cities. Thus, we need to use a temperature reference for each city. Meteo-France temperature data, which are sampled every 60 minutes, is linearly interpolated every 15 minutes in order to have a sampling frequency identical to the one of all the other stations used for calibration or verification purpose.

Meteo-France stations are also used to characterize the meteorological conditions. Six physical quantities, measured every hour, are utilized for this work : temperature and relative humidity (1.5 m height), wind speed and direction (10 to 11 m height), solar radiation and nebulosity. The sampling method utilized for the measurement differs from one physical quantity to another. Wind speed and direction as well as nebulosity are averaged during the last ten minutes of each hour. Air temperature and relative humidity are sampled every hour. Each solar radiation measurement is the average value of the previous hour. Two Météo-France stations are used for the urban area of Nantes : nebulosity is measured at the station called *MF\_b* whereas the five other physical quantities are measured at the station called *MF\_nse* (figure 1a). For the Angers and La Roche-sur-Yon stations, nebulosity is rarely measured during night-time, thus night-time values are estimated using linear interpolation from day-time val-



ues. More informations concerning each reference station may be found on the Météo-France web site <sup>1 2 3 4</sup>.

### 2.3. Geographical data

Because the empirical models developed in this study must be applied to any other french city, data used to characterize the geographical context of each measurement station must be standardized. Those data are presented table 3. The BDTopo®(version 2011), produced by the french IGN (National geographical and forest institute), is a vector-based database. It is used to quantify urban morphology (building footprint and building height are given) as well as urban land cover (at least water surfaces). To characterize urban land cover further in detail, RapideEye images are have been acquired at noon on May, 23rd 2011 in Nantes, on May 21st 2011 in LRY and on May, 23rd in Angers. They have been provided by the Equipex Geosud project<sup>5</sup>. Orthorectification had been performed by the IGN (according to Benosa (2014)). They have a 5 m horizontal resolution and they are composed of five spectral-bands (blue, green, red, red-edge, near infrared).

Table 3: Geographical data description

Name	Provider	Type	City	characteriza- tion	Acquisition date or version name	Attributes
BD Topo	IGN	Vector (topo- graphical database)	Urban	morphology and urban land type (exclusively water surfaces)	Version 2011	Building footprint building height water surface foot- print
Rapideye image	Equipex Geosud	Raster (5 m resolu- tion)	Urban	land type	2011-05-23 or 2011-05- 21 at noon depending on the city	Reflectance for 5 spectral band (blue, green, red, red-edge, near infrared)

### 3. Methods

The methodology is only applied for the calibration network. In order to have both an equal day-time and night-time duration all along a year, they have to be transformed in a dimensionless time variable. Each day and each night are divided by their duration. Each day starts at 0 and ends at 1, each night starts at

<sup>1</sup>MF\_b (Nantes): [https://donneespubliques.meteofrance.fr/metadonnees\\_publicques/fiches/fiche\\_44020001.pdf](https://donneespubliques.meteofrance.fr/metadonnees_publicques/fiches/fiche_44020001.pdf) accessed in July 2017

<sup>2</sup>MF\_nse (Nantes): [https://donneespubliques.meteofrance.fr/metadonnees\\_publicques/fiches/fiche\\_44110002.pdf](https://donneespubliques.meteofrance.fr/metadonnees_publicques/fiches/fiche_44110002.pdf) accessed in July 2017

<sup>3</sup>MF\_a (Angers) : [https://donneespubliques.meteofrance.fr/metadonnees\\_publicques/fiches/fiche\\_49020001.pdf](https://donneespubliques.meteofrance.fr/metadonnees_publicques/fiches/fiche_49020001.pdf) accessed in July 2017

<sup>4</sup>MF\_lry (La Roche-sur-Yon) : [https://donneespubliques.meteofrance.fr/metadonnees\\_publicques/fiches/fiche\\_85191003.pdf](https://donneespubliques.meteofrance.fr/metadonnees_publicques/fiches/fiche_85191003.pdf) accessed in July 2017

<sup>5</sup><https://www.theia-land.fr/en/projects/geosud>

1 and ends at 2. The aim of this study is to model the air temperature differences within the urban area from geographical and meteorological conditions.

### 3.1. Identification of an air temperature calculation period

Temperature varies temporally and spatially during night-time. The temperature of the urban area is often quite homogeneous at the sunset and strongly heterogeneous few hours after sunset. We define temperature spatial dispersion as the standard deviation of the temperature between all measurement stations. We are interested in heterogeneous situations, i.e. when the UHI phenomenon is maximum, highlighting high temperature differences between all parts of a city. Then interesting time-periods to investigate are the ones characterized by high and steady temperature spatial dispersion. Those periods are identified using four steps. First, for each time step  $t$  of each day  $d$ , air temperature dispersion  $\sigma_{T_d}(t)$  is calculated as the standard deviation of the temperature between all measurement stations.

$$\sigma_{T_d}(t) = \sqrt{\frac{\sum_i (T_{i,d} - \bar{T}_d)^2}{n_i - 1}} \quad (1)$$

where  $n_i$  the number of stations,  $\bar{T}_d = \frac{\sum_i T_{i,d}}{n_i}$  the urban average temperature at the sample  $t$ .

Second, for each day, day-time and night-time are scaled by their length according to Runnalls and Oke (2000). Third, the evolution of the night-time dispersion is averaged for each season  $S$  (figure 2a) :

$$\sigma_{T_s}(t) = \frac{\sum_{d \in S} T_d(t)}{n_{d,S}} \quad (2)$$

where  $n_{d,S}$  the number of days used during season  $S$ .

These first three steps are applied to temperature time-derivative ( $\dot{T}$ ) in order to investigate the stability of the temperature spatial dispersion (figure 2b). The last step consists in the period identification, characterized by high standard deviation of the temperature (high temperature spatial dispersion) and low standard deviation of the temperature derivative (stable temperature spatial dispersion). This period ( $p_{UHI}$ ), corresponding to the time interval [1.4 - 1.8], is identified on figure 2.

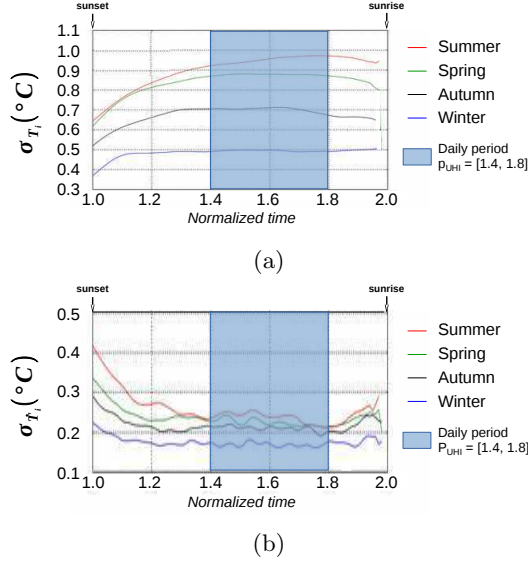


Figure 2: Mean night-time temperature dispersion (a) and mean high-time cooling rate dispersion (b) for each season represented along a normalized time axis

### 3.2. Cross modeling of spatial and temporal variations

Temperature difference  $\Delta T_{i,d}$  between each station  $i$  and the reference station  $ref$  is calculated. First, temperature is averaged during the period  $p_{UHI}$  for each day  $d$  :

$$\overline{T}_{i,d} = \frac{\sum_{t \in p_{UHI}} T_{i,d}(t)}{n_{t \in p_{UHI}}} \quad (3)$$

where  $n_{t \in p_{UHI}}$  is the number of samples in period  $p_{UHI}$ .

Then temperature difference between each station  $i$  and the reference station  $ref$  is calculated for each day  $d$  :

$$\Delta T_{i,d} = \overline{T}_{i,d} - \overline{T}_{ref,d} \quad (4)$$

For a given measurement station  $i$ ,  $\Delta T_{i,d}$  value varies between days. Those variations may be explained by statistical models using meteorological conditions (wind speed, solar radiations, etc.) as explanative variables. The problem is that the average and the standard deviation of the variations may differ a lot depending on the station analysed. In this condition, one model can not be used to explain the temporal variation for any location of the city. To overcome this issue, temporal variations are normalized ( $\widetilde{\Delta T}_d$ ) according to equation 5 :

$$\widetilde{\Delta T}_d = \alpha_i \frac{\Delta T_{i,d} - \overline{\Delta T}_i}{\sigma_{\Delta T_i}} \quad (5)$$

where  $\overline{\Delta T_i} = \frac{\sum^d \Delta T_{i,d}}{n_d}$  is the average temperature difference for station  $i$ ,  $n_d$  is the number of days,  $\sigma_{\Delta T_i} = \sqrt{\frac{\sum^d (\Delta T_{i,d} - \overline{\Delta T_i})^2}{n_d - 1}}$  is the standard deviation of its temporal variations,  $\alpha_i$  is a coefficient equal to 1 if  $\Delta T_{i,d} \geq 0$  and equal to -1 if  $\Delta T_{i,d} < 0$ .

This normalization allows to decompose  $\Delta T_{i,d}$  into spatial and temporal indicators:

- temporal variations are represented by  $\widetilde{\Delta T_d}$  and will be modeled according to meteorological variable  $V_m$  :

$$\widetilde{\Delta T_d} = F(V_m) \quad (6)$$

- spatial variations are represented by both  $\overline{\Delta T_i}$  and  $\sigma_{\Delta T_i}$  and will be modeled according to geographical indicators ( $I_g$ ), representing the geographical context of each measurement station :

$$\overline{\Delta T_i} = G(I_g) \quad (7)$$

$$\sigma_{\Delta T_i} = H(I_g) \quad (8)$$

The two spatial models (equations 7 and 8) as well as the temporal model (equation 6) which are finally obtained are gathered under the name of model group and may be used to calculate  $\Delta T_{i,d}$  for any urban areas :

$$\Delta T_{i,d} = \frac{\sigma_{\Delta T_i} \widetilde{\Delta T_d}}{\alpha_i} + \overline{\Delta T_i} \quad (9)$$

For reading clarity concern, indices  $d$  and  $i$  will be dropped for  $\Delta T_{i,d}$ ,  $\sigma_{\Delta T_i}$ ,  $\widetilde{\Delta T_d}$  and  $\overline{\Delta T_i}$  in the following sections.

### 3.3. Temporal variations modeling

In order to build the  $\widetilde{\Delta T}$  model, the following meteorological variables are used : solar radiation (K), wind speed (U), air temperature (T), relative humidity (RH) and nebulosity (N). Three main issues are associated to this objective :

- To which meteorological variables is  $\widetilde{\Delta T}$  sensitive ?
- During which part of the day (period  $p_p$ ) have these variables the greatest influence on  $\widetilde{\Delta T}$  ?
- What kind of function  $f_m$  relate  $\widetilde{\Delta T}$  to each variable  $V_m(p_p)$  ?

To answer these questions, several multiple linear regressions using meteorological variables are tested to explain  $\widetilde{\Delta T}$  variations (equation 10).

$$\widetilde{\Delta T} = a_0 + \sum_m a_m \cdot f_m(V_m(p_p)) \quad (10)$$

where  $a_0$  is the regression intercept,  $a_m$  is the regression coefficient associated to the explanatory variable  $f_m(V_m)$ .

The test procedure is composed of five steps presented in figure 3.

1. Day-time and night-time meteorological data are temporally normalized;
2. Each meteorological variable is averaged during two day-time periods ( $p_C = [0.7 - 1]$  and  $p_D = [0 - 1]$ ) and two night-time periods ( $p_A = [1.4 - 1.8]$  and  $p_B = [1 - 1.4]$ );
3. Several algebraic functions (exponential, inverse, linear, logarithmic, quadratic) are tested to linearize the relationship between  $\widetilde{\Delta T}$  and each combination of meteorological variable / period  $p_p$ . Then simple linear regressions are performed.
4. For each meteorological variable, the function  $f_{m_{fin}}$  and the period  $p_{fin}$  that maximise the simple linear regressions determination coefficient ( $R^2$ ) are selected for each meteorological variable  $V_m$ ;
5. Finally, multiple linear regressions are calculated for all combinations of explanatory variables  $f_{m_{fin}}(V_m(p_{m_{fin}}))$ . A combination is excluded when any of its regression coefficients is insignificant ( $p > 0.05$ ). The combination that maximizes the  $AIC_c$  criteria (equation 11) is finally chosen.

$$AIC_c = -2\ln(\max(f_L)) + 2nb_{x_n} + \frac{2nb_{x_n}(nb_{x_n} + 1)}{nb_{ech} - nb_{x_n} - 1} \quad (11)$$

where  $AIC_c$  is the corrected Akaike information criterion,  $\max(f_L)$  is the likelihood function maximum,  $nb_{x_n}$  is the explanatory variables number,  $nb_{ech}$  is the sample size used for regression (here the number of days  $n_d$ ).

### 3.4. Spatial variation modeling

Two models are built to explain the spatial distribution of the UHI. The first is used to model UHI mean value ( $\widetilde{\Delta T}$  - equation 7) and the second the UHI standard deviation value ( $\sigma_{\Delta T}$  - equation 8). They are estimated from linear regressions using geographical indicators as explanative variables, calculated according to the procedure described figure 4. In order to use geographical indicators that could explain the UHI phenomenon, we associate each heat flux involved in the energy budget of an urban volume to one or several geographical indicators. Those indicators are calculated thanks to the OrbisGIS platform (Bocher and Petit, 2013) from the formulae presented in table 4 and according to the data processing illustrated figure 4.

The energy budget of an urban volume is :

$$Q^* + Q_F = Q_H + Q_E + \Delta Q_A + \Delta Q_S \quad (12)$$

where  $Q^*$  is the net all-wave radiation flux density,  $Q_F$  the anthropogenic heat flux density due to combustion,  $Q_H$  the turbulent sensible heat flux density,  $Q_E$  the turbulent latent heat flux density,  $\Delta Q_A$  the horizontal energy transport in the air per unit horizontal area and  $\Delta Q_S$  the sub-surface heat flux density.

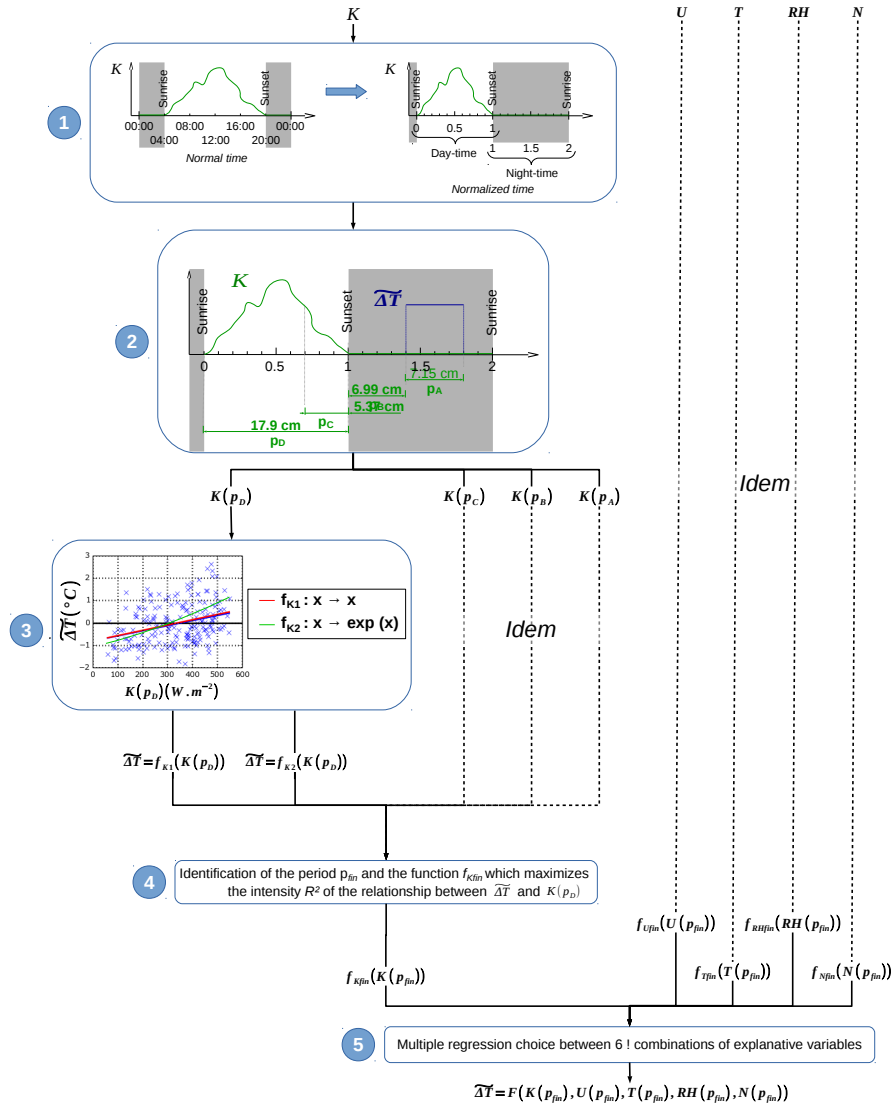


Figure 3: Test procedure to identify the relationship between UHI intensity and meteorological conditions

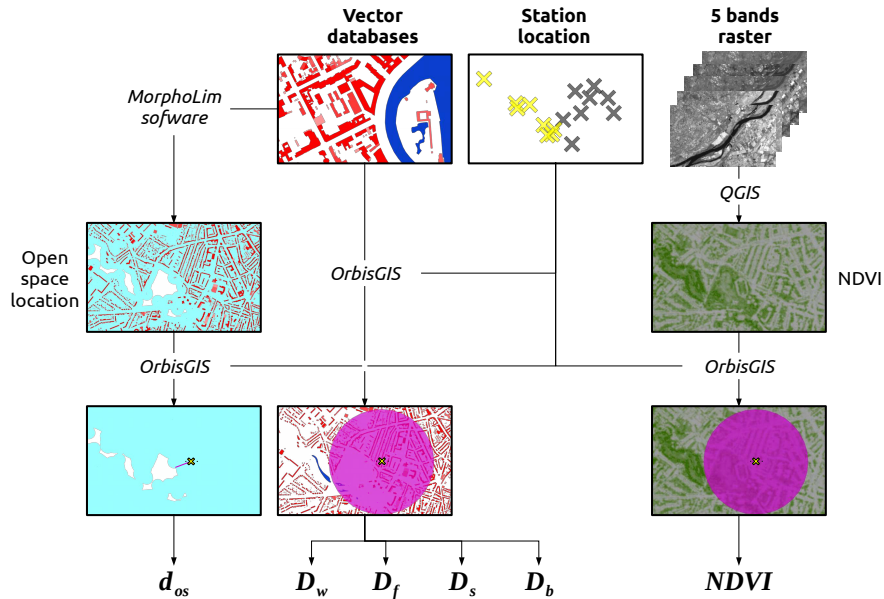


Figure 4: Geographical database process for geographical indicator calculation

According to Maby (2003) definition, geographical indicators may be divided into thematic and spatial indicators. The thematic indicators are constructed averaging a geographical information within a 500 m radius buffer circle centered on a measurement station - the value of 500 m is arbitrarily chosen as the median value taken from a literature review (Krüger and Givoni, 2007; Yan et al., 2014; Gal et al., 2009; Yokobori and Ohta, 2009; Steeneveld et al., 2014; László and Szegedi, 2012; Fenner et al., 2014; Erell and Williamson, 2007; Petralli et al., 2014; Chen et al., 2012; Balázs et al., 2009; Hjort and Marmion, 2008; Hart and Sailor, 2009; Suomi et al., 2012). The spatial indicators measure a distance separating a measurement station from a specific object. In this study, these geographic features are the open spaces, defined from urban morphology informations according to a fractal approach using the software MorphoLim (Tannier et al., 2011).

Table 4: Description of the geographical indicators used to characterize the UHI spatial variations

Heat flux	Indicator symbol $I_g$	Indicator name	Geographical data used	Basic indicators or informations used for indicator calculation	Formula	Relationship between heat flux and geographical indicator
$Q^*$	$D_f$	Facade density	BDTopo $\textcircled{R}$	$A_0$ : Area of the 500 m radius circle $A_f$ : total facade area contained in $A_0$	$\frac{A_f}{A_0+A_f}$	Net short wave (Bernabé et al., 2015) and long wave radiations are correlated to the sky view factor which is directly linked to facade density
$Q_F$	$D_b$	Building density	BDTopo $\textcircled{R}$	$A_0$ : Area of the 500 m radius circle $A_b$ : total roof area contained in $A_0$	$\frac{A_b}{A_0}$	We consider that higher the building surfaces density, higher the potential anthropogenic heat flux ( $\frac{A_b+A_f}{A_0} = D_b + D_s - 1$ )
	$D_s$	Urban surface density	BDTopo $\textcircled{R}$	$A_0$ : Area of the 500 m radius circle $A_f$ : total facade area contained in $A_0$ , $A_b$ : total roof area, contained in $A_0$	$\frac{A_0+A_f}{A_0}$	
$Q_H$ and $Q_E$	$NDVI$	Normalized Difference Vegetation Index	Rapideye image	$R_{red_p}$ : Reflectance of the pixel $p$ in the red band $R_{nir_p}$ : Reflectance of the pixel $p$ in the near infrared band, $A_p$ : Area of the pixel $p$ $A_0$ : Area of the 500 m radius circle	$\frac{\sum_{p \in A_0} A_p \frac{R_{nir_p} - R_{red_p}}{R_{nir_p} + R_{red_p}}}{\sum_{p \in A_0} A_p}$	A linear relationship between $Q_E$ and $NDVI$ is proposed by Pelgrum and Bastiaanssen (1996)
$\Delta Q_S$	$D_s$	Urban surface density	BDTopo $\textcircled{R}$	$A_0$ : Area of the 500 m radius circle $A_f$ : total facade area contained in $A_0$ $A_b$ : total roof area contained in $A_0$	$\frac{A_f}{A_0}$	Higher the urban surface density, higher the potential for heat flux storage
$\Delta Q_A$	$lnd_{op}$	Logarithm of the distance from open space	Open space zones defined from BD-Topo $\textcircled{R}$ and the software MorphoLim	$d_{op}$ : distance from open space	$lnd_{op}$	Open spaces may be sources of cooling that may drift to their close environment (Balázs et al., 2009; Spronken-Smith and Oke, 1999)



All linear combinations of these 6 indicators are tested to explain the variables  $\overline{\Delta T}_i$  and  $\sigma_{\Delta T_i}$  using multiple regression relationships (equations 13 and 14).

$$\overline{\Delta T} = b_0 + \sum_g b_g \cdot x_g \quad (13)$$

$$\sigma_{\Delta T} = c_0 + \sum_g c_g \cdot x_g \quad (14)$$

where  $b_0$  and  $c_0$  are the regression intercepts,  $b_g$  and  $c_g$  are the regression coefficients associated to the explanative variable  $x_g$ ,  $x_g = I_{g_i} - I_{g_{ref}}$  is the geographical indicator  $I_g$  difference between the site  $i$  and the site *ref*.

The step 5 of the procedure used for temporal model choice (section 3.3) is applied again in order to identify the relationship that best explains each of the variables ( $n_{ech} = n_i$  the number of stations used).

### 3.5. Variants to the general methodology

At the end of the general methodology application, one model group is obtained (figure 5). The performance of this model has to be assessed. However, in order to test the performances of more models, five variants to the general methodology are proposed. They are illustrated in figure 5 by dash arrows :

1. *Wind direction sorting* : the source area which contributes to the value measured by the temperature sensor strongly depends on the wind direction (Oke, 2004). In order to consider this “footprint” phenomenon, meteorological data and the buffer circle used for the thematic indicator calculation are divided into six parts. Each of them correspond to the source area of the temperature sensor when the wind blows from a certain range of direction. Each circle part is considered as a proper station and has its proper temperature data, multiplying artificially the number of stations by 6.
2. *Seasonal sorting* : depending on the season, geographical indicators (related to heat fluxes) may have a different weighting contribution to  $\overline{\Delta T}_i$  and  $\sigma_{\Delta T_i}$  values (equations 13 and 14). Meteorological measurement are then splitted by season according to World Meteorological Organization (WMO) definition for the North hemisphere. One model group is then proposed for each season.
3. *Most sensible station use* : for certain stations which have a geographical context close to the one of the reference station,  $\Delta T$  is low and does not fluctuate much when meteorological conditions change. In this case, their temporal variations are not well explained by meteorological conditions because measurement uncertainty becomes predominant. To overpass this potential lack of accuracy, only the three stations having the highest temporal variability (standard deviation  $\sigma_{\Delta T}$ ) are used for the temporal model calibration.

4. *Robust normalization* : in order to separate temporal from spatial components,  $\Delta T_{i,d}$  has been normalized using mean and standard deviation values (equation 5). Yet, these indicators are very sensitive to outlier values. Then they are respectively replaced by median and median absolute deviation (MAD) which are robust indicators to outlier values.
5. *Robust regression* : multiple linear regressions used to establish temporal and spatial models are calibrated according to least square method, which is again very sensitive to outlier data points. The Huber M-estimator, robust to outliers is then used for regression.

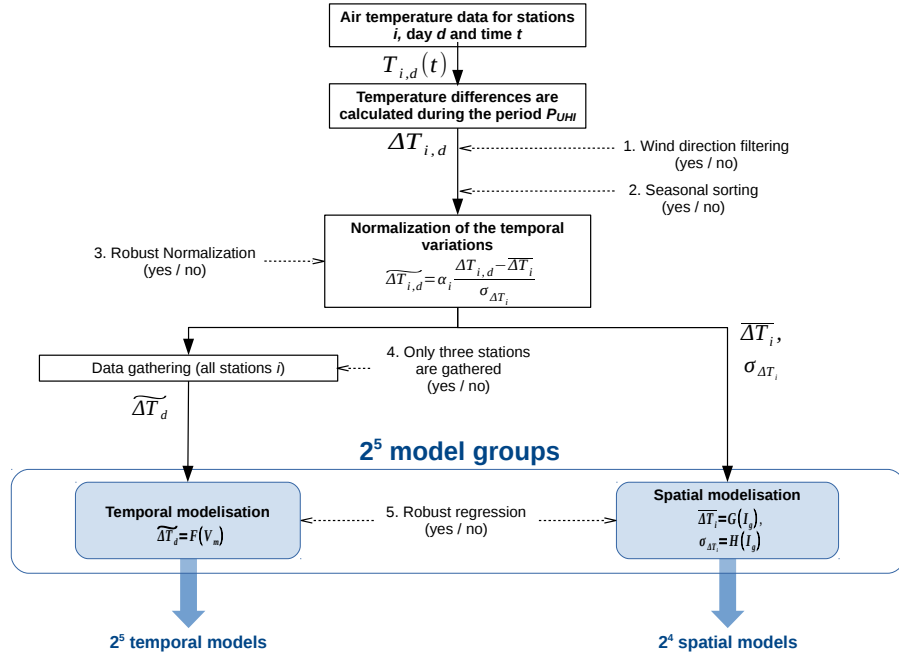


Figure 5: General methodology and variants (dash arrows) used to model spatial and temporal UHI variations

The combination of these 5 variants leads to the calibration of 32 model groups, composed of 32 temporal models ( $2^5$ ) and only 16 spatial models ( $2^4$ ), the variant 3 having no effect on spatial modeling (figure 5). In order to choose the best model, each of them is used to estimate the temperature difference  $\Delta T$  between each station of the 3 verification networks (ID 2, 3 and 4) and their reference station. The accuracy of these estimations is evaluated by calculating a normalized root mean square error (NRMSE), defined as the ratio between  $RMSE_i$  (calculated from  $\Delta T_i$  estimations) and  $RMSE_{i=ref}$  (calculated considering that each station is at the reference temperature value i.e.  $\Delta T_i = 0 \forall i$ )

:

$$NRMSE = \frac{RMSE_i}{RMSE_{i=ref}} = \sqrt{\frac{\frac{1}{nb_d} \frac{1}{nb_i} \sum_d^{nb_d} \sum_i^{nb_i} (\Delta T_{obs_i}(j) - \Delta T_{est_i}(j))^2}{\frac{1}{nb_d} \frac{1}{nb_i} \sum_d^{nb_d} \sum_i^{nb_i} (\Delta T_{obs_{i,d}} - \Delta T_{obs_{ref,d}})^2}} \quad (15)$$

where  $\Delta T_{obs_{i,d}}$  and  $\Delta T_{obs_{ref,d}}$  are observed  $\Delta T$  values for the day  $d$  respectively for a station  $i$  and for the reference station  $ref$ ,  $\Delta T_{est_{i,d}}$  is the  $\Delta T$  value estimated according to the model for the day  $d$  and the station  $i$ ,  $nb_d$  the number of days used for verification purpose.

The lower the NRMSE, the better the performance of the model. Then for each season, the 32 model groups are classified by increasing magnitude of NRMSE for each network data set. A same model group that minimizes the NRMSE for all data sets is considered as the best model.

## 4. Results

### 4.1. Model group selected

The following spatial and temporal models are finally selected.

$$\widetilde{\Delta T}^{(aut)} = 0.0016K(P_C) + 6.3 \frac{1}{U(P_A) + 3.0} + 0.01T(P_C) - 0.095 \exp\left(\frac{N(P_B) - 6}{6}\right) - 1.1 \quad (16)$$

$$\widetilde{\Delta T}^{(win)} = 0.081 \exp(-U(p_B) + 3) - 0.12N(P_A) + 0.64 \quad (17)$$

$$\widetilde{\Delta T}^{(spr)} = 0.0011K(P_C) + 11.0 \frac{1}{U(P_A) + 3.0} + 0.033T(P_D) - 0.12N(P_A) - 2.3 \quad (18)$$

$$\widetilde{\Delta T}^{(sum)} = -0.0083RH(P_D) - 0.36U(P_A) - 0.13T(P_A) + 3.2 \quad (19)$$

19

$$\overline{\Delta T}^{(aut)} = 5.3x_{D_f} + 0.15 \ln(d_{op} + 1) - 5.7x_{NDVI} - 5.5x_{D_b} - 0.39 \quad (20)$$

$$\overline{\Delta T}^{(win)} = 2.0x_{D_s} + 0.12 \ln(d_{op} + 1) - 2.7x_{NDVI} - 4.0x_{D_b} - 0.2 \quad (21)$$

$$\overline{\Delta T}^{(spr)} = 3.3x_{D_w} - 5.1x_{NDVI} + 0.031 \quad (22)$$

$$\overline{\Delta T}^{(sum)} = 12.0x_{D_w} - 4.5x_{NDVI} + 0.23 \quad (23)$$

$$\sigma_{\Delta T}^{(aut)} = 3.2x_{D_f} + 0.06 \ln(d_{op} + 1) - 2.7x_{NDVI} - 2.9x_{D_b} + 0.38 \quad (24)$$

$$\sigma_{\Delta T}^{(win)} = -3.0x_{D_f} + 1.4x_{D_s} + 0.065 \ln(d_{op} + 1) + 0.5 \quad (25)$$

$$\sigma_{\Delta T}^{(spr)} = 1.7x_{D_s} + 0.12 \ln(d_{op} + 1) - 2.9x_{NDVI} - 3.4x_{D_b} + 0.41 \quad (26)$$

$$\sigma_{\Delta T}^{(sum)} = -6.4x_{D_f} + 3.0x_{D_s} + 4.8x_{D_w} + 0.15 \ln(d_{op} + 1) + 0.86 \quad (27)$$

In order to compare the weight of each explanatory variables  $ev$  used in a same equation, their regression coefficient  $rc_{ev}$  is multiplied by an indicator of the explanative variable  $ev$  amplitude ( $amp_{ev}$  - defined as the difference between the percentile 95 and the percentile 5 of the explanatory variable  $ev$ ). The result is defined as the normalized regression coefficient  $RC_{ev}$  :

$$RC_{ev} = |rc_{ev}amp_{ev}| \quad (28)$$

In the case of temporal models, normalized regression coefficients are gathered in table 5. The period of calculation of the explanatory variable is also indicated as well as the tendency of the UHI evolution when the explanative variable increases.

In the case of spatial models, the calibration network (ID 1) is composed of only 10 stations and may not be representative of all the geographical contexts observed in the urban area. To overcome this issue, each explanatory variable is calculated according to the methodology described in section 3.4 for 200 points randomly chosen in the urban area territory.  $amp_{ev}$  is then calculated from this new data set. For each explanatory variable, normalized regression coefficients (equation 28) and the tendency of the UHI evolution when the explanatory variable increases are finally gathered in tables 6 and 7.

Table 5: Normalized regression coefficients (equation 28 with notation  $A = RC$ ), period of calculation and  $UHI$  direction of change for each meteorological condition variable of  $\Delta T$  models

Season	$A_K$			$A_{RH}$			$A_U$			$A_T$			$A_N$		
	value	period $p_p$	direction of change	value	period $p_p$	direction of change	value	period $p_p$	direction of change	value	period $p_p$	direction of change	value	period $p_p$	direction of change
Autumn	0.54	$P_C$	↗	-	-	-	0.73	$P_A$	↘	0.22	$P_C$	↗	0.93	$P_B$	↘
Winter	-	-	-	-	-	-	0.01	$P_B$	↘	-	-	-	0.93	$P_A$	↘
Spring	0.34	$P_C$	↗	-	-	-	1.43	$P_A$	↘	0.42	$P_D$	↗	0.95	$P_A$	↘
Summer	-	-	-	0.33	$P_D$	↘	1.29	$P_A$	↘	1.23	$P_A$	↘	-	-	-

Table 6: Normalized regression coefficients (equation 28 with notation  $B = RC$ ), period of calculation and  $UHI$  direction of change for each geographical variable of  $\Delta T$  models

Season	$B_{D_f}$		$B_{D_s}$		$B_{D_w}$		$B_{d_{op}}$		$B_{NDVI}$		$B_{D_b}$	
	value	direction of change	value	direction of change	value	direction of change	value	direction of change	value	direction of change	value	direction of change
Autumn	1.51	↗	-	-	-	-	0.86	↗	2.23	↘	2.43	↘
Winter	-	-	0.79	↗	-	-	0.72	↗	1.06	↘	1.75	↘
Spring	-	-	-	-	1.20	↗	-	-	2.00	↘	-	-
Summer	-	-	-	-	4.17	↗	-	-	1.77	↘	-	-

Table 7: Normalized regression coefficients (equation 28 with notation  $C = RC$ ), period of calculation and  $UHI$  direction of change for each geographical variable of  $\sigma_{\Delta T}$  models

Season	$C_{D_f}$		$C_{D_s}$		$C_{D_w}$		$C_{d_{op}}$		$C_{NDVI}$		$C_{D_b}$	
	value	direction of change	value	direction of change	value	direction of change	value	direction of change	value	direction of change	value	direction of change
Autumn	0.91	↗	-	-	-	-	0.35	↗	1.07	↘	1.30	↘
Winter	0.86	↘	0.55	↗	-	-	0.38	↗	-	-	-	-
Spring	-	-	0.66	↗	-	-	0.70	↗	1.13	↘	1.50	↘
Summer	1.83	↘	1.21	↗	1.72	↗	0.90	↗	-	-	-	-

According to table 5, wind speed and nebulosity are the predominant variables for most of the seasons, which is consistent with previous literature results (Svensson et al., 2002; Kim and Baik, 2002; Runnalls and Oke, 2000; Yan et al., 2014). However, wind speed influence is negligible during winter-time whereas nebulosity is replaced by air temperature summer-time and its effect on UHI development is unexpectedly negative : higher rural air temperature, lower temperature differences within the urban area. Night-time values of nebulosity and wind speed contribute to UHI development whereas for other variables such as solar radiation or temperature, this is mainly day-time values which explain the UHI values.

Both  $\overline{\Delta T}$  and  $\sigma_{\Delta T}$  have almost the same set of variables explaining their variations (respectively table 6 and table 7).  $\overline{\Delta T}$  being easier to physically understand, we analyze its equations. The contribution of each geographical variable is analyzed through the urban heat fluxes it is related to. Thus,  $D_w$  and  $NDVI$  (related to latent heat flux) are the only variables to drive the UHI development both during spring and summer. During winter,  $D_w$  and  $NDVI$  influence is much lower whereas  $D_b$  and  $D_s$  (related to anthropogenic heat flux) become predominant. During autumn,  $D_w$  and  $NDVI$  (related to latent heat flux) and  $D_f$  (related to radiative heat flux) may be identified from equation 20 but none is predominant (table 6). The indicator  $d_{os}$  (related to advection heat flux) is used in winter and autumn equations but has the lowest influence within all variables.

#### 4.2. Model performance analysis

The performances of the selected model group are presented table 8.

Table 8: NRMSE values of the final model for each season and each network

Season	Network 1	Network 2	Network 3	Network 4
Autumn	0.58	-	1.4	0.71
Winter	0.81	-	1.4	0.87
Spring	0.42	0.53	1.4	0.87
Summer	0.46	0.54	1.2	0.76

When the NRMSE is lower than 1, we may consider that using the model estimation is more accurate than assuming that the whole urban area is at the same temperature as the reference station. Only the network 3 exceeds this threshold for winter, spring and summer condition. The low performances of the model is only due to some of the stations. Spring is the season for which we have both the highest number of observed days and the lowest model performances for almost all networks. In order to identify model limitations, a detailed analysis of the results obtained for each station of the networks used for verification purpose (networks 2, 3 and 4) is performed at this season.

For 64% of the Nantes (network 2) and Angers (network 3) stations, the temperature estimation is considered as accurate since the scatter is close to the line  $\Delta T_{est} = \Delta T_{obs}$  (an example is given figure 6 for the station N02). For

all these stations, the root mean square error value ( $RMSE$ ) is included in the range  $0.36^{\circ}\text{C}$  and  $1.38^{\circ}\text{C}$  and for each of them, the  $RMSE$  value is always lower than both the observed median  $\Delta T$  and the observed median absolute  $\Delta T$  deviation.

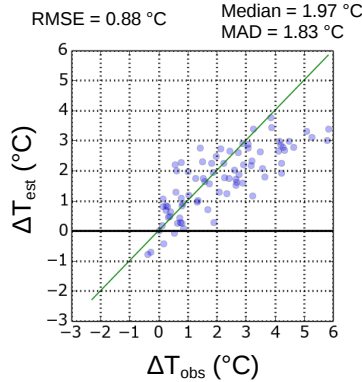


Figure 6: Estimated  $\Delta T$  values *versus* observed  $\Delta T$  values for station N02 - network 2

For the other stations, the limited performances of the modeling may have three causes. The first is the lack of geographical variations representativeness of the network 1 (used for calibration purpose) when compared with the entire urban area. Figure 7 shows the distribution of  $NDVI$  values for both the stations used for calibration and for verification in the same city (respectively networks 1 and 2). Two stations of the verification network (N14 and N15) are characterized by large  $NDVI$  value ( $\geq 0.5$ ) but none of the calibration network. This may explain why the model is inefficient to estimate the temperature observed for stations having a high  $NDVI$  value (figure 8).

This problem is also observed for stations located near large water areas. Estimations are considered incorrect when  $D_w \geq 0.2$ .

Temporal models also show performance limitations when observed UHI is high. This is illustrated figure 9. When observed UHI is lower than  $3^{\circ}\text{C}$ , observed and estimated values are quite similar whereas when the observed UHI is included in a  $[3 - 5]^{\circ}\text{C}$  range, estimated UHI is always below  $3.1^{\circ}\text{C}$ .

Another source of uncertainty is the buffer size (500 m radius) chosen for geographical indicator calculation. If a station has a strong geographical heterogeneity within 500 m radius circle, geographical indicator values may differ a lot whether it is calculated in a 200 m radius circle or in a 500 m radius circle. If the source area which contributes to the measured temperature is actually closer to a 200 m than a 500 m radius circle, models may be erroneous. This is the case for the network 4 reference station ( $MF_{l_{ry}}$ ) regarding the  $NDVI$  value that increases when the buffer size decreases (figure 10). If the 200 m radius buffer size is used for the calculation of the geographical indicators at the reference station, the estimation of the UHI for all network 4 stations is



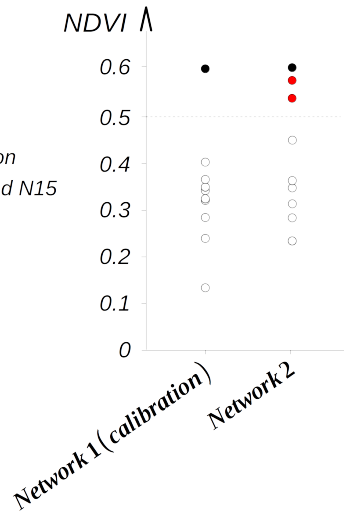


Figure 7: NDVI distribution of the network 1 stations (calibration purpose) and the network 2 stations (verification purpose)

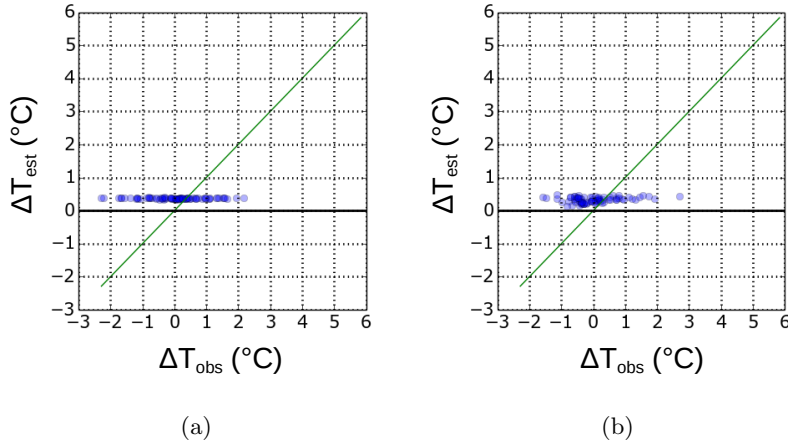


Figure 8: Estimated  $\Delta T$  values *versus* observed  $\Delta T$  values for station N14 (a) and station N15 (b) - network 2

more accurate since estimations are close to the line  $\Delta T_{est} = \Delta T_{obs}$  (figure 11).

### 5. Model application

The model group showed good performances when applied to the territory where it has been calibrated (Nantes conurbation). It is used here to map the UHI distribution in the city of Nantes under a calm summer night (2013, June

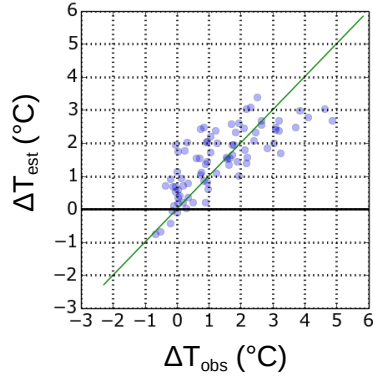


Figure 9: Estimated  $\Delta T$  values *versus* observed  $\Delta T$  values for station *N05* - network 2

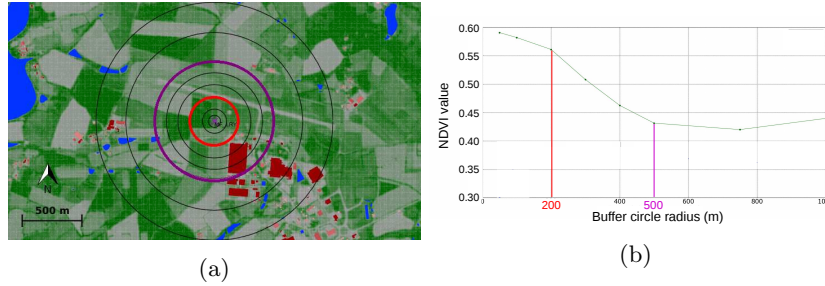


Figure 10: Geographical context (a) and NDVI value calculated for different buffer circle sizes (b) for the station *MF\_lry* (network 4)

13th). Temporal model is applied according to equation 19 and the associated meteorological conditions observed during this day (mean air temperature for the few first hours of the night ( $p_A$ ) is 8.7 °C, mean wind speed is 0.93 m/s, and mean relative humidity during the last quarter of the day ( $p_D$ ) is 83 %). Spatial models (equations 23 and 27) are applied for a set of 855 points regularly distributed every 300 m within the city of Nantes. The reference station used for calculation is the same that the one used for the calibration step (*MF\_nse*). Regarding the previous highlighted limitations of the spatial model, points where the water density is higher than 0.2 or the NDVI higher than 0.5 are excluded. UHI values are then linearly interpolated according to a Delaunay triangulation method (figure 12).

As expected, zones having a high NDVI value (but lower than 0.5) such as parks have almost the same temperature as the rural reference (less than 1°C higher). The city center, consisting in mineral surfaces and tall buildings is much warmer (more than 2.5°C higher than the reference station). Two areas, although they are located outside the city core, are also warm. On the south west, the neighborhood is an industrial zone without vegetation. On the north east, the corresponding warm area is a combination of industrial and commercial

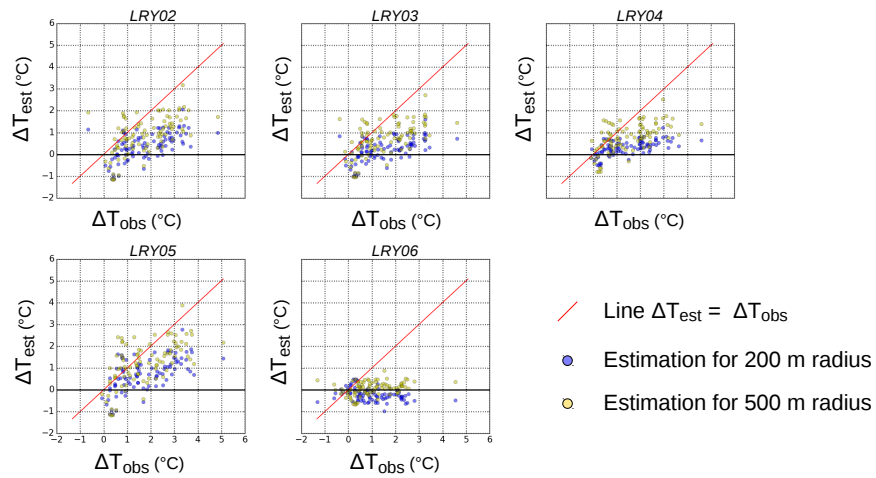


Figure 11: Estimated  $\Delta T$  values *versus* observed  $\Delta T$  values for each station of network 4 for geographical indicator calculated in a buffer circle of 200 m radius and 500 m radius for the reference station

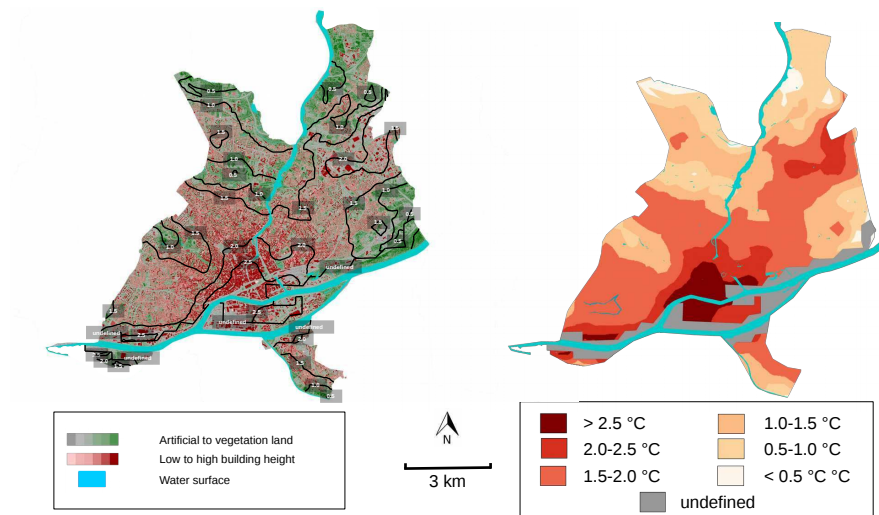


Figure 12: UHI distribution within the city of Nantes territory under the 2013 June 13th night

buildings, also composed of a majority of mineral surfaces.

## 6. Conclusion and discussion

Four main contributions to UHI empirical modeling have been proposed.

- First, we have identified a night-time period of UHI modeling that maximizes the spatial dispersion of air temperature and minimizes the spatial distribution of the cooling rate. The beginning and the end of this period ( $p_{UHI} = [1.4, 1.8]$ ) are defined as a ratio of the night-time duration, and thus varies seasonally in the local time system.
- Second, UHI value may be estimated for any night and any location in a city according to a normalization method which consists in separating temporal from spatial variations. UHI temporal variations then have been modeled according to meteorological conditions observed during the day of interest whereas spatial variations have been modeled according to the geographical context of the area of interest.
- Third, the period when a meteorological variable impacts the night-time UHI has been investigated for each meteorological variable. Wind speed and nebulosity are the predominant meteorological variables in the temporal models. Their night-time low value are responsible for the night-time urban heat island development.
- Finally, the geographical indicators proposed to explain the UHI spatial variations are related to one or several heat fluxes, which makes them easy to interpret. The NDVI, which is related to latent heat flux, is predominant during spring and summer season. Indicators such as building density  $D_b$  or surface density  $D_s$ , related to the anthropogenic heat flux, are predominant during winter time and affect only the variability of the UHI ( $\sigma_{\Delta T}$ ) during the rest of the year.

The model group (composed of temporal and spatial models) was calibrated and then applied to other cities to verify its ability to extrapolation. Good results were found for most of the verification stations located in the conurbations of Nantes and Angers. For the city of Nantes, UHI map has been produced, showing the good potential of such a work : the air temperature may be estimated for any day and any location in those cities according to basic geographical informations and meteorological conditions observed near or inside the urban area. This map may be combined to other spatial informations to help decision making : for example data regarding at-risk population may be cross-used in order to elaborate a strategy to protect vulnerable people during heat-waves. Furthermore, the calculation is fast and easy to implement which makes the method usable to take into account urban heat island phenomenon in dynamic thermal simulation for urban buildings. An interesting aspect of this work is that both meteorological and geographical data used to explain respectively temporal and spatial temperature variations are homogeneous and available for the entire French territory. The method proposed here can then be applied for any French city. However, it would be interesting to identify similar data which are available at world scale.

Improvements may be performed to better estimate urban heat island temporal and spatial variations. Three main limitations have been highlighted :

first, they are not appropriate to estimate the temperature in areas characterized by large fraction of vegetation or water. Calibration should be performed with new set of data more representative of the geographical context diversity of a city. Second, the temporal model is not accurate to estimate high UHI values whereas it is fundamental to have an accurate estimation of those specific values since they usually correspond to heat-wave periods, that are important regarding health issues. Weighting regressions could be applied to better estimate high UHI values, which will however imply lower accuracy regarding lower UHI values. Third, buffer size for geographical indicator calculation has been set according to values proposed in the literature but seems inappropriate in certain cases (e.g. LRY reference station). A detailed analysis must be performed to identify a more suitable buffer size. Finally, the model group which will be produced after all these modifications should be tested on several other French cities in order to verify its extrapolation performances. If the results show a dependency between the accuracy of the estimation and the size of the city, further investigation would be needed to take into account this factor (such as weighting the spatial model equations by any indicator of urban size). Investigation should also be carried out to integrate the sea breeze influence on UHI intensity in case it is a predominant factor explaining the temporal model bias. A final improvement would be to replace the existing geographical data sources by world-wide databases in order to make the model usable for any city in the world. also be investigated in order to generalize the method to any city in the world.

### Acknowledgements

This work was supported by public funds received in the framework of GEO-SUD, a project (ANR-10-EQPX-20) of the program "Investissements d'Avenir" managed by the French National Research Agency. It was also supported by the French Environment and Energy Management Agency (ADEME) and the french region "Pays de la Loire".

### References

- J. Armstrong. Temperature differences between two ground-level sites and a roof site in southampton. *Meteorological Magazine*, 103:360–368, 1974.
- T. Asaeda and V. T. Ca. The subsurface transport of heat and moisture and its effect on the environment: a numerical model. *Boundary-Layer Meteorology*, 65(1):159–179, 1993.
- B. Balázs, J. Unger, T. Gál, Z. Sümegehy, J. Geiger, and S. Szegedi. Simulation of the mean urban heat island using 2d surface parameters: empirical modelling, verification and extension. *Meteorological Applications*, 16(3):275–287, 2009.
- L. Benosa. Ortho-sat, descriptif de contenu. Technical report, Institut national de l'information géographique et forestière, 2014.

- A. Bernabé, M. Musy, H. Andrieu, and I. Calmet. Radiative properties of the urban fabric derived from surface form analysis: A simplified solar balance model. *Solar Energy*, 122:156–168, 2015.
- J. Bernard. *Geographic and weather signature of the spatial and temporal air temperature variations within an urban area*. Theses, École Centrale Nantes, Jan. 2017. URL <https://hal.archives-ouvertes.fr/tel-01449935>.
- E. Bocher and G. Petit. *ORBISGIS: Geographical Information System Designed by and for Research*, pages 23–66. John Wiley & Sons, Inc., 2013. ISBN 9781118561928. doi: 10.1002/9781118561928.ch2. URL <http://dx.doi.org/10.1002/9781118561928.ch2>.
- R. Britter and S. Hanna. Flow and dispersion in urban areas. *Annual Review of Fluid Mechanics*, 35(1):469–496, 2003.
- L. Chen, E. Ng, X. An, C. Ren, M. Lee, U. Wang, and Z. He. Sky view factor analysis of street canyons and its implications for daytime intra-urban air temperature differentials in high-rise, high-density urban areas of hong kong: a gis-based simulation approach. *International Journal of Climatology*, 32(1):121–136, 2012.
- J. Clarke and W. Bach. Comparison of the comfort conditions in different urban and suburban microenvironments. *International journal of biometeorology*, 15(1):41–54, 1971.
- J. F. Clarke. Some effects of the urban structure on heat mortality. *Environmental research*, 5(1):93–104, 1972.
- C. de Munck, G. Pigeon, V. Masson, F. Meunier, P. Bousquet, B. Tréméac, M. Merchat, P. Poeuf, and C. Marchadier. How much can air conditioning increase air temperatures for a city like paris, france? *International Journal of Climatology*, 33(1):210–227, 2013.
- E. Erell and T. Williamson. Intra-urban differences in canopy layer air temperature at a mid-latitude city. *International Journal of Climatology*, 27(9):1243–1256, 2007.
- D. Fenner, F. Meier, D. Scherer, and A. Polze. Spatial and temporal air temperature variability in berlin, germany, during the years 2001–2010. *Urban Climate*, 10:308–331, 2014.
- T. Gal, F. Lindberg, and J. Unger. Computing continuous sky view factors using 3d urban raster and vector databases: comparison and application to urban climate. *Theoretical and applied climatology*, 95(1-2):111–123, 2009.
- C. Grimmond, M. Blackett, M. Best, J. Barlow, J. Baik, S. Belcher, S. Bohnenstengel, I. Calmet, F. Chen, A. Dandou, K. Fortuniak, M. Gouvea, R. Hamdi, M. Hendry, T. Kawai, Y. Kawamoto, H. Kondo, E. Krayenhoff, S. Lee, T. Loridan, A. Martilli, V. Masson, S. Miao, K. Oleson, G. Pigeon, A. Porson,

- Y. Ryu, F. Salamanca, L. Shashua-Bar, G. Steeneveld, M. Tombrou, J. Voogt, D. Young, and N. Zhang. Urban surface energy balance models: model characteristics and methodology for a comparison study. In *Meteorological and Air Quality Models for Urban Areas*, pages 97–123. Springer, 2009.
- M. Hart and D. Sailor. Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. *Theoretical and Applied Climatology*, 95(3-4):397–406, 2009.
- J. Hjort and M. Marmion. Effects of sample size on the accuracy of geomorphological models. *Geomorphology*, 102(3):341–350, 2008.
- X.-M. Hu and M. Xue. Influence of synoptic sea-breeze fronts on the urban heat island intensity in dallas-fort worth, texas. *Monthly Weather Review*, 144(4):1487–1507, 2016.
- Y.-H. Kim and J.-J. Baik. Maximum urban heat island intensity in seoul. *Journal of Applied Meteorology*, 41(6):651–659, 2002.
- E. Krüger and B. Givoni. Outdoor measurements and temperature comparisons of seven monitoring stations: Preliminary studies in curitiba, brazil. *Building and environment*, 42(4):1685–1698, 2007.
- M. Lacombe. Results of the wmo intercomparison of thermometer screens/shields and hygrometers in hot desert conditions. TECO-2010–WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation, Helsinki, Finland, 30 August–1 September 2010, <http://www.wmo.int/pages/prog/www/IMOP/publications/IOM-104> TECO-2010/3 4 Lacombe France. pdf, 2010.
- M. Lacombe, D. Bousri, M. Leroy, and M. Mezred. Wmo field intercomparison of thermometer screens/shields and humidity measuring instruments, ghardaia, algeria, november 2008–october 2009. Technical report, World Meteorological Organization, 2011.
- E. László and S. Szegedi. Impacts of some surface parameters on urban heat island development. 2012.
- S. A. Lowe. An energy and mortality impact assessment of the urban heat island in the us. *Environmental Impact Assessment Review*, 56:139–144, 2016.
- J. Maby. Approche conceptuelle et pratique des indicateurs en géographie. *Objets et indicateurs géographiques*, Avignon, A. Barthélémy, pages 16–41, 2003.
- L. Malys, M. Musy, and C. Inard. Microclimate and buildings energy consumption: sensitivity analysis of coupling methods. In *Eighth International Conference on Urban Climate and 10th Symposium on the Urban Environment*, 2012.

- Y. Nakamura and T. Oke. Wind, temperature and stability conditions in an east-west oriented urban canyon. *Atmospheric Environment (1967)*, 22(12): 2691–2700, 1988.
- K. Neil and J. Wu. Effects of urbanization on plant flowering phenology: A review. *Urban Ecosystems*, 9(3):243–257, 2006.
- T. Oke. The distinction between canopy and boundary-layer urban heat islands. *Atmosphere*, 14(4):268–277, 1976.
- T. Oke. *Initial guidance to obtain representative meteorological observations at urban sites*, volume 81. World Meteorological Organization Geneva, 2004.
- T. R. Oke. *Boundary layer climates. 2nd.* Methuen, 289p, 1987.
- M. C. Peel, B. L. Finlayson, and T. A. McMahon. Updated world map of the köppen-geiger climate classification. *Hydrology and earth system sciences discussions*, 4(2):439–473, 2007.
- H. Pelgrum and W. Bastiaanssen. An intercomparison of techniques to determine the area-averaged latent heat flux from individual in situ observations: A remote sensing approach using the european field experiment in a desertification-threatened area data. *Water Resources Research*, 32(9):2775–2786, 1996.
- M. Petralli, L. Massetti, G. Brandani, and S. Orlandini. Urban planning indicators: useful tools to measure the effect of urbanization and vegetation on summer air temperatures. *International Journal of Climatology*, 34(4): 1236–1244, 2014.
- G. Pigeon, A. Lemonsu, N. Long, J. Barrié, V. Masson, and P. Durand. Urban thermodynamic island in a coastal city analysed from an optimized surface network. *Boundary-Layer Meteorology*, 120(2):315–351, 2006.
- D. Revel, H.-M. Füßel, and A. Jol. Climate change, impacts and vulnerability in europe 2012. Technical report, European Environmental Agency, 2012.
- K. Runnalls and T. Oke. Dynamics and controls of the near-surface heat island of vancouver, british columbia. *Physical Geography*, 21(4):283–304, 2000.
- D. Sailor. A review of methods for estimating anthropogenic heat and moisture emissions in the urban environment. *International Journal of Climatology*, 31(2):189–199, feb 2011. ISSN 08998418. doi: 10.1002/joc.2106.
- M. Santamouris. Cooling the cities—a review of reflective and green roof mitigation technologies to fight heat island and improve comfort in urban environments. *Solar Energy*, 103:682–703, 2014.
- R. Spronken-Smith and T. Oke. Scale modelling of nocturnal cooling in urban parks. *Boundary-Layer Meteorology*, 93(2):287–312, 1999.



- G. Steeneveld, S. Koopmans, B. Heusinkveld, and N. Theeuwes. Refreshing the role of open water surfaces on mitigating the maximum urban heat island effect. *Landscape and Urban Planning*, 121:92–96, 2014.
- J. Suomi, J. Hjort, and J. Käyhkö. Effects of scale on modelling the urban heat island in turku, sw finland. *Clim Res*, 55:105–118, 2012.
- M. Svensson, I. Eliasson, and B. Holmer. A gis based empirical model to simulate air temperature variations in the goteborg urban area during the night. *Climate Research*, 22(3):215–226, 2002.
- R. Taesler. *Studies of the development and thermal structure of the urban boundary layer in Uppsala*. Meteorologiska institutionen (Uppsala universitet), 1980.
- C. Tannier, I. Thomas, G. Vuidel, and P. Frankhauser. A fractal approach to identifying urban boundaries. *Geographical Analysis*, 43(2):211–227, 2011.
- H. Yan, S. Fan, C. Guo, F. Wu, N. Zhang, and L. Dong. Assessing the effects of landscape design parameters on intra-urban air temperature variability: The case of beijing, china. *Building and environment*, 76:44–53, 2014.
- T. Yokobori and S. Ohta. Effect of land cover on air temperatures involved in the development of an intra-urban heat island. *Climate research (Open Access for articles 4 years old and older)*, 39(1):61, 2009.