Assessing Traffic's Impact on Urban Heat Islands: An Agent-Based Approach Using Real-World and Simulated Data

Aygun Ismayilova[1,2] and Manon Prédhumeau[1][0000−0002−2003−1432]

[1] Université Toulouse Capitole, IRIT, Toulouse, France
{firstname.surname}@ut-capitole.fr
[2] Université Grenoble Alpes, LIG, Grenoble, France

Abstract. Urban Heat Islands (UHIs) — where cities are significantly warmer than surrounding rural areas — are driven by urbanization and human activities, including road traffic. Understanding traffic-UHI interactions requires high-resolution spatio-temporal data to capture the local and dynamic characteristics of both systems. While urban temperature sensors are increasingly deployed, fine-scale traffic data remain difficult to obtain and are lacking in most cities. To address this gap, we propose a framework combining MATSim-based traffic simulated data and real-world open data. A synthetic population is used to generate individual vehicular trips, which are then validated against observed traffic data. The approach integrates outputs from multi-agent traffic simulation, with urban structure and land-use variables, as well as meteorological measurements, for a full spatio-temporal analysis. Results suggest diurnal and seasonal patterns in traffic–temperature relationships, highlighting the need to capture dynamic human–environment interactions in UHI research. This study shows how combining simulated and real-world data can enhance our understanding of traffic-induced heat effects.

Keywords: Agent-Based Modelling · MATSim Simulation · Urban Climate · Real-World Data · Synthetic Population

1 Introduction

Global cities are experiencing rapid urbanization, contributing to the formation of Urban Heat Islands (UHIs). The UHI phenomenon refers to significantly higher temperatures in cities than in surrounding rural areas [11]. Intensified by climate change, UHIs pose growing environmental and public health risks by increasing heat exposure, thermal discomfort, and mortality during heat waves [9]. This underscores the urgent need for targeted UHI mitigation strategies.

UHI is a well-documented phenomenon driven by interacting socio-ecological factors, including meteorology, land use, urban form, materials, and human activities. Transportation, and road traffic in particular, contributes to UHI through engine waste heat and pollutant emissions that inhibit heat dissipation [23]. Multiple studies identified vehicular heat as a major contributor to UHIs [6,10,15,25], highlighting the need to include traffic in mitigation strategies.

However, the exact role of road traffic in UHI remains under-explored, as its spatio-temporal variability and interactions with environmental factors make it difficult to isolate and quantify. Prior studies identify several challenges. Isolating traffic impacts from other interacting UHI drivers is difficult [11], as dense road networks and high-rise buildings limit heat dispersion and complicate attribution [25]. In addition, traffic dynamics vary in speed and flow, making their influence difficult to characterize [6]. Finally, the limited availability and resolution of traffic data restrict fine-scale analyses [15]. A framework that integrates high-resolution dynamic traffic data with static urban form and environmental factors to explore the contribution of road traffic to UHIs is missing.

While street-level temperature sensors are increasingly deployed [1], fine spatio-temporal traffic data remain scarce in most cities. In this context, agent–based modelling and simulation (ABMS) offers a promising alternative to real-world data. ABMS is well-suited for representing complex systems such as urban mobility, where heterogeneous agents interact over space and time. ABMS has been successfully employed in many applications, such as assessment of road traffic noise exposure [14]. Still, no prior studies integrated synthetic and real-world data to examine the relationship between road traffic dynamics and UHI variations. Our main contribution is methodological, demonstrating how large-scale transport ABMS can support urban climate analysis in data-scarce contexts. To address this scarcity, we use the eqasim and MATSim (Multi-Agent Transport Simulation) frameworks to generate and simulate daily traffic patterns. After calibration and validation, the simulated traffic data are combined with observed UHI temperatures and urban characteristics in a correlation analysis.

The remainder of the paper is structured as follows: section 2 defines the UHI phenomenon and reviews related works. Section 3 describes the proposed approach and section 4 applies the approach to the case study of Toulouse, France. Section 5 presents preliminary results. Finally, Section 6 summarizes the main contributions and outlines future directions.

2 Background

2.1 UHI Concept, Effects and Factors

The UHI phenomenon refers to higher temperatures in urban areas compared to nearby rural regions. UHIs are common in densely built cities, and its magnitude is quantified by the Urban Heat Island Intensity (UHII), defined as the average temperature difference between urban and nearby rural areas [12,18], in °C :

$$\mathrm{UHII} = T_{\mathrm{urban}} - T_{\mathrm{rural}} \tag{1}$$

where T_{urban} and T_{rural} are average temperatures of the urban area and a nearby rural reference, measured concurrently. UHIIs can reach up to 12°C [18], and previous studies show that high UHII severely affects human comfort and health. Extreme heat and UHI exposure are linked to a 1.5% increase in cardiovascular hospitalizations in the US [4] and a 45% median increase in mortality risk across 85 European cities [9], particularly affecting vulnerable populations [4].

The emergence and intensification of UHIs are influenced by multiple factors [11] , including meteorological conditions (e.g. wind and temperature), land use and surface materials (e.g. greenery, water, and surface albedo), urban structure (e.g. street width and building density), and anthropogenic activities (buildings, industry, road traffic, and human metabolism). Three main sources of anthropogenic heat are population density, building density, and traffic density [12], with traffic playing a key role through heat emissions. Traffic is a major source of urban heat, contributing up to 30% of anthropogenic heat emissions in Singapore [20], with high traffic density identified as the main driver of UHII across seasons in Chiang Mai, Thailand [12]. Similarly, simulations of an 80% traffic reduction to model the COVID-19 pandemic reported near-surface temperature decreases of about 1°C in high-traffic areas and up to 20% fewer hours above 30°C [21]. Another study found that slow, heavy, congested traffic can raise local temperatures by up to 7 °C [25]. Still, most UHI research focuses on urban structure and land use, and the role of traffic remains relatively underexplored.

2.2 Related Work

To study traffic impacts on UHIs, a common approach is spatiotemporal analysis using real-world data. For example, Welegedara et al. studied neighbourhood-level UHII changes, using data on land surface temperature (LST), land-use, average weekday traffic volume, and population characteristics [24]. Pearson's correlation analysis was used to examine seasonal and diurnal variations. Lee and Berkelhammer assessed congestion effects on UHII using satellite LST and Chicago bus speeds in generalized additive models [15]. Li et al. [16] estimated anthropogenic heat in a stepwise regression based on LST data, mean traffic data and energy consumption data. Similarly, Husni et al. collected vehicular flow data with IoT sensors and cameras and analyzed correlation with temperatures through a principal component analysis and regression [10].

Such data-driven studies reveal local UHII variations and vulnerable areas, but often rely on remotely sensed LST, which is less relevant than air temperature for health and comfort. Using average traffic volumes or speeds can miss congestion patterns, and while traffic sensors are more accurate [10], citywide networks are rare due to cost and maintenance. Overall, high-resolution data capturing spatiotemporal variability in traffic and temperature remain difficult to obtain. A review of European cities confirms that open, spatiotemporally compatible datasets are scarce, with temperature data often limited in coverage and traffic data missing or overly aggregated (Table 1, Supplementary Material).

Another approach uses numerical simulations to model heat fluxes and assess contributing factors. Haddad et al. [6], and Quah and Roth [20] formulated transport-related heat emissions in equations to simulate their impact on UHI. Zhu et al. applied a cell-transmission model to simulate vehicular heat accumulation and correlate it with air temperature [25]. These models enable scenario testing for UHI mitigation, but rely on numerous physical equations and empirical parameters, making calibration time-consuming. Moreover, they often overlook heterogeneous driving behaviours, vehicle types, and urban interactions.

Traffic simulation is well-established, and both relevant and promising for environmental studies. Leveraging growing urban temperature datasets, we propose a new hybrid approach that combines open data on urban structure, land-use, and temperatures with simulated traffic data to study UHI–traffic dynamics.

3 Proposed Method

This section presents the proposed framework to analyze traffic–UHI dynamics in real urban environments. The proposed method considers the city as a complex socio-ecological system where mobility interacts with urban form and microclimate. The approach relies solely on open data and is generic enough to be applied to any case study with sufficient data. Air temperature, the key metric, requires direct measurement because it is influenced by many interacting factors, and is difficult to model accurately. In contrast, activity- and agent-based approaches have proven effective in reproducing road traffic dynamics [7]. Agent-based simulation is thus used to capture the complexity of urban mobility. It explicitly models behavioral heterogeneity, with agents representing travelers whose daily activities and interactions produce emergent traffic phenomena such as congestion and flow changes. The proposed framework combines several complementary approaches, depicted in the workflow graph presented in Figure 1.

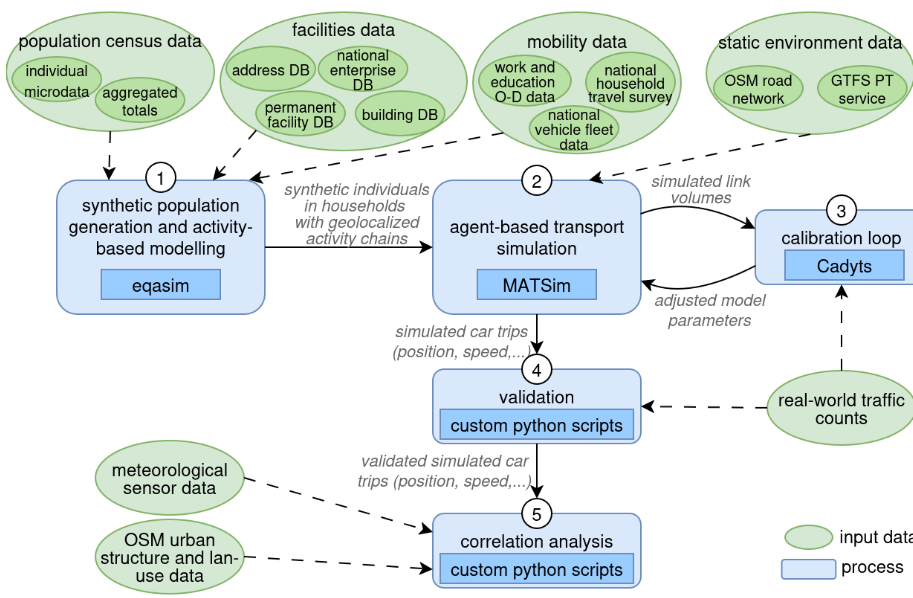


Fig. 1: Proposed methodological framework in five steps: 1) generation of a synthetic population and activity-based model, 2) agent-based simulation, 3) calibration, 4) validation and 5) correlation analysis with real-world data.

3.1 Synthetic Population Generation and Activity-Based Modelling

Agent-based transportation models increasingly use synthetic populations and activity-based models (AcBM) to represent travellers and travel demand. Synthetic populations are artificial sets of individuals and households reflecting real-world demographics. Based on a synthetic population, an AcBM represent individual mobility as a sequence of daily activities, i.e. where, when, and how people travel. These can be generated via methods like iterative proportional fitting, combinatorial optimization, or microsimulation, combined with rule-, constraint- or utility-based approaches, with several ready-to-use tools available.

Eqasim is an open source pipeline for generating synthetic populations with sociodemographic attributes and geolocated daily activity patterns [7]. It enables a realistic representation of individual mobility while preserving privacy. Eqasim is well generalized and covers any region in France, California, Bavaria, Sao Paulo, and Switzerland, using mostly open-data [7]. The detailed procedure [7] includes two main steps: collecting heterogeneous datasets on population, facilities, and mobility, and running the eqasim pipeline to generate the population and AcBM. The synthesis starts by creating individuals and assigning them to households through direct sampling of the census data. Next, sociodemographic attributes and household residence locations are determined. Household's private vehicles, including their engine types (combustion or electric), are generated based on national data. Each individual is then given an activity chain using statistical matching with the national travel survey data, along with work or education locations. Lastly, secondary activities locations (shopping, leisure,...) are selected. The resulting synthetic population and AcBM include individuals with sociodemographic attributes, related into households, each following a daily activity chain detailing activity types, times, locations, and connecting trips.

3.2 Agent-Based Transport Simulation

The eqasim pipeline is built as a generic Python package but is also integrated with MATSim, an open-source agent-based framework for large-scale transport simulations [8]. MATSim is widely used and simulates individual travellers behaviour and their interactions. The synthetic population along with OpenStreetMap (OSM) and General Transit Feed Specification (GTFS) data are fed as input to MATSim to simulate dynamic traffic conditions.

MATSim uses an evolutionary algorithm where agents iteratively adjust their plans to maximize utility. Each iteration, agents execute their activity chains (i.e. plan), choose routes based on previous network conditions, and interact to produce emergent traffic dynamics. After each iteration, plans are scored, and some agents modify their mode, route, or timing to produce a new plan while others select their highest-scoring plan. Over iterations, the simulation converges to an equilibrium producing realistic travel demand. To speed up convergence, a discrete mode-choice model was executed during replanning, allowing agents to select pertinent modes instead of choosing randomly. Eqasim includes a convergence module that monitors mode-share evolution and stops iterations once the modal share is stable. The simulation produces detailed spatio-temporal outputs, including agent positions, vehicle speeds, and link traffic volumes.

Simulating millions of travellers is computationally expensive, so MATSim commonly relies on population downscaling. By simulating only a fraction of the population and scaling the network capacity accordingly, it still captures the diversity of agents and their interactions. While a 5% sample can reproduce realistic congestion patterns, sampling rates below 10% introduce biases in key indicators [2], thus simulating at least 20% of the population is recommended.

3.3 Calibration Loop

After setup, the simulation is calibrated using real-world observations to enhance realism. While MATSim simulations can be calibrated using various objectives, the proposed framework focuses on achieving realistic traffic flow distribution by aligning simulated volumes with observed traffic counts. Cadyts (Calibration of Dynamic Traffic Simulations) is a calibration tool commonly used with MATSim to adjust the model parameters so that simulated traffic flows match observed counts [5]. It iteratively adjusts plan utilities during simulation to improve the match between simulated and observed traffic counts, improving the realism of the traffic distribution while preserving individual travel behaviour. Such calibration is a long process, requiring accurate and representative travel counts and the identification of optimal configuration parameters (minimum standard deviation of flow, variance scale, etc.) according to the local context.

3.4 Validation

The validation step aims to evaluate how accurately the MATSim model reflects real-world mobility patterns. The focus is on traffic volume and speed, as these metrics are critical for understanding transport-related contributions to UHI.

A first validation ensures that network-wide volumes are accurately modelled over time, including the timing and magnitude of morning and evening rush-hour peaks. The model's ability to reproduce spatiotemporal flow patterns is further assessed by comparing simulated hourly volumes with observed counts at selected locations. The coefficient of determination (R^2) measures the proportion of variance in observed flows explained by the simulation. It must indicate a moderate to good fit (e.g., $R^2 > 0.5$). In addition to R^2, the GEH statistic is computed for individual link-hour pairs to detect local discrepancies. The GEH statistic is a widely used goodness-of-fit measure in traffic modelling:

$$GEH = \sqrt{\frac{2(M-C)^2}{M+C}} \tag{2}$$

where M is the modelled traffic flow and C is the observed traffic flow. While a well-calibrated model should achieve GEH < 5 for most link-hour pairs, GEH values below 10 are considered acceptable.

Finally, simulated speeds are compared with observed speeds to check that congestion is realistically captured. Good agreement in both volumes and speeds indicates that the model can be confidently used for further analysis.

3.5 Correlation Analysis

A correlation analysis is finally conducted to assess relationships between meteorological data (specifically, UHII), urban factors, and simulated traffic data.

First, meteorological data are collected and preprocessed. Sensors (or weather stations) are classified as urban or rural based on geographic coordinates and municipal boundaries. Data for the period of interest (e.g. 2023) are downloaded via API. Raw meteorological data usually contain sensor errors, noise, extremes values, and duplicates, requiring thorough cleaning. For this, a rolling mean and standard deviation can be used to filter out temperature spikes exceeding three standard deviations over 48 hours. After cleaning, temperatures from the N rural stations are averaged at each timestamp to obtain a rural reference. For each urban station j at time t, the UHII is computed as the difference between the temperature recorded at j, and the average rural reference at time t.

Second, urban structure and land-use data are collected to include static factors in the correlation analysis. OSM provides open, crowd-sourced geospatial data, allowing consistent analysis of urban features across cities. Relevant OSM tags are used to compute building, vegetation, pedestrian, and water surfaces within a 300 m buffer around each station. Average building height and number of trees within the buffer are also computed. Additional metrics such as population density and number of enterprises can be included if local data are available.

Third, the simulated traffic data are used to derive three hourly link-level features: traffic volume, average speed, and fuel-type-specific vehicle counts. Custom Python scripts and Apache Spark are used to process MATSim outputs and generate these metrics. Each weather station is assigned its nearest road link to capture local traffic effects on UHII. Buffer-based approaches could also be applied but averaging metrics across multiple links might smooth the data and obscure local patterns. More advanced techniques, such as distance-weighted aggregation of traffic metrics from nearby links, could also be applied.

Finally, a Pearson correlation between all static features, dynamic traffic features and UHII measurements is applied, to assess the linear relationship between each pair of variables. More advanced correlation analysis that can detect non-linear relationships could also be used at this step.

4 Case Study

This case study applies the proposed method to examine the relationship between traffic and UHI variations in Toulouse. Toulouse, the fourth-largest city in France (1.4 million inhabitants in 2021 [22]), is located in the Haute-Garonne department, with an urban area exerting influence over most of the department. The city experiences pronounced UHI effects and high traffic density, with private cars accounting for 55% of trips in 2023 [22]. Toulouse Metropole operates a network of weather stations across the metropolitan area, providing high-resolution historical meteorological data via its Open Data platform [19]. While the portal also offers average daily traffic counts, these are too coarse for UHI analysis. The combination of high-resolution weather data and lack of fine-scale traffic data makes Toulouse an ideal case study for demonstrating the proposed framework.

A synthetic population of 287516 individuals (141712 households), representing 20% of Haute-Garonne's population, was generated using eqasim 1.3. Mobility is modelled across Haute-Garonne to capture traffic interactions between Toulouse and its surrounding area, ensuring realistic network dynamics. Each household is assigned vehicles and income, while each individual has sociodemographic attributes (age, sex, socio-professional class, employment status, driving license) and geolocated daily activities and trips. Activity types include home, work, education, shop, leisure, and other, with trip modes of car, walk, bike, car passenger, or public transit. Data sources include the 2021 population census, 2022 French buildings database (BD TOPO), French addresses database (BAN), facility, enterprise and school censuses (SIRENE, BPE 2023, MOBPRO 2021, MOBSCO 2021), and the 2008 National Transport and Travel Survey (ENTD). Vehicle fleet data comes from the 2021 national private vehicle stock (French Ministry of Ecological Transition), while spatial and transit data were obtained from OSM via Geofabrik and GTFS feeds from the local operator TISSEO.

Agent-based simulations were run in MATSim, with 287516 agents over 55 iterations until modal share converged. The traffic model (639710 car trips) was then calibrated over 200 iterations using Cadyts and observed traffic counts from 477 detectors for 2008, sourced from UTD19, a major publicly available urban traffic datasets [17]. Cadyts parameters were manually tuned through multiple simulations to balance trust between observed counts and prior simulation demand, with measured flow variance set to 3 and minimal count standard deviation to 10% to account for measurement error. Simulation and calibration took 4 h 51 min and 19 h 15 min, respectively, on a standard laptop.

Simulated traffic patterns were validated against observed data. Simulated mobility for Haute-Garonne was compared with 2023 Toulouse urban area statistics [22], as no departmental-level statistics are available. Validation shows 3.39 trips per agent versus 3.48 reported, average car trip lengths of 10.35 km versus 9.1 km reported, and a simulated car mode share of 65% versus 55% reported. This is consistent since simulated trips outside the urban area (not included in the 55% reported by statistics) are mainly by car. Morning (8–9 am) and evening (5–6 pm) peak hours are well reproduced, consistent with TomTom traffic data [3]. Hourly simulated counts compared with UTD19 data [17] yielded $R^2 = 0.56$, with average absolute errors of 200–300 trips (30–50% relative error). This indicates moderate but acceptable simulated flow realism, considering that the UTD19 data date back to 2008. GEH statistics were below 5 for 35.9% of link-hours and below 10 for 61.4%, indicating mostly good or acceptable matches. Simulated speeds also reflected slower traffic in the city centre and during rush hours, aligning with observations; more detailed speed validation is planned. Simulation results can be explored via SimWrapper [3] web visualization at `https://maprdhm.github.io/simwrapper/` (select "Public Data Folder").

[3] https://www.tomtom.com/traffic-index/city/toulouse

For the correlation analysis, static factors data were processed following an existing pipeline developed by a previous study [13]. Meteorological data were obtained via API from the Toulouse Open Data platform [19]. Of 77 weather stations, only 27 met data quality requirements, including 3 rural stations, yielding 144223 cleaned records for analysis. OSM data were used to calculate the surface area occupied by each land use type within a 300 m buffer around each station. Additional metrics (average building height, number of trees, population density, and number of enterprises from Toulouse Open Data), were also computed within the buffer. Simulated traffic data provided hourly traffic volume, average speed, and fuel-type-specific vehicle counts for 984416 road links. Each of the 24 urban stations was linked to its nearest road link and attributes. To assess seasonal and diurnal variability of traffic impacts, correlation analyses were performed using 36783 records in summer (25681 daytime, 11102 nighttime), 35652 in autumn (24828 daytime, 10824 nighttime), 37356 in winter (26085 daytime, 11271 nighttime), and 34432 in spring (24220 daytime, 10212 nighttime).

5 Preliminary Results

This section presents preliminary results from the correlation analysis on the Toulouse case study. Toulouse has a warm-temperate climate characterized by hot summers, mild winters, and noticeable seasonal and daily temperature variations. Results thus examine UHI across seasonal variations and daily cycles. Figure 2 shows the correlation matrix for winter season; matrices for the other seasons are provided in Figures 3 to 6 in the Supplementary material.

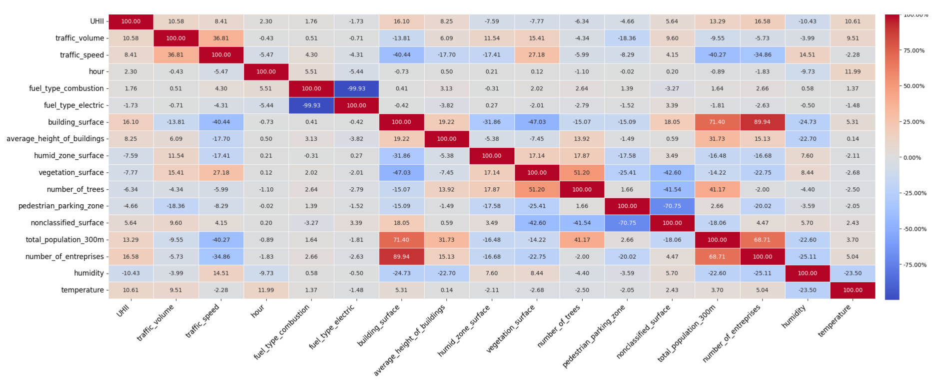


Fig. 2: Correlation matrix of static and traffic factors with UHII — Winter.

Across all seasons and both day and night, UHII is correlated with the static environment. UHII increases with building surface, population density, and enterprise density, and decreases with green and blue surfaces. Unlike static urban structure and land-use factors, traffic variables show strong seasonal and diurnal variability. In most seasons, traffic–UHII correlations are stronger during the day than at night, with correlation direction varying seasonally. In spring and winter, traffic volume and speed show weak to moderate positive correlations with UHII, especially during daytime. Higher traffic volumes and speeds, such as on main roads and highways, are linked to increased UHII. During daytime, the strongest UHII drivers are traffic volume and building surface, highlighting the role of anthropogenic heat and solar-heated infrastructure. At night, with traffic absent, static features dominate by releasing the heat they accumulated during the day. Winter shows the strongest positive correlations. This suggests that under reduced solar exposure, traffic heat emissions are more clearly detectable.

In autumn and summer, traffic volume and speed show a weak to moderate negative correlation with UHII, meaning higher traffic and faster speeds are linked to lower UHII. The summer results, showing moderate negative correlations during daytime, indicate that traffic effects are likely overshadowed by dominant meteorological drivers such as solar radiation and atmospheric conditions. In these seasons, traffic volume seems to act primarily as a proxy for urban functional intensity rather than a direct heat driver. High-traffic and fast traffic areas tend to coincide with lower building surface ratio, lower population density and greater vegetation cover. Vegetation and surface might exert stronger control on UHII during these seasons, their cooling influence can outweigh the thermal contribution of traffic, leading to a negative traffic–UHII correlation.

Across all seasons, combustion vehicle share shows a weak positive correlation with UHII, while electric vehicle share is weakly negative. However, the small magnitudes prevent firm conclusions about vehicle-type impacts on UHI.

These findings highlight the limitations of simple bivariate correlations for this study. The seasonal sign reversals suggest that 1) traffic impact on UHII depends on the surrounding conditions, 2) solar radiation and climate conditions influence how clearly traffic effects can be observed, 3) linear correlations are not sufficient to separate the influence of the build environment from dynamic effects of traffic. The variables investigated are highly interrelated. For example, temperature naturally peaks at midday, while traffic volumes vary over the day and influence vehicle speeds. Traffic speed is ambiguous: it is zero with no vehicles, high with few vehicles, and lower under congestion. Moreover, speed limits are lower in dense urban centres than in suburban areas. These dependencies call for a multivariate regression approach, to isolate the independent contribution of traffic while controlling the urban morphology and meteorological factors.

6 Conclusions

We proposed a new hybrid approach to investigate how road traffic patterns contribute to the spatiotemporal variability of UHI. The approach consists of generating a synthetic population with daily activity chains from open data, simulating traffic with an agent-based model, and calibrating and validating it against observations. A correlation analysis is then conducted between simulated traffic data and observed UHII data at matching spatial and temporal frames.

Preliminary results for Toulouse indicate that this approach effectively generates synthetic traffic data to complement real-world observations. Results reveal clear diurnal and seasonal patterns in traffic–temperature relationships, as previously observed by [12] and [25]. Traffic's impact on UHII is strongly modulated by seasonal weather, highlighting the need for targeted seasonal and diurnal mitigation strategies rather than uniform year-round measures. Moreover, simple linear correlations are insufficient to infer traffic impacts on UHI, as both temperature and traffic vary over time, and heat effects can be non-linear and lagged, which Pearson correlation cannot capture. Further research is required to refine the first conclusions of this study. First, the simulation covers only a single typical day, and ignore seasonal demand fluctuations, such as higher traffic during tourist periods and lower traffic on holidays. Observed seasonal differences in correlations reflect meteorological effects rather than traffic variability, and including seasonal and weekly traffic variations would provide a more complete picture. Second, a more rigorous validation will be considered for traffic speed, and using more recent open data provided by Toulouse platform. Third, the robustness and explanatory power of analyses could be improved with advanced statistical models that can detect non-linear relationships such as multivariable regression or generalized additive models. Additionally, aggregating traffic data over one hour may miss finer dynamics affecting temperatures; analysing shorter intervals and focusing on rush hours could yield new insights. Finally, future work should adopt more advanced spatial matching between traffic features and sensors. Each meteorological station was linked to its nearest road link, which proved adequate for an initial integration. Incorporating neighbouring links using distance-weighted methods will better capture local traffic influence while preserving correlations. These findings pave the way for more sophisticated data-driven strategies to mitigate traffic-induced urban heat.

Acknowledgments. This work has been funded by the Occitanie Region as part of the MIDOC Key Challenge for Smart and Sustainable Mobility in Occitanie.

Disclosure of Interests. The authors have no competing interests to declare.

References

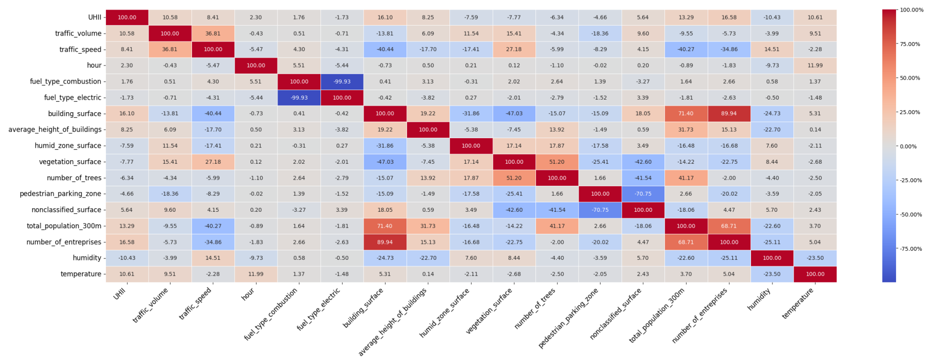
1. Anet, J., Burger, M., Spirig, C., Suter, I.: Improving urban temperature measurements and two applications. City and Environment Interactions **23** (2024)
2. Ben-Dor, G., Ben-Elia, E., Benenson, I.: Population downscaling in multi-agent transportation simulations: A review and case study. Simul. Model. Pract. Theory **108** (2021)
3. Charlton, W.: Simwrapper (2024), `https://simwrapper.app`
4. Cleland, S.E., Steinhardt, W., Neas, L.M., Jason West, J., Rappold, A.G.: Urban heat island impacts on heat-related cardiovascular morbidity: A time series analysis of older adults in us metropolitan areas. Environment International **178** (2023)
5. Flötteröd, G., Chen, Y., Nagel, K.: Behavioral calibration and analysis of a large-scale travel microsimulation. Networks and Spatial Economics **12**, 481–502 (2012)

6. Haddad, L., Aouachria, Z., Haddad, D.: Impact of the transport on the urban heat island. Int. Journal for Traffic and Transport Engineering **5**(3), 252–263 (2015)
7. Hörl, S., Balac, M.: Introducing the eqasim pipeline: From raw data to agent-based transport simulation. Procedia Computer Science **184**, 712–719 (2021)
8. Horni, A., Nagel, K., Axhausen, K. (eds.): Multi-Agent Transport Simulation MATSim. Ubiquity Press, London (2016)
9. Huang, W.T.K., Masselot, P., Bou-Zeid, E., Fatichi, S., Paschalis, A., Sun, T., Gasparrini, A., Manoli, G.: Economic valuation of temperature-related mortality attributed to urban heat islands in European cities. Nat. Commun. **14**(1) (2023)
10. Husni, E., Prayoga, G.A., Tamba, J.D., Retnowati, Y., Fauzandi, F.I., Yusuf, R., Yahya, B.N.: Microclimate investigation of vehicular traffic on the urban heat island through IoT-based device. Heliyon **8**(11) (2022)
11. Jabbar, H.K., Hamoodi, M.N., Al-Hameedawi, A.N.: Urban heat islands: A review of contributing factors, effects and data. IOP Conf. Ser.: Earth Environ. Sci. **1129**(1) (2023)
12. Kammuang-Lue, N., Sakulchangsatjatai, P., Sangnum, P., Terdtoon, P.: Influences of population, building, and traffic densities on urban heat island intensity in Chiang Mai City, Thailand. Thermal Science **19**(suppl. 2), 445–455 (2015)
13. Lakehal, S., Bravo, J., Cazeneuve, C., Audevard, B.: ICU_Toulouse: Study of the UHI - Toulouse Metropolis (2025), https://github.com/Tiben31/ICU_Toulouse
14. Le Bescond, V., Can, A., Aumond, P., Gastineau, P.: Open-source modeling chain for the dynamic assessment of road traffic noise exposure. Transp. Res. D: Transp. Environ. **94** (2021)
15. Lee, J., Berkelhammer, M.: Evaluating the influence of traffic congestion on surface urban heat island intensity. Geophysical Research Letters **52**(16) (2025)
16. Li, J., Sun, R., Liu, T., Xie, W., Chen, L.: Prediction models of urban heat island based on landscape patterns and anthropogenic heat dynamics. Landsc. Ecol. **36**(6) (2021)
17. Loder, A., Ambühl, L., Menendez, M., Axhausen, K.W.: Understanding traffic capacity of urban networks. Scientific Reports **9**, 16283 (2019)
18. Memon, R.A., Leung, D.Y., Liu, C.H.: An investigation of urban heat island intensity (UHII) as an indicator of urban heating. Atmos. Res. **94**(3), 491–500 (2009)
19. Open Data Toulouse Métropole: Open data portal of Toulouse metropolis. https://data.toulouse-metropole.fr/pages/accueil/, access: 21/01/2026
20. Quah, A.K., Roth, M.: Diurnal and weekly variation of anthropogenic heat emissions in a tropical city, Singapore. Atmospheric Environment **46**, 92–103 (2012)
21. Teufel, B., Sushama, L., Poitras, V., Dukhan, T., Bélair, S., Miranda-Moreno, L., Sun, L., Sasmito, A.P., Bitsuamlak, G.: Impact of COVID-19-Related Traffic Slowdown on Urban Heat Characteristics. Atmosphere **12**(2) (2021)
22. Tisséo Collectivités, Cerema: Enquête Mobilité Certifiée Cerema (EMC$^{\mathbf{2}}$) – Report, Greater Toulouse area. Tech. rep. (2024), https://tisseo-collectivites.fr/publications/enquete-mobilite-certifiee-cerema-emc2-rapport-final
23. Wang, A., Ren, C., Wang, J., Feng, Z., Kumar, P., Haghighat, F., Cao, S.J.: Health assessment and mitigating solutions to heat-pollution induced by urban traffic. Journal of Cleaner Production **434** (2024)
24. Welegedara, N.P., Agrawal, S.K., Lotfi, G.: Exploring spatiotemporal changes of the urban heat Island effect in high-latitude cities at a neighbourhood level: A case of Edmonton, Canada. Sustainable Cities and Society **90**, 104403 (2023)
25. Zhu, R., Wong, M.S., Guilbert, E., Chan, P.W.: Understanding heat patterns produced by vehicular flows in urban areas. Scientific Reports **7**(1) (2017)

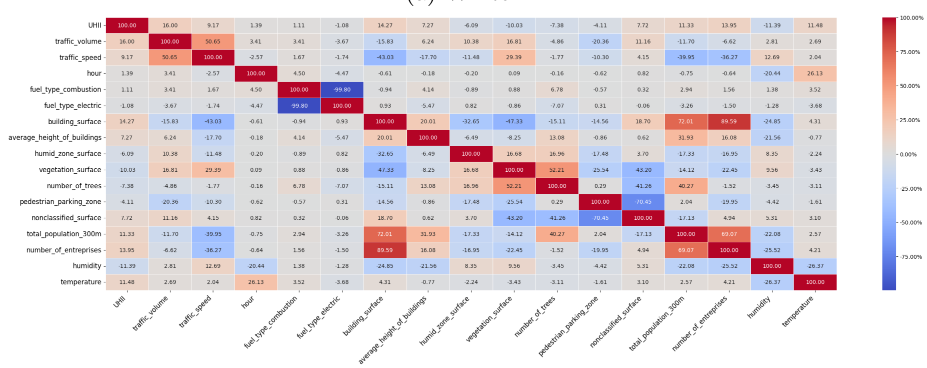
A Supplementary material

City	Traffic data	Weather data	Source	Can be analysed?
Barcelona	yes(> 2017)	yes***	Opendata BCN, Sentilo BCN	No - Limited urban coverage for weather data
Berlin	yes(2015-25)	yes***	Berlin Open Data, Urban Climate Observatory, FU Berlin Microclimate Network	No - Limited urban coverage for weather data
Bordeaux	yes(2015-24)*	no	DataHub Bordeaux Metropole	No - Missing weather data
Dijon	no	no	Dijon Metropole Open Data, MUSTARDijon	No - Missing traffic data, weather data but not open
Grenoble	yes* (<2022)	yes	Data Metropole Grenoble, CASSANDRE, EaSy Data	No - Traffic data too aggregated
Lyon	yes (2022-25)	yes***	Data Grand Lyon	No - Limited urban coverage for weather data
Madrid	yes(2022-24)	yes***	Mendeley Data, MateMAD project	No - Limited urban coverage for weather data
Marseille	yes* (2020)	no	Data Metropole Aix-Marseille Provence	No - Missing weather data
Montpellier	yes* (> 2012)	no	Open Data Montpellier Méditerranée Métropole	No - Missing weather data
Paris	yes(2020-25)	no	Paris Data	No - Missing weather data
Rennes	yes**	no	Rennes Metropole OpenData, Rennes Urban Network	No - Weather data but not open
Toulouse	yes* (2018-23)	yes	Toulouse Metropole OpenData	No - Traffic data too aggregated
Vienna	yes* (<2024)	yes	Open Government Data Austria portal, Luftdaten.info	No - Traffic data too aggregated
Konstanz	no	yes	Open Data Konstanz	No - Missing traffic data
Freiburg	yes***	yes	Open Data for Baden-Württemberg, uni-Weather app	No - Limited urban coverage for traffic data

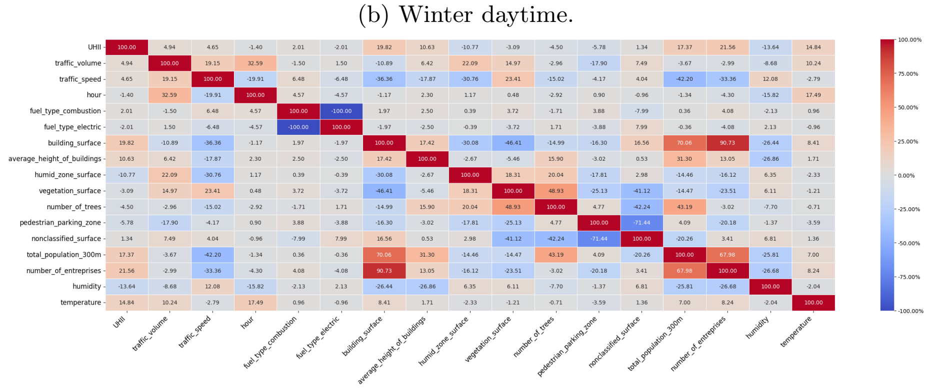
Table 1: Review of publicly available datasets for 15 European cities. *: very aggregated. **: no historical data, only real-time data. ***: very few stations.



(a) Winter.

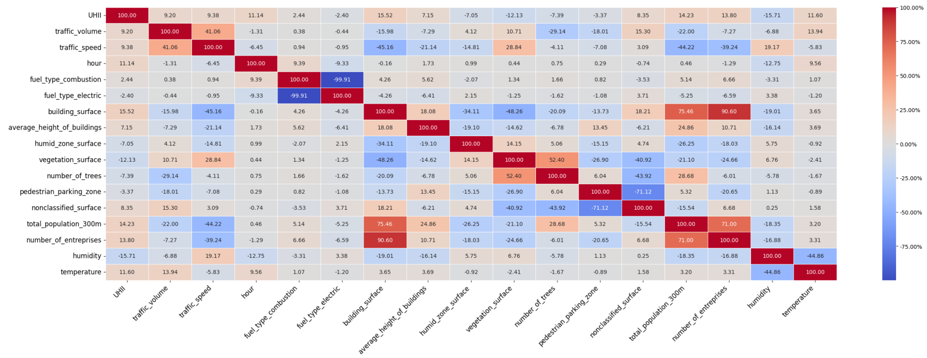


(b) Winter daytime.

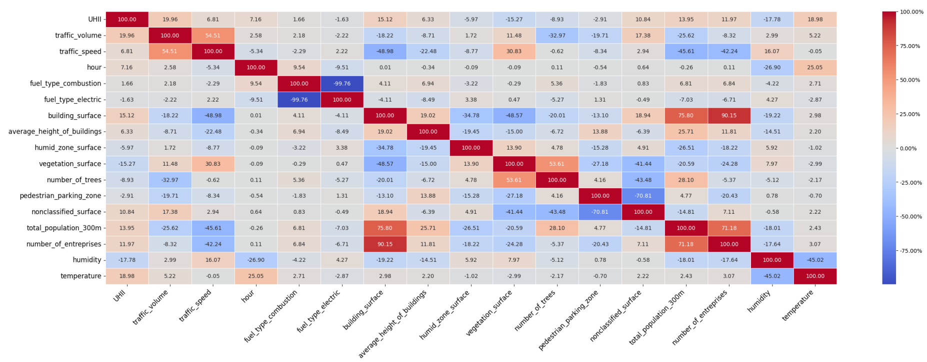


(c) Winter nighttime.

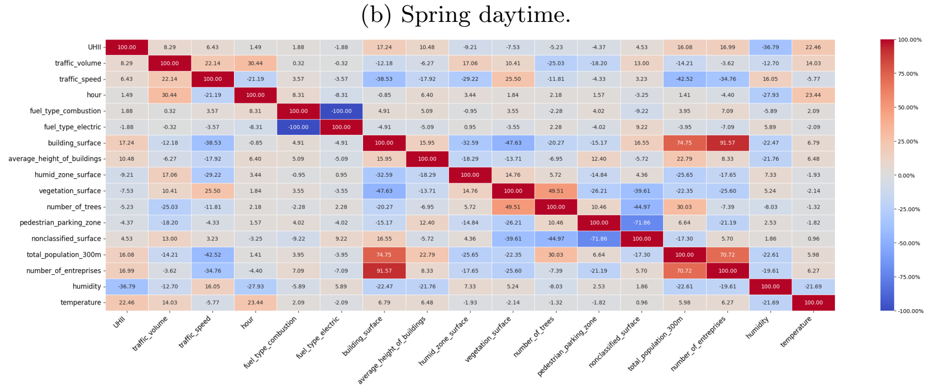
Fig. 3: Correlation matrices of static and traffic factors with UHII.



(a) Spring.

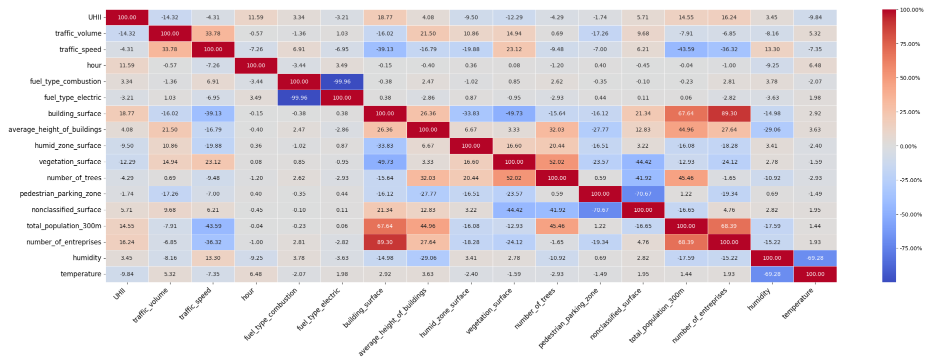


(b) Spring daytime.

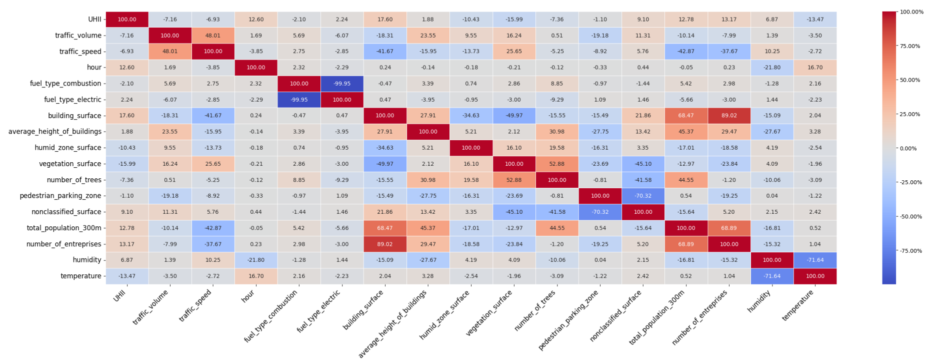


(c) Spring nighttime.

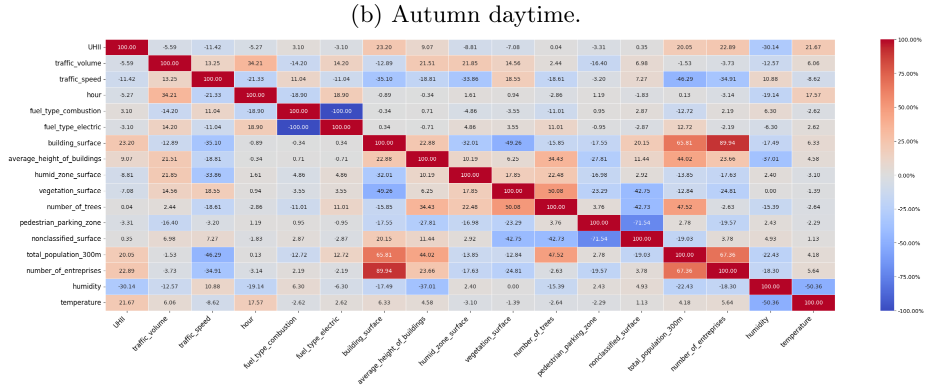
Fig. 4: Correlation matrices of static and traffic factors with UHII.



(a) Autumn.

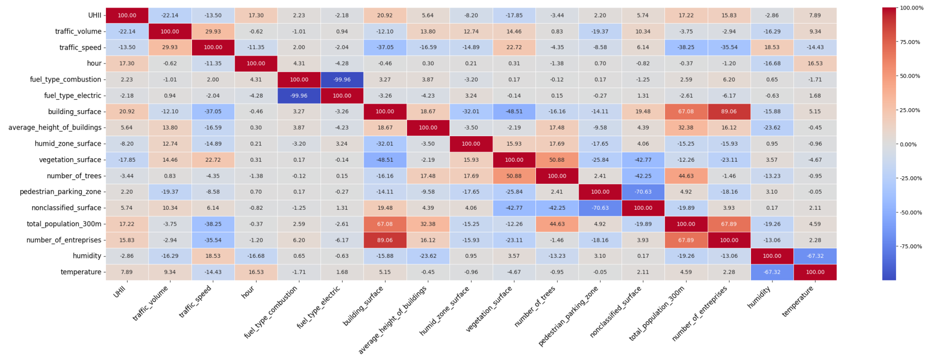


(b) Autumn daytime.

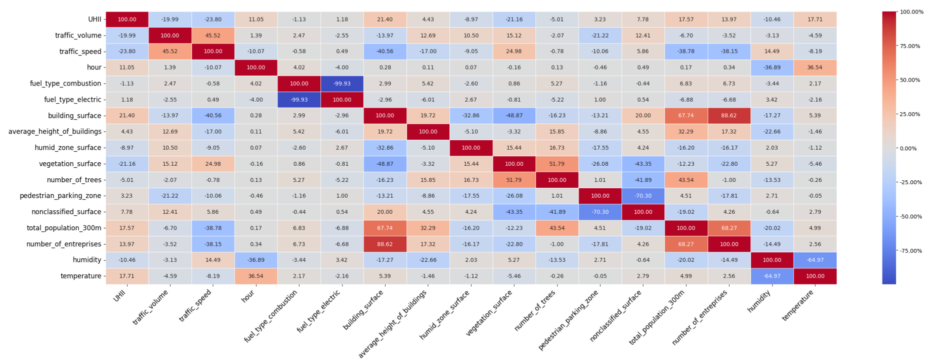


(c) Autumn nighttime.

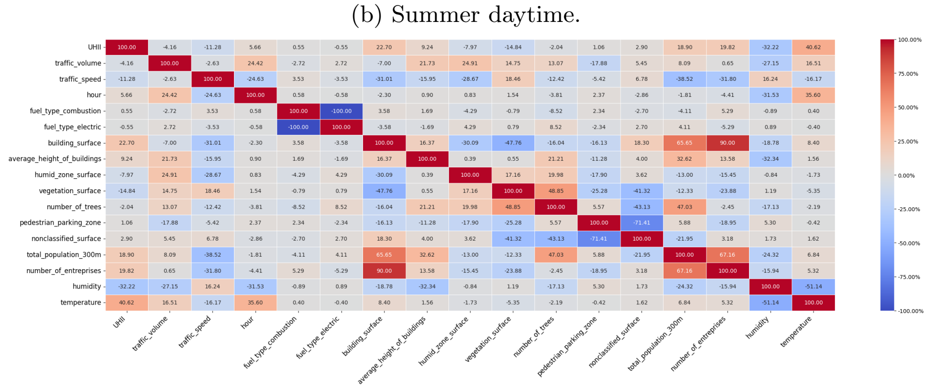
Fig. 5: Correlation matrices of static and traffic factors with UHII.



(a) Summer.



(b) Summer daytime.



(c) Summer nighttime.

Fig. 6: Correlation matrices of static and traffic factors with UHII.