

Article

Getting Real: The Challenge of Building and Validating a Large-Scale Digital Twin of Barcelona's Traffic with Empirical Data

Javier Argota Sánchez-Vaquerizo 

Computational Social Science, D-GESS, ETH Zürich, 8092 Zürich, Switzerland; javier.argota@gess.ethz.ch

Abstract: Large-scale microsimulations are increasingly resourceful tools for analysing in detail citywide effects and alternative scenarios of our policy decisions, approximating the ideal of 'urban digital twins'. Yet, these models are costly and impractical, and there are surprisingly few published examples robustly validated with empirical data. This paper, therefore, presents a new large-scale agent-based traffic microsimulation for the Barcelona urban area using SUMO to show the possibilities and challenges of building these scenarios based on novel fine-grained empirical big data. It combines novel mobility data from real cell phone records with conventional surveys to calibrate the model comparing two different dynamic assignment methods for getting an operationally realistic and efficient simulation. Including through traffic and the use of a stochastic adaptive routing approach results in a larger 24-hour model closer to reality. Based on an extensive multi-scalar evaluation including traffic counts, hourly distribution of trips, and macroscopic metrics, this model expands and outperforms previous large-scale scenarios, which provides new operational opportunities in city co-creation and policy. The novelty of this work relies on the effective modelling approach using newly available data and the realistic robust evaluation. This allows the identification of the fundamental challenges of simulation to accurately capture real-world dynamical systems and to their predictive power at a large scale, even when fed by big data, as envisioned by the digital twin concept applied to smart cities.

Keywords: traffic microsimulation; real data; validation; SUMO; urban digital twin; smart cities; large-scale; mobile phone data



Citation: Argota Sánchez-Vaquerizo, J. Getting Real: The Challenge of Building and Validating a Large-Scale Digital Twin of Barcelona's Traffic with Empirical Data. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 24. <https://doi.org/10.3390/ijgi11010024>

Academic Editor: Wolfgang Kainz

Received: 7 November 2021

Accepted: 26 December 2021

Published: 30 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cities are complex systems [1] that feature non-stationary behaviour [2], that is, they are in constant change and transformation. To increase the quality and resilience of urban areas, new forms of mobility, together with different patterns of (land) use, and a redesign of spaces are proposed, which are typically characterized by as much chaos as order. This makes them unsuitable for traditional control approaches, as the involved conditions and problems are always evolving. This complexity of urban settings makes it particularly difficult to assess and evaluate the feedback, side and cascading effects of policies, decisions, and behaviours. Unfortunately, testing in real-life settings is costly, risky, sometimes unethical, and often unfeasible. Even in the case of piloting, the possibility of extrapolating and scaling results is questionable.

Alternatively, it is possible to develop computer-based simulation models that try to capture the functioning of a city as accurately as possible [3]. The development of such virtual scenarios allows testing, assessing, and measuring the impact of change at reduced cost and risk, with higher flexibility, and a capacity to control the conditions better than possible in real life. These so-called digital twins [4] are fed with (partly real-time) smart city data [5] and are the goal of many cities [6–11], thereby pushing the long tradition of urban simulation [12–15] to the technological cutting edge. The progressive spread of

cyber-physical systems [16,17] becomes a powerful opportunity for real-time management, for anticipating possible outcomes of planning and policy decisions before implementations are made, and for fostering co-creation and citizen participation [5].

However, many of the envisioned changes are on a small and high-frequency scale [3,18]. That is, on the scale of people's perception and experiences, such as drivers getting around, new forms of mobility, the spatial configuration of buildings, streets, and public spaces, the design of infrastructures, and the emergence of new smart digitally-enabled systems. Therefore, the use of simulations as planning tools needs the high level of interpretability and detail obtained by microsimulations capabilities [19]. Despite all the incremental details, depth, and richness of information becoming available, it is needed to account for necessary simplification [3], non-measurable qualities, as well as limitations and uncertainties when it comes to the behaviour of people and socio-cultural systems, which need to be included [20].

Therefore, the challenge is to build and validate realistically a microsimulation of a large urban area, to study on a small scale of detail citywide policies. This paper presents a pipeline to build a digital twin for the traffic flow in the large, dense, and congested urban area of Barcelona, Spain, based on novel cell-phone mobility data and publicly available databases using the SUMO agent-based microsimulation framework [21]. The paper is organised as follows. Section 1.1 introduces existing large-scale urban microsimulation scenarios. Section 2 describes the pipeline to build the proposed model, including the pre-calibration of parameters. Section 3 explains the evaluation of the results of the simulation scenarios based on a robust multi-variable comparison with real measured data, namely with respect to various metrics on different spatio-temporal scales. Finally, Section 4 expands on limitations, possible applications, and potential future research.

1.1. Existing Large-Scale Urban Microsimulation Scenarios

Agent-based microsimulations follow the logic that larger-scale effects observed in the physical reality can be obtained by the aggregation of emulated individual properties and behaviours of agents at a small scale (such as the individual position, speed, and acceleration) [22]. However, they require large amounts of detailed disaggregated data and long processing times, needing considerable computational resources both for calibrating and for running them. Determining the many parameters describing individual behaviours and the interactions of numerous agents, are among the main challenges in building, calibrating, and evaluating these models. This results in compromises in terms of realism, while one needs to consider that minor errors in any of these components may propagate and cumulate, thereby producing deviations or even incorrect results [23].

The relatively few published examples of complete, comprehensive, and validated large-scale urban scenarios using microsimulations reflect this complexity. Urban settings are even more challenging to calibrate and build than equivalent interurban settings [24,25], because of very high levels of mobility demand concentrated in small areas with complex, dense, and intricate networks.

The numerous examples of building, calibrating, and validating traffic simulations in scientific literature and policymaking [11,26–30] show this complexity trade-off between size and level of detail: large-scale models are more suitable for macro and mesoscopic approaches, while detailed microsimulations focus on smaller scenarios [31]. This bounds their operational possibilities and limits the detailed study of citywide interventions and policies [32]. Precisely, this paper focuses on this gap in the urban simulation literature: the lack of large-scale urban scenarios that use traffic microsimulation [33–35], and missing robustly quantitative validation with real data [33,36–41] (see Table A1). Reviewing the few scenarios combining a large-scale scope with a microscopic approach supports the design of the proposed new data pipeline for building, calibrating, and validating the model while being able to highlight and address the existing challenges:

- Large urban microscopic scenarios are complicated and slow, which obstructs microsimulations over extended periods of time, as well as their calibration (requiring multiple runs with different parameter sets);
- This causes compromises between efficiency and realism;
- Validation is often based on a qualitative assessment of results or lacks a quantitative comparison with empirical measurements (requiring a different data set than used for calibration);
- The empirical data needed are frequently incomplete or inaccurate.

However, beyond the different combinations of data sources and strategies for demand creation used by these urban simulations, their most relevant differences are in their size and validation methods. The lack of commonly established approaches and the differences in available observational data in each location for validation makes it difficult to assess the results and consequently to compare them regarding size, duration, and realism.

Therefore, this paper analyzes existing examples to develop further large-scale microscopic urban simulations with a wide multi-scalar realistic validation supported by observational data. It aims to provide a feasible path to investigate their new operational capabilities and limitations, both in applicability and developments steps within the increasing interest of the so-called 'digital twins'.

Hence, the proposed microsimulation for the large urban area of Barcelona combines novel mobility data from cell phone records and publicly traditional surveys with detailed urban network data extracted from OpenStreetMap. The quantitative validation is based on a different empirical data set over 24 h using various metrics. As a result, reportedly, this model is one of the most complete published large-scale urban traffic microsimulations regarding the scope, network size, duration, and realism (validation). Hence, this is probably the most detailed simulation scenario based on continuous-in-time car-following and route choice models currently known, which allows us to judge the potentials, challenges, and limitations of the "digital twin" approach including the technical feasibility of the concept, its safety, societal impact, and predictability power (see Section 4).

2. Materials and Methods: Building of the Scenario

The construction of a new microsimulation model for the urban area of Barcelona is based on the use of traditional data sources, including publicly available geographical information and conventional surveys, as well as new big data, specifically anonymized mobility data from mobile phones geographic position records. The general pipeline (Figure 1) shows the three main steps for building the scenario. The first step involves the generation of the transport system network and the travel demand. In the second step, the travel demand is adapted to find an efficient spatial distribution of routes through microsimulations. Finally, the third step compares the results of the resulting microsimulations from stage two with real-world data to evaluate how well the model fits the empirical ground truth.

The microsimulation software SUMO [21] is particularly suited for this purpose because of its open-source design, offering microscopic and mesoscopic capabilities able to conveniently provide detailed descriptions of interactions [19] and great versatility for scenario building [42,43]. Among others, it supports different data formats, demand modelling based on synthetic and real populations, different assignment approaches and route choice methods [44,45], as well as routing algorithms [46–48].

The construction of an urban model to simulate mobility patterns requires two basic components: (i) a network, as the virtual environment where the agents will move, and (ii) a demand pattern that will generate the flows between the different locations of the mentioned network. In this case, the focus of the study is on the interaction between the (public) road infrastructure and the (private) vehicle traffic. Over here, public transportation and freight flows (~5% of traffic flow) are not taken into account.

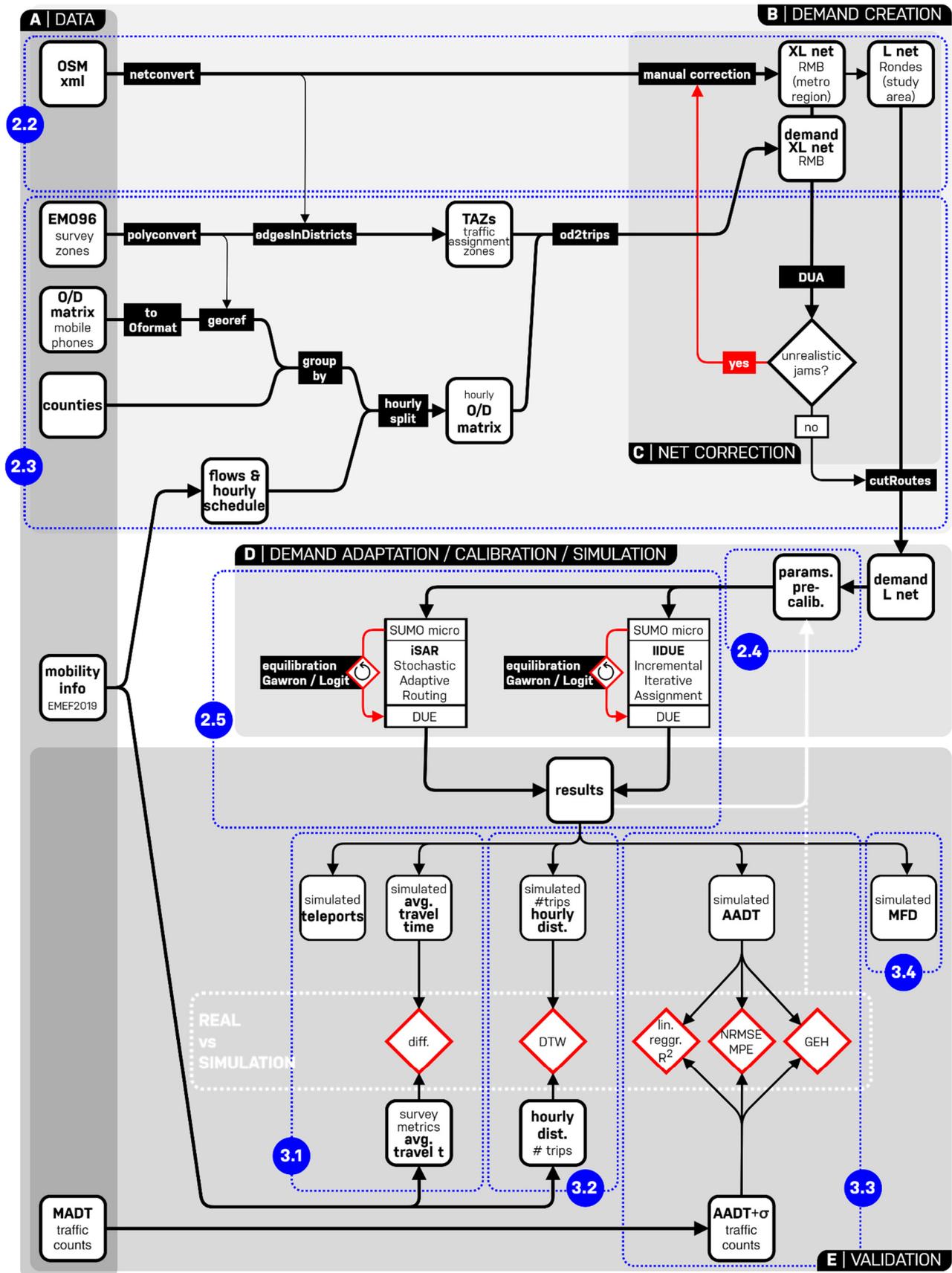


Figure 1. Proposed data pipeline to build the microsimulation scenario. In blue, the sections of this paper are referenced.

2.1. Geographical Scope

The implementation of the microsimulation model covers more than the city of Barcelona itself (Figure 2). The selection of the Metropolitan Region of Barcelona (*Regió Metropolitana de Barcelona* or RMB) as the extended initial area of the model, larger than the city itself, allows including the relevant through traffic for the core area. The RMB has a size of 3126 km² and contains about 5,103,000 inhabitants (~40% of the surface and ~90% of the population of the province of Barcelona). It includes the counties (*comarques*) of Alt Penedès, Baix Llobregat, Barcelonès, Garraf, Maresme, Vallès Occidental, and Vallès Oriental. Additionally, surveys use alternative zoning for mobility statistics, based on concentric metropolitan subareas or belts (city of Barcelona, rest of the inner Metropolitan Area, outer Metropolitan Area, and rest of the RMB) (Figure 2a). Altogether, the vehicular trips between Barcelona and any of the locations within the RMB represent 96% of all the daily car trips measured in the city (i.e., less than 4% of car trips within the city limits of Barcelona are related to locations out of the RMB) [49] (Table 1).

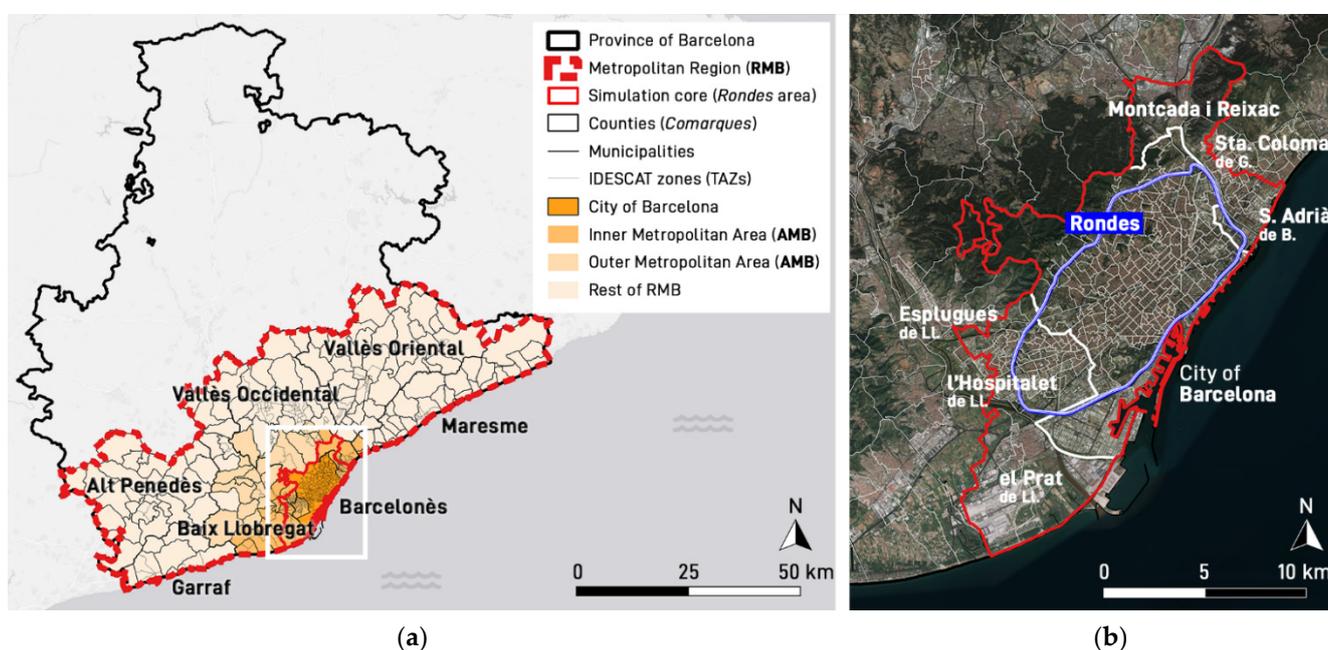


Figure 2. Geographical domains covered by the microsimulation: (a) Location of the RMB within the province of Barcelona, including the different geographical divisions used in the model; (b) Core area of the detailed microsimulation (in red) defined by the city ring roads (Rondes, in blue) and corresponding Traffic Assignment Zones (TAZs, in white).

Table 1. Characterization of main metropolitan domains.

	Province of Barcelona	RMB (Extended Area)	Rondes Area (Simulation Core)	City of Barcelona
Area (km ²)	7726	3126.2	182.55	101.35
Population (2019)	5,664,579	5,151,263	2,411,755	1,636,762
Total trips (any mode)	19,259,471	17,430,628	8,176,511	5,682,214
Trips in private vehicles	6,946,355	5,927,332	2,063,177	1,653,183

Thus, this expanded spatial scope allows capturing the complete traffic demand in the denser part of the urban area of Barcelona. Differently than in previously published microsimulation scenarios, where the network size and complexity in extended urban regions are frequently cropped or topologically simplified [50,51], in this research, the extended metropolitan region (RMB) is initially simulated with full detail to provide the complete spatial distribution of trips within the core area. This core area (see Figure 2b) is

not based on administrative divisions but on the functional and typological characterization defined by the main ring roads of the city (i.e., *Ronda Dalt* and *Ronda Litoral*, hence its name, *Rondes*). It includes the built-up area of the city of Barcelona and parts of a few surrounding municipalities (l'Hospitalet de Llobregat, Cornellà de Llobregat, Esplugues de Llobregat, el Prat de Llobregat, Sant Adrià de Besòs, Santa Coloma de Gramenet, and Moncada i Reixac). This compromise between realism and simplification allows preserving a realistic representation of the whole traffic demand of a very large urban area without simplifications while keeping the complexity of the network reasonably well by focusing on the part of the urban area that suffers the highest levels of congestion (i.e., *Rondes*).

2.2. Transport Network

For the creation of the urban network, the use of collaborative platforms offers open-source geographical data with a high degree of precision. OpenStreetMap (OSM) [52] is probably not only the best-known example of volunteered geographic information (VGI) [53], it is also used by corporations as a mapping alternative to technologies such as Google Maps [54]. Additionally, SUMO provides convenient tools to transform OSM data into networks suitable for microsimulations.

Given the large area to be modelled, a script based on OSMNX [55] is adapted for downloading OpenStreetMap data [56] in XML format using the Overpass API [57], preserving all the relevant features for a microsimulation, including access and turning restrictions, speed limits, priority and class of roads, directions of circulation, traffic signs, and the number of lanes. SUMO's *netconvert* tool [58], allows transforming this OSM XML file into an appropriate directed road network that encodes all the relevant features in the SUMO network format. This process relies on a good number of heuristics that try to translate the original OSM attributes into connectivity features, particularly regarding the layout of intersections, rights of access, and the positioning and coordination of traffic lights settings.

Note that the frequently inconsistent quality of OSM data [53] directly affects the creation of the network [59]. Misspecifications in mapping features such as intersections layout, number of lanes, speed limits, or traffic restrictions lead to unrealistic simulation outcomes that can considerably affect the performance of the model, for example, a lower traffic throughput (see Figure 3).

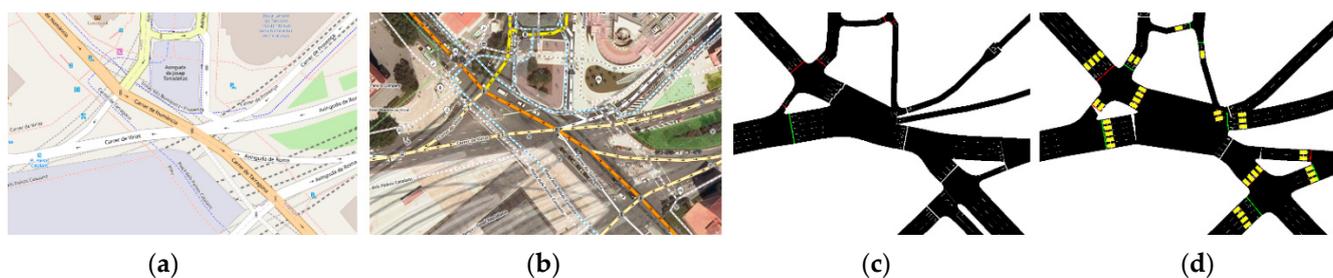


Figure 3. Process of OSM data acquisition and edit for network generation: (a) OSM data; (b) OSM data correction based on third-party satellite imagery; (c) SUMO network with conversion errors (e.g., misspecifications of the lanes number, traffic lights, etc.); (d) final corrected SUMO network used for the simulation including virtual induction loops to measure traffic flows.

To prioritize the vast task of correcting the large road network, whether due to errors caused by mapping misspecifications, wrong outcomes of SUMO conversion heuristics, or a combination of both, two complementary approaches are applied:

- Short simulation runs using direct assignments of the traffic demand are employed to identify major errors hindering traffic performance.
- Satellite [60–63] and street-level [64,65] imagery are used as independent third-party reference (ground truth) for correction.

Most of the needed corrections were done directly on OSM servers, contributing to the collaborative aim of the platform, by improving and updating the quality of the map in the area. Some fine-tuning has been performed directly in the definitive SUMO net file, using SUMO's *netedit* tool [66], correcting errors caused by *netconvert*'s conversion heuristics, particularly regarding intersections and traffic lights. Traffic light plans are a very relevant feature. In the current model, the real traffic light system used by the city was not available to us and, hence, not utilized. To simulate a realistic, simple, and efficient management of regulated intersections, the *actuated* option of SUMO is chosen for all the traffic lights plans [67]. As a result, the network in the extended simulated area (RMB) covers a mix of urban, suburban, and rural areas. Differently, the core area (Rondes) has a very urban and dense character (Figure 4 and Table 2 for networks details).

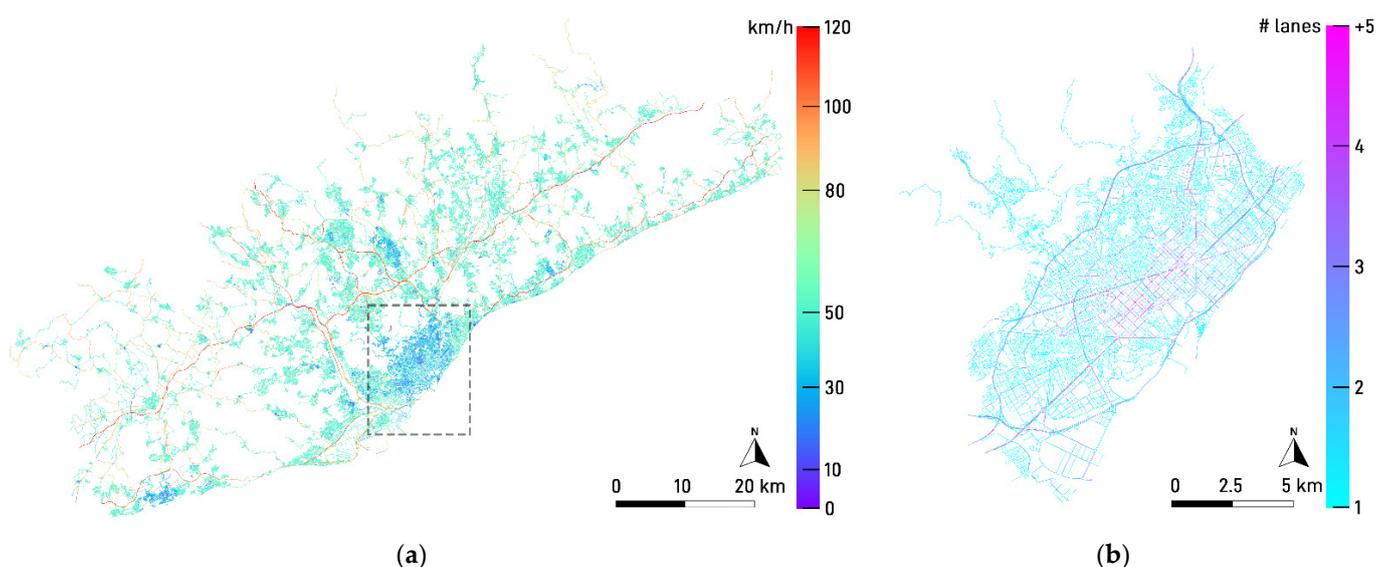


Figure 4. Road network: (a) Network of the extended simulation area (RMB) showing speed limits in km/h; (b) Network of the core simulation area (Rondes) used for validating the detailed microsimulation, where different colours represent the number of lanes.

Table 2. Features of the road networks.

	RMB (Extended Area)	Rondes Area (Simulation Core)
Area (km ²)	7726	182.55
Roads length (km)	24,153.12	2506.03
Lanes length (km)	27,794.13	3673.17
Number of links	204,935	25,307
Number of intersections	98,752	13,993
Traffic lights	3567	2035
Traffic Assignment Zones (TAZs)	577	296
Virtual induction loops	10,496	10,496

2.3. Demand Creation

The approach followed in the demand generation aims to avoid underestimating real traffic flow, as it seems to happen in many microsimulations. In general, traffic within an area is expected to be caused by trips that have their origin or destination in that zone and also by through traffic originating and ending elsewhere. To deal with this situation, a basic spatial and temporal distribution of traffic is created for the extended simulation area (RMB) to identify the fraction of traffic that directly affects the core urban area (Rondes), which provides more realistic results than directly overlooking through traffic (see Section 3 for comparing results without considering through traffic).

Two sources are used complementarily for the generation of traffic demand:

1. Originally, traffic patterns are extracted from the origin/destination (O/D) matrices estimated from cell phone data [68,69]. This method provides a direct estimation of flows between smaller areas (here areas refer to statistical zones used by the Catalan Institute of Statistics (IDESCAT). They are larger than census tracts, but smaller than districts or neighborhoods (see Figure 2)) with finer granularity and larger sampling, overcoming limitations of estimations from self-reported surveys [70]. Given the focus on vehicular traffic, only the O/D matrix specific for daily private vehicle trips during a working day is considered, resulting from raw cell phone data processed based on the modal split from mobility surveys and public transit ridership data;
2. Then, conventional mobility surveys [49,71] are used complementarily to expand and correct the O/D matrix, in particular, to fill some gaps such as the actual hourly distribution of trips. It has been observed that the method to infer mobility patterns based on mobile phone location records tends to overestimate trips [68].

Thus, (1) the mobile phone data determine the detailed proportions of flows between the areas with a granularity unfeasible for survey data, while (2) the survey data determine the overall number of trips (which do not count partial trips separately, as cell phone data often do).

These O/D flows are spatially aggregated at the level of IDESCAT areas (see Figure 2). These areas considered as Transport or Traffic Analysis Zones (TAZs) are used for georeferencing the trips. SUMO provides the *edgesInDistricts* script [72] to reference network edges to TAZs in such a way that origins and destinations of trips are linked to defined positions in the transportation network.

Initially, traffic demand is modelled for the entire extended area (RMB) using default parameters in SUMO. Later, all the O/D flows, which potentially may affect the core simulation area (Rondes), are kept, but peripheral traffic flows between the outer AMB and the outer RMB are ignored. This is also beneficial for computational performance.

As remarked, the main issues of these private vehicles' O/D matrices from cell phone data are the overestimation of measured trips and the lack of disaggregated hourly distributions. Hence, inspired by data fusion approaches, which combine cell phone data with other sources [73], this model uses mobility surveys for complementing the O/D data. Particularly, these surveys provide hourly histograms of trips separated by RMB transport subareas and mode of transportation [49]. This newly implemented method for O/D data correction and expansion is performed in two steps:

1. TAZs are grouped by transport RMB subarea. Then, following the indications of the researchers in charge of the original analysis of the cell phone data, O/D flows are linearly scaled such that they match the number of trips for each of these subareas measured by the yearly mobility surveys [49,71], which are considered as accurate references in transport modelling and planning [73]. Each subarea requires a different scaling, as the errors in the estimation of trips from cell phone data vary with trip length, population density, and the concentration of antennas [68,70];
2. The scale-corrected and hourly disaggregated O/D data obtained in the first step assume a symmetric number of trips between the different areas (i.e., outbound and inbound trips are assumed to be equal in each hour of the day). This is unrealistic, of course (e.g., residential areas tend to emit more outbound trips in the morning towards working areas, while they tend to have more inbound trips in the afternoon and evening for returning commuters). To account for unbalanced flows between areas, the hourly histogram of trips is corrected by a factor based on skew-normal distributions for the peak hours of inbound and outbound trips in the morning and the afternoon/evening [74], which are computed using the Attraction and Emission Ratio (*Ràtio d'atracció i emissió*, RAE) [49]. Due to the lack of more disaggregated data, the trips staying within the same transport subarea are assumed to be symmetrically balanced throughout the day.

As a result, 16 different O/D matrices are created for the pairwise relations between the four transport RMB subareas, accounting for inner, inbound, and outbound trips with their distinct hourly distributions. For instance, areas with lots of working places tend to have an inbound morning peak, and outbound afternoon or evening peaks.

In the next stage, all this time-series information is merged in SUMO, using *od2trips* [75] with a timeline argument to convert raw O/D matrices into trips files referenced to the transportation network with the appropriate hourly distribution. The initial travel demand created by O/D matrices is defined by the time and edge of departure, and the destination edge in the network for each vehicle or trip. Unfortunately, this does not provide the whole path needed to identify which trips cross the core area. Therefore, *duarouter* is used to route all trips with the Direct User Assignment (DUA) method [76]. Although this simple assignment approach only provides the optimum path when the network is empty (i.e., without considering changing traffic conditions and travel times throughout the day), it is used as a reference to create the initial distribution, which is adapted and optimized later.

Once the DUA routing is performed for the extended area (RMB), SUMO's *cutRoutes* tool [77] is used for cropping the resulting paths, which traverse the smaller but much denser network of the core area (Rondes). This is the scenario that is used for detailed analyses and simulations. Out of the 3,185,185 private vehicle trips initially computed for the whole extended RMB area, 2,063,177 trips turn out to cross some section of the core urban area of Barcelona (Rondes) (see Table 2). These are included in the final travel demand model.

2.4. Calibration of Simulation Parameters

The numerous parameters involved in microscopic models and their long runtimes (around 30 h in this case) are troublesome for calibration [26], making the use of the conventional simulation-based optimization [78] paradigm intractable [79]. This leads to compromises in realism. Initial tests identified that the aspects with the greatest impact on the simulation performance were the time step defining the temporal resolution, the car-following model chosen, and the procedure of rerouting:

- The time step of the simulation needed to be reduced from 1 to 0.25 s to reproduce the continuous change of the traffic state sufficiently well [80];
- Otherwise, the default parameters of Krauss' car-following [81] and Erdmann's lane-changing [82] models in SUMO are used, adjusting the length of the vehicle to 4 m. This is closer to the average size of vehicles in Spain [83] and also accounts for the exceptionally high proportion of motorbikes in Barcelona [49,84];
- Two parameters linked to routing are analysed: the probability for a vehicle to update its path during the simulation (*device.rerouting.probability*) [85] and the factored priority of roads (*weights.priority-factor*) as encoded in OSM data. To find their best values, different parameter combinations are explored by an algorithm that iteratively runs simulations for the *device.rerouting.probability* with values between 0.6 and 1.0 in steps of 0.1 and for the *weights.priority-factor* with values between 0 and 110 in steps of 10. The results are then compared based on the metrics used for the general evaluation of the model: number of teleports ("Teleporting" is a mechanism that SUMO uses for avoiding agents, i.e., vehicles or pedestrians, to get indefinitely stuck in the simulation, by moving them to their following route edge, if they collided or are stopped for longer than a defined time period [86]. This ensures that minor specification errors do not mess up the entire simulation.), regression coefficient, R^2 , RMSE, NRMSE, GEH for traffic counts, and DTW for the demand time series, see Section 3 for the explanation of the metrics. Consequently, the method obtains the minimum value for the errors ($RNMSE_{traffic\ counts} \sim 0.385$, $DTW_{hourly\ demand} \sim 3.5$) and maximum correlation coefficient for the linear regression (*coefficient* ~ 0.91 , $R^2 \sim 0.81$) for *device.rerouting.probability* = 1 and *weights.priority-factor* = 100 (see Figure 5).

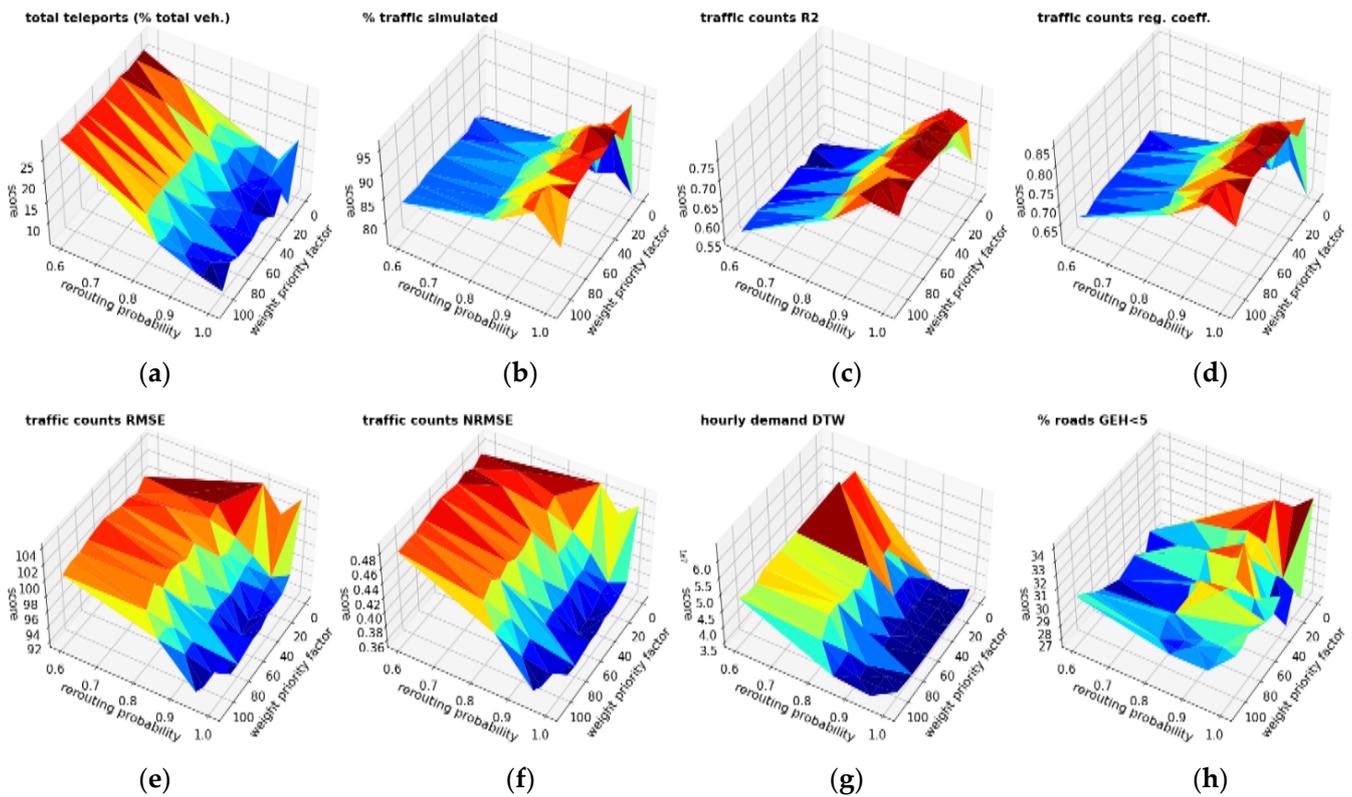


Figure 5. Evaluation of the *weights.priority-factor* and *device.rerouting.probability*. Lower values for RMSE, NRMSE, and teleports are preferred, while higher values of the R^2 , regression coefficient, and GEH indicate better performance: (a) Total teleports expressed in percentage of total vehicles loaded; (b) percentage of total demand effectively loaded into the simulation; (c) R^2 coefficient between simulated and real traffic counts; (d) regression coefficient between simulated and real traffic counts; (e) RMSE between simulated and real traffic counts; (f) NRMSE between simulated and real traffic counts; (g) distance between theoretical temporal traffic demand curve and simulated based on DTW metric; (h) percentage of monitored links with a GEH statistic below 5.

2.5. Demand Adaptation

O/D pairs are the only source for the creation of traffic demand. It means that no ground truth regarding traffic counts, turning ratios, or other similar data is used to force the demand such that it matches reality. Instead, the results of the unconditioned adaptation are compared to the empirical “ground truth” to measure the quality of the model (see Section 3).

The initial demand creation based on a Direct User Assignment (DUA) is simplistic because: (i) it does not account for the changing traffic conditions over the day (i.e., the fastest paths change depending on the congestion levels caused by other vehicles in the network), and otherwise (ii) it is found to be somewhat insufficient for large complex traffic scenarios. The process of demand adaptation tries to redistribute traffic realistically and efficiently in a time-dependent way. Frequently, in traffic simulations, this process is performed by Dynamic User Equilibrium (DUE) algorithms [45], which expand Wardrop’s user equilibrium principle under the assumption that traffic tends to distribute spatially such that no single user can improve the travel time by changing the route [87]. However, making route choices is mostly considered to be a probabilistic process, which rejects the assumption that drivers have a perfect and complete knowledge of the network to optimize their travel cost [88].

SUMO provides the *dualIterate* tool to approximate the DUE [45] through an iterative optimization process that minimizes the overall travel time (or any other given cost function). In each run, a DUA is used to perform a whole simulation that allows computing

the travel costs of edges, considering the changing congestion of the network loaded with traffic. Then, this is used to compute lower-cost paths, which is the basis of the DUA in the next iteration. By repeating this process many times, the algorithm tends to get closer to the equilibrium solution, which is imagined to be the best. Therefore, this paper extends for larger scenarios from previous approaches on heuristic dynamic assignment for achieving a stochastic user equilibrium [89] by proposing two iterative optimization approaches using *dualIterate*:

- In-Simulation Adaptive Rerouting (iSAR) uses the capabilities of SUMO to allow vehicles to update automatically their paths during the simulation to find a faster route based on existing traffic conditions. This approach does not seek directly for a DUE but tries to adapt traffic efficiently, using a very fast adaptive strategy. Additionally, integrating this method into the *dualIterate* routine allows checking the stability of the results. It might be possible to improve the performance of the traffic assignment even further by identifying underlying structures in the evolution of the travel costs of the network, although it is expected that reactive rerouting during simulation could address many of these effects;
- Iterative Incremental Dynamic User Equilibrium (IIDUE) depends exclusively on *dualIterate* to run the simulation and the DUA repeatedly and to adapt the traffic distribution based on newly found optimal travel costs in each iteration. However, due to the high level of congestion in this scenario, it is not possible to simulate the whole traffic demand from the beginning. The oversaturation and gridlock caused by the initial, trivial DUA do not generate informative travel cost values for edges suitable for the optimization process. To overcome this limitation, this approach implements an incremental assignment strategy. Starting with 2% of the overall traffic demand, an additional 2% of trips are added in each iteration to allow the algorithm to gradually account for changing travel costs, before reaching unrecoverable and uninformative congestion. On the other hand, this process is time-consuming, that is, slow and computationally expensive.

To compute the probability of choosing a new route based on cost optimization in every single iteration, *dualIterate* offers two different methods, which are compared: **Gawron** and **Logit**. Both approaches take as inputs a weight or cost function w for the edges, a set of routes R , and for each of these routes r , an old cost c_r , an old probability p_r , and a new cost c'_r and probability p'_r [45]. The two methods differ on how they update these new values in every simulation iteration:

Logit is simpler. It uses an explicit analytical formula to compute the new probability directly by only considering the sum of the costs of all edges e , which are part of the route r in the last iteration:

$$c'_r = \sum_{e \in r} w(e) \quad (1)$$

Then, the probabilities are updated, using an exponential distribution containing a parameter θ , and normalized by the sum over all routes s in the set R :

$$p'_r = \frac{\exp(\theta c'_r)}{\sum_{s \in R} \exp(\theta c'_s)} \quad (2)$$

However, in this paper, a generalization of the logit model named *c-Logit* [90] is used, which is defined by

$$p'_r = \frac{\exp(\theta(c'_r + cf'_r))}{\sum_{s \in R} \exp(\theta(c'_s + cf'_s))} \quad (3)$$

This gives more realistic route choice probabilities, which consider a commonality term cf'_r that accounts for the overlap of alternative paths (this is particularly relevant

given the regular grid of Barcelona that may lead to many similar overlapping alternatives), specified by

$$cf'_r = \beta \ln \sum_{s \in R} \left(\frac{L_{rs}}{\sqrt{L_r} \sqrt{L_s}} \right)^\gamma \quad (4)$$

Herein, L_{rs} is the length of edges shared by routes r and s , respectively, with total lengths L_r and L_s . β and γ are parameters (when $\beta = 0$, c -logit becomes the simpler logit model). The models in this paper use a value of $\beta = 0.15$ and $\gamma = 1$ (the default in SUMO).

Gawron computes the probability for choosing a new route from a set of alternatives, based on the travel time (or any other cost function) of the route chosen in the previous iteration. It further considers the sum of travel times of these alternatives and the previous probability of choosing a route [91]. After each simulation iteration, travel times are updated according to

$$\tau'_d(x) = \begin{cases} \tau_s(x) & \text{if route } x \text{ was simulated} \\ \beta \tau_r(x) + (1 - \beta) \tau_d(x) & \text{otherwise} \end{cases} \quad (5)$$

Herein, for a given route x , $\tau_s(x)$ is the simulated travel time, $\tau_r(x)$ is the estimated travel time from the simulation for other routes not used in the iteration, and β is a remembering factor for scaling the impact of costs of past unused routes. Then, the probabilities are updated according to

$$p'_d(r) = \frac{p_d(r)(p_d(r) + p_d(s)) \exp\left(\frac{\alpha \delta_{rs}}{1 - \delta_{rs}^2}\right)}{p_d(r) \exp\left(\frac{\alpha \delta_{rs}}{1 - \delta_{rs}^2}\right) + p_d(s)} \quad (6)$$

and

$$p'_d(s) = p_d(r) + p_d(s) - p_d(s) \quad (7)$$

where $p_d(x)$ and $p'_d(x)$ are the prior and new probabilities to use route x , r is the route used in the previous simulation iteration, s another route in the set of alternatives R , and δ_{rs} is the relative cost difference between routes r and s defined by

$$\delta_{rs} = \frac{\tau_d(s) - \tau_d(r)}{\tau_d(s) + \tau_d(r)} \in [-1, 1] \quad (8)$$

$\tau_d(s)$ represents the travel time for driver d to complete route x . In this scenario, SUMO's default values are used for Gawron's parameters, namely $\alpha = 0.5$ and $\beta = 0.9$.

3. Results Validation

It is not possible to formally prove the convergence of the proposed heuristic dynamic traffic assignment approaches as there is no analytical solution for the loading process in microsimulations [89,92]. Therefore, an extensive multi-variable evaluation over the 24 h of a generic weekday allows test empirically many of the different aspects of the goodness-of-fit between simulation and reality at different levels and scales. Based on the reviewed literature regarding traffic simulation evaluation, the chosen performance metrics can be classified according to their different levels of spatial and temporal aggregation (Figure 6). However, as already pointed out in previous research, the possibilities of evaluation are limited by the availability of real data serving as ground truth and different from the ones used for demand generation.

3.1. Large-Scale Aggregated Metrics: Teleports and Average Travel Time

Let us start by looking at the most aggregated metrics: "Teleporting" has no correspondence to a real measurement, but constitutes a first sanity check of the simulation, as a large proportion of vehicles teleporting means a poor model performance. In general,

iSAR shows a reasonable number of teleports (around 5%), with lower variability than the teleports resulting from the IIDUE method (see Figure 7a,b).

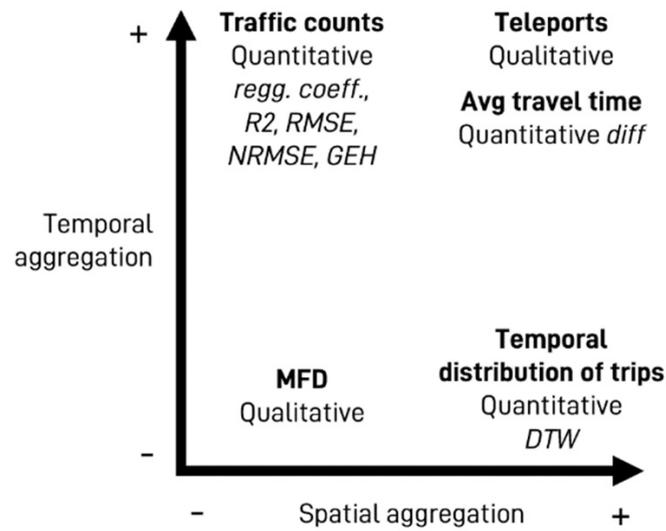


Figure 6. Classification of evaluation variables and metrics based on their spatio-temporal aggregation (based on the framework proposed in [93] (Figure 1)).

On the highest level of aggregation, the average travel times between simulation and reality are compared. Mobility surveys provide information regarding the average duration of trips for different subareas of the RMB and separated by mode of transport. The inner Metropolitan Area, which roughly corresponds to the core area of the simulation (Rondes), has an average travel time for private vehicles of 23.5 min (or 1410 s) [49]. This value is very similar to the results of the simulations based on the iSAR approach (around 1400 s); while the average travel times resulting from the IIDUE method deviate considerably (see Figure 7c).

3.2. Temporally Disaggregated Metrics: Hourly Distribution of Trips

With a higher level of temporal disaggregation, the hourly distribution of trips throughout the day is as well compared (see Figure 8). For the ground truth, mobility surveys provide 24-h temporal histograms of when trips begin, differentiated by RMB subareas and mode of transportation [49] (p. 59).

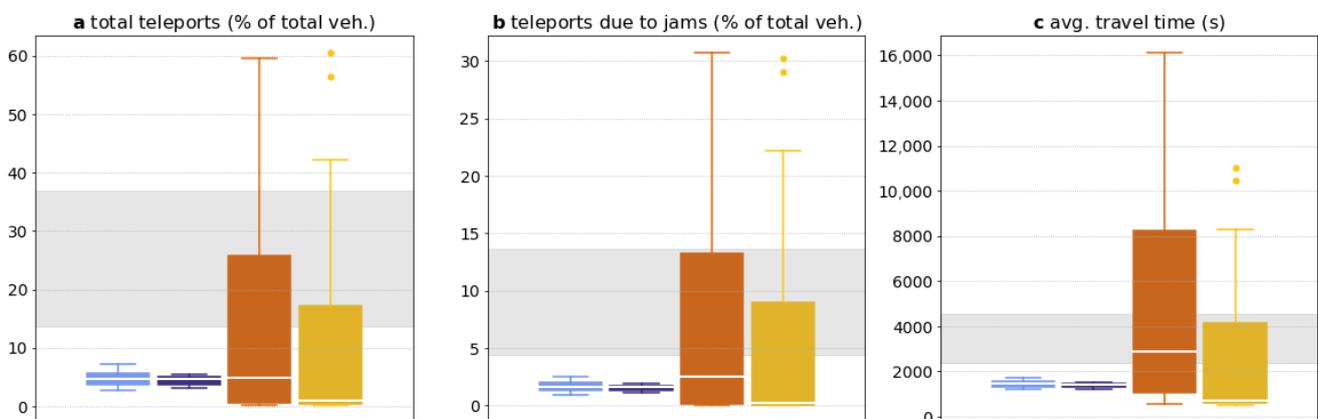


Figure 7. Cont.

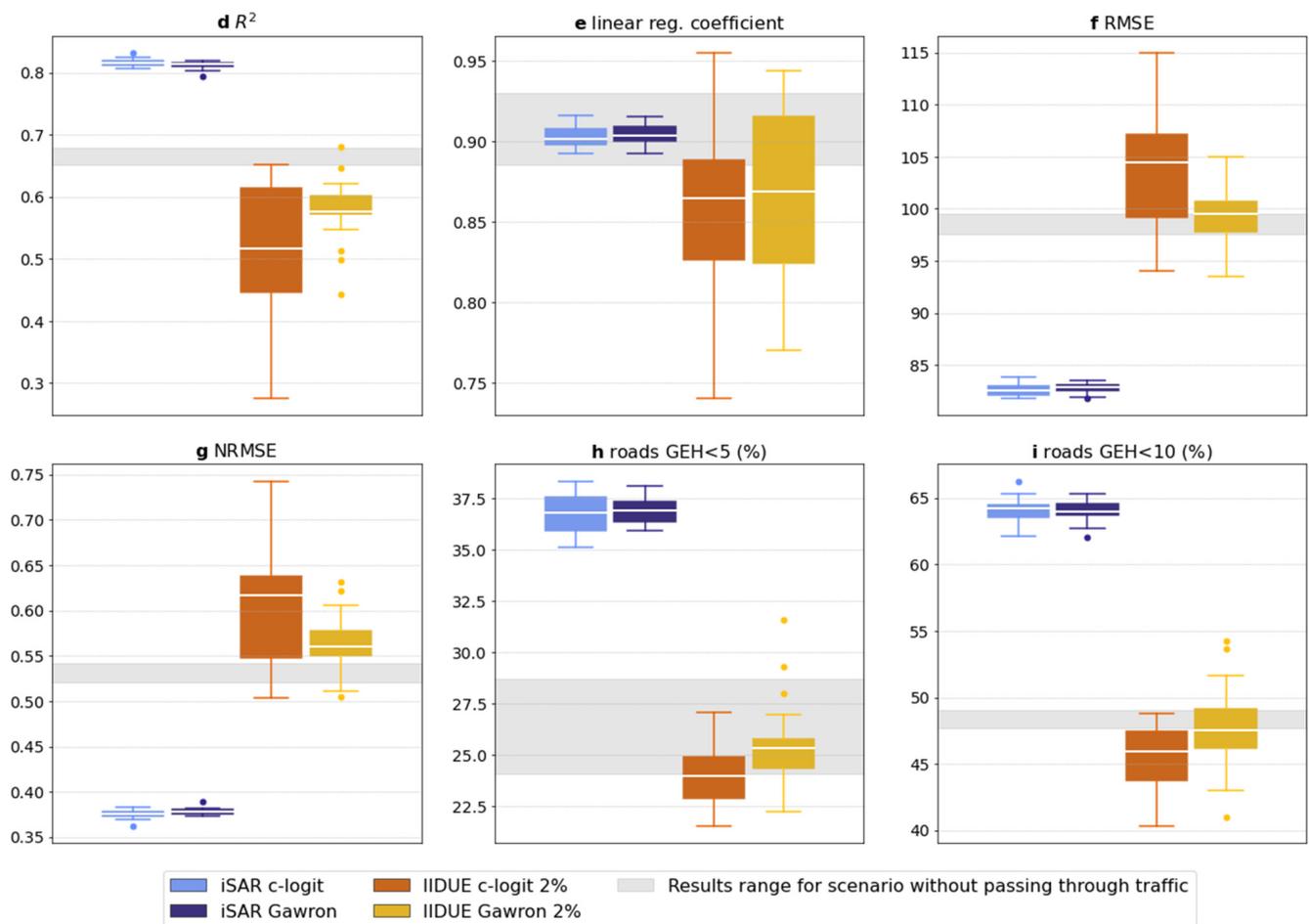


Figure 7. Performance of the analysed models (for the core simulation area, Rondes). iSAR outperforms IIDUE results: it shows more stable results, closer to measured average travel time, better goodness of fit, and lower error measures. Shaded ranges show the model's underperformance when not considering passing through traffic (i.e., crossing Rondes core area): (a) Total teleports expressed in percentage of total vehicles loaded; (b) Teleports caused by jams, expressed in percentage of total vehicles loaded; (c) Average travel times for all vehicles in seconds; (d) R^2 coefficient between simulated and real traffic counts; (e) regression coefficient between simulated and real traffic counts; (f) RMSE between simulated and real traffic counts; (g) NRMSE between simulated and real traffic counts; (h) percentage of monitored links with a GEH statistic below 5; (i) percentage of monitored links with a GEH statistic below 10.

When scenarios are close to congestion, programmed trips may not find enough space in the transport network to be inserted, such that they are delayed or eventually even skipped. This can potentially lead to different demand curves or an overall lower processed demand (see the uncalibrated simulated demand in Figure 7). Dynamic Time Warping (DTW) [94] with the *FastDTW* implementation [95] is used to compare real and simulated temporal distributions. This technique allows finding the similarities between two time series by non-linearly distorting the time axis, which becomes very useful when divergences are caused by different speeds, frequencies, and phasing as they may particularly occur in the case of network (over)saturation. Overall, there is a good match between simulation and ground truth, with low variability, showing clearly the observed morning and early evening peaks. The DTW distance value stays very stable across the successive iterations with a value around 19,000. However, in general, we can consistently observe a slightly lower simulated demand, particularly in peak hours, most likely due to high levels of

congestion that prevents some vehicles to be inserted in the network at the expected time, which are eventually skipped.

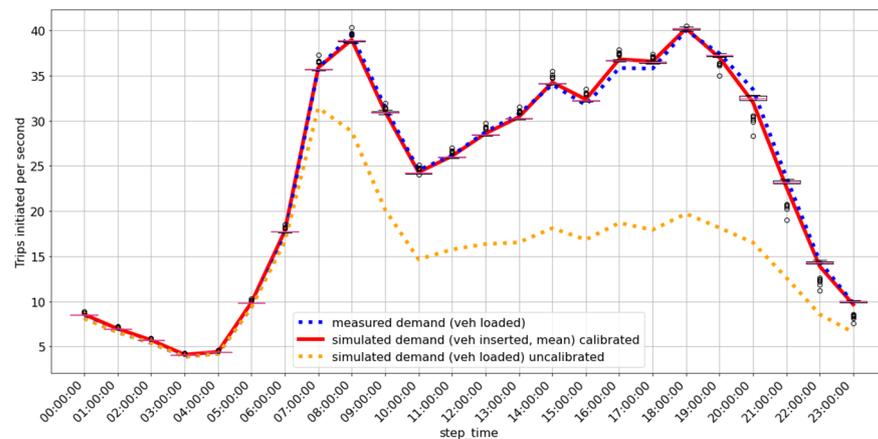


Figure 8. Temporal distribution of trips showing the measured demand in surveys (“ground truth”, dotted blue curve) as compared to the simulated demand (red solid line) after parameter calibration (Section 2.4) and adaptation (Section 2.5), with boxplots showing variations across iterations. The dotted orange line shows the results of the simulation without calibration.

3.3. Spatially Disaggregated Metrics: Traffic Counts

Traffic counts provide a spatially disaggregated evaluation. The city of Barcelona has 488 permanent traffic counting stations (see Figure 9a). However, only Monthly Average Daily Traffic (MADT) counts are publicly available [96]. Other metrics such as temporal distributions of trips for each detector [85] and average speeds [29,97], which have been used effectively for more detailed assessments in other scenarios, were not available.

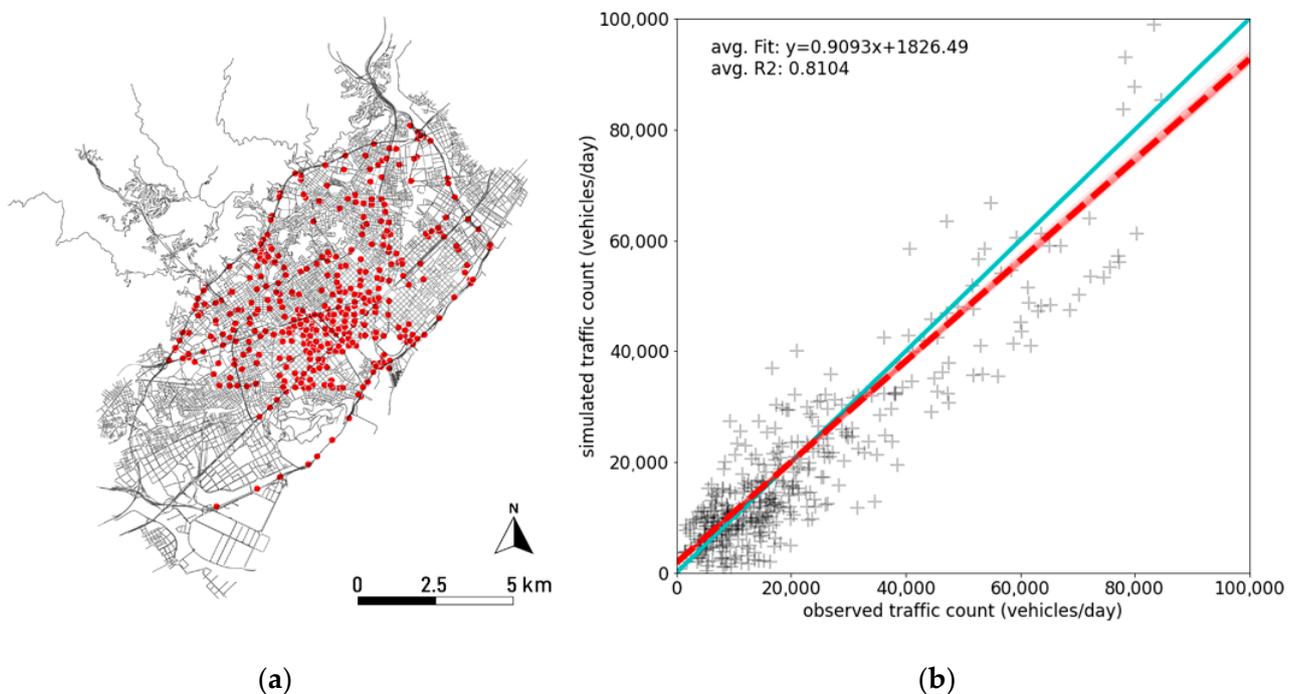


Figure 9. Traffic counts: (a) Location of the 488 AADT detectors managed by the city of Barcelona, which are used to assess the quality of the model calibration; (b) Regression plot comparing real and simulated traffic counts (averaged over all the iSAR simulations using the c-Logit and Gawron methods). $R^2 > 0.8$ is considered a good fit.

The large variability of observed monthly averages throughout the year in some of the control points (see Figure 10) creates a wide range of possible correct matchings with the ground truth, particularly as the timestamps of the cell phone tracks used for generating the traffic demand are unknown. To address this limitation, the monthly averages are aggregated to annual averages (AADT). Values within 1 standard deviation of this average are considered to be a good fit.

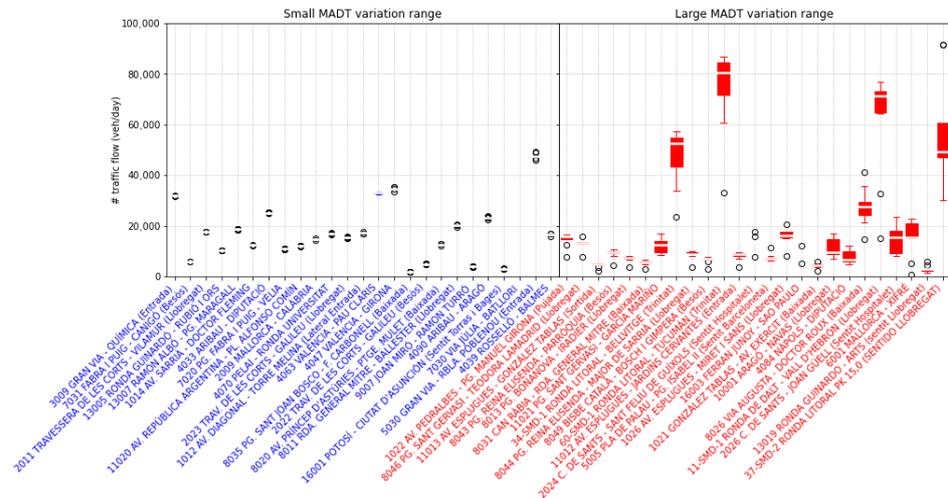


Figure 10. Measured traffic counts at selected locations showing Monthly Average Daily Traffic counts (MADT) ranges with the 25 control points where the normalized variability ($\text{normalized range}_{\text{MADT}} = \frac{\text{abs}(\text{max}_{\text{MADT}} - \text{min}_{\text{MADT}})}{\text{MADT}}$) is lowest (left, in blue) and highest (right, in red).

Linear regression is used to evaluate the general goodness-of-fit between simulated and real measurements of traffic counts (see Figure 9b). iSAR simulations results show a slope of 0.91 and an R^2 value of 0.81 (see Figure 7d,e), which can be considered a good match [98]. Again, the results from the IIDUE method reach worse values for both metrics.

A second set of metrics regarding traffic counts offers a more detailed evaluation. Among the many available model performance statistics reflecting different aspects of the discrepancies between predictions and ground truth, three of the most common ones across recent literature are selected [26,97] (see Table A1): the general Root Mean Square Error (RMSE), a Normalised Root Mean Square Error (NRMSE), and the Geoffrey E. Harvers statistic (GEH).

The RMSE and NRMSE are specified a

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^n (S_i - O_i)^2}{n}} \tag{9}$$

$$\text{NRMSE} = \frac{\sqrt{\frac{\sum_{i=0}^n (S_i - O_i)^2}{n-1}}}{\frac{\sum_{i=0}^n O_i}{n}} \tag{10}$$

where n is the number of observation points, and O_i and S_i are the real observed and simulated values of traffic counts at location i , respectively. RMSE is sensitive to scale, so it allows the comparison of both assignment methods for the two route choice methods as they use the same traffic demand. The iSAR method shows lower and more stable RMSE values, which is better than the IIDUE approach, with no relevant differences between the c -Logit and Gawron methods (see Figure 7f).

RMSE normalization allows for a more general and scale-independent (i.e., data-independent) comparison between models. The mean of the real observations is chosen as

the normalization criterion for the NRMSE approach. It reports values in the range [0,1], so multiplying it by 100 generates a percentage-like score (known as well as NRMSE or %RMSE). According to the established literature [98–100], any value below 0.3 (or 30%) can be considered good. In the case of the presented models, the iSAR simulations outperform again the results of the IIDUE method (see Figure 7g) but even in the best scenario, the values are above 0.3, so the model performance is limited.

The previously considered metrics (i.e., regression's coefficient, R^2 , RMSE, and NRMSE values), which are commonly used for the performance evaluation of (simulation and prediction) models in many fields, does not work particularly well for the case of complex transportation networks with highly variable network elements [26]. To overcome this problem, the Geoffrey E. Harvers statistic (GEH), which is a modified chi-squared metric, is frequently used in traffic studies. It leverages relative and absolute differences to compute percentage errors regarding the mean value of observed and simulated traffic counts. The GEH statistic is defined by

$$GEH = \sqrt{\frac{(S - O)^2}{(S + O)/2}} \quad (11)$$

where S is the simulated traffic count, and O is the real observed traffic count [101,102]. GEH values below 5 are considered a good fit, values between 5 and 10 are ok, but need further checks, and GEH values over 10 are a bad match. Ideally, 85% or more of the observed links should have a GEH value below 5. None of the resulting models reaches this score, which is consistent with the analysed previous large-scale urban microsimulations. However, iSAR performs better, having 38% of the analysed links with a GEH smaller than 5, where the differences between the Gawron and c -Logit approaches are negligible (see Figure 7h,i).

3.4. Macroscopic Metrics: MFD

Finally, macroscopic fundamental diagrams (MFD) [103] provide an additional perspective for the evaluation of the generated models. However, in contrast to the previous quantitative tests, this check is qualitative. The lack of empirical data and the discrepancies in estimations for real scenarios [104] are still a challenge for their use. Nonetheless, one can make an additional assessment with them. The obtained MFD (Figure 11) features the characteristic inverse- λ form [19], with capacity values compatible with measured capacities for different roadways typologies [105].

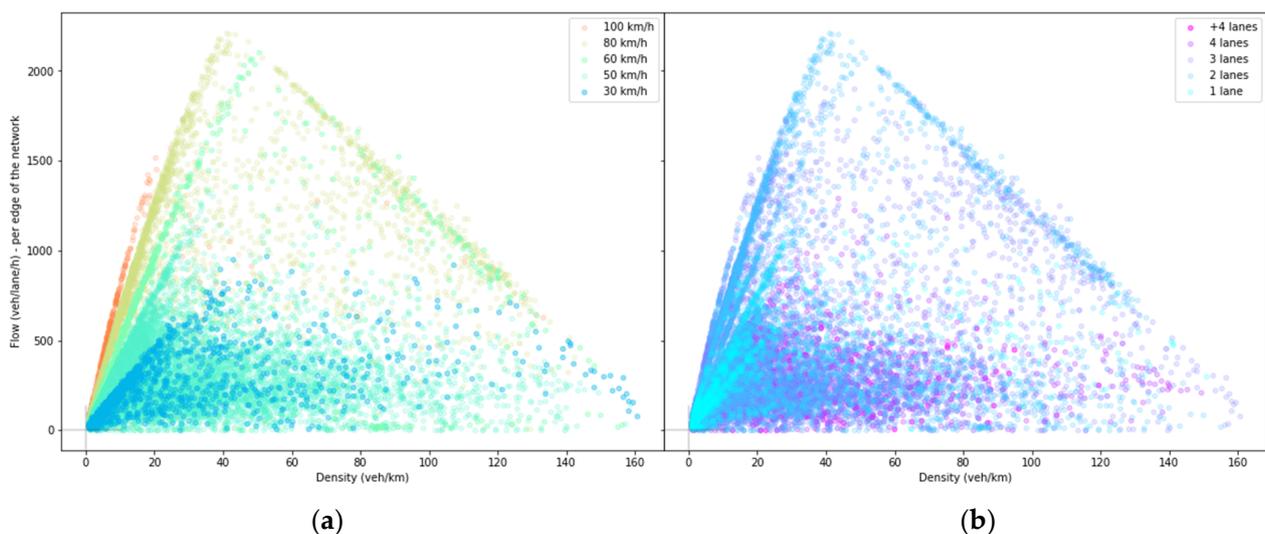


Figure 11. Obtained Macroscopic Fundamental Diagrams values with the microsimulation: (a) coloured based on the speed limit of the road section; (b) coloured based on the number of lanes of the road section.

4. Discussion and Outlook

In this paper, a new large-scale traffic microsimulation for the urban area of Barcelona using the software SUMO is presented. It combines novel and more precise origin-destination matrices estimated from anonymized mobile phone records with traditional mobility surveys on a network extracted from OpenStreetMap. The resulting model shows a good level of correspondence with reality (see Figure 12), outperforming previous examples, after being widely validated by metrics on different scales. It expands the state-of-the-art of traffic simulation by pushing microscopic approaches into large-scale scenarios supported by real data to obtain an operationally realistic and efficient model that can assess citywide impacts at fine-grained detail of policy decisions. Simultaneously, it allows identifying the limitations of these modelling approaches when considered within the concept of “urban digital twins”.

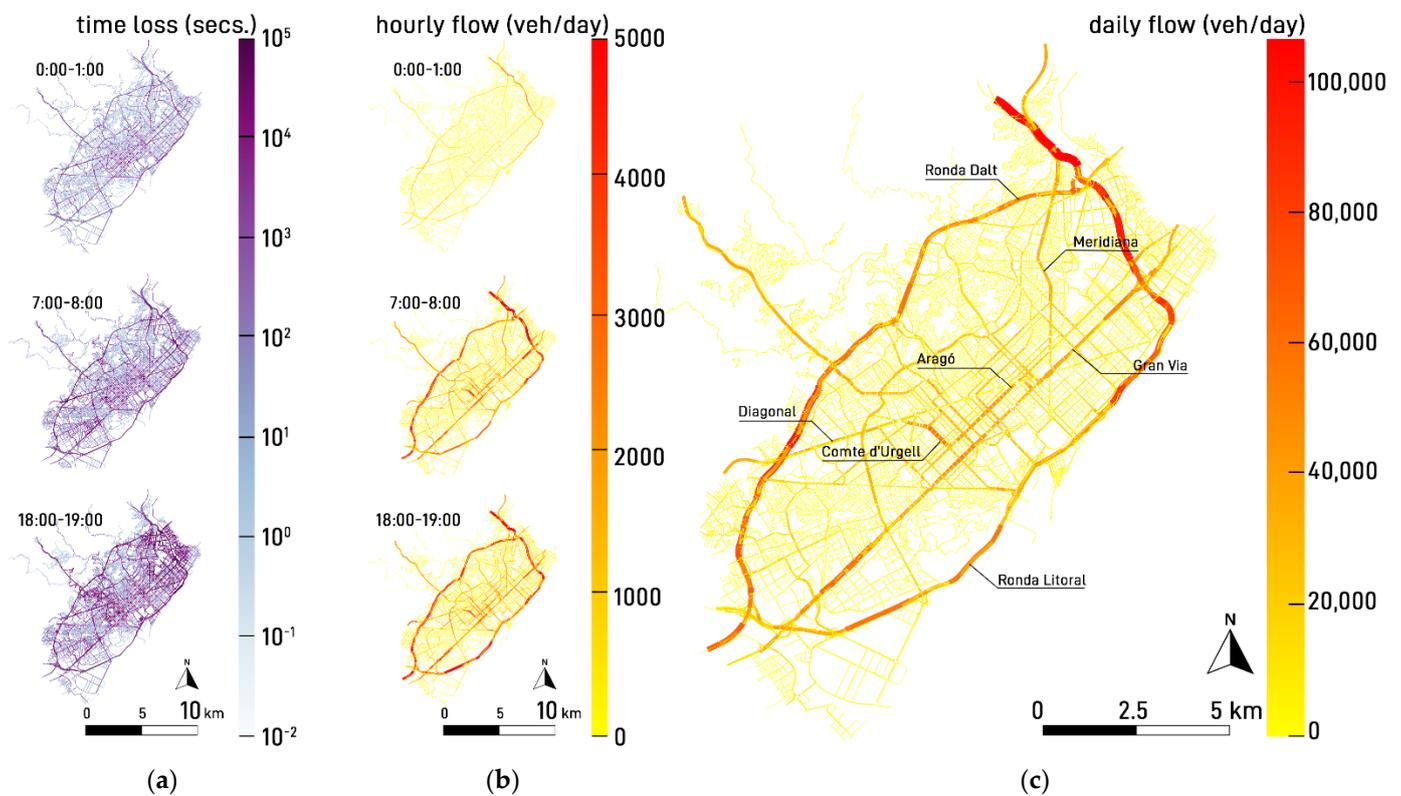


Figure 12. Simulation results: (a) Congestion level measured by aggregated loss of time in seconds in 3 relevant intervals (off-peak plus morning and evening peak hour); (b) Congestion level measured by aggregated traffic counts in 3 relevant intervals (off-peak plus morning and evening peak hour); (c) 24h traffic counts.

The calibration of the model identified the simulation time step, the probability for the vehicles to reroute during the simulation, and the influence of the edge priority as the most relevant factors without having to perform an exhaustive exploration of the huge parameter space, which often makes the calibration and use of these models impractical.

Two approaches (in-Simulation Adaptive Rerouting, iSAR, and Iterative Incremental Dynamic User Equilibrium, IIDUE), under two route choice algorithms (Gawron and *c*-Logit), were compared to tackle the challenge of distributing a high traffic demand in a dense and complex large urban network. The lack of analytical form in this microsimulation approach requires an extensive multi-variable empirical evaluation to check their performance comparing real and simulation data on different spatial and temporal scales. These different quantitative metrics included the average trip duration, the similarity of temporal distributions of trips using DTW, and the comparison of simulated and real traffic

counts based on the linear regression coefficient, R^2 , RMSE, NRMSE, and GEH statistics. Additionally, two qualitative tests based on simulation “teleports” and MFDs are included.

iSAR outperforms IIDUE in all the metrics without relevant differences between the c -Logit and Gawron’s approaches. Computationally, iSAR is more efficient than IIDUE in approximating the prospective equilibrium state. The best results obtained with the iSAR approach match the real measured average trip duration of about 23.5 min. The performance of these models in terms of fitting the traffic counts is also good, having values over 0.9 (and below 1.1) for the regression coefficient and an R^2 value above 0.8. However, error measures represented by the NRMSE and GEH statistics are less impressive, with values of 0.38 and 38%, respectively. This demonstrates that accurately reproducing trip distributions is challenging. In agreement with previous publications, the presented simulations also tend to underestimate the empirically measured flows, but considerably less so as compared to other studies. One possible explanation for the remaining problem is that microsimulations distribute larger volumes of traffic over more secondary streets, such that the count points on major roads are bypassed.

4.1. Contributions

The resulting overall goodness-of-fit matches outperform the relatively few existing validated continuous-in-time microsimulation studies published on large-scale urban traffic flow scenarios, even considering that the presented model checks a larger number of metrics than these previous scenarios. Moreover, comparatively, the geographical scope, size, and complexity of this model for Barcelona make this large scenario particularly challenging (Table 3). This research mastered this by a data processing pipeline and modelling approach, which allows the handling of large and complex microsimulation urban scenarios more easily (i.e., simulating a considerably extended spatial area makes it possible to reproduce through traffic more accurately than through direct simplification). Thereby, the proposed simulation approach got closer to reality. It could also handle saturated flows much better even in a large-scale, complex, and saturated setting.

Table 3. Comparison of performance between selected large-scale urban simulations quantitatively validated. In bold, proposed scenario in this paper.

Year/Ref.	Location/Scenario	Metrics					
		Area (Extended)	Net Length (Extended)	R^2	Regression Coeff.	NRMSE	# Roads GEH < 5 (GEH < 10)
2020 [85]	Ingolstadt (InTAS)	52 km ²	717.23 km	-	-	0.3343	-
2021 [51]	Bologna	50 km ² (3703 km ² , simplified)	(3316 km, simplified)	0.61	0.98		31% (59%)
2021 [29]	Barcelona, Barcelona Virtual Mobility Lab (Macroscopic)	332 km ² (~6000 km ²)	4767 km	0.77	0.88	0.35	-
2021 this paper	Barcelona	183 km² (3126 km²)	2506 km (24,153 km)	0.81	0.91	0.38	37% (64%)

From the perspective of the research on high fidelity modelling, this paper tackles the challenge of providing an operational and effective approach to build a large-scale traffic microsimulation from novel, more granular, empirical mobility data that is robustly validated with real observations at different spatio-temporal scales. There are few realistically validated large-scale urban microsimulations (see Table A1). It is shown how cell phone tracking data and a faster adaptive stochastic approach provide more realism. In this sense, the research focuses more on connecting innovative modelling approaches with real data

both for feeding and evaluation steps. As a simulation that gets as close as possible to the real system, it is a fundamental component for building a traffic digital twin. In fact, it matches the early definitions of digital twins [4,106,107] and it is closer to the concept of digital shadow [108]. Overall, simulation is a fundamental component of digital twins as it enables the bidirectional interaction between the physical environment and the virtual model [109,110].

4.2. Limitations

Despite the impressive body of previous work, it is concluded that the development of highly accurate and realistic models of city-related dynamics at a microscopic scale remains a key challenge. There is room for improvement to collect more granular data for modelling and validation. It is possible to add detail to the virtual environment. Scanning exhaustively the whole parameters space for the sake of calibration is unfeasible because of limited computational power. For example, the theoretical assumptions used to explain how people choose their routes and how they plan their trips seem to be insufficient to fully grasp the underlying interdependencies and complexity when we scale up to a large system. Thus, the one-dimensional utilitarian assumption that people tend to minimize travel times, which is behind the common route choice models, is not fully supported by empirical data [111–113]. This research shows that using a *weight.priority-factor* accounting for the hierarchy of roads affects routing and improves model performance. This suggests the existence of additional underlying factors influencing route choice, which are likely related to the observed challenges of realistically distributing traffic flow and avoiding underestimation of counts.

In more general terms, it is possible to pinpoint some existing challenges and limitations of digital twins [108]:

- Ambiguous, non-consensual definition.
- Non-existing functional full-scale examples of digital twins.
- Lack of common data models [114].
- Heterogeneous digital twins environments, data types, and sources [115].
- Potential use of Artificial Intelligence to improve digital twins performance and applications (by now it is only used at a small scale).
- Need to increase security, decentralization, and sharing capabilities among stakeholders [116], to avoid misuse, cyber threats, and privacy issues [117].
- Need of common ways of sharing data between devices, stakeholders, and environments.
- Need to foster distributed schemes to increase reliability, accuracy, and performance.

4.3. Future Research

Despite these limitations, this work has significantly contributed to the progress and realism of large-scale microsimulations, as shown above. Therefore, in line with others before this paper, this research supports the idea that microsimulations can be a powerful tool to inform researchers, urban planners, policymakers, and citizens, allowing them to explore alternative scenarios that could not be tested in real life. In this sense, simulations are valuable for their prospective power, introducing alternatives of how things could be more than just describing existing systems. However, this is only possible if we ensure their realism through developing robust and multi-variable evaluation procedures and if we analyse critically our modelling assumptions beyond equilibrium-seeking approaches.

It became clear that a “digital twin”, which reproduces the traffic of a city exactly, is still a long way to go. It will probably never be possible unless all degrees of freedom are eliminated from the system, which would not be desirable, however. In contrast to modelling the infrastructure in a city, which can be digitally represented with a high degree of accuracy, the traffic flow dynamics in a city is a “wicked” and complex problem. It is determined by non-linear dynamics and network interactions, self-organization effects (such as emergent traffic jams), and large variability. This implies a considerable level of uncertainty, which is partly due to randomness and non-linear feedbacks, but also due

to the human element. In fact, the latter has been often neglected in digital twins, even though it is omnipresent in cities, for example, through decision-making, social interactions and dynamics, culture, politics, and markets. The issue here is not just the lack of data. Important qualities relevant for humans, such as consciousness, dignity, creativity, freedom, meaning, values, and being social are hardly quantifiable, if at all [118].

The identified limitations suggest that it is necessary to expand the research on real-time data flow between the digital twin and the real system [110], particularly applied to cities [108] and involving human behaviour [114]. The following research directions would contribute not only to improving the results of simulations but also more broadly to the accomplishment of the concept of digital twins. They cover aspects of interaction (1), data (2–4), and collaboration (5–6):

1. The concept of the digital twin has evolved from simply mirroring as accurately as possible a physical system [4] to incorporating a bidirectional real-time interaction between the virtual and physical sides that affects and informs each other [109,110]. How can this virtual-physical interaction be implemented? [116].
2. Digital twins require large amounts of low-latency data [108]. How can more granular urban real-time data be collected while taking into account legal, ethical, and privacy issues?
3. Digital twins use massive amounts of data for getting as close as possible to the real system. However, stochasticity, perhaps amplified by complexity effects, together with the lack of connection between functional and physical processes to socio-economic systems, and even the difficulty to quantify many aspects of city life limit their predictability. Are more data providing more accurate digital mirrors? [3,117]. How can big data help to provide better human behavioural models?
4. To perform accurately, digital twins require to integrate simulation and optimization in real-time based on high-frequency and low-latency large amounts of data [108,119,120]. How can it be optimized? How scalable, efficient, and environmentally sustainable it can be?
5. Digital twins can facilitate the coordination of self-organized bottom processes and top-down governance enabling participation, self-determination, and democracy [117]. How can digital twins be used to enhance collaboration between different stakeholders, including hybrid settings (i.e., human-human, human-machine, and machine-machine)? How can the information be shared securely and effectively among stakeholders? [116].

5. Conclusions

While all of this matters when understanding and modelling city life, modelling traffic flows a less demanding and possible with a considerable degree of accuracy and realism. However, human nature still shines through. Precisely, the use of empirical big data regarding the movement of people and goods allows developing, testing, and validating new modelling approaches that eventually would lead to a better capturing of reality.

In general, a purely data-driven approach may need to be complemented by a data science and complexity perspective to provide better predictions. Although this paper focuses on traffic simulation, the same principles apply to the different interrelated realms in the built environment such as land use, climate, energy, or economic activity. The prospective and exploratory power of digital twins in combination with the inclusion of social and behavioural values can become a facilitator of governance and co-creation of cities, expanding human decision-making capabilities by using computation.

Funding: This research was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, grant number 833168.

Data Availability Statement: The raw mobility data based on cell phone records that support the findings of this study are available from the Generalitat de Catalunya. Restrictions apply to the

availability of these data, which were used under license for this study. Data are available from the authors with the permission of the Generalitat de Catalunya.

Acknowledgments: I acknowledge using map data from OpenStreetMap contributors and available from <https://www.openstreetmap.org> (accessed on 25 March 2021). I would like to thank Francesc Calvet (Generalitat de Catalunya) for providing the novel O/D matrices estimated from mobile phone records. Moreover, I want to thank Nino Antulov-Fantulin for suggestions on how to validate traffic counts, Dirk Helbing for providing feedback and recommendations particularly regarding validation and comparative scenarios, and suggesting the discussion of digital twins, and Sachit Mahajan for helping to address the final comments.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Table A1. Published large-scale urban traffic microsimulations characterized by size, demand creation and adaptation, traffic modes, and validation technique used which are reviewed.

Year/Ref.	Location/Net/Size	Demand			Validation
		Creation	Adaptation/Calibration	Modes	
2011 [36]	Cologne, OSM, 400 km ²	Activity generation from TAPAS surveys [121]. 1.2 million trips	Microscopic iterative optimization (Dynamic Use Equilibrium, DUE)	Car	Qualitative based on macroscopic measurements
2013 [33]	Ottawa, OSM, about 1 km ²	AADT and turning counts at intersections	Routing by turn probabilities	Car	The used demand adaptation algorithm tends to generate a traffic distribution that matches the input real data
2014 [34]	Vila Real, OSM, 7 km ²	Activity generation from synthetic pop. based on census and survey data. (24,023 people, 10,143 trips)	Microscopic direct assignment. (DUA)	Car, bus (public transport)	No validation
2015 [35]	Bologna, VISSIM. 2 neighbourhoods, about 2 km ²	AADT and turning counts at intersections	Direct microscopic assignment estimated from counting. (Dynamic Use Assignment, DUA)	Car, bus	Quantitative comparison of traffic flow between simulated and measured. Only limited to the morning peak hour
2017 [122] (LuST)	Luxemburg, OSM, 156 km ² , 930 km of roads.	Activity generation from a synthetic population	Microscopic direct assignment (DUA) & iterative optimization (DUE) based on time loss, rerouting, and speed metrics	Car, bus	Quantitative comparison based on speed distribution in selected roadways
2018 [37] (MoST)	Monaco, OSM, 74 km ² , 587 km of roads, with topography data.	Activity generation from a synthetic population	Microscopic direct assignment (DUA)	Car, bikes, pedestrians, public transit	Qualitative validation only for the morning peak hour
2019 [38] (TuST)	Turin, OSM, 500 km ² , 6500 km of roads	O/D matrices from real observational data	Macroscopic iterative assignment	Car	Qualitative comparison with total vehicles running and expected average length of trips

Table A1. Cont.

Year/Ref.	Location/Net/Size	Demand			Validation
		Creation	Adaptation/Calibration	Modes	
2019 [39]	Berlin, around 4500 km ²	O/D matrices from TAPAS surveys [120]	First, faster macroscopic iterative assignment, followed by slower and more detailed microscopic iterative assignment (DUE)	Car	Not specified. Only metrics regarding the speeding up of the assignment process
2020 [85] (InTAS)	Ingolstadt, OSM, 52 km ² , 717.23 km of roads	Activity generation from a synthetic pop. (109.090 people)	Microscopic iterative optimization (DUE), stochastic rerouting in simulation	Car	Quantitative: absolute error and NRMSE on times series of total trips in the city and AADT counts
2020 [40]	Ireland. Urban, rural & motorway. (Dublin city centre, 17.5 km ²)	AADT counting-based demand	Direct microscopic assignment (DUA) with stochastic rerouting	Car (with autonomous vehicles)	Qualitative and limited to short periods
2020 [50]	Zurich, OSM simplified net based on centrality metrics. (~900 km ²)	O/D matrices	Surrogate iterative models to optimize global mesoscopic simulation parameters	Car	Quantitative validation based on DTW and RMSE between simulated and measured MFD
2020 [41]	Porto. SHP file. 42 km ²	O/D matrices	Direct microscopic assignment (DUA), iterative assignment (DUE), and incremental assignment	Car	Qualitative validation based on correlation coefficient, and RMSE only applied to 6 edges in two periods of the day
2021 [51]	Bologna, OSM, 50 km ² detailed core area, 3703 km ² extended urban area with simplified network	Activity-based, disaggregation of OD matrices, GPS traces, GTFS	Mode choice, and iterative optimization (DUE)	Car, bus, motorcycle, bike, pedestrians	Quantitative validation based on comparing AADT between simulation and reality (regression Coeff., R ² , GEH). Only morning peak hour (7–8 am)
2021 this paper	Barcelona, OSM, 182.55 km ² detailed core area with 2506 km of roads, 3126 km ² extended urban area for whole demand creation with 24,153 km of roads.	O/D matrices from mobile phone records + aggregated data from mobility surveys. 3,185,285 trips for the whole urban area cropped to 2,063,177 trips for the core simulated area	Initial direct microscopic assignment (DUA), then refined by stochastic rerouting and iterative incremental assignment (DUE)	Car	Quantitative validation on comparing simulation vs. reality of AADT counts (regression Coeff., R ² , RNMSE, NRMSE, GEH), temporal distributions of trips (DTW), and macroscopic diagrams

References

1. Portugali, J. *Self-Organization and the City*; Springer: Berlin/Heidelberg, Germany, 2000; ISBN 9783642084812.
2. Gershenson, C. Living in Living Cities. *Artif. Life* **2013**, *19*, 401–420. [CrossRef]
3. Batty, M. Digital twins. *Environ. Plan. B Urban Anal. City Sci.* **2018**, *45*, 817–820. [CrossRef]
4. Grieves, M. Digital Twin: Manufacturing Excellence through Virtual Factory Replication. *White Pap.* **2014**, *1*, 1–7.
5. White, G.; Zink, A.; Codecá, L.; Clarke, S. A digital twin smart city for citizen feedback. *Cities* **2021**, *110*, 103064. [CrossRef]

6. Schrotter, G.; Hürzeler, C. The Digital Twin of the City of Zurich for Urban Planning. *PFG—J. Photogramm. Remote Sens. Geoinf. Sci.* **2020**, *88*, 99–112. [[CrossRef](#)]
7. Ayuntamiento de Madrid. Madrid Actualiza su Simulador de Movilidad e Incluye Todos los Sistemas de Transporte para Reducir las Afecciones al Tráfico. In *Diario de Madrid*; Ayuntamiento de Madrid: Madrid, 2021. Available online: <https://diario.madrid.es/blog/notas-de-prensa/madrid-actualiza-su-simulador-de-movilidad-e-incluye-todos-los-sistemas-de-transporte-para-reducir-las-afecciones-al-trafico/> (accessed on 17 March 2021).
8. Government of Singapore. *Virtual Singapore*; Digital Twin, Computer Software; Government of Singapore: Singapore, 2018.
9. European Commission. Rotterdam’s Digital Twin Redefines Our Physical, Digital, & Social Worlds. Available online: <https://smart-cities-marketplace.ec.europa.eu/news-and-events/news/2019/rotterdams-digital-twin-redefines-our-physical-digital-social-worlds> (accessed on 4 August 2021).
10. Airaksinen, E.; Bergstrom, M.; Heinonen, H.; Kaisla, K.; Lahti, K.; Suomisto, J. *The Kalasatama Digital Twins Project*; City of Helsinki: Helsinki, Finland, 2019.
11. Hourdos, J.; Michalopoulos, P. *Twin Cities Metro-wide Traffic Micro-Simulation Feasibility Investigation*; Minnesota Department of Transportation: St. Paul, MN, USA, 2008.
12. Lowry, I.S. *A Model of Metropolis*; RAND Corporation: Santa Monica, CA, USA, 1964.
13. Forrester, J.W. *Urban Dynamics*; Productivity Press: Portland, OR, USA, 1969; ISBN 1883823390.
14. Acheampong, R.A.; Silva, E.A. Land use–transport interaction modeling: A review of the literature and future research directions. *J. Transp. Land Use* **2015**, *8*, 11–38. [[CrossRef](#)]
15. Wegener, M. Land-use transport interaction models. In *Handbook of Regional Science*; Fischer, M.M., Nijkamp, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 741–758; ISBN 9783642234309.
16. Tomko, M.; Winter, S. Beyond digital twins—A commentary. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 395–399. [[CrossRef](#)]
17. Batty, M. A map is not the territory, or is it? *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 599–602. [[CrossRef](#)]
18. Wilfire, C. How Can We Spearhead City’scale Digital Twins? *Infrastructure Intelligence*. 9 May 2018. Available online: <http://www.infrastructure-intelligence.com/article/may-2018/how-can-we-spearhead-city-scale-digital-twins> (accessed on 3 March 2021).
19. Treiber, M.; Kesting, A. *Traffic Flow Dynamics: Data, Models and Simulation*; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 9783642324604.
20. Mathias, J.D.; Debeljak, M.; Deffuant, G.; Diemer, A.; Dierickx, F.; Donges, J.F.; Gladkykh, G.; Heitzig, J.; Holtz, G.; Obergassel, W.; et al. Grounding Social Foundations for Integrated Assessment Models of Climate Change. *Earth’s Futur.* **2020**, *8*, e2020EF001573. [[CrossRef](#)]
21. Lopez, P.A.; Behrisch, M.; Bieker-Walz, L.; Erdmann, J.; Flotterod, Y.P.; Hilbrich, R.; Lucken, L.; Rummel, J.; Wagner, P.; Wiebner, E. Microscopic Traffic Simulation using SUMO. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2011; pp. 2575–2582.
22. Helbing, D. *Quantitative Sociodynamics: Stochastic Methods and Models of Social Interaction Processes*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2010; ISBN 9783642115455.
23. Heppenstall, A.J.; Crooks, A.T.; See, L.M.; Batty, M. *Agent-Based Models of Geographical Systems*; Springer: Berlin/Heidelberg, Germany, 2012; ISBN 9789048189274.
24. Kastner, K.H.; Keber, R.; Pau, P.; Samal, M. Real-Time Traffic Conditions with SUMO for ITS Austria West. In *Proceedings of the Simulation of Urban Mobility, SUMO 2013 Conference, Berlin, Germany, 15–17 May 2013*; Lecture Notes in Computer Science; Behrisch, M., Krajzewicz, D., Weber, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8594, pp. 146–159; ISBN 9783662450796. [[CrossRef](#)]
25. Urquiza-Aguilar, L.F.; Bautista, P.B.; Gómez, W.C.; Calderón, X. Comparison of traffic demand generation tools in SUMO. Case study: Access highways to Quito. In Proceedings of the PE-WASUN 2019—16th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, New York, NY, USA, 25 November 2019; Volume 19, pp. 15–22.
26. Balakrishna, R.; Antoniou, C.; Ben-Akiva, M.; Koutsopoulos, H.N.; Wen, Y. Calibration of microscopic traffic simulation models: Methods and application. *Transp. Res. Rec.* **2007**, 198–207. [[CrossRef](#)]
27. Oketch, T.; Carrick, M. Calibration and validation of a micro-simulation model in network analysis. In Proceedings of the 84th TRB Annual Meeting, Washington, DC, USA, 9–13 January 2005.
28. Bartin, B.; Ozbay, K.; Gao, J.; Kurkcu, A. Calibration and validation of large-scale traffic simulation networks: A case study. *Procedia Comput. Sci.* **2018**, *130*, 844–849. [[CrossRef](#)]
29. Rodríguez-Rey, D.; Guevara, M.; Linares, M.P.; Casanovas, J.; Salmerón, J.; Soret, A.; Jorba, O.; Tena, C.; Pérez García-Pando, C. A coupled macroscopic traffic and pollutant emission modelling system for Barcelona. *Transp. Res. Part D Transp. Environ.* **2021**, *92*, 102725. [[CrossRef](#)]
30. Bassolas, A.; Ramasco, J.J.; Herranz, R.; Cantú-Ros, O.G. Mobile phone records to feed activity-based travel demand models: MATSim for studying a cordon toll policy in Barcelona. *Transp. Res. Part A Policy Pract.* **2019**, *121*, 56–74. [[CrossRef](#)]
31. Saidallah, M.; El Fergougui, A.; Elalaoui, A.E. A comparative study of urban road traffic simulators. In *Proceedings of the MATEC Web of Conferences, ICTTE 2016 2016 5th International Conference on Transportation and Traffic Engineering, Lucerne, Switzerland, 6–10 July 2016*; Figueira, M., Guo, Z., Eds.; EDP Sciences: Les Ulis, France, 2016; Volume 81, p. 05002.

32. Hardy, M.; Dodge, L.; Smith, T.; Váscónez, K.C.; Wunderlich, K. Evacuation management operations modeling assessment: Transportation modeling inventory. In Proceedings of the 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting, Washington, DC, USA, 16–20 November 2008; Volume 4, pp. 2648–2664.
33. McKenney, D.; White, T. Distributed and adaptive traffic signal control within a realistic traffic simulation. *Eng. Appl. Artif. Intell.* **2013**, *26*, 574–583. [[CrossRef](#)]
34. Soares, J.; Lobo, C.; Vale, Z.; De Moura Oliveira, P.B. Realistic traffic scenarios using a census methodology: Vila real case study. In Proceedings of the IEEE Power and Energy Society General Meeting, National Harbor, MD, USA, 27–31 July 2014.
35. Bieker, L.; Krajzewicz, D.; Morra, A.P.; Michelacci, C.; Cartolano, F. Traffic simulation for all: A real world traffic scenario from the city of Bologna. In *Proceedings of the 2nd SUMO Conference 2014, Berlin, Germany, 15–16 May 2014*; Modeling Mobility with Open Data. Lecture Notes in Mobility; Behrisch, M., Weber, M., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 13, pp. 47–60.
36. Uppoor, S.; Fiore, M. Large-scale urban vehicular mobility for networking research. *IEEE Veh. Netw. Conf. VNC* **2011**, 62–69. [[CrossRef](#)]
37. Codecá, L.; Härri, J. Monaco SUMO Traffic (MoST) Scenario: A 3D Mobility Scenario for Cooperative ITS. *Epic Ser. Eng.* **2018**, *2*, 43–55. [[CrossRef](#)]
38. Rapelli, M.; Casetti, C.; Gagliardi, G. TuST: From Raw Data to Vehicular Traffic Simulation in Turin. In Proceedings of the 2019 IEEE/ACM 23rd International Symposium on Distributed Simulation and Real Time Applications (DS-RT), Cosenza, Italy, 7–9 October 2019. [[CrossRef](#)]
39. Behrisch, M. *Connecting Macroscopic and Microscopic Traffic Assignment*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 47–54.
40. Guériau, M.; Dusparic, I. Quantifying the impact of connected and autonomous vehicles on traffic efficiency and safety in mixed traffic. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITS, Rhodes, Greece, 20–23 September 2020.
41. Pereira Machado, D. Microscopic Traffic Assignment under Dynamic Demand. Master's Thesis, Faculdade de Engenharia, Universidade do Porto, Porto, Portugal, 2020.
42. Diallo, A.O.; Lozenguez, G.; Doniec, A.; Mandiau, R. Comparative evaluation of road traffic simulators based on modeler's specifications: An application to intermodal mobility behaviors. In Proceedings of the ICAART 2021—13th International Conference on Agents and Artificial Intelligence, Online, 4–6 February 2021; Volume 1, pp. 265–272.
43. Urquiza-Aguiar, L.; Coloma-Gómez, W.; Barbecho Bautista, P.; Calderón-Hinojosa, X. Comparison of SUMO's vehicular demand generators in vehicular communications via graph-theory metrics. *Ad Hoc Netw.* **2020**, *106*, 102217. [[CrossRef](#)]
44. Gawron, C. An Iterative Algorithm to Determine the Dynamic User Equilibrium in a Traffic Simulation Model. *Int. J. Mod. Phys. C* **1998**, *9*, 393–407. [[CrossRef](#)]
45. German Aerospace Center (DLR) and Others Dynamic User Assignment—SUMO Documentation. Available online: https://sumo.dlr.de/docs/Demand/Dynamic_User_Assignment.html (accessed on 2 May 2021).
46. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numer. Math.* **1959**, *1*, 269–271. [[CrossRef](#)]
47. Hart, P.E.; Nilsson, N.J.; Raphael, B. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Trans. Syst. Sci. Cybern.* **1968**, *4*, 100–107. [[CrossRef](#)]
48. Geisberger, R.; Sanders, P.; Schultes, D.; Delling, D. Contraction hierarchies: Faster and simpler hierarchical routing in road networks. In Proceedings of the 7th International Workshop on Experimental Algorithms (WEA 2008), Provincetown, MA, USA, 30 May–2 June 2008; Lecture notes in Computer Science. McGeoch, C.C., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5038, pp. 319–333. [[CrossRef](#)]
49. Institut d'Estudis Regionals i Metropolitans de Barcelona (IERMB). Enquesta de Mobilitat en Dia Feiner 2019 (EMEF 2019). In *La mobilitat a l'àmbit del Sistema Integrat de Mobilitat Metropolitana de Barcelona*; IERMB: Bellaterra, Spain, 2020. Available online: https://web.archive.org/web/20210922230839/https://observatorimobilitat.atm.cat/docs-observatori/enquestes-de-mobilitat/Enquestes_ambit_ATM/EMEF/2019/EMEF_2019_Informe_SIMMB.pdf (accessed on 23 September 2021).
50. Tilg, G.; Bogenberger, K.; Ambühl, L.; Menéndez, M. Generating and calibrating large-scale, mesoscopic SUMO network. In Proceedings of the SUMO User Conference, Online, 26–28 October 2020.
51. Schweizer, J.; Poliziani, C.; Rupi, F.; Morgano, D.; Magi, M. Building a Large-Scale Micro-Simulation Transport Scenario Using Big Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 165. [[CrossRef](#)]
52. Haklay, M.; Weber, P. OpenStreet map: User-generated street maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [[CrossRef](#)]
53. Haklay, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environ. Plan. B Plan. Des.* **2010**, *37*, 682–703. [[CrossRef](#)]
54. Anderson, J.; Sarkar, D.; Palen, L. Corporate editors in the evolving landscape of OpenStreetMap. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 232. [[CrossRef](#)]
55. Boeing, G. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* **2017**, *65*, 126–139. [[CrossRef](#)]
56. OpenStreetMap Contributors. Planet Dump for the Metropolitan Area of Barcelona (RMB). Available online: <https://planet.osm.org> (accessed on 25 March 2021).

57. OpenStreetMap Contributors. Wiki OpenStreetMap Overpass API/Overpass QL [Computer Software]—OpenStreetMap Wiki. 2021. Available online: https://wiki.openstreetmap.org/w/index.php?title=Overpass_API/Overpass_QL&oldid=2180487 (accessed on 3 August 2021).
58. German Aerospace Center (DLR) and Others Netconvert—SUMO Documentation. Available online: <https://sumo.dlr.de/docs/netconvert.html> (accessed on 5 August 2021).
59. Dingil, A.E.; Schweizer, J.; Rupi, F.; Stasiskiene, Z. Road network extraction with OSMNx and SUMOPy. *EPiC Ser. Eng.* **2018**, *2*, 111–117.
60. Maxar. Maxar (c) Premium Imagery for the Area of Barcelona. Background Layer Available for OpenStreetMap, 2021. Available online: <https://www.openstreetmap.org/edit> (accessed on 25 March 2021).
61. Instituto Geográfico Nacional de España (IGN); Centro Nacional de Información Geográfica (CNIG). Plan Nacional de Ortofotografía Aérea (PNOA); Background Layer Available for OpenStreetMap, 2018. Available online: <https://www.openstreetmap.org/edit> (accessed on 25 March 2021).
62. Esri. Esri (c) World Imagery for the Area of Barcelona. Sources: Esri; DigitalGlobe; GeoEye; I-cubed; USDA FSA; USGS; AEX; Getmapping; Aerogrid; IGN; IGP, swisstopo, and the GIS User Community. Background Layer Available for OpenStreetMap, 2021. Available online: <https://www.openstreetmap.org/edit> (accessed on 25 March 2021).
63. Mapbox. Mapbox (c) Satellite Imagery for the Area of Barcelona. Sources: NASA; USGS; Other Commercial Providers. Background Layer Available for OpenStreetMap, 2021. Available online: <https://www.openstreetmap.org/edit> (accessed on 25 March 2021).
64. Grab Holdings. OpenStreetCam (KartaView) for the Area of Barcelona. Photo Overlay Data Available for OpenStreetMap, 2021. Available online: <https://www.openstreetmap.org/edit> (accessed on 25 March 2021).
65. Mapillary and Contributors. Mapillary (c) Imagery for the Area of Barcelona. Photo Overlay Data Available for OpenStreetMap, 2021. Available online: <https://www.openstreetmap.org/edit> (accessed on 25 March 2021).
66. German Aerospace Center (DLR) and Others Netedit—SUMO Documentation. Available online: <https://sumo.dlr.de/docs/Netedit/index.html> (accessed on 5 August 2021).
67. Koonce, P.; Rodegerdts, L. *Traffic Signal Timing Manual*; U.S. Department of Transportation: Washington, DC, USA, 2008.
68. Calvet, F.; Torre, M. *Obtenció De Matrius Origen-Destinació (OD) a Partir de Dades de Telefonía Mòbil. M2019*; Generalitat de Catalunya: Barcelona, Spain, 2020.
69. Cáceres, N.; Wideberg, J.P.; Benitez, F.G. Deriving origin-destination data from a mobile phone network. *IET Intell. Transp. Syst.* **2007**, *1*, 15–26. [[CrossRef](#)]
70. Cáceres, N.; Romero, L.M.; Benitez, F.G. Exploring strengths and weaknesses of mobility inference from mobile phone data vs. travel surveys. *Transp. A Transp. Sci.* **2020**, *16*, 574–601. [[CrossRef](#)]
71. Institut d'Estudis Regionals i Metropolitans de Barcelona (IERMB). *Enquesta de Mobilitat en Dia Feiner 2019 (EMEF 2019). La mobilitat a Barcelona*; IERMB: Bellaterra, Spain, 2020. Available online: https://web.archive.org/web/20210923095118/https://observatorimobilitat.atm.cat/docs-observatori/enquestes-de-mobilitat/Enquestes_ambit_ATM/EMEF/2019/EMEF_2019_Informe_BCN.pdf (accessed on 23 September 2021).
72. German Aerospace Center (DLR) and Others District—SUMO Documentation. Available online: <https://sumo.dlr.de/docs/Tools/District.html> (accessed on 5 August 2021).
73. Montero, L.; Ros-Roca, X.; Herranz, R.; Barceló, J. Fusing mobile phone data with other data sources to generate input OD matrices for transport models. *Transp. Res. Procedia* **2019**, *37*, 417–424. [[CrossRef](#)]
74. Melakessou, F.; Derrmann, T.; Engel, T. Asymmetry analysis of inbound/outbound car traffic load distribution in Luxembourg. In *Proceedings of the MobiWac 2015—13th ACM International Symposium on Mobility Management and Wireless Access*, Cancun, Mexico, 2–6 November 2015; pp. 5–12.
75. German Aerospace Center (DLR) and Others Od2trips—SUMO Documentation. Available online: <https://sumo.dlr.de/docs/od2trips.html> (accessed on 6 August 2021).
76. German Aerospace Center (DLR) and Others Duarouter—SUMO Documentation. Available online: <https://sumo.dlr.de/docs/duarouter.html> (accessed on 6 August 2021).
77. German Aerospace Center (DLR) and Others Routes—SUMO Documentation. Available online: <https://sumo.dlr.de/docs/Tools/Routes.html> (accessed on 6 August 2021).
78. Ciuffo, B.; Punzo, V.; Torrieri, V. Comparison of simulation-based and model-based calibrations of traffic-flow microsimulation models. *Transp. Res. Rec.* **2008**, 36–44. [[CrossRef](#)]
79. Vavasis, S.A. *Nonlinear Optimization: Complexity Issues*; The International Series of Monographs on Computer Science 8; Oxford University Press: New York, NY, USA, 1991; ISBN 0195072081.
80. Lieberman, E.; Rathi, A.K. *Traffic Simulation*. In *Traffic Flow Theory*; US Federal Highway Administration: Washington, DC, USA, 1996.
81. Krauss, S.; Wagner, P.; Gawron, C. Metastable states in a microscopic model of traffic flow. *Phys. Rev. E—Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **1997**, *55*, 5597–5602. [[CrossRef](#)]
82. Erdmann, J. SUMO's Lane-changing model. In *Proceedings of the 2nd SUMO Conference 2014, Berlin, Germany, 15–16 May 2014; Modeling Mobility with Open Data. Lecture Notes in Mobility*; Behrisch, M., Weber, M., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 13, pp. 105–123.

83. Mock, P.; Díaz, S. (Eds.) European vehicle market statistics. In *Pocketbook 2020/21*; International Council on Clean Transportation Europe: Berlin, Germany, 2020.
84. Marquet, O.; Miralles-Guasch, C. City of Motorcycles. On how objective and subjective factors are behind the rise of two-wheeled mobility in Barcelona. *Transp. Policy* **2016**, *52*, 37–45. [[CrossRef](#)]
85. Lobo, S.C.; Neumeier, S.; Fernandez, E.M.G.; Facchi, C. InTAS—The Ingolstadt traffic scenario for SUMO. *arXiv* **2020**, arXiv:2011.11995.
86. German Aerospace Center (DLR) and Others Why Vehicles Are Teleporting—SUMO Documentation. Available online: https://sumo.dlr.de/docs/Simulation/Why_Vehicles_are_teleporting.html (accessed on 6 August 2021).
87. Wardrop, J.G. Some theoretical aspects of road traffic. *Proc. Inst. Eng.* **1952**, *1*, 325–362.
88. Daganzo, C.F.; Sheffi, Y. On Stochastic Models of Traffic Assignment. *Transp. Sci.* **1977**, *11*, 253–274. [[CrossRef](#)]
89. Barceló, J.; Casas, J. Stochastic Heuristic Dynamic Assignment Based on AIMSUN Microscopic Traffic Simulator. *Transp. Res. Rec. J. Transp. Res. Board* **2006**, *1964*, 70–80. [[CrossRef](#)]
90. Cascetta, E.; Nuzzolo, A.; Russo, F.; Vitetta, A. A modified Logit Route Choice Model Overcoming Path Overlapping Problems. Specification and Some Calibration Results. In Proceedings of the Transportation & Traffic Theory: 13th International Symposium on Transportation & Traffic Theory, Lyon, France, 24–26 July 1996.
91. Gawron, C. Simulation-Based Traffic Assignment: Computing User Equilibria in Large Street Networks. Ph.D. Thesis, Mathematisch-Naturwissenschaftlichen Fakultät, Universität zu Köln, Cologne, Germany, 1999.
92. Batty, M. *The New Science of Cities*; MIT Press: Cambridge, MA, USA; London, UK, 2013; ISBN 9780262019521.
93. Kang, J.Y.; Aldstadt, J. Using multiple scale spatio-temporal patterns for validating spatially explicit agent-based models. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 193–213. [[CrossRef](#)]
94. Bellman, R.; Kalaba, R. On adaptive control processes. *IRE Trans. Autom. Control* **1958**, *4*, 1–9. [[CrossRef](#)]
95. Salvador, S.; Chan, P. Toward accurate dynamic time warping in linear time and space. *Intell. Data Anal.* **2007**, *11*, 561–580. [[CrossRef](#)]
96. Ajuntament de Barcelona Mobility Gauging Detail of the City of Barcelona—Datasets. Available online: <https://opendata-ajuntament.barcelona.cat/data/en/dataset/aforaments-detall> (accessed on 16 May 2021).
97. Yu, M.; Fan, W. Calibration of microscopic traffic simulation models using metaheuristic algorithms. *Int. J. Transp. Sci. Technol.* **2017**, *6*, 63–77. [[CrossRef](#)]
98. Ministerio de Fomento. *NS 5/2.014. Prescripciones Y Recomendaciones Técnicas Para La Realización De Estudios De Tráfico De Los Estudios Informativos, Anteproyectos Y Proyectos De Carreteras*; Ministerio de Fomento: Madrid, Spain, 2014.
99. He, W.; Yang, J.Y.; Drury, C.F.; Smith, W.N.; Grant, B.B.; He, P.; Qian, B.; Zhou, W.; Hoogenboom, G. Estimating the impacts of climate change on crop yields and N₂O emissions for conventional and no-tillage in Southwestern Ontario, Canada. *Agric. Syst.* **2018**, *159*, 187–198. [[CrossRef](#)]
100. Li, Z.T.; Yang, J.Y.; Smith, W.N.; Drury, C.F.; Lemke, R.L.; Grant, B.; He, W.T.; Li, X.G. Simulation of long-term spring wheat yields, soil organic C, N and water dynamics using DSSAT-CSM in a semi-arid region of the Canadian prairies. *Nutr. Cycl. Agroecosyst.* **2015**, *101*, 401–419. [[CrossRef](#)]
101. Department for Transport (UK Government). *TAG UNIT M3.1: Highway Assignment Modelling*; Department for Transport: London, UK, 2020.
102. Dowling, R.; Skabardonis, A.; Alexiadis, V. *Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software*; US Federal Highway Administration: Washington, DC, USA, 2004.
103. Geroliminis, N.; Daganzo, C.F. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transp. Res. Part B Methodol.* **2008**, *42*, 759–770. [[CrossRef](#)]
104. Ambühl, L.; Loder, A.; Menendez, M.; Axhausen, K.W. Empirical Macroscopic Fundamental Diagrams New insights from loop detector and floating car data. In *TRB 96th Annual Meeting Compendium of Papers*; Transportation Research Board: Washington, DC, USA, 2017. [[CrossRef](#)]
105. Hall, F.L. Traffic Stream Characteristics. In *Traffic Flow Theory*; US Federal Highway Administration: Washington, DC, USA, 1996.
106. Shafto, M.; Conroy, M.; Doyle, R.; Glaessgen, E.; Kemp, C.; LeMoigne, J.; Wang, L. *DRAFT Modelling, Simulation, Information Technology & Processing Roadmap—Technology Area 11*; National Aeronautics and Space Administration (NASA): Washington, DC, USA, 2012.
107. Rosen, R.; Von Wichert, G.; Lo, G.; Bettenhausen, K.D. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine* **2015**, *48*, 567–572. [[CrossRef](#)]
108. Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access* **2020**, *8*, 108952–108971. [[CrossRef](#)]
109. Wang, P.; Yang, M.; Peng, Y.; Zhu, J.; Ju, R.; Yin, Q. Sensor Control in Anti-Submarine Warfare—A Digital Twin and Random Finite Sets Based Approach. *Entropy* **2019**, *21*, 767. [[CrossRef](#)]
110. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2021**, *58*, 346–361. [[CrossRef](#)]
111. Lima, A.; Stanojevic, R.; Papagiannaki, D.; Rodriguez, P.; González, M.C. Understanding individual routing behaviour. *J. R. Soc. Interface* **2016**, *13*, 20160021. [[CrossRef](#)]
112. Alessandretti, L. A new computational model for human navigation. *Nat. Comput. Sci.* **2021**, *1*, 642–643. [[CrossRef](#)]

113. Bongiorno, C.; Zhou, Y.; Kryven, M.; Theurel, D.; Rizzo, A.; Santi, P.; Tenenbaum, J.; Ratti, C. Vector-based pedestrian navigation in cities. *Nat. Comput. Sci.* **2021**, *1*, 678–685. [[CrossRef](#)]
114. Hudson-Smith, A.; Signorelli, V.; Dawkins, O.; Batty, M. More Than One Twin: An Ecology of Model Applications in East London. In *Digital Twins for Smart Cities—Conceptualization, Challenges & Practices*; Wan, L., Nochta, T., Tang, J., Schooling, J., Eds.; Institution of Civil Engineers: London, UK, 2022.
115. Tao, F.; Qi, Q. Make more digital twins. *Nature* **2019**, *573*, 490–491. [[CrossRef](#)]
116. Sepasgozar, S.M.E. Differentiating digital twin from digital shadow: Elucidating a paradigm shift to expedite a smart, sustainable built environment. *Buildings* **2021**, *11*, 151. [[CrossRef](#)]
117. Arcaute, E.; Barthelemy, M.; Batty, M.; Caldarelli, G.; Gershenson, C.; Helbing, D.; Moreno, Y.; Ramasco, J.J.; Rozenblat, C.; Sánchez, A. Future Cities: Why Digital Twins Need to Take Complexity Science on Board [Preprint]. 2021. Available online: https://www.researchgate.net/publication/354446988_Future_Cities_Why_Digital_Twins_Need_to_Take_Complexity_Science_on_Board (accessed on 10 December 2021).
118. Helbing, D.; Fanitabasi, F.; Giannotti, F.; Hänggli, R.; Hausladen, C.I.; van den Hoven, J.; Mahajan, S.; Pedreschi, D.; Pournaras, E. Ethics of Smart Cities: Towards Value-Sensitive Design and Co-Evolving City Life. *Sustainability* **2021**, *13*, 11162. [[CrossRef](#)]
119. Francisco, A.; Asce, S.M.; Mohammadi, N.; Asce, A.M.; Taylor, J.E.; Asce, M. Smart City Digital Twin-Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking. *J. Manag. Eng.* **2019**, *36*, 04019045. [[CrossRef](#)]
120. Ruohomaki, T.; Airaksinen, E.; Huuska, P.; Kesaniemi, O.; Martikka, M.; Suomisto, J. Smart City Platform Enabling Digital Twin. In Proceedings of the 9th International Conference on Intelligent Systems 2018: Theory, Research and Innovation in Applications, Funchal, Portugal, 25–27 September 2018; pp. 155–161.
121. Varschen, C.; Wagner, P. Mikroskopische Modellierung der Personenverkehrsnachfrage auf Basis von Zeitverwendungstagebüchern. In Proceedings of the 7. Aachener Kolloquium “Mobilität und Stadt” (AMUS 2006), Aachen, Germany, 7–9 September 2006; Integrierte Mikro-Simulation von Raum- und Verkehrsentwicklung. Theorie, Konzepte, Modelle, Praxis. Institut für Stadtbauwesen und Stadtverkehr, RWTH Aachen: Aachen, Germany, 2006; Volume 81, pp. 63–69.
122. Codecá, L.; Frank, R.; Faye, S.; Engel, T. Luxembourg SUMO Traffic (LuST) Scenario: Traffic Demand Evaluation. *IEEE Intell. Transp. Syst. Mag.* **2017**, *9*, 52–63. [[CrossRef](#)]