https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

Transport modelling for dynamic urban climate studies: MATSDA-roads v2.0

Tiancheng Ma[1], Denise Hertwig[1], Megan McGrory[1], Matthew Paskin[1], and Sue Grimmond[1]

[1]Department of Meteorology, University of Reading, Reading, RG6 6ET, UK

Correspondence to: Sue Grimmond (c.s.grimmond@reading.ac.uk)

Abstract. Representing dynamic patterns of people's movement is crucial for modelling high-resolution urban systems with feedback to emissions (e.g., anthropogenic heat, pollutants) and exposure of individuals to environmental stressors. We have developed the transport model MATSDA (Movement And Transport Simulations using Dijkstra's Algorithm) consists of MATSDA-roads and MATSDA-metro that iteratively computes optimised routes through a simplified nodal representation of urban transport networks. The approach can be used for different transport modes with routes being represented as a sequence of journeys between linked nodes. MATSDA-roads v2.0's input network has a hierarchy of road types (motorway to local roads, UK) connected by junctions and intersections. MATSDA-roads v2.0 is assessed in London (UK) with data collected in real-time from the Google Maps Directions API over several weeks to capture day of week, time of day, route direction (outbound v. inbound) and route alternatives (total of >87,000 reference routes). A high level of road network detail, notably carriageway types (e.g., dual carriageways, slip roads), is critical to obtain travel times accurately to within 5-10 minutes during daytime, particularly for longer journeys. Model parameter choices are shown to impact model performance, with effective length of local roads and junction delay penalties increasing the modelled travel time. MATSDA-roads v2.0 captures the diurnal variability of urban traffic through its input data, including morning and evening rush hours but travel times are systematically underestimated late at night (between 22 h and 4 h). The model exhibits high skill at identifying major travel corridors (Fractions Skill Score ~0.7 at 500 m grid resolution), indicating its route choices are spatially realistic. This work provides a valuable tool for transport research, urban climate modelling and environmental exposure assessment that require dynamic human movement patterns.

1. Introduction

With more than 45% of the world's population (approximately 3.7 billion people) living in urban areas (UNPD 2025), the reliance on vast and complex transportation systems is greater than ever. The efficiency of these systems directly governs travel times, which in turn influences critical aspects of urban life, including sustainable development and economic vitality (Louf and Barthelemy 2014). However, urban transport systems are also a major source of environmental pressure (Afshari et al. 2018). They are significant contributors to greenhouse gas emissions and energy consumption (Blunn et al. 2024), and their environmental impacts are both direct and indirect. In addition to air pollutant emissions, there are anthropogenic heat emissions from urban transport systems which modify the urban microclimate. Vehicles with combustion engines directly

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

warm the air by releasing waste heat. The amount of heat released depends on vehicle type, fuel source and speed (e.g., Grimmond 1992, Iamarino et al. 2012; Sailor and Lu 2004). It varies with neighbourhood road network density, road types (e.g., affecting speed limits), time of day (e.g., rush hour), day types (weekday/weekend; e.g., Qian et al. 2024), and seasonal variations in fuels (Grimmond 1992). Indirectly, transport infrastructure (e.g., paved surfaces) absorb and store incoming solar radiation and emit it as heat (Ivanchev and Fonseca 2020). Furthermore, transport dynamics influence other sources of anthropogenic heat (Capel-Timms et al. 2020; Hertwig et al. 2025b). For example, transport behaviour (e.g., driven by preferences, policies, short-term traffic disruptions, reductions/increase of remote work) can impact daytime building energy demand (Liu et al. 2022). Traffic congestion can alter commute times (Thabet and Zengin 2024), thereby shifting building occupancy patterns and affecting indoor heat gains from energy consumption and associated heat emissions from buildings. Transport also impacts human exposure to urban environmental stressors like air pollution and heat through duration of stay in different transport environments (e.g., subway, train, car) and in different locations of the city with higher/lower exposure (e.g., Yang et al. 2009, Smith et al. 2016).

The need to understand these complex interactions is amplified by advancements in environmental modelling. As numerical weather prediction (NWP) and climate models increase their resolution (O(100 m)) over urban areas (Lean et al. 2024; Hall et al. 2024), the dynamic behaviour of the urban population becomes a critically important factor. Urban environments are distinct from rural areas in their surface properties, energy exchange patterns and anthropogenic activities (Oke et al. 2017). At sub-kilometre scale, human activities like commuting and leisure travel are no longer sub-grid phenomena but can become resolvable processes that can significantly impact local energy fluxes and therefore other atmospheric variables (e.g. stability, wind speed, temperature) and pollutant dispersion.

Here, we introduce the urban transport model MATSDA (Movement And Transport Simulations using Dijkstra's Algorithm, Sect. 2) for the dynamic modelling of travel routes through a city's transport network. The framework (Fig. M1.2[1] in the User Manual: Ma et al. 2025a) has two distinct components: MATSDA-roads, designed for road-based movement (driving, cycling, walking), and MATSDA-metro, designed for public transport networks (e.g., bus, rail, waterways). While established routing platforms like Google Maps or Apple Maps provide highly accurate navigation for individuals based on real-time or historical data, they are not optimized for the specific requirements of integrated urban climate research. Commercial platforms often function as 'black boxes' with proprietary algorithms, that are computationally and financially prohibitive to query at the scale required for city-wide flux calculations. Whereas, MATSDA can support urban climate modelling activities such as the derivation of anthropogenic heat emissions from traffic sources but also provides valuable spatio-temporal data for exposure assessments. We demonstrate MATSDA-roads v2.0 applied to the UK's largest metropolitan area, Greater London (~2,300 km^2, Fig. 1a). The detailed road network (Sect. 3.1) for the model domain includes the M25 orbital motorway. The choice and analysis of independent reference data, as well as the set-up for seven configurations are evaluated (Sect. 3.2). The performance

[1] Cross-referenced Sections, Tables or Figures with their relevant numbers in this paper have no prefix used, whilst for the Supplementary material a S or the Manual a M prefix is used.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

of MATSDA-roads' route-finding algorithm for road traffic by vehicles is evaluated in time and space (Sect. 4). Specifically, we address the following research questions:

- What level of detail is needed to represent the road network for realistic travel route modelling (assessed using independent data, e.g., Google Maps)?
- What are the drivers of model uncertainty and controls on bias (e.g., path or travel duration)?
- Given MATSDA-roads' v2.0 input data and road-network representation, what degree of variability is captured in the modelled transport routes and travel durations (e.g., differences by day type (weekday, weekend), time of day (e.g., rush-hour) and road types)?

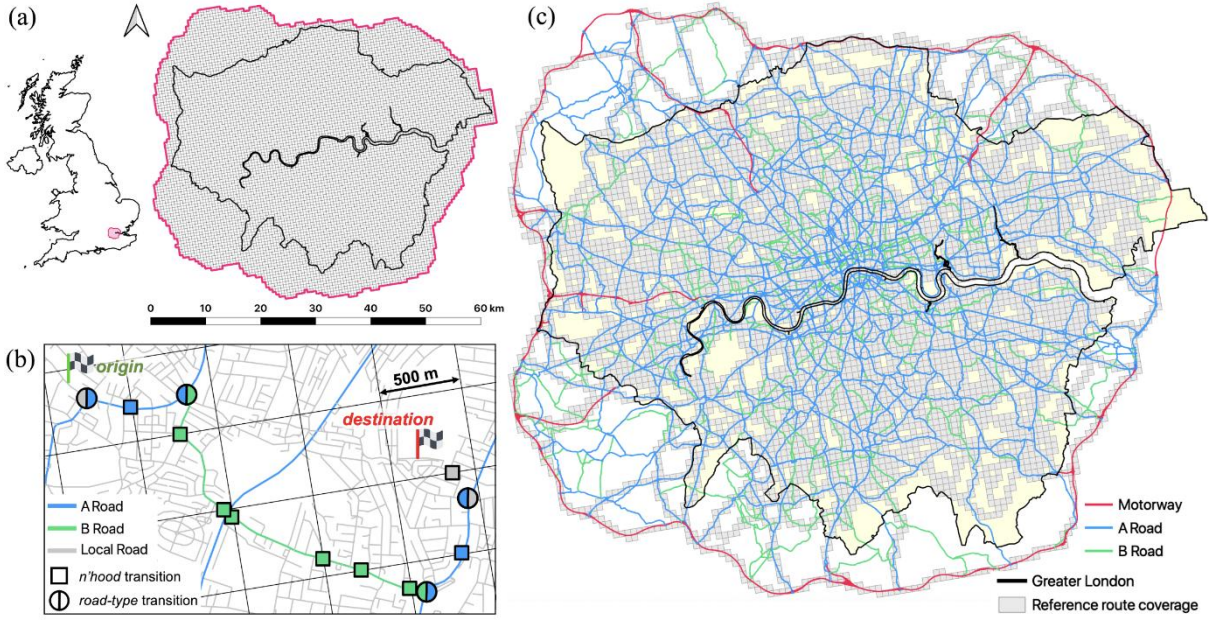


Figure 1: Aspects of the MATSDA-roads v2.0 model domain shown at different scales: **(a)** full extent of the Greater London model domain (inset: UK) out to the M25 motorway (pink border) with 500-m grid length neighbourhoods shown, **(b)** schematic of a MATSDA-roads v2.0 route with its sequence of varying road types (nodes) connected by journeys that are grid-boundary (black) aware, and **(c)** all major traffic arteries within the model domain plus the independent reference data route coverage (grid-cell level) used to evaluate the model (Section 3).

2. MATSDA-roads model description

MATSDA-roads is an urban transport model that derives optimal travel routes (e.g., to minimise travel duration, travel cost or air pollution exposure) between origin–destination (O-D) pairs in space (Fig. 1b) for different transport modes such as driving by car, bicycling or walking. The model is applicable across scales from local neighbourhoods to large metropolitan areas and can respond to accessibility changes or choices (e.g., road closures, avoidance of low emission or congestion charge zones). MATSDA-roads/-metro travel routes are usable by the agent-based modelling system DAVE (Dynamic Anthropogenic actiVities and feedback to Emissions; Hertwig et al. 2025b) to allow people to move between (and within) city neighbourhoods.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

MATSDA uses Dijkstra's (1959) pathfinding algorithm, a method widely applied for determining the optimal shortest-path for many network types, including roads (Casey et al. 2014, Bing and Lai 2022), communication (Kaur et al. 2020), and social media (Mensah et al. 2020). While effective, this approach is computationally intensive (Cormen et al. 2009), as algorithm run time and memory requirements grow with network size and level of detail used to capture the system of interest (Ahuja et al. 1993). Therefore, to model large road networks of mega-cities, MATSDA implements a balance between accuracy of network representation and computational costs (details in Sect. 3.1.1).

Fundamentals of the MATSDA-roads approach and terminology used (Table 1) are introduced in Sect. 2.1, followed by general model input requirements (Sect. 2.2) and features of the pathfinding implementation (Sect. 2.3). Information on this specific MATSDA-roads v2.0 application is given in Sect. 3.1.

2.1 Approach and terminology

MATSDA-roads takes the road network as input across a model domain of neighbourhood areas (O(~100 m) horizontal extent) containing intra- and inter-neighbourhood connected transport nodes (Fig. 1, Table 1). Thus, the spatial discretisation requirements for urban climate modelling (e.g., use of grid-cells) are addressed. With MATSDA-roads v2.0, a high-resolution representation of the sub-grid scale street network, needed for realistic traffic routing, is introduced (Sect. 2.2).

Fundamental to both of MATSDA's modelling components (road, metro) is a graph-structured travel database (Table 1; Sect. 2.2) with the intra- and inter-neighbourhood transport network represented as nodes. Different node types exist, such as for roads (MATSDA-roads) or public transport lines via bus, tram, subway, train and etc. (MATSDA-metro). Nodes connect (journeys) to other nodes via junctions within a neighbourhood and can connect to other nodes at the neighbourhood border (Fig. 1b; road leading from one neighbourhood to another). Each journey from one node to another directly accessible node has one or more weight attributes that characterise the connection (e.g., travel time, travel cost). Journey weights can vary by transport mode, node type used (e.g., motorway cf. local road), day type, time of day (time period applicable; Table 1), service schedule (public transport timetables) and node length. The generation of the complete nodal network, referred to as MATSDA's travel database, is done with MATSDA's Node Creator (Ma et al. 2025a, Fig. M1.2). In a second modelling step, the travel database is used to compute the complete travel route with MATSDA's Pathfinder (Ma et al. 2025a, Fig. M1.2) as the optimal sequence of journeys (minimising the weight total) between nodes from origin to destination using Dijkstra's (1959) method.

Table 1: Terminology used in MATSDA modelling and its specific application as part of MATSDA-roads v2.0 in this study, with Dijkstra's (1959) paper terms (*).

Terminology	General description	This study (MATSDA-roads v2.0)
Neighbourhood	Area where travel begins/ends/passes through (order ~ 1 ha)	500 m x 500 m grid-cells covering Greater London (Fig. 1)
Travel database (graph*)	Transport network (e.g., roads) with nodes (neighbourhood-node-type pairs) connected by journeys	Greater London travel database (Sect. 2.2)

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

Node	Sub-neighbourhood scale transport network component with information on possible connections (journeys) to other nodes	Uniquely identifiable based on neighbourhood number and road number/name (Fig. 2a)
Node type	Category influencing traffic speed, capacities and connectivity (e.g., road type, public transport lines/services)	UK road types (DfT 2012) (1) major: motorway, A Road, B Road; (2) minor: local
Junction	Point where nodes intersect, enabling transitions (journeys) between, e.g., road types, transport modes	Major road junctions, roundabouts, connections between adjacent neighbourhoods (Fig. 1b)
Journey (edge*)	Connection between two nodes with information on travel time, travel cost, etc.	Connections of road types and neighbourhoods (Fig. 1b)
Journey weight	A non-negative attribute (e.g., travel time, distance, cost) characterising each journey	Travel time
Time period applicable	Journey weights can vary by time of day and day of week (e.g., traffic rush-hour, public transport schedules)	Traffic flow speeds for different day types and time intervals: Weekday (h): 4-7, 7-9, 9-12, 12-14, 14-16, 16-19, 19-22, 22-4; Weekend (h): 4-7, 7-10, 10-14, 14-19, 19-22, 22-4 (Sect 2.2, Fig. 2a; Table S8)
Travel route (shortest path*)	Complete sequence of journeys between nodes from origin to destination obtained by optimisation considering journey weights (Dijkstra 1959)	Travel time minimised

2.2 MATSDA-roads v2.0 travel database

The MATSDA-roads v2.0 Road-Node Creator 2.0 tool generates a travel database for road transport based on user-processed geospatial road network input data (Sect. M3.1, M3.2 in Ma et al. 2025a). To construct the travel database, three general steps are required:

1. Geospatial analysis to define road-nodes in each neighbourhood
2. Aggregation of neighbourhood road lengths and derivation of travel times
3. Extraction of journeys available between nodes within and across neighbourhoods

For step 1, geospatial road network data are needed to derive road lengths, road type and/or number/name. Road segments within each neighbourhood are extracted and aggregated by a common identifier (road_ID) based on road names, numbers and/or carriageway types. Combined with a neighbourhood identifier (nhood_ID), this creates unique identifiers for all nodes in a neighbourhood (<nhood_ID>_<road_ID>; example in Fig. 2a).

Each node has a list of journeys to other nodes and associated travel times (weights; Fig. 2a). To derive these travel times (step 2), local (e.g., road-type dependent) traffic flow speeds are needed (e.g., based on observations or speed limits) together with road lengths (step 1). Two transition cases exist (Fig. 1b): *i.* node-type transition: junction between different road-node types within the same neighbourhood (split circle, Fig. 2b), and *ii.* neighbourhood transition: roads connect to an adjacent neighbourhood. These are derived from geospatial road junction locations and from mapping intersections between roads and neighbourhood boundaries (square, Fig. 2b). For the inter-node transition, the model allows specification of a time penalty to characterise the anticipated time loss from changing roads. The impact of this is discussed in Section 4.1. The derivation of

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

the MATSDA-roads v2.0 travel database within the Greater London modelling domain (Fig. 1c), used in this study, is given in Section 3.1.

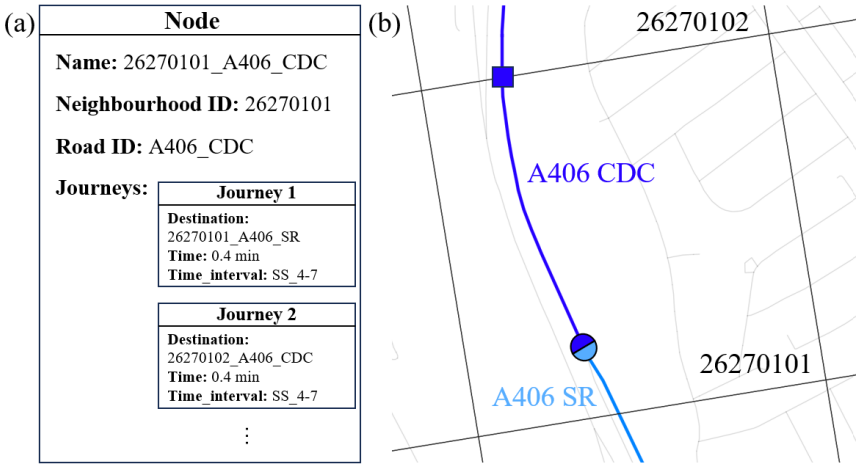


Figure 2: Example MATSDA-roads v2.0 node information for a case with a major road (A406, CDC: collapsed dual carriage way) connected to two other nodes (journey destinations): **(a)** data structure and attributes, and **(b)** map of nodes and junctions (connection symbols as in Fig. 1b). Journey 1: connects to a different road type (SR: slip road) within the same neighbourhood through a junction; Journey 2: connects to the same road but in a different neighbourhood). Travel time (min) is for a specific time interval or period (Saturday/Sunday (SS), 4:00-7:00). Additional journeys during other time periods exist.

2.3 MATSDA Pathfinder 2.0

The MATSDA Pathfinder 2.0 (Fig. 3) is designed to determine an optimal path (based on journey weights, Table 1) from a start node (origin) to all other nodes in the road network (Table 1), and uses the same generic approach shown in Fig. 3 for both road and metro model components. To describe the Pathfinder logic, we use a road transport example, with travel time as the journey weight. Other weights (e.g., travel cost, air quality concentration) are possible, including combining weights. A critical optimization in the MATSDA's pathfinding framework is the calculation of the complete shortest-path tree for each origin node in a single execution, rather than computing routes for individual origin-destination (O-D) pairs at runtime (Sect. M3.3 in Ma et al. 2025a).

MATSDA-roads v2.0 travel routes use network connections (Fig. 2) to provide intra- and inter-neighbourhoods movement. To ensure realistic behaviour, we require travel routes to originate and terminate on local roads, based on three assumptions. First, most real-world travel patterns begin and end at home or workplaces that are frequently situated on smaller, local roads rather than directly on major roads (e.g., motorway, A/B road; Hertwig et al. 2024). Second, major roads (e.g. motorways or A roads) have limited and controlled access points (junctions; Fig. S1 in Supplementary Information). Starting on a local road requires the Pathfinder to first find a suitable junction to access these faster routes, avoiding logical inconsistencies of starting a journey on a major road (e.g., a motorway) in a neighbourhood where no direct access is possible. Third, having only one local origin node per neighbourhood enhances computational feasibility for a large-scale network. For origin and/or destination neighbourhoods with no local roads, we designate the nearest neighbouring grid-cell with the longest total local road length as

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

a parking node. A fixed 5-min walking time is added to the trip to account for the transition between the intended origin/destination and this parking node, ensuring network connectivity between all neighbourhoods in the domain.

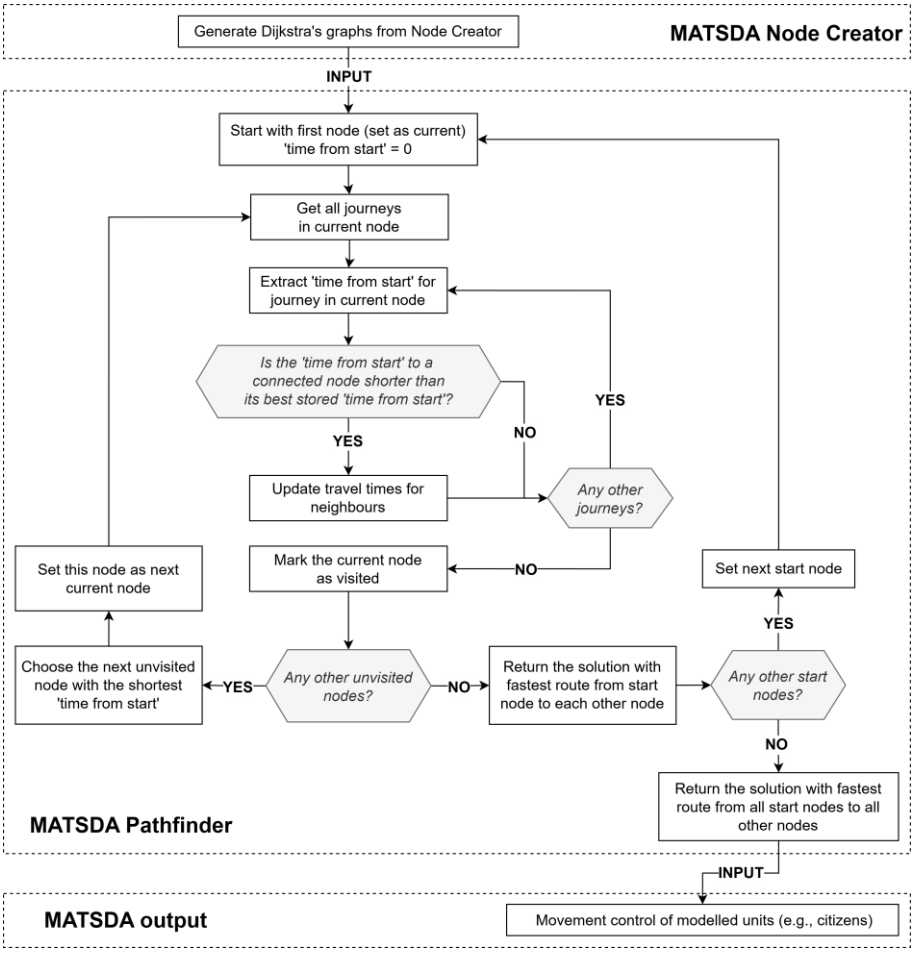


Figure 3: MATSDA Pathfinder v2.0 computational logic with arrows showing processing directions.

The Pathfinder iteratively processes MATSDA's travel database by continuously updating the 'shortest travel time' parameter from a given start node to all other network nodes (Fig. 3). Once a connection to a node has been analysed, the node is marked as 'visited.' All possible journeys from each start node are evaluated for each time period applicable (Table 1), as travel times across the network can be subject to temporal variations, e.g., reflecting typical diurnal changes of traffic flow speeds. The 'fastest route' from the start node to every end node is stored with the associated 'fastest time' giving the optimal route solution. Solutions are stored for analysis and can be used as input data to different urban modelling systems, such as agent-based models like DAVE (Hertwig et al. 2025b). Details for MATSDA-roads v2.0 Pathfinder v2.0 output are presented in Sect. S2.2.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

3. MATSDA-roads v2.0 setup and evaluation cases for Greater London

Here, setup and evaluation of MATSDA-roads v2.0 is optimised for road (driving) travel-time in Greater London (Fig. 1a,c). However, the general methodology is applicable to elsewhere (i.e., independent of city or country) and also for other transport types (e.g., MATSDA-metro), although data processing differences may exist (Sect. 3.1) depending on local differences in road-type classifications. Note that hereafter, 'MATSDA' refers specifically to MATSDA-roads v2.0 unless otherwise stated. We evaluate both modelled travel duration and spatial routing using reference data extracted through the Google Maps Directions API (Google 2025, Sect. 3.2). All notation used in this section can be found in Appendix A.

3.1 Greater London travel database

3.1.1 Domain and road network

The model domain in this study is the region enclosed by the M25 orbital motorway, encompassing the Greater London metropolitan area. The domain is subdivided into 9,416 grid-cells or neighbourhoods (500 m x 500 m resolution, Fig. 1a,c). Details of MATSDA-roads v2.0 Pathfinders v2.0 output are given in Sect. S2.2. The grid-cell is the fundamental spatial unit of the model, but a Cartesian grid is not a model requirement (i.e., neighbourhoods can be represented in other ways, such as census or administrative regions).

A hierarchy is used to capture the road network to balance computational feasibility and modelling accuracy. There are two representations for the main traffic arteries (Fig. 1c) of major roads (motorways, A roads, B roads; DfT 2012) used in MATSDA's travel database: *i.* only using the road number (e.g., A406) or *ii.* adding different carriageway type information (e.g., dual carriageway, single carriageway, slip road, roundabout) available from OS Open Roads (2021). The major road (numbered) and carriageway type combinations create unique identifiers that are used to define road nodes (Fig. 2).

Within each neighbourhood, major roads are clipped by the 500-m grid boundaries to precisely calculate the road-node length, ℓ_n (m). This length is derived from corresponding road segments, l_s (m), of the OS Highways (2024) input data. This level of detail is critical as road type impacts travel speeds and hence routing choices, with motorways often given priority for long-distance trips and local roads for intra-neighbourhood movement (DfT 2012). It is computationally unfeasible to apply the same level of detail to the complex network of local (minor) roads in the domain. Instead, local roads within each neighbourhood are treated as an aggregated entity and represented only by a single node. As local road journeys within a neighbourhood typically only use a subset of the available local road network, an effective local road length is derived. We consider two approaches to calculate this effective length: *i.* a constant fraction of total local road length per neighbourhood, and *ii.* a parameterisation based on O-D (origin-destination) straight-line distance. These are evaluated in Section 4.1.1.

The neighbourhood road network complexity is reflected by the number of nodes within the MATSDA travel database. Across the model domain there are a total of 20,273 nodes (database size is 21.2 MB in JSON format, Ma et al. 2025b), with a maximum of 13 nodes occurring in a single neighbourhood. Computational details of the travel database generation are presented in Sect. S1.2-S1.3.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

3.1.2 Traffic flow speed

To compute travel times for journeys between nodes, travel speeds are derived from average traffic speed data for individual roads (Digimap Pilot Collection 2024). The dataset is based on observed (GPS) vehicle speeds over a six-month period, for weekdays (Monday–Friday; MF), weekends (Saturday/Sunday; SS) and different times of day (MF (h): 4-7, 7-9, 9-12, 12-14, 14-16, 16-19, 19-22, 22-4; SS (h): 4-7, 7-10, 10-14, 14-19, 19-22, 22-4). Thus, diurnal variations in travel times across day types can be modelled in MATSDA (Sect. M3.1 in Ma et al. 2025a; Sect. S1.2.3).

For bidirectional roads, Digimap Pilot Collection (2024) distinguishes vehicle speeds for both flow directions but without information on the geospatial direction (e.g., from or towards the city centre) or road lane. Spatial variability of the neighbourhood traffic speeds across the model domain shows clear differences by road type and time (day type, time of day; Fig. 4). The impact of selecting either minimum or maximum speeds from the two flow directions on the MATSDA route outcome is discussed in Sect. 4.2.

Speed information is evaluated per neighbourhood by mapping the data to OS Highways (2024) road segments and creating average speeds by road and carriageway type (Sect. S1.2.3). The 'aggregated' local roads are assigned the median of the observed local road speeds within each grid-cell neighbourhood. As a single major road object (road number-carriageway type) may be represented by several line segments in the OS Highways (2024) geospatial data, the average speed (V, m min^{-1}) per unique major road-node in a neighbourhood is calculated per time period (t) as the segment length (λ_s, m) weighted mean of all $s \geq 1$ road segment speeds ($V_{t,s}$, m min^{-1}):

$$V_t = \frac{\Sigma(\lambda_s V_{t,s})}{\Sigma\lambda_s} . \quad (1)$$

Based on the total road-node lengths (ℓ_n, Sect. 3.1.1) and traffic speeds (Eq. 1), the journey duration ($\tau_{n,t}$, min) per time period t is:

$$\tau_{n,t} = \frac{\ell_n}{V_t} . \quad (2)$$

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

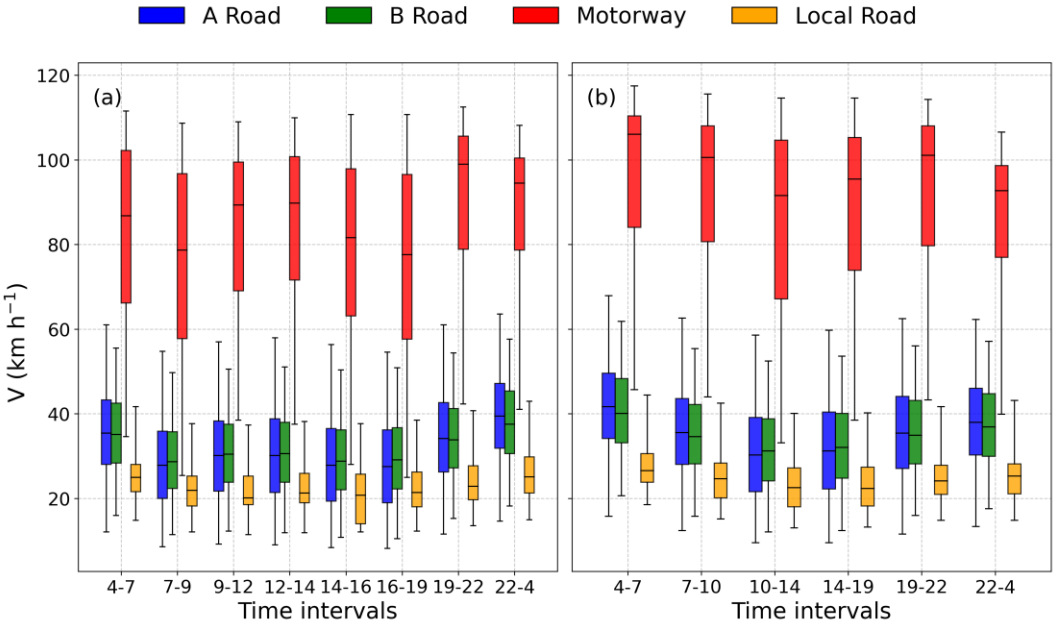


Figure 4: Variability (median, interquartile range, 5th and 95th percentiles) of all model domain neighbourhoods (Fig. 1c) traffic flow speeds (maximum from raw data; minimum shown in Fig. S2; Digimap Pilot Collection 2024) by road type (colour) and time of day for: **(a)** weekdays and **(b)** weekend.

3.2 Model evaluation

To assess MATSDA's performance we use Google Maps routing model data (GM; Google 2025), focusing on ability to realistically model route choices and travel times under diverse conditions. The goal of the evaluation is to demonstrate that MATSDA can capture traffic patterns in time and space realistically through its simplified approach when compared to the routing from a more complex and detailed transport model. The evaluation methodology requires: *i.* selection of representative O-D pairs (Sect. 3.2.1), *ii.* collection and processing of GM data (Sect. 3.2.2), and *iii.* definition of both evaluation cases (Sect. 3.2.3) and metrics (Sect. 3.2.4, Sect. 3.2.5).

3.2.1 Route selection

To select O-D neighbourhoods a stratified sampling strategy is used considering realistic travel scenarios distributed across the model domain (Fig. 1a,c) and covering multiple distances. The Greater London (administrative boundary Fig. 1a,c) population data (ONS 2011) are used to select the O-D neighbourhoods.

Origin neighbourhoods are selected based on the local driving population (Fig. S4) which is estimated from the residential population and car commuter fraction (ONS 2011). Neighbourhoods are stratified into five car commuters' frequency classes (*cf.* other modes): 1-10%, 11-25%, 26-45%, 46-70%, >71%. From the five classes, a total of 75 unique origin neighbourhoods

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

are randomly selected using a weighting of 5, 10, 15, 20, 25, respectively, ensuring the sample covers a varying reliance on private vehicles (Fig. 5b).

Destination neighbourhoods are selected for each origin using five non-overlapping, straight-line distance categories (Short/S: [1.5,5), Short-Medium/SM: [5,12), Medium/M: [12,20), Medium-Long/ML: [20,28) and Long/L: [28,47] km), excluding destinations <1.5 km. For the [5,12) km distance there are 60 (rather than 75) origin neighbourhoods. For all distance bands, we identify the most likely workplace neighbourhood based on data from Hertwig et al. (2024, 2025a) to create 391 unique O-D pairs, representing realistic residence-workplace pairs across Greater London (Fig. 5).

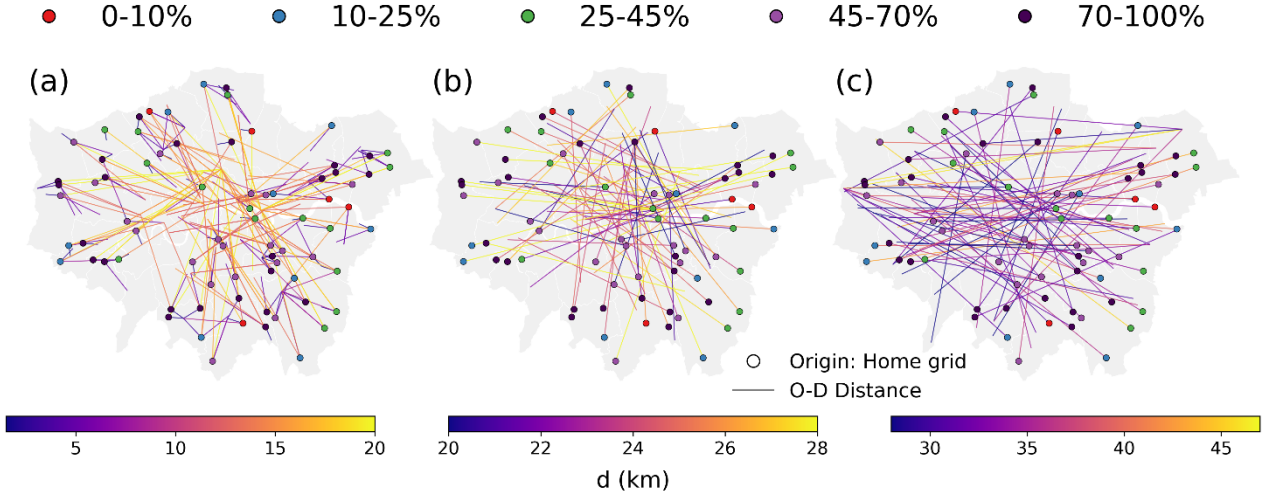


Figure 5: Spatial distribution of the 391 unique origin-destination (O-D) pairs in Greater London, faceted by five straight-line distance categories: **(a)** S: [1.5-5) km, SM: [5-12) km, M: [12-20) km; **(b)**: ML: [20-28) km; **(c)**: L: [28-47] km). All 75 origin points (60 for SM) are shown in each panel, coloured by the driving population percentage. The lines connect origins (home grid-cell) to their corresponding destinations (workplace grid-cell).

3.2.2 GM reference data

For each O-D pair (Sect. 3.2.1) the Google Maps Directions API query uses the O-D neighbourhoods' centroids to return a 'best guess' (primary) route plus up to two alternatives from GM's traffic model which synthesizes historical traffic patterns (based on time of day and day of week) with current live traffic conditions (Google 2025). To capture different traffic conditions, queries are undertaken for different day types (weekdays with school: WD_S, weekdays during school holidays: WD_{NS}, and weekends/public holidays: WE/H) and times (Table S11). To avoid API congestion and capture representative traffic, requests are initiated 5 min past the hour (e.g., 08:05) and staggered over 15–20 min. Simultaneously, a second set of API calls are made for reverse O-D pairs (i.e. workplace to residence). Data are analysed by time intervals (e.g., 08:05 → 07:00–09:00; Table 1 and Sect. 3.1.2).

Of the 156,370 (154,697) retrieved GM routes, 44,841 (44,187) are unique home→work (work→home) routes. Each unique route is assigned an ID capturing the O-D neighbourhood pair, geospatial routing data, day type and time of day. Non-unique

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

routes, differ only because of day collected (i.e., same ID as all other attributes are identical) or have minor geospatial variations (e.g., functionally identical as they traverse the exact same sequence of neighbourhoods). The retrieved GM data includes geospatial line segments with length and travel time attributes, allowing segment-specific traffic speeds to be derived. To enable like-for-like comparisons to MATSDA road-types (Sect. 3.1.1), we assign each segment OS Highways (2024) data attributes. Thus, direct comparison of travel times, routing, and road types can occur. All processing steps of the GM data are given in Sect. S3.2.

3.2.3 Evaluation cases

To evaluate MATSDA with the Greater London and GM database (Sect. 3.1, Sect. 3.2.2), we use:

- Spatially consistent routes (*SCR*): the GM retrieved routes (grey grids; Fig. 1c) known to MATSDA's Pathfinder are used to evaluate travel time, with spatial uncertainty removed as similar routing and road-types are used.
- Full: uses all GM database cases assessing MATSDA's spatial pathfinding as the full domain road network is available.

The GM O-D pairs are for two directions (home→work, work→home) to capture asymmetries in travel times and congestion: work→home GM routes are evaluated during the evening rush hour period (16-19 h), whereas home→work GM routes are evaluated for all other periods (morning rush hour and off-peak). This direction-specific filtering uses the GM direction that best represents the relevant period, which fits the data collection design (Sect. 3.2.2).

Within the GM dataset (Section 3.2.2) there are 44,102 home→work and 43,293 work→home routes that cover a wide range of neighbourhoods (Fig. 1c) and are fully within the model domain (Fig. 1a). The 739 (home→work) and 894 (work→home) unique routes that use roads outside our model domain are excluded. A further 7,466 (8,782) GM routes are not assessed in the *SCR* case as small geospatial resolution discrepancies exist between the GM routes and MATSDA's OS Open Roads (2021)-based input data prevent MATSDA finding a continuous path through the same neighbourhoods (details given in Sect. S3.3). For the *SCR* assessment there are 36,636 (34,511) GM routes. These are randomly split using a ~6:4 ratio into training (used in Sect. 4.1) and evaluation (used in Sect. 4.2 and 4.3) datasets (Table 2), with stratification for day type and time of day. The *full* evaluation uses the *SCR* subset plus the routes removed with small geospatial inconsistencies (Table 2). Table S13 summarise data availability by day type, time of day, route distance and dominant road type.

Table 2: Number of routes (*N*) by travel direction in each GM route dataset. More details given in Table S13. Training dataset is used to determine model parameters.

GM data set use	*N: home→work*	*N: work→home*
Evaluation (temporal) – *spatially consistent routes (SCR)*	14,667	13,815
Evaluation (spatial) – *full* domain	22,133	22,597
Training – *subset* domain	21,969	20,696

Parameters (Sect. 4.1) are determined using the training dataset (Table 2), prior to use in the independent model evaluations (Table 2). To investigate the impacts of model input data (e.g., traffic flow speed, road description) and parameters (i.e., effective local road length fraction, junction time penalty) seven MATSDA runs are undertaken (Table 3).

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

The effective local road length fraction ($\mathcal{F}_L$) is estimated using the local road length per grid-cell (λ_L; m) from the GM training dataset (Table 2):

$$\mathcal{F}_L = \frac{\lambda_L}{l_L}, \quad (3)$$

where l_L (m) is the corresponding local road length of MATSDA's travel database (Sect. 3.1.1). The road-type (k) usage ratio $f_{k,GM}$ ($f_{k,MATSDA}$) is defined using the distance Λ_k ($\mathcal{L}_k$) travelled on a particular road type (in the UK context, i can be A/B/Motorway (M)/Local (L)), as part of the entire route from all datasets (Table 2):

$$f_{k,GM}(f_{k,MATSDA}) = \frac{\Lambda_k\,(\mathcal{L}_k)}{\Lambda_{tot}\,(\mathcal{L}_{tot})} * 100\%, \quad (4)$$

where Λ_{tot} ($\mathcal{L}_{tot}$) is the total GM (MATSDA) route length, so that sum of the road-type usage fractions (Eq. 4) is:

$$f_A + f_B + f_M + f_L = 1, \quad (5)$$

The constant values and parameterizations for $\mathcal{F}_L$ (Eq. 3) as well as the junction time penalty (*JTP*) used in the seven MATSDA runs are derived from sensitivity analyses performed on the training dataset detailed in Section 4.1. A range of constant values (0.1-0.3) is tested for #3 (Table 3), with $\mathcal{F}_L$ =0.16 selected as the optimal choice, as it provides the best overall performance across metrics (MAE, MBE, and Hit Rate, Sect. 3.2.4) for typical journeys. A similar sensitivity analysis is conducted for *JTP* (#5, Table 3), testing penalties from 0.1 to 3 minutes. A value of 0.25 minutes is chosen as the most suitable compromise, as it improves model performance without significantly degrading the accuracy for daytime commuting. Cases with varying $\mathcal{F}_L$ configurations (#4,5,7, Table 3) use a distance-based choice that selects an appropriate precomputed MATSDA travel database for each O-D pair based on its straight-line distance (see Sect. 4.1.1 and Sect. S4.2.2 for details). The distance-based selection, rather than a single $\mathcal{F}_L$, allows the model to have a local-road weighting depending on route length.

Table 3: Seven MATSDA run types (#), with some having multiple cases as parameters are varied. Runs are used to evaluate different aspects of the model by using only the spatially consistent routes (SCR) or the full dataset (Table 2). Model configurations use two traffic speeds (min: minimum, max: maximum, Sect. 3.1.2), two representations of the roads (RN: road number, CWT: carriageway type, Sect. 3.1.1), multiple effective local road length fractions ($\mathcal{F}_L$, Eq. 3), and multiple junction time penalties (Sect. 2.2).

#	GM	Speed	Major road-nodes	$\mathcal{F}_L$	JTP (min)	Cases
1	*SCR*	min	RN	0.16	0	1
2	*SCR*	max	RN	0.16	0	1
3	*SCR*	max	RN + CWT	0.1→0.3	0	9
4	*SCR*	max	RN + CWT	0.16 (0-14 km) 0.14 (>14 km)	0	1
5	*SCR*	max	RN + CWT	0.16 (0-14 km) 0.14 (>14 km)	0.1→3	16
6	*full*	max	RN + CWT	0.16	0	1
7	*full*	max	RN + CWT	0.16 (0-14 km) 0.14 (>14 km)	0.25	1

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

3.2.4 Temporal evaluation

Standard evaluation metrics are used to compare predicted ($T_{i,t,MATSDA}$) and observed ($T_{i,t,GM}$, GM data) travel duration values for all $i = 1, \dots, N$ routes available in each model run (Table 3) at specific time periods t, including:

- **Mean Absolute Error (MAE):** A lower MAE indicates better overall agreement, with a value of 0 representing a perfect prediction (Willmott and Matsuura 2005). MAE has units of the variable being evaluated.

$$MAE(t) = \frac{1}{N}\sum_{i=1}^{N}\left|T_{i,t,MATSDA} - T_{i,t,GM}\right|, \quad (6)$$

- **Mean Bias Error (MBE):** A measure of the average systematic error, identifying whether the model tends to overestimate (MBE>0) or underestimate (MBE<0). A perfect model has MBE=0, but caution is needed as large positive and negative errors can cancel each other (Britter and Schatzmann 2007). MBE has units of the variable being evaluated.

$$MBE(t) = \frac{1}{N}\sum_{i=1}^{N}(T_{i,t,MATSDA} - T_{i,t,GM}), \quad (7)$$

- **Hit Rate (HR):** Ideally the fraction of predictions that fall within a specified tolerance (W) of the observed values (e.g., Schlünzen and Katzfey 2003) would be 100%. The tolerance threshold can be set with an absolute (e.g., within ±X min) or as a fraction of the reference data. This provides insight into the proportion of predictions that are 'acceptably close' to the reference.

$$HR(t) = \frac{1}{N}\sum_{i=1}^{N} I\left(\left|T_{i,t,MATSDA} - T_{i,t,GM}\right| \le W\right) * 100\%, \quad (8)$$

where I is the indicator function (1 if the condition is met, 0 otherwise). In this study, different thresholds W are used depending on the route length: 5 min (HR_5) for short routes (≤14 km), 10 min (HR_{10}) for long routes (>14 km).

3.2.5 Spatial evaluation

To complement the evaluation of travel duration, an assessment of the spatial accuracy of MATSDA's route finding is essential. We use the Fractions Skill Score (FSS; Roberts and Lean 2008) as a neighbourhood-based evaluation approach to assesses skill across a range of spatial scales. The FSS is conceptually based on treating the spatial distribution of an event (e.g., a neighbourhood being traversed by a GM or MATSDA travel route) within a local area probabilistically rather than deterministically. For a given neighbourhood size, a perfect model outcome has the same fractional coverage of the event as the observations, regardless of the exact placement of the event within that neighbourhood. The FSS is defined based on the mean squared error (MSE) of the event fractions within neighbourhoods, normalized by a reference MSE corresponding to the case of no spatial correspondence between model and observation. The score is calculated as:

$$FSS = 1 - \frac{\Sigma_{i=1}^{X}\Sigma_{j=1}^{Y}\left(P_{MATSDA}(i,j) - O_{GM}(i,j)\right)^2}{\Sigma_{i=1}^{X}\Sigma_{j=1}^{Y}\left(P_{MATSDA}(i,j)^2 + O_{GM}(i,j)^2\right)}, \quad (9)$$

where $P_{MATSDA}(i,j)(i,j)$ and $O_{GM}(i,j)$ are the fractions of grid-cells within a neighbourhood centred at grid point (i,j) that contain the modelled (MATSDA) and observed (GM) events. The domain (X, Y) is the area enclosed by the M25 orbital

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

motorway (Fig. 1a,c). The FSS ranges from 0 (complete mismatch) to 1 (perfect match in fractional coverage) at a specific scale.

A neighbourhood is assigned a value of 1 if any segment of the route passes through it and 0 otherwise. This creates two binary fields for each O-D pair. The binary 'event' for the FSS calculation is defined as the presence of a travel route within a given 500 m x 500 m neighbourhood.

In this study, rather than evaluating the FSS over a range of increasing spatial scales (neighbourhood sizes), the analysis focuses on the neighbourhood grid-cells (500 m), which is equivalent to a direct, point-wise comparison of the MATSDA and GM routes. The analysis is performed iteratively by applying a progressively stricter 'grid count threshold'. Initially, the FSS is calculated for the entire set of routes (a threshold of 1, i.e., as long as a grid has an 'event' happen once). The threshold is then increased, restricting the analysis to only those grids that appear more frequently in the GM reference dataset. This process reveals whether the model's spatial skill is consistent across all journeys or if it performs differently when comparing common transport corridors versus less-frequented routes.

4. Model demonstration and evaluation results

The partitioned GM dataset (Sect. 3.2.3) is used for training and evaluation (Table 2). First, we use the training dataset to explore two key model parameters (Sect. 4.1): *i.* a spatially uniform effective local road length fraction versus its parametrization as a function of O-D straight-line distance (Sect. 4.1.1); *ii.* a time penalty applied when changing road-type nodes through junctions (Sect. 4.1.2). Second, MATSDA's route-finding travel times and temporal variations (Sect. 4.2) are evaluated for seven cases (Table 3) using the evaluation datasets (Table 2). Third, the spatial representation of MATSDA's route-finding is evaluated (Sect. 4.3) across the *full* domain using the best-performing parameter configurations (Table 3).

4.1 Sensitivity to model parameters

4.1.1 Effective local road length fraction ($\mathcal{F}_L$)

The effective local road length fraction ($\mathcal{F}_L$, Eq. 3), determined from the training dataset (Table 2) grid-cell data routes, is <0.3 for 80% of the grid-cells (Fig. 6). Hence, only a small fraction of the available total local road length is typically used for trips traversing a neighbourhood. We assess the impact of assigning a fixed $\mathcal{F}_L$ using values in the range 0.1 to 0.3.

Sensitivity of MATSDA's travel time *cf.* GM using a constant $\mathcal{F}_L$ is assessed to find the optimal value, using detailed carriageway information (Sect. 3.1.1, #3 Table 3). The performance of different $\mathcal{F}_L$ values is evaluated across three time periods (day: 07:00-19:00, night: 19:00-07:00, and full 24 h) and three day types (WD_S, WD_{NS}, WE/H; Sect. 3.2.2) with results shown in Fig. 7a-c.

During weekdays (WD_S, WD_{NS}) the best daytime hit rate, based on a ±10 min threshold (HR_{10}), occurs when $\mathcal{F}_L$=0.14 (0.16 for WE/H; Fig. 7c). Although at night HR_{10} continues to improve up to $\mathcal{F}_L$=0.26, this does not offset the daytime deterioration

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

beyond $\mathcal{F}_L$=0.16 for WD_S and WD_{NS}. Thus, using $\mathcal{F}_L$=0.16 gives the best HR_{10} across all day types. Using the MAE metric, a consistent result is obtained at night, decreasing MAE (*cf.* HR, Fig. 7c) from $\mathcal{F}_L$=0.1 to 0.26. The daytime MAE is minimised when $\mathcal{F}_L$ is 0.14 during weekdays (WD_S WD_{NS}) and is again slightly larger for weekend or holiday trips ($\mathcal{F}_L$=0.18).

For all day types and time periods, the model transitions from under- to over-predicting travel times as $\mathcal{F}_L$ increases (Fig. 7b), with the daytime overprediction occurring at smaller $\mathcal{F}_L$ values (*cf.* night). Again, we see the steady improvement at night for all day types as $\mathcal{F}_L$ increases from 0.1 to 0.3, with the MBE changing from ~-15 min to 0 min, whilst, the daytime MBE is near zero when $\mathcal{F}_L$=0.14 (WD_S, WD_{NS}) and $\mathcal{F}_L$=0.2 (WE/H; Fig. 7b). For weekdays, a $\mathcal{F}_L$value of 0.16 gives the best overall metrics (small MAE, high HR_{10} and small MBE, Fig. 7a-c).

Typically, the road-type usage ratio (f_k, Eq. 4) has small variations when $\mathcal{F}_L \geq$0.16 (Fig. 8 for WD_{NS}; other results in Fig. S8, S9). Good agreement is found both qualitatively (hierarchy of road-types used) and quantitatively with the GM road-type for typical usage scenarios (25^{th} to 75^{th} percentiles). The extreme (5^{th} and 95^{th} percentile) f_k cases are challenging for standalone MATSDA as it is time-optimizing, whereas the GM reference data includes routes shaped by real-world constraints impacted by complex human behaviours (Sect. 4.3). This analysis confirms that the $\mathcal{F}_L$=0.16, selected as the optimal choice based on the temporal performance metrics (Fig. 7a-c), also produces realistic road-type fractions for these typical scenarios. Based on this combination of results, $\mathcal{F}_L$=0.16 is used in subsequent model evaluations (Runs #1-3, and #6, Table 3).

We also explored a potential relationship between $\mathcal{F}_L$ and the number of junctions per grid-cell (Fig. S6), investigating if $\mathcal{F}_L$ could be parameterized based on local road network complexity. However, this analysis did not yield a robust parameterization and was not used in the final model configuration (further details available in Sect. S4.2.1).

Based on the hypothesis that travel behaviour and the relative importance of local roads might vary with route length, an alternative parameterization is explored. This approach, used in model runs #4, #5, and #7 (Table 3), assumes a distance-based variation with values derived from the training data giving a split between $\mathcal{F}_L$=0.16 for shorter journeys (O-D straight-line distance d≤14 km) and $\mathcal{F}_L$=0.14 for longer journeys (d>14 km). Testing of this more complex parameterization to determine if it could capture more behavioural nuance, indicates its overall performance (discussed in Sect. S4.2) did not show a significant advantage over the simpler, constant $\mathcal{F}_L$=0.16. Therefore, the constant $\mathcal{F}_L$=0.16 is considered the most robust and recommended parameterization for the London case, as the added complexity of the distance-varying approach did not yield a sufficiently useful improvement.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

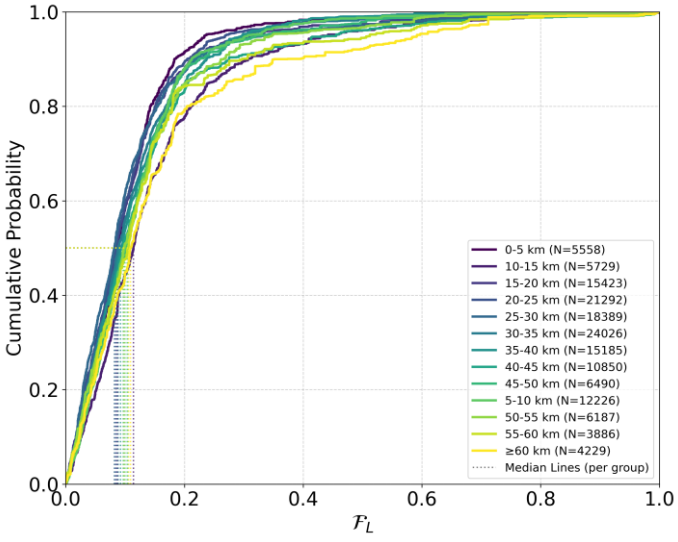


Figure 6: Cumulative frequency distribution of effective local road length fraction ($\mathcal{F}_L$, Eq. 3) grouped by total route length (colour; with number of samples; N), with median $\mathcal{F}_L$ indicated by route length. $\mathcal{F}_L$ calculated with the GM training dataset (Table 2, home→work) and corresponding MATSDA travel database information (Sect. 3.1).

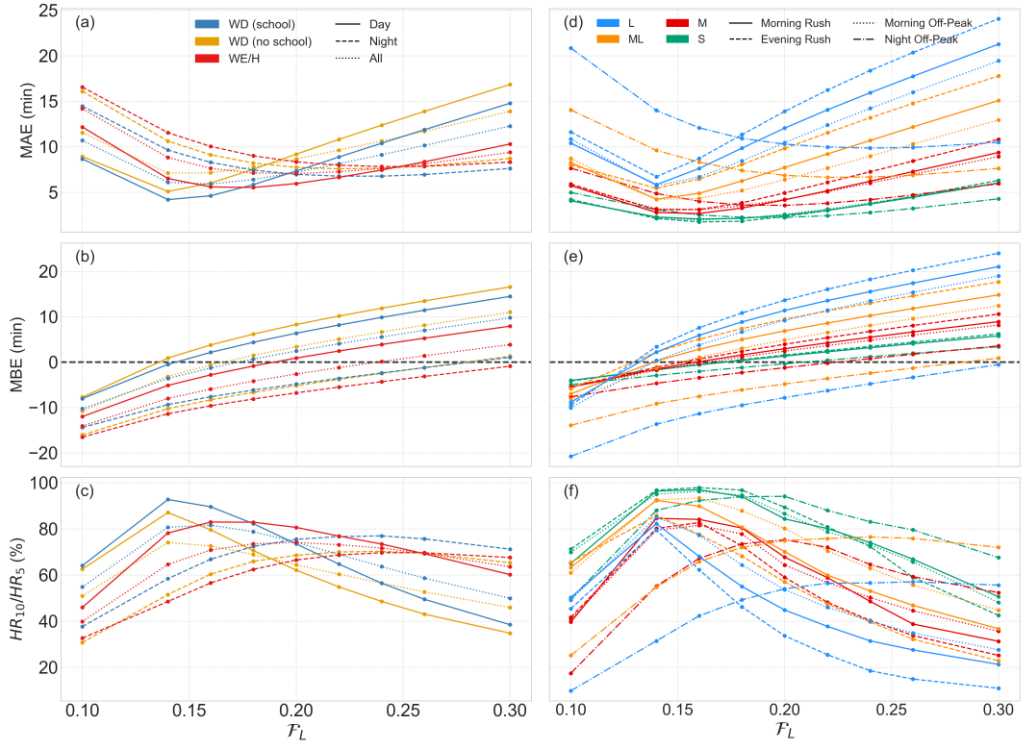


Figure 7: Evaluation of MATSDA travel time (#3, Table 3, Sect. 3.2.3) with effective local road length fraction ($\mathcal{F}_L$ 0.1→0.3) for different **(a-c)** day types (colour) by time of day (line type; 07:00–19:00 day, 19:00–07:00 night, 24 h all) and **(d-f)** length of trip (colour; S: 0-15, M: 15-30, ML: 30-60, L: >60 min) by time of day (lines type; morning rush (07:00–9:00), morning off-peak (9:00-16:00, evening rush (16:00–

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

19:00) and night off-peak (19:00-07:00), assessed with **(a, d)** mean absolute error (MAE, Eq. 6), **(b, e)** mean bias error (MBE, Eq. 7), and hit rate with threshold of **(c)** ±10 min (HR_{10}, Eq. 8) and **(f)** ±10 min for L and ML and ±5 min (HR_5, Eq. 8) for S and M routes.

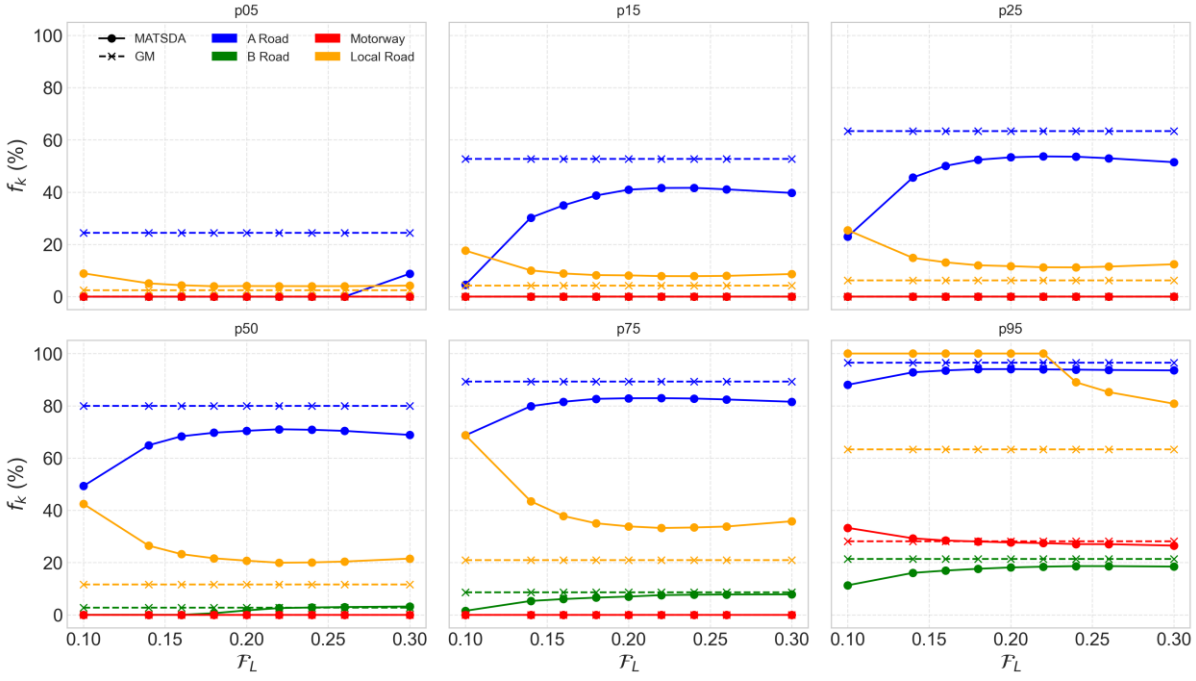


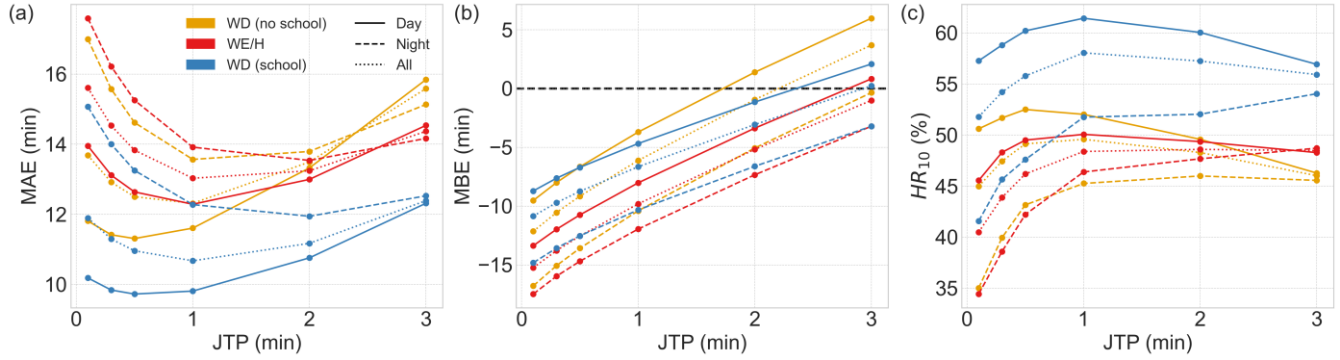
Figure 8: Road-type usage ratio ($f_{k,MATSDA}$ and $f_{k,GM}$, Eq. 4) as a function of prescribed effective local road length fraction ($\mathcal{F}_L$) for different road types (colour; A/B/L/M) from two sources (lines type; GM/MATSDA) for WD_{NS} day type (run#3 Table 3, section 3.2.3) with both data sorted to show behaviour at six percentiles of the distribution (subplots p: 5, 15, 15, 50, 75, 95). Other day types are shown in Fig. S8 and S9.

4.1.2 Junction time penalty (*JTP*)

A junction-based time penalty parameter (*JTP*) is introduced to account for delays when road-type changes occur (e.g. local⟶A-road transition) associated with probable slowing down and subsequent acceleration. Assigning a time penalty should decrease road type transitions enroute for a Dijkstra algorithm optimising the overall travel time. To assess impact of *JTP,* a series of constant values (0.1 to 3 min) are assigned to all road-type transitions (e.g. local↔major, major↔major), so not altering the road type mix between simulations (#5, Table 3, Sect. 3.2.3).

Longer *JTP* reduces MATSDA's overall tendency to underestimate travel times (Fig. 9b), with a *JTP* around 2 min minimizing MBE across day types and times of day. However, other metrics suggest the optimal *JTP* may vary by time of day (Fig. 9), with the lowest MAE and highest HR at 0.25 to 0.75-min during the daytime. Increasing the *JTP* beyond 0.75-min degrades daytime performance, but improves nighttime accuracy. To capture typical commuting patterns, daytime performance is crucial. A non-uniform *JTP* should consider extra time for busier roads. However, for the uniform case a *JTP* of 0.25-min is selected as the most suitable value.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026



425 **Figure 9:** As Fig. 7a-c, but for junction time penalty (*JTP*; #5 Table 3) with values between 0.1 and 3 min (points), assessed with (a) mean absolute error (MAE, Eq. 6), (b) mean bias error (MBE, Eq. 7), and (c) hit rate (threshold: ±10 min, HR_{10}; Eq. 8) for different day types (colour) and times of day (solid: day (07:00–19:00), dashed: night (19:00–07:00), and dotted: all times).

4.2 Temporal variability

An independent evaluation dataset (Table 2) is used to assess car travel times across three scenarios. First, we assess the impact 430 of the speed choice (Fig. 10d). This assessment is necessary because the MATSDA input data for London provides only an average speed for two flow directions, as geographical or lane information is unavailable (Sect. 3.1.2). When the maximum speed (#2, Table 3) is used, MATSDA's large overestimation of travel times is reduced *cf.* using the minimum speed (#1), with smaller MBE and MAE at all times of day (Fig. 11a,b).

Second, including detailed carriageway information (Sect. 3.1.1) for all major roads in MATSDA's nodal network 435 description further improves accuracy (#3, Fig. 10e) compared to only considering road numbers (#1, #2; Fig. 10d). MATSDA's median travel times for #3 (Fig. 10e) are centred around the 1:1 line with the GM reference, partially due to compensating errors from overestimating (underestimating) travel times during the daytime (night) rush hours (Fig. 11d). However, the daytime overall MBE is very low.

Third, the impact of effective local road length fraction ($\mathcal{F}_L$) on overall travel times are explored, with two $\mathcal{F}_L$ choices: (a) 440 constant (Eq. 3; #3), and (b) distance-varying (#4). MATSDA's median travel times for #4 are systematically slower than the reference times, especially for longer trips (Fig. 10e). As smaller $\mathcal{F}_L$ values (Sect. 4.1.1) result in less use of local roads, shorter travel times can occur even at the same speed, creating a more evident bias at night, whilst the daytime MBE is centred on 0-min (Fig. 11d). For #5 the underestimated travel times (i.e., MATSDA is too fast) is partially compensated for using a non-zero junction time penalty (*JTP*) when transitioning between road-nodes. This improves the night-time MAE and MBE (Fig. 445 11d,e), demonstrating MASTDA's capacity for more complex physical representations, allowing it to realistically model time delays for transitions.

For #6 and #7 (Table 3, Fig. 10f) the full travel database domain is used in MATSDA's Pathfinder (*cf.* #3 and #5, respectively). When not constrained to the GM paths, MATSDA selects its own optimal routes through neighbourhoods in the domain. Despite this freedom, #6 and #7 maintain a reasonable level of accuracy regarding travel durations for both shorter 450 and longer trips, but with a clear systematic bias towards faster travelling (travel-time underestimated; Fig. 10f). This is

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

particularly evident at night, when MATSDA is up to 25-min faster for longer routes (Figure 13g), whereas daytime is well within ± 5-min for shorter routes (≤14 km) or around ±10 min for longer routes (>14 km).

MATSDA's travel time metrics, for all modelling cases considered, have a distinct diurnal pattern (Fig. 11). Compared to the GM reference, MATSDA has longer travel times during the morning and evening rush hours and faster travel during off-peak periods, especially at night, aligning with real-world congestion dynamics that reflect the diurnal traffic speed patterns of the input data (Fig. 4). GM vehicle speeds (derived from extracted route segment lengths and travel durations) have much smaller diurnal variations, notably without distinct differences between off-peak and rush hours (see Fig. S11). This difference in the assumed vehicle speeds needs to be considered when interpreting MATSDA and GM comparisons. MATSDA's MBE has positive peaks (slower travel) during the morning (~07:00–09:00) and evening (~16:00-18:00) rush hours (Fig. 11a,d). However, for shorter trips (≤14 km) during daytime the MBE magnitude is small (within ± 5 min) for #2-#5 and within ±10 min for the remaining cases. Overnight (~22:00-04:00) MATSDA consistently underestimates travel times (faster travel than predicted by GM; MBE < 0-min; Fig. 11a,d,g). This temporal variation in bias directly impacts MATSDA's hit rate (HR, Fig. 11c,f,i) which is lowest (i.e., poorest) for all cases at night when the bias is largest (especially for longer trips).

MATSDA's time- and neighbourhood dependent input speed data (Digimap Pilot Collection 2024; Fig. 4) distinguishes weekday and weekend (WE/H), but not different weekday types (with and without school; WD_S and WD_{NS}). Performance metric differences between WD_{NS} and WD_S therefore arise from the GM data, with shorter daytime travel times on WD_{NS} than WD_S, but relatively small differences at night. The diurnal peaks of MBE for weekends and bank holidays (WE/H) are less pronounced, probably linked to less congestion on non-workdays.

With different periods of data collection (GM route: 12/2024-03/2025 vs vehicle speed: 09/2023-02/2024), results can differ within neighbourhoods because of local disruptions, changes and/or activities (e.g., construction, road closures, speed limit changes; TfL 2023). For the London road network, the GM data (Fig. S11) and the Digimap Pilot Collection (2024) have good agreement across all boroughs for speed limits (Fig. S10). Discrepancies are apparent from local policy changes, such as speed limit adjustments implemented in 2024 for some TfL A-roads (TfL 2021). These changes exist in the GM data, as do other real-time traffic adjustments.

While systematic biases exist, MATSDA demonstrates a strong capability to model the temporal dynamics of urban traffic realistically, reflecting the relative changes in travel time throughout different times of day and across different day types. The model could reflect more detailed day types when such data are available.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

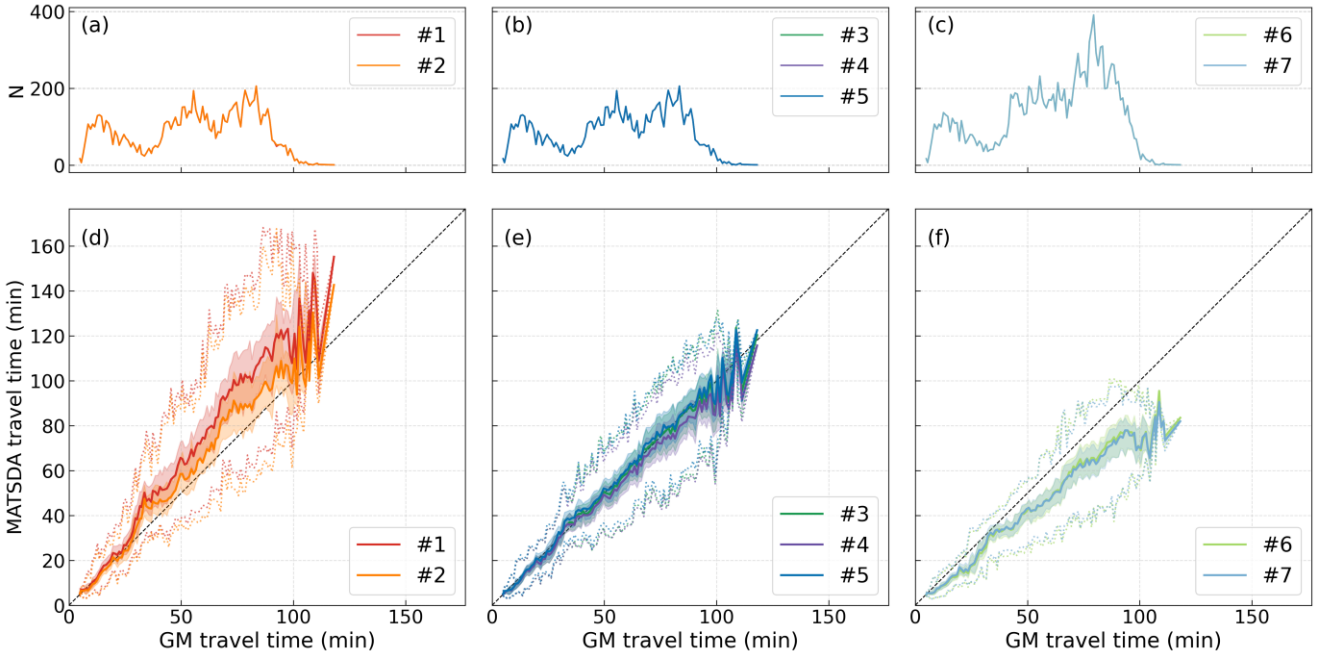


Figure 10: Evaluation of MATSDA travel time with GM for #1-#7 (colours; Table 3) shown for 1-min bins of **(a-c)** varying sample sizes (***N***), indicating median (solid line), interquartile range (shading), full range (dashed) and 1:1 (black) for **(d)** baseline runs (#1, #2): comparing input speeds (min vs. max) using the road number (RN) network, **(e)** parameter sensitivity runs (#3-5): comparing effective local road length fraction ($\mathcal{F}_L$) and junction time penalty (*JTP*), all using the detailed carriageway type (RN + CWT) network on spatially consistent routes (SCR), **(f)** full domain runs (#6, #7): Extending from *SCR* to *full*, unconstrained domain (*cf.* #3, #5).

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

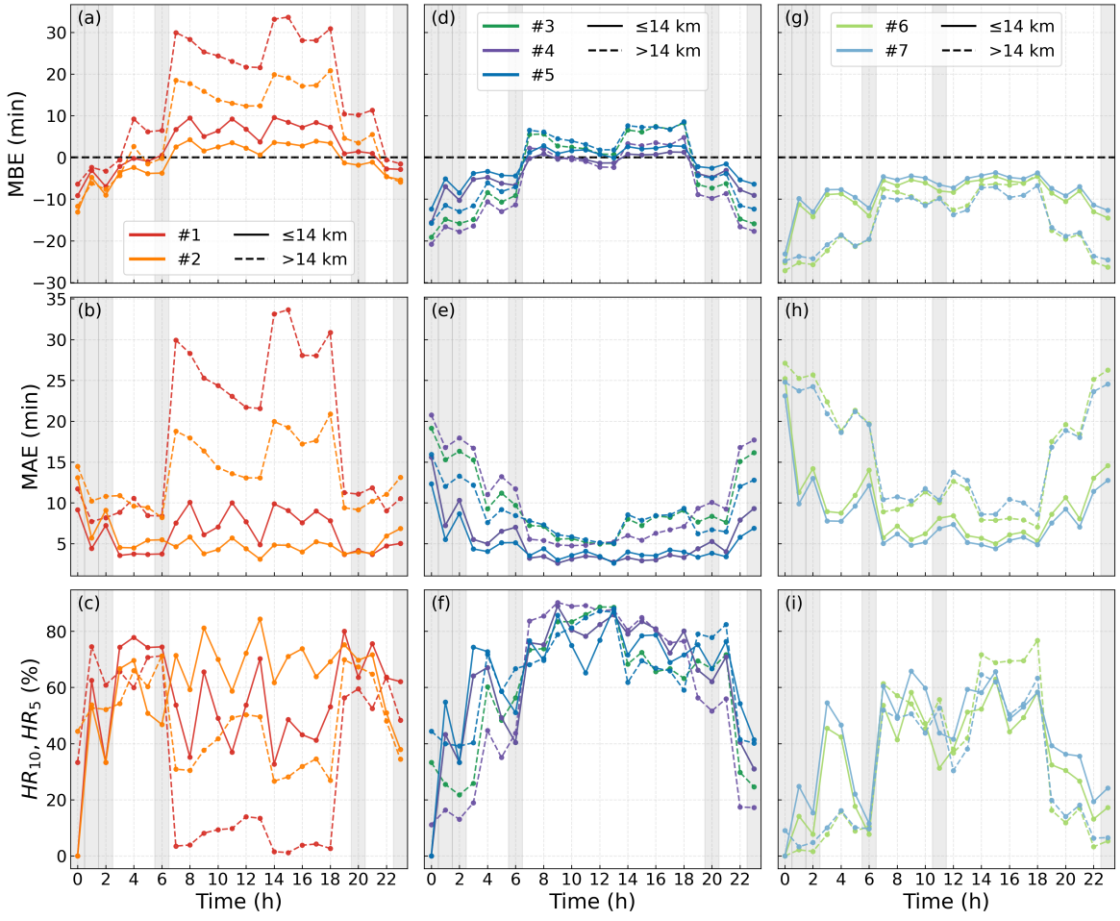


Figure 11: Evaluation of MATSDA travel time by time of day for the 7 run types (colours, Table 3) shown for two O-D distances (solid: ≤14 km; dashed: >14 km): **(a, d, g)** MBE, **(b, e, h)** MAE, **(c, f, i)** 5 min (HR_5) for short routes (≤14 km) and 10 min (HR_{10}) for long routes (>14 km). Hours with low sample sizes (< 25 % of ***N***) indicated by grey shading.

4.3 Spatial variability

The spatial accuracy of MATSDA's unconstrained pathfinding evaluated using the full dataset (Table 2) is assessed using the FSS (Eq. 9) for #6 and #7 (Table 3).

The FSS analysis begins by including all 500-m grid-cell neighbourhoods traversed by at least one route in the GM full dataset (Fig. 13a), which corresponds to a minimum route count per grid of 1 (Sect. 3.2.5). At this least restrictive threshold, both configurations demonstrate high spatial skill (Fig. 12). The analysis then proceeds iteratively, increasing this minimum threshold to focus only on grid-cells that are part of more frequently used routes (i.e., higher counts in the GM full dataset). Excluding extreme thresholds >1000, the mean FSS is ~0.69 and 0.62 for #6 and #7, respectively. While for thresholds <10, FSS is steadily ~0.76 and 0.74 for #6 and #7, respectively, i.e., well above the 0.5 threshold often used to indicate a 'useful' modelling result (Roberts and Lean 2008). This shows that MATSDA's pathfinding results in routes that are spatially

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

comparable to the GM reference for a majority of O-D pairs. As the grid count threshold increases (Fig. 12), the evaluation focuses on areas that are traversed by a larger number of routes, the FSS for both MATSDA cases gradually declines. This trend suggests that while MATSDA is skilled at selecting plausible and spatially similar routes overall, it does not always show the same routing preferences along certain traffic arteries as evident in GM. However, the FSS for both cases still is >0.6 at threshold around 100. For a threshold of ~1000, the FSS for #6 increases back to 0.7, which can also be seen as the hot spots in the spatial maps shown in Fig. 13. Across almost the entire range of thresholds, #6 consistently has a higher FSS than #7. This indicates that adding a junction time penalty (#7) results in improved accuracy of travel time predictions (Fig. 10e), but is associated with a spatial change in the routing (Fig. 13b,c). The simpler model configuration (i.e., #6) produces routes that are, on average, in better spatial agreement with the GM reference data.

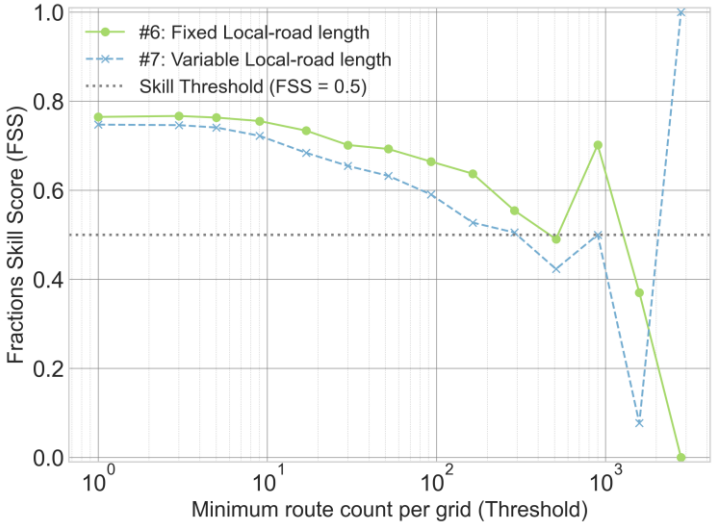


Figure 12: Fractional skill scores (FSS; Eq. 9) for MATSDA #6 and #7 (Table 3, Sect. 3.2.3) as a function of the grid count threshold. The analysis is progressively restricted to routes that appear at least ***N*** times (Threshold) in the GM evaluation-full reference dataset (Table 2). The dotted line at FSS = 0.5 indicates the threshold for 'useful skill' (Roberts and Lean 2008).

Deeper insights into these routing differences can be gained from a visual comparison of the route density maps (Fig. 13). First, MATSDA optimizes routes based solely on travel duration. It does not account for policy-driven route preferences, such as avoiding London's Congestion Charge Zone (CCZ) or Ultra Low Emission Zone (ULEZ). In contrast, GM's travel model incorporates these cost factors, leading it to route traffic around the city centre more often. This is shown in Fig. 13a-c, where MATSDA directs more routes through the central area (Fig. 13b,c) compared to GM (Fig. 13a), which shows a more diffuse pattern that avoids the very centre. Second, MATSDA's pathfinding is based on a static road network and average travel speeds, whereas GM's traffic model has historical data and real-time information on road closures, construction, and accidents. Furthermore, MATSDA's network does not currently capture road directionality (i.e., one-way streets). This simplification can lead to significant routing differences, particularly in dense urban areas where GM avoids certain routes that MATSDA incorrectly assumes are efficient or bidirectional corridors. These differences are also apparent in certain regions. For example,

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

the difference maps (Fig. 13d, e) reveal that MATSDA generates more travel routes in South-West London compared to GM. This is likely because MATSDA identifies major A-roads in this area as the theoretically fastest paths, while GM's model, informed by real-world data, accounts for local congestion and other factors that make these routes less optimal in practice. Finally, the initial choice of MATSDA's start node (local roads) can have a cascading impact on the entire route. For example, the GM model might initiate a route by snapping a start coordinate to the nearest road segment, independent of its type, potentially placing a journey onto a major road immediately. MATSDA, by contrast, is constrained to begin on a local road, which can lead to different initial pathing decisions and, consequently, a different overall route.

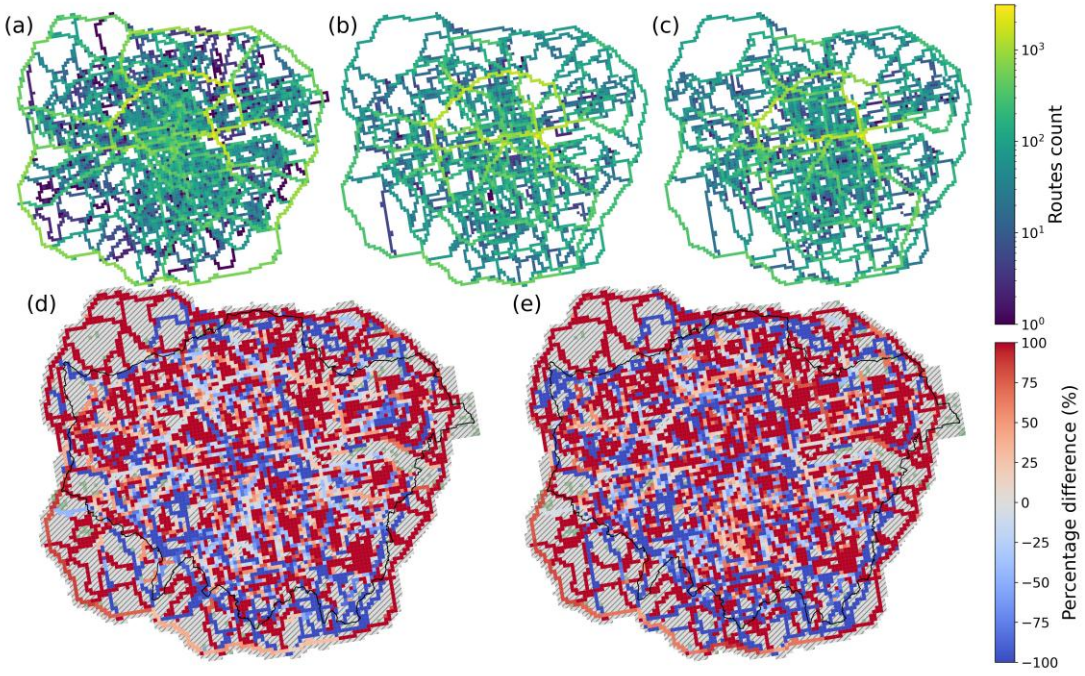
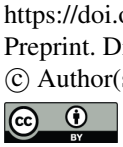


Figure 13: (a-c) Counts of routes traversing through each 500 m grid-cell neighbourhood of the Greater London model domain (Fig. 1a, c) for **(a)** GM evaluation-full reference routes (Tabel 2, Sect. 3.2.3), **(b)** MATSDA #6 (Table 3), **(c)** MATSDA #7. Note that route counts are plotted on a logarithmic scale. **(d-e)** Count percentage differences calculated as (GM – MATSDA) / 0.5* (GM + MATSDA) for **(d)** #6 and **(e)** #7. Grid-cells through which none of the routes led are white in **(a-c)** and hatched in **(d, e)**, with neighbourhoods without any roads highlighted in green in **(d, e)**. Colour bar in log scale for **(a-c)** and in percentage for **(d-e)**.

Figure 14 compares probability distributions for the fraction of distance travelled on each road type (Motorway, A, B, and Local) between MATSDA cases (Table 3, Sect. 3.2.3) and the GM reference data. The GM data establishes a baseline for typical long-distance car journeys in London (Fig. 14a). It shows that most journeys use a relatively small fraction of local and B roads, a dominant fraction of A roads, and a distinct bimodal usage of Motorways (i.e., a route either relies heavily on a motorway for most of its length or not at all). The inclusion of carriageway types in #3 marks a critical improvement. This added detail allows MATSDA to create a more closely aligned usage profile of A- and local roads *cf.* the GM reference (Fig. 14d). When MATSDA is allowed to find its own optimal route across the full domain (#6 and #7; Fig. 14g,h), it exhibits a stronger preference for B roads, especially for #7 with a corresponding decrease in A-road usage. In addition, #7 also has a

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026



higher usage of local roads than the other cases, which indicates that MATSDA's pure travel-time optimization will select the theoretically fastest road class for long journeys. However, #6 is showing a road-usage profile that is well-aligned with the GM reference, which is consistent with the previous spatial assessments (Figs. 12, 13).

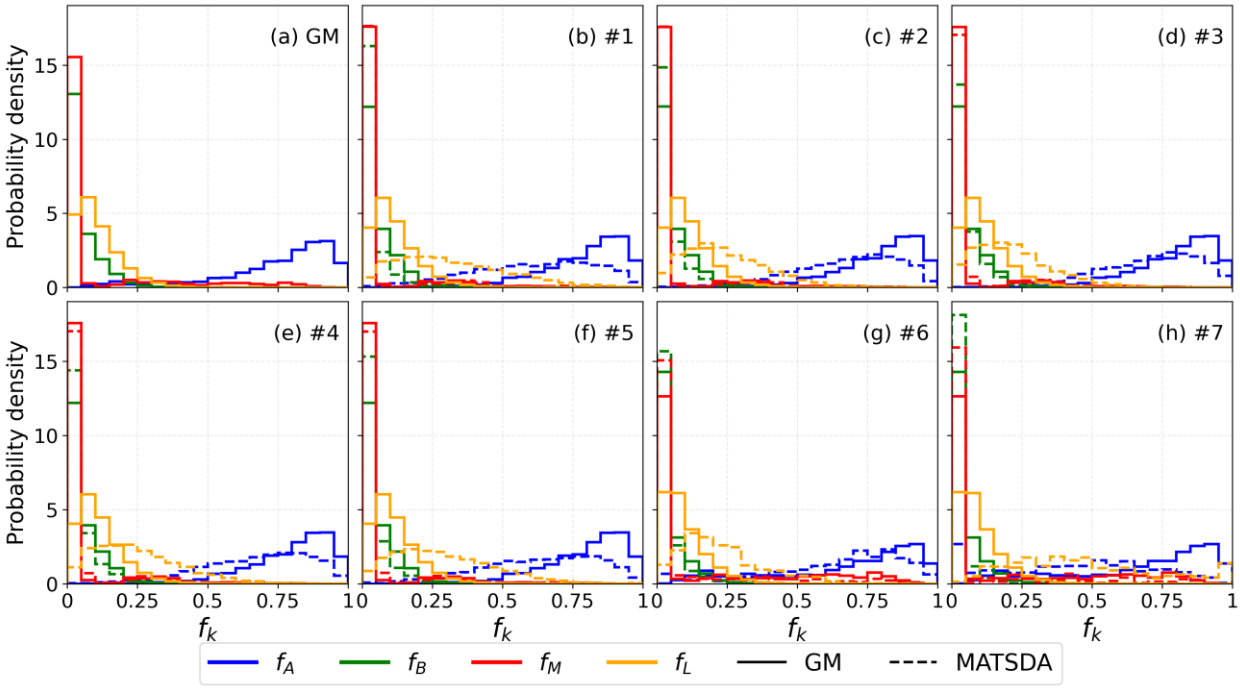


Figure 14: Probability density of road-type usage fractions ($\boldsymbol{f_k}$ with $\boldsymbol{k}$=A, B, M, L, Eq. 4) of A ($\boldsymbol{f_A}$, blue), B ($\boldsymbol{f_B}$, green), local ($\boldsymbol{f_L}$, orange) roads and motorways ($\boldsymbol{f_M}$, red) derived from: **(a)** the GM reference data and **(b-h)** MATSDA #1-7 (Table 3, Sect. 3.2.3) for journeys >14 km. Line styles distinguish between GM (solid) and MATSDA cases (dashed). For journeys ≤ 14 km see Fig. S5.

5. Conclusions

To balance realism with computational cost, the MATSDA-roads v2.0 travel database represents the complex road network using a nodal system. Major roads within this network are defined not just by their number (e.g., A406) but are further distinguished by their specific carriageway type (e.g., dual carriageway, slip road, roundabout). This level of detail is essential because it prevents the artificial averaging of speeds from different road designs that may share the same road number, such as a fast dual carriageway and a slower single carriageway, which could otherwise introduce compensating errors. Furthermore, the results demonstrate that the inclusion of distinct carriageway types is crucial for achieving accurate travel time estimations, particularly for longer journeys. This is most evident in the performance improvement from #2 (using only road numbers for major roads) to #3 (adding carriageway type information, Table 3, Sect. 3.2.3). The latter configuration significantly reduces both MAE (Eq. 6) and MBE (Eq. 7). This finding underscores that the computational and data processing costs associated with implementing detailed road network data are justified by the substantial gains in model fidelity, which can be served as a guideline for future model development and data collection works in other metropolitan areas. However, #1 and #2 (less road

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

resolution) are found performing well for short time travel (e.g., below 30 min), which can be a good benchmark for future different cities' transport road network database generation (i.e., it is possible to sacrifice road resolution for short time travel if computationally expensive or high-resolution road data not available).

The analysis identified the representation of effective local road lengths and the delays at junctions as two primary factors for model uncertainty. The parameterization of these elements not only improved model accuracy but also yield quantitative insights into real-world travel behaviours. The empirically derived constant effective local road length fraction ($\mathcal{F}_L$) value of 0.16 suggests that for a typical journey traversing a 500 m neighbourhood, only about 16% of the total available local road length constitutes the effective, through-traffic path. This reflects the inherent efficiency of urban travel, where drivers utilize a small subset of the dense local network to connect to faster arterial roads. With longer journeys, $\mathcal{F}_L$ is decreasing. MATSDA is further refined by implementing an O-D straight-line distance-based parameterization for $\mathcal{F}_L$, which captures the behavioural nuance that drivers prioritize high-speed roads more strongly on longer trips, thereby minimizing travel on local roads.

Spatially, the model demonstrates high skill in identifying major travel corridors. The unconstrained pathfinding cases achieve FSS values ~0.7, indicating that MATSDA's route choices are spatially consistent with the GM reference data. Given the resolution of the model, minor positional errors of MATSDA are expected but do not negate the practical value of the route choice. In addition, performances difference between #6 and #7 reveal that a simplified configuration may produce more spatially aligned routes.

5.1 Limitations and future directions

1. *Local road representation:* MATSDA's aggregation of local roads into a single node in each neighbourhood is introduced to keep both the generation of the travel database and running of the model computationally feasible when applied to large cities with a vast and diverse road network. This greatly reduces the detail with which the local road network is represented and requires the introduction of an effective local road length in the simulation to model travel times. In this study, the magnitude of the effective road length is derived from analysis of reference data for a single city, but differences between cities are likely making case-by-case assessments necessary. Furthermore, if the model were applied to smaller cities/towns or very suburban neighbourhoods, the prevalence of local roads *cf.* major roads is much higher compared to the London network analysed here. In such cases, treating the local roads as a continuum would be unsuitable and the generation of a detailed nodal representation of local roads is advised. Given the smaller overall spatial extent of the road network in small cities/towns this should be computationally feasible.
2. *Uniform junction time penalty*: Time penalties for the change of road-nodes in the current system apply uniformly to all junction types (e.g., simple T-junctions, multi-lane roundabouts, motorway junctions). In future model versions, a junction-type dependent functionality can be implemented to model the impact of time-delays more realistically. This could address the observed trade-off between spatial and temporal accuracy (Sect. 4.2, 4.3), allowing the model to make more human-like decisions at road network transitions.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

3. *Road directionality*: Depending on the nature of the road-network input data, information on whether roads are one-way or bidirectional may not be available, as is the case in the present application of MATSDA in London. This can contribute to routing and travel duration differences. While the current implementation of MATSDA's nodal network does not consider directionality of roads, this aspect can be implemented if input data allow for this.
4. *Path weighting:* While in this study MATSDA's pathfinding algorithm optimises routes based on travel-time weights of the nodes in the travel database, other or additional weights can be used. The spatial analysis of the travel-time optimised MATSDA routes (Sect. 4.3) showed spatial difference of travel corridors in the city centre and some of the outskirts of the city when compared to the reference data. This suggests that other factors influence route choice. Incorporating criteria such as travel cost (e.g., congestion or low-emission zone charges), fuel consumption, or proxies for route simplicity (e.g., a penalty for frequent road transitions or a preference for staying on major roads) can make the modelled routes more realistic and better reflect complex human decision-making.
5. *Dependence on input data:* The nature and detail of the input data (geospatial road and junction data, traffic speed data) play a crucial role for the level of realism that can be captured in the MATSDA travel database (Sect. 2.2, 3.1) and hence in the Pathfinder results. For data-rich cities like London, reconstructing a realistic representation of the major traffic arteries becomes possible, but adjustments to the detail of MATSDA's nodal network may need to be made in case of less data detail.
6. *Transport modes:* In this study, MATSDA is demonstrated for car travel within London's road network. Other transport modes (e.g., walking, bicycling) can be modelled in the same way and require the generation of additional travel databases that reflect different travel speeds and roads available for certain modes (e.g., no walking/cycling allowed on motorways, etc.). Similarly, modelling public transport routes with MATSDA-metro (e.g., bus, tube/subway, train) requires the nodal model system to reflect the connections of different lines and services between stops/stations on the road and rail network.

MATSDA's modelled travel routes can be used as input for other types of models, such as agent-based models that represent human behaviour and its impact on urban climate variables. In a next step, MATSDA will be linked to the model DAVE (Hertwig et al. 2025b) to establish a dynamic, two-way feedback loop. MATSDA provides realistic travel times that influence agent decisions on mode choice or departure time within DAVE. In turn, the resulting spatial and temporal distribution of citizens in different modes of transport from DAVE can be used to create dynamic traffic flow patterns that modify the travel time weights within MATSDA's network. This deeper integration creates a versatile tool for exploring complex scenarios, such as assessing the city-wide impact of remote work policies or new transport infrastructure on traffic congestion, energy consumption, and urban heat emissions.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

Appendix A: List of symbols and acronyms

	Description	Sect.	Eq./Fig.
Acronyms			
A	A road (UK)	3.1.1	Fig. 1b,c
B	B road (UK)	3.1.1	Fig. 1b,c
CCZ	Congestion Charge Zone	4.3	
CWT	Carriageway Type (road network representation)	3.2.3	Table 3
FSS	Fractions Skill Score, a measure of spatial accuracy	3.2.5	Eq. 9, Fig. 12
GM	Google Maps	3.2	
HR	Hit Rate, the percentage of predictions within a given tolerance W	3.2.4	Eq. 8, Fig. 7,11,13
HR_5	5 min tolerance for hit rate	3.2.4	Eq. 8
HR_{10}	10 min tolerance for hit rate	3.2.4	Eq. 8
I	Indicator function (equals 1 if a condition is met, 0 otherwise)	3.2.4	Eq. 8
JTP	Junction time penalty, a fixed delay added when changing between road-nodes	3.2.3	Fig. 9
L	Local road	3.1.1	Fig. 1b
M	Motorway	3.1.1	Fig. 1c
MAE	Mean Absolute Error, a measure of the average magnitude of error	3.2.4	Eq. 6, Fig. 7,11,13
MATSDA	Movement And Transport Simulations using Dijkstra's Algorithm	2	Fig. 3
MBE	Mean Bias Error, a measure of the average systematic error	3.2.4	Eq. 7, Fig. 7,11,13
O-D	Origin-Destination	2	Fig. 5
RN	Road Number (road network representation)	3.2.3	Table 3
SCR	Spatially Consistent Routes	3.2.3	Table 3
ULEZ	Ultra Low Emission Zone	4.3	
W	Tolerance threshold for the Hit Rate calculation	3.2.4	Eq. 8
Variables			
d	Straight-line distance between origin and destination	4.1.1	Fig. 5
$f_{k,GM}$	Fraction of a route's total length travelled on a specific road type k (GM)	3.2.3	Eq. 4, Fig. 8,16
$f_{k,MATSDA}$	Fraction of a route's total length travelled on a specific road type k (MATSDA)	3.2.3	Eq. 4, Fig. 8,16
$\mathcal{F}_L$	Effective local road length fraction, representing the proportion of local roads used in grid-cell / neighbourhoods for a route $\mathcal{F}_L = \lambda_L / l_L$	3.2.3	Eq. 3, Fig. 6,7,10
i	Index for routes (for temporal evaluation)	3.2.4	Eq. 6-8
(i,j)	Indices for grid points (for FSS)	3.2.5	Eq. 9
k	Index for road types, UK, London case: M (motorway), A-, B-, L(local)- roads	3.1.1	Eq. 4,5, Fig. 1b,c
λ_L	Total local road length used in a neighbourhood (GM) $\lambda_L = \Sigma\lambda_{L,s}$	3.2.3	Eq. 3
λ_s	Length of a road segment in a grid-cell / neighbourhood (GM)	3.1.2	Eq. 1
Λ_k	Total distance travelled on a specific road type k for a given route (GM) $\Lambda_k = \Sigma\lambda_k$	3.2.3	Eq. 4
Λ_{tot}	Total length of a complete travel route (GM) $\Lambda_{tot} = \Sigma\lambda_n$	3.2.3	Eq. 4
l_L	Total available local road length in a grid-cell / neighbourhood (input data) $l_L = \Sigma l_{L,s}$	3.2.3	Eq. 3
l_s	Length of a road segment in a grid-cell / neighbourhood (geospatial input data)	3.1.1	
ℓ_n	Length of a MATSDA road-node type within a grid-cell / neighbourhood (except for Local road) $\ell_n = \Sigma l_{n,s}$ $(\ell_L = \Sigma l_{L,s} * \mathcal{F}_L)$	3.1.1	Eq. 2
$\mathcal{L}_k$	Total distance travelled on a specific road type k for a given route (MATSDA) $\mathcal{L}_k = \Sigma\ell_k$	3.2.3	Eq. 4
$\mathcal{L}_{tot}$	Total length of a complete travel route (MATSDA) $\mathcal{L}_{tot} = \Sigma\ell_n$	3.2.3	Eq. 4
n	Index for a road-node (e.g., A1_R, B1_CDC, M1_SR, L…)	3.1.1	Eq. 2, Fig. 2
N	Total number of routes in an evaluation or training dataset	3.2.3	Table 2
s	Index for different road segments	3.1.2	Eq. 1
$\tau_{n,t}$	Travel duration of a road-node for a specific time period	3.1.2	Eq. 2

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

t	Index for different time periods	3.1.2	Eq. 1,2
T_t	Total travel duration for a route for a specific time period $T_t = \Sigma\tau_{n.t}$	3.2.4	Eq. 6-9
$T_{i,t,GM}$	Observed (reference) travel duration for a route at specific time period t from Google Maps	3.2.4	Eq. 6-9
$T_{i.t.MATSDA}$	Predicted (modelled) travel duration for a route at specific time period t from MATSDA	3.2.4	Eq. 6-9
$V_{t,s}$	Speed on a specific road segment for a time period	3.1.2	Eq. 1, Fig. 4
V_t	Average speed for a road-node for a specific time period	3.1.2	Eq. 1,2

Code and data availability

MATSDA-roads v2.0 model codes and User Manual can be found at https://doi.org/10.5281/zenodo.17736682 (Ma et al. 2025a). Processing code together with model data (MATSDA-roads v2.0 travel database) used in this study and code for processing of the Google Maps (GM) reference routes are available at https://doi.org/10.5281/zenodo.17521112 (Ma et al. 2025b). Model outputs and corresponding plotting scripts for each figure used in the main paper are at https://doi.org/10.5281/zenodo.17736562 (Ma et al. 2025c). Further details are provided in the Supplementary Material (S), which includes S1 (MATSDA-roads v2.0 travel database generation), S2 (MATSDA-roads v2.0 pathfinding output), S3 (GM reference data processing and dataset split) and S4 ($\mathcal{F}_L$ parametrization methodology).

Author contribution

MM designed and implemented the initial MATSDA framework. TM further developed MATSDA-roads v2.0, performed all analyses, and wrote the original paper draft. DH & SG provided supervision throughout the whole project, with MM providing additional supervision early in the project. DH & SG provided methodological and scientific input at all stages of the project. MM & MP contributed to the methodological input and analysis of model-related results early in the project. The project was originally conceived by DH, MM and SG. All co-authors reviewed the manuscript and provided feedback.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments and financial support

We gratefully acknowledge funding by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 855005; project *urbisphere*) and through the UKRI Natural Environment Research Council (NERC) projects APEx (NE/T001887/2) and ASSURE (NE/W002965/1). We would like to thank Dr. Sylvia Bohnenstengel (Met Office@Reading) for her valuable contributions to the manuscript, including her helpful comments and review.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

References

Afshari, A., Schuch, F., and Marpu, P.: Estimation of the traffic related anthropogenic heat release using BTEX measurements – A case study in Abu Dhabi., Urban Clim., 24, 311–325, https://doi.org/10.1016/j.uclim.2017.02.001, 2018.

Ahuja, R. K., Magnanti, T. L., and Orlin, J. B.: Network Flows: Theory, Algorithms, and Applications., Prentice Hall, Upper Saddle River, 1993.

Bing, H. and Lai, L.: Improvement and Application of Dijkstra Algorithms., Acad. J. Comput. Inf. Sci., 5, 97–102, https://doi.org/10.25236/AJCIS.2022.050513, 2022.

Blunn, L., Xie, X., Grimmond, S., Luo, Z., Sun, T., Perera, N., and Ratnayake, R.: Spatial and temporal variation of anthropogenic heat emissions in Colombo, Sri Lanka., Urban Clim., 54, 101828, https://doi.org/10.1016/j.uclim.2024.101828, 2024.

Britter, R. and Schatzmann, M. (Eds.): Model evaluation guidance and protocol document., Technical report of COST Action 732, University of Hamburg, Germany, 28 pp., 2007.

Capel-Timms, I., Smith, S. T., Sun, T., and Grimmond, S.: Dynamic Anthropogenic activitieS impacting Heat emissions (DASH v1.0): development and evaluation., Geosci. Model Dev., 13, 4891–4924, https://doi.org/10.5194/gmd-13-4891-2020, 2020.

Casey, B., Bhaskar, A., Guo, H., and Chung, E.: Critical Review of Time-Dependent Shortest Path Algorithms: A Multimodal Trip Planner Perspective., Transport Reviews, 34, 522–539, https://doi.org/10.1080/01441647.2014.921797, 2014.

Cleveland, W. S.: Robust Locally Weighted Regression and Smoothing Scatterplots., J. Am. Stat. Assoc., 74, 829–836, https://doi.org/10.1080/01621459.1979.10481038, 1979.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C.: Introduction to algorithms., MIT press, 2009.

DfT: Guidance on road classification and the primary route network., [online] Available from: https://www.gov.uk/government/publications/guidance-on-road-classification-and-the-primary-route-network/guidance-on-road-classification-and-the-primary-route-network (Accessed 2 July 2025), 2012.

Digimap Pilot Collection: Average Speeds, Scale 1:10000, Tiles: GB, Updated: 31 August 2023, Basemap, Using: EDINA Pilot Digimap Service, [online] Available from: https://digimap.edina.ac.uk (Accessed 1 March 2024), 2024.

Dijkstra, E. W.: A note on two problems in connexion with graphs., Numer. Math., 1, 269–271, https://doi.org/10.1007/BF01386390, 1959.

Google: Google Maps Directions API., [online] Available from: https://developers.google.com/maps/documentation/directions (Accessed 23 December 2024), 2025.

Grimmond, C. S. B.: The suburban energy balance: Methodological considerations and results for a mid-latitude west coast city under winter and spring conditions., Int. J. Climatol., 12, 481–497, https://doi.org/10.1002/joc.3370120506, 1992.

Hall, T. W., Blunn, L., Grimmond, S., McCarroll, N., Merchant, C. J., Morrison, W., et al.: Utility of thermal remote sensing for evaluation of a high-resolution weather model in a city., Q. J. R. Meteorol. Soc., 150, 1771–1790, https://doi.org/10.1002/qj.4669, 2024.

Hertwig, D., McGrory, M., Paskin, M., Liu, Y., Lo Piano, S., Llanwarne, H., Smith, S. T., and Grimmond, S.: Multi-scale harmonisation Across Physical and Socio-Economic Characteristics of a City region (MAPSECC): London, UK., Zenodo, https://doi.org/10.5281/zenodo.12190341, 2024.

Hertwig, D., McGrory, M., Paskin, M., Liu, Y., Lo Piano, S., Llanwarne, H., Smith, S. T., and Grimmond, S.: Connecting physical and socio-economic spaces for multi-scale urban modelling: a dataset for London., Geosci. Data J., https://doi.org/10.1002/gdj3.289, 2025a.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

Hertwig, D., McGrory, M., Liu, Y., Ma, T., Paskin, M., Smith, S. T., and Grimmond, S.: DAVE (Dynamic Anthropogenic actiVities and feedback to Emissions) Documentation and Manual (2025_1.0)., Zenodo, https://doi.org/10.5281/zenodo.11369893, 2025b.

Iamarino, M., Beevers, S., and Grimmond, C. S. B.: High-resolution (space, time) anthropogenic heat emissions: London 1970-2025., Int. J. Climatol., 32, 1754–1767, https://doi.org/10.1002/joc.2390, 2012.

Ivanchev, J. and Fonseca, J.: Anthropogenic Heat Due to Road Transport: A Mesoscopic Assessment and Mitigation Potential of Electric Vehicles and Autonomous Vehicles in Singapore., Technical report, ETH Zurich, 2020.

Kaur, G. and Tripathi, N.: Application of Dijkstra's Algorithm in Wireless Communication System., Int. J. Res. Appl. Sci. Eng. Technol., 8, https://ssrn.com/abstract=4331916, 2020.

Lean, H. W., Theeuwes, N. E., Baldauf, M., Barkmeijer, J., Bessardon, G., Blunn, L., et al.: The hectometric modelling challenge: Gaps in the current state of the art and ways forward towards the implementation of 100-m scale weather and climate models., Q. J. R. Meteorol. Soc., 150, 4671–4708, https://doi.org/10.1002/qj.4858, 2024.

Liu, Y., Luo, Z., and Grimmond, S.: Revising the definition of anthropogenic heat flux from buildings: role of human activities and building storage heat flux., Atmos. Chem. Phys., 22, 4721–4735, https://doi.org/10.5194/acp-22-4721-2022, 2022.

Louf, R. and Barthelemy, M.: How congestion shapes cities: from mobility patterns to scaling., Sci. Rep., 4, 5561, https://doi.org/10.1038/srep05561, 2014.

Ma, T., Hertwig, D., McGrory, M., Paskin, M., & Grimmond, S., MATSDA-roads (Movement And Transport Simulations using Dijkstra's Algorithm - roads) (v2.0). Zenodo. https://doi.org/10.5281/zenodo.17736682, 2025a.

Ma, T., Hertwig, D., McGrory, M., Paskin, M., & Grimmond, S., Model input and processing codes for "Transport modelling for dynamic urban climate studies: MATSDA-roads v2.0" (v1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17521112, 2025b.

Ma, T., Hertwig, D., McGrory, M., Paskin, M., & Grimmond, S., Model output and plotting codes for "Transport modelling for dynamic urban climate studies: MATSDA-roads v2.0" (v1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17736562, 2025c.

Mensah, D. N. A., Gao, H., and Yang, L. W.: Approximation Algorithm for Shortest Path in Large Social Networks., Algorithms, 13, 36, https://doi.org/10.3390/a13020036, 2020.

Mhmood Thabet, A. S. and Zengin, A.: Optimizing Dijkstra's Algorithm for Managing Urban Traffic Using Simulation of Urban Mobility (Sumo) Software., Ann. Math. Phys., 7, 206–213, https://doi.org/10.17352/amp.000124, 2024.

Oke, T. R., Mills, G., Christen, A., and Voogt, J. A.: Urban Climates., Cambridge University Press, Cambridge, https://doi.org/10.1017/9781139016476, 2017.

ONS: Office for National Statistics 2011: Location of usual residence and place of work (with outside UK collapsed) (OA/WPZ level)., [online] Available from: https://www.nomisweb.co.uk/census/2011/wf02ew (Accessed 10 March 2023), 2011.

Ordnance Survey: OS MasterMap Highways Network [GML3 geospatial data], Scale 1:2500, Tiles: GB, Updated: 26 May 2023, Using: EDINA Digimap Ordnance Survey Service, [online] Available from: https://digimap.edina.ac.uk (Accessed 9 April 2024).

OS Open Roads: Ordnance Survey Open Roads dataset. Version date: 2021-04., [online] Available from: https://osdatahub.os.uk/downloads/open/OpenRoads (Accessed 16 March 2024), 2021.

Qian, J., Zhang, L., Schlink, U., Meng, Q., Liu, X., and Janscó, T.: High spatial and temporal resolution multi-source anthropogenic heat estimation for China., Resour. Conserv. Recycl., 203, 107451, https://doi.org/10.1016/j.resconrec.2024.107451, 2024.

https://doi.org/10.5194/egusphere-2025-6060
Preprint. Discussion started: 19 February 2026

Roberts, N. M. and Lean, H. W.: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events., Mon. Weather Rev., 136, 78–97, 2008.

Sailor, D. J. and Lu, L.: A top-down methodology for developing diurnal and seasonal anthropogenic heating profiles for urban areas., Atmos. Environ., 38, 2737–2748, 2004.

Schlünzen, K. H. and Katzfey, J. J.: Relevance of sub-grid-scale land-use effects for mesoscale models., Tellus A, 55, 232–246, https://doi.org/10.3402/tellusa.v55i3.12095, 2003.

Smith, J. D., Mitsakou, C., Kitwiroon, N., Barratt, B. M., Walton, H. A., Taylor, J. G., et al.: London hybrid exposure model: improving human exposure estimates to NO2 and PM2.5 in an urban setting., Environ. Sci. Technol., 50, 11760–11768, https://doi.org/10.1021/acs.est.6b01817, 2016.

TfL: Mayor of London and Transport for London: Vision Zero action plan: Post-2020 progress report., [online] Available from: https://content.tfl.gov.uk/vision-zero-action-plan-progress-report-2021.pdf 2021.

TfL: Transport for London: London Digital Speed Limit Map., [online] Available from: https://content.tfl.gov.uk/london-digital-speed-limit-map.pdf, 2023.

United Nations: 2025 Revision of World Urbanization Prospects., Department of Economic and Social Affairs, [online] Available from: https://www.un.org/development/desa/pd/world-urbanization-prospects-2025 (Accessed 1 December 2025), 2025.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance., Clim. Res., 30, 79–82, https://doi.org/10.3354/cr030079, 2005.

Yang, W. and Omaye, S. T.: Air pollutants, oxidative stress and human health., Mutat. Res./Genet. Toxicol. Environ. Mutagen., 674, 45–54, https://doi.org/10.1016/j.mrgentox.2008.10.005, 2009.