

We would like to thank Reviewer 2 for the careful and constructive review of our manuscript. We appreciate the positive assessment of the model description and dataset, and the clear guidance on contribution, scope, and evaluation limitations.

Please find below our point-by-point responses to each comment.

- Location of manuscript changes in blue
- Changes in the manuscript in response to issues raised are indicated in red
- Tracked changed version of manuscript will be submitted once the review period has ended.

REVIEWER 2 COMMENTS

GENERAL	Overview
	This manuscript introduces MATSDA-roads v2.0, a Dijkstra's-algorithm-based transport model that computes optimised road-travel routes through a simplified, hierarchical nodal representation of a city's road network. The model is demonstrated for car driving in Greater London, evaluated against a large set of reference routes retrieved from the Google Maps (GM) Directions API on a gridded network representation. Local roads within each grid cell are aggregated into a single node, with F_L and JTP selected through sensitivity analysis (Sect. 4.1). Origin-destination sampling is designed around realistic commuter journeys (residence-workplace pairs stratified by car-commuting share and straight-line distance, Sect. 3.2.1). Route-realism factors not captured by the model, such as charging and low-emission zones, and other limitations are discussed in Sect. 5.1.
	The model description and evaluation are clearly presented, and the underlying dataset is substantial. Before publication, though, the manuscript should clarify more explicitly what MATSDA specifically contributes relative to existing routing approaches, over what scope its London-calibrated parameters are intended to apply, and what the comparison against Google Maps can and cannot show about the model's performance. These points, together with a few smaller clarifications, are discussed below; none require new analyses or climate-model coupling.
	Recommendation
	Major Revision The manuscript documents the model clearly and evaluates it against a substantial reference dataset, with code and data openly archived. Publication, however, depends on three points being clarified: (i) what MATSDA specifically contributes relative to existing routing approaches (Comment 1); (ii) the scope over which the London-calibrated parameters apply and what adaptation would be needed elsewhere (Comment 2); and (iii) what the comparison against Google Maps can and cannot show, together with a clearer account of the overnight-bias implications (Comments 3-4). The framing of the climate-study scope (Comment 5) should also be brought in line with what is demonstrated. None of these points require new analyses, additional data collection, or climate-model coupling - they concern how the existing material is framed, discussed, and cross-referenced.

REVIEWER 2 COMMENTS		
KEY COMMENT	R2.K1	Positioning and contribution relative to existing routing approaches MATSDA-roads uses a standard shortest-path algorithm over an aggregated, hierarchical representation of the road network. The manuscript should explain how MATSDA differs in purpose from routing engines such as OSRM and GraphHopper and traffic simulators such as SUMO. The manuscript motivates MATSDA mainly by contrasting it with commercial 'black-box' routers on grounds of transparency and query cost (Sect. 1), but does not discuss how it relates to these open-source alternatives. It would help readers to know what MATSDA is intended to provide for urban-climate applications that these related tools do not already offer - for example, a particular network-aggregation and calibration workflow, specific climate-relevant outputs, or another property tied to the intended use case. This is a request to clarify the intended distinction and scope of contribution, not to demonstrate equivalence or superiority through a benchmark against these tools.
	R2.K2	Transferability of the London-calibrated parameters F_L and the associated calibration (JTP, F_L = 0.16) are derived from London data (Sect. 4.1). The authors note that 'differences between cities are likely making case-by-case assessments necessary' (Sect. 5.1, item 1), but the manuscript does not elaborate on what this means in practice. Please clarify which city characteristics (e.g., block size, local-road density, network topology) would be expected to most affect F_L, and what adaptation
Author responses		
We clarify the relation of MATSDA to open-source routing tools and traffic simulators (e.g. OSRM, GraphHopper, and SUMO). The revised text states that Dijkstra's algorithm is standard and identifies MATSDA's contribution as its neighbourhood-based aggregation and calibration workflow, together with time-dependent outputs for urban modelling. We do not claim superior performance and have not benchmarked MATSDA against these tools. Text added to Section 1 Introduction, lines 65-74: Existing open-source transport tools include routing engines that calculate routes on detailed road graphs, such as OSRM (Luxen and Vetter, 2011) and GraphHopper (2026), and traffic simulators that represent individual vehicles and network interactions, such as SUMO (Alvarez Lopez et al. 2018). Like OSRM and GraphHopper, MATSDA uses an established shortest-path algorithm for routing. MATSDA is intended to support dynamic urban weather, climate and exposure modelling, using a transparent workflow while allowing spatial aggregation of a hierarchical road network to city neighbourhoods that can be informed by, and easily aligned with, typical urban meteorological spatial modelling units (e.g., grid-based). MATSDA assigns time-dependent travel weights, calibrates representation of local-roads and junctions, and stores origin-to-all-destination routes plus travel times by time interval. MATSDA outputs are intended for use in subsequent agent-based, emissions, exposure, and urban weather/climate modelling. We note (Section 5.1) that one local-road representation will not transfer directly to other cities. We expand this discussion to clarify what this means in practice as F_L and JTP should be treated as local calibration parameters, with the sensitivity analysis repeated using local reference routes, plus selected values need to be independently evaluated. We also condense the distance-dependent F_L conclusion.		

REVIEWER 2 COMMENTS

	- re-calibration, a repeated sensitivity analysis, or something else - would be needed before applying MATSDA to a different city. This would help define the intended scope of applicability without requiring validation in another city within this manuscript. Separately, the distance-varying F_L parameterisation (Eq. 3) did not outperform the constant F_L = 0.16 (Sect. 4.1.1); the explanation for this is spread across two sentences and could be tightened into a single concluding statement.	Clarification added to Section 5.1, lines 669-676: The calibration results for the London network (F_L= 0.16 and JTP = 0.25 min, for 500 m grid-cells) should not be treated as transferable model defaults. F_L is expected to be most sensitive to neighbourhood scale, local-road density and connectivity, and access to major roads. JTP depends on the mix of junction types and control arrangements (e.g., traffic signals and priority rules), as well as local delay conditions. For another city with a local travel database constructed, will require sensitivity analyses using local reference routes followed by independent evaluation of the assigned values for separate routes. The current structure is intended for major cities with a clear hierarchy of major and minor roads, in which major roads form the principal traffic arteries. If this hierarchy is weak, assuming one aggregated local-road node per neighbourhood would need assessment. Revised concluding sentence in Section 4.1, lines 396-397: Because a distance-dependent F_L does not materially improve overall performance, we retain the simpler constant F_L = 0.16 for the London application.
R2.K3	Google Maps as an evaluation reference GM is used as the primary evaluation reference throughout Sect. 4, while Sect. 1 notes that commercial routers are opaque about their underlying routing logic. GM is a useful, widely available comparison point, but - by the manuscript's own framing - it is not a fully transparent ground truth. Please clarify what the comparison against GM can and cannot show about MATSDA's performance, and consider noting the absence of independent validation against other data sources (e.g., TfL loop-detector counts, TomTom/Waze data) as a scope limitation of the current evaluation rather than only as future work.	We agree GM is a benchmark, rather than ground truth. We clarify Section 3.2 to indicate what we assess, and replace 'observed' and 'accuracy' when referring to GM only, plus indicate the absence of independent trajectory or journey-time evaluation (Section 5.1). Traffic counts are not used as a direct evaluation measure of individual routes because MATSDA does not assign travel demand itself (as this would come from a coupled model; e.g., agent-based). Clarification added to Section 3.2 Model evaluation, lines 252-256: GM is used as an external routing benchmark rather than an observed ground truth. We assess agreement between MATSDA and GM of travel times, road-type use, diurnal variation and route corridors for the sampled O-D pairs and query times. Because GM's network, route-choice cost function and use of historical or live traffic are proprietary, agreement does not establish absolute accuracy and differences cannot be attributed to MATSDA alone. Limitation added to Section 5.1, lines 681-685: The present study does not evaluate MATSDA against independent observed trajectories or journey times. Transport for London's manual or automatic traffic-counter data are not directly comparable because the current numerical experiments do not include travel demand or estimate traffic volumes. However, these are relevant when MATSDA is coupled to a demand model such as DAVE (Hertwig et al. 2025b). Independent mobility observations and additional routing datasets are therefore priorities for future evaluation of coupled MATSDA.
R2.K4	Overnight bias 22:00-04:00 MATSDA systematically underestimates travel times overnight across all configurations (Fig. 11a, d, g; MBE < 0), reaching up to approximately 25 min faster than GM for longer routes. Sect. 4.2 already attributes part of this to the flatter diurnal	As noted, Section 4.2 and Fig. S11 both report the overnight bias and flatter diurnal profiles than GM-derived speeds. We now bring these points together when the bias is first discussed and add a statement on uncertainty. Possible effects of night-time sampling and vehicle composition cannot be tested because the Digimap Pilot Collection (2024) does not include source observation counts or vehicle classes.

REVIEWER 2 COMMENTS

	speed profile in the GM data (Fig. S11), which is a reasonable starting point, but this explanation only appears in the latter part of the section. Consolidating the discussion in one place, and adding a clearer account of the remaining implications (e.g., whether GPS data sparsity at night or night-time traffic-composition changes also contribute, and what this means for 24 h applications), would help.	Consolidated discussion in Section 4.2, lines 499-504: Across all cases, MATSDA underestimates GM overnight travel times by up to ~ 25 min for longer full-domain journeys (22:00-04:00; Fig. 12a,d,g). Larger diurnal variation in MATSDA's input traffic speeds, compared to those inferred from the GM routes (Fig. S11), plus different data-collection periods may contribute to this bias. The broad six-hour night-time interval and unquantified differences in GPS sampling or vehicle composition may also contribute. Unfortunately, the speed data product contains neither source-observation counts nor vehicle classes with which to test these effects, requiring the assumption these overnight results are more uncertain. Editorial corrections in Section 4.2: The cross-reference 'Figure 13g' is corrected to 'Fig. 12g', and 'MASTDA' is corrected to 'MATSDA'.
R2.K5	Framing of the dynamic urban climate studies scope The title and abstract frame the work within 'dynamic urban climate studies,' while the manuscript itself is a transport-model description and evaluation study; coupling via DAVE is deferred to future work (Sect. 5.1, item 6). No climate-model coupling or additional analysis is needed here - rather, please adjust the title/abstract/introduction so that the framing and intended use match what is actually demonstrated in this paper, i.e., that this paper validates the transport component as a step toward the stated climate application.	We modify the title to better reflect the actual paper content. The Abstract and Introduction wording are also clarified – to indicate we develop and evaluate the road-routing component for London. We note that demand, emissions, exposure, and climate coupling are not evaluated but mention the intended use of MATSDA (see also response to R2.K1). Section 5.1 presents the link to the agent-based model DAVE as future work. Revised title: MATSDA-roads v2.0: development and evaluation of road-routing intended for urban climate and exposure applications Scope statement added to the Abstract, line 12-14: We assess MATSDA-roads v2.0 using real-time routing benchmark data from the Google Maps Directions API obtained over several weeks with day of week, time of day, route direction (outbound v. inbound) and route alternatives (total captured for >87,000 reference routes in London, UK). Scope statement added to the Abstract, lines 20-23: MATSDA-roads v2.0 provides time-dependent routes and travel times intended as inputs to subsequent urban activity, emissions, exposure and climate modelling. This work provides a valuable tool for transport research, urban climate modelling and environmental exposure assessment that require dynamic human movement patterns, but coupling to these models is not evaluated here. Scope statement added to the Introduction, included in R2.K1, lines 71-73: MATSDA assigns time-dependent travel weights, calibrates representation of local-roads and junctions, and stores origin-to-all-destination routes plus travel times by time interval. MATSDA outputs are intended for use in subsequent agent-based, emissions, exposure, and urban weather/climate modelling. Coupling to these models remains future work.

REVIEWER 2 COMMENTS

SECONDARY COMMENT R2.S1	Source speed-data description and licensing The Digimap Pilot Collection (2024) GPS speed data lack basic summary metadata - observation counts, spatial coverage across boroughs, and temporal density per time bin. Where full disclosure is restricted by licensing, a short statement plus any shareable high-level summary (e.g., counts or coverage per borough or per time bin) would help readers judge how well the averaged speeds represent actual traffic conditions, particularly at night (see Comment 4 above).	Sections 3.1.2 and S1.2.3 report the six-month averaging period, weekday and weekend intervals, and spatial summaries. We add a concise road-link inventory and clarify the product neither provides underlying GPS observation counts nor sampling density and has source-data licensing restrictions. Clarification added to Section 3.1.2, lines 222-223: Within the 9,416 grid-cells London model domain (Fig. 1a,c), 526,688 road segments intersect the domain, representing approximately 27,734 km of road after clipping. Clarification added to Section 3.1.2, lines 227-228: Note, the quality or number of GPS observations underlying these averages is not documented within the data set. Neither, the traffic speed nor the Ordnance Survey source data can be redistributed under their licences.
R2.S2	Rationale for the Hit Rate thresholds Sect. 3.2.4 uses a +/-5 min tolerance (HR5) for short routes (≤ 14 km) and a +/-10 min tolerance (HR10) for long routes (> 14 km; Eq. 8). A brief justification for these specific values would help readers judge how stringent the metric is.	We add a brief rationale. HR5 and HR10 are pragmatic descriptive tolerances scaled to journey length, not independently calibrated acceptance criteria. The 14 km division follows the O-D distance breakpoint used in the F_L analysis, and HR, MAE and MBE are assessed. Clarification added to Section 3.2.4 Temporal evaluation, lines 358-360: The 14 km division arises from the O-D straight-line-distance breakpoint identified in the F_L sensitivity analysis (Sect. 4.1.1; Sect. S4.2.2).
R2.S3	Notation clarity The lambda and Fraktur F notations in Appendix A should be made more consistent and clear.	We correct the notation. These editorial corrections do not affect the calculations or results. Equation (1) used λ_s for the input-road segment length, whereas Section 3.1.1 defines the same quantity as l_s λ_s changed to l_s in Eq. (1)
R2.S4	Charging and low-emission zones and spatial route realism MATSDA routes more journeys through central London than Google Maps, which the authors attribute in part to the omission of charging and low-emission zones (Sect. 4.3, Fig. 13a-c). This may affect route realism; a brief quantification would help.	We revise Section 4.3 to indicate charging-zone treatment as one possible factor. The largest grid-cell differences are shown in Fig. 13d,e, but this does not support a zone-wide or causal attribution. Revised paragraph in Section 4.3, lines 560-563: The largest negative differences ($\leq -100\%$, Fig. 14d,e) indicates grid-cells where the MATSDA route count is at least three times the nonzero GM count. It also includes grid-cells traversed by MATSDA routes but by no GM routes. As this grid-cell comparison does not account for the impacts from zone (e.g. CCZ or ULEZ) or road (e.g. one-way) restrictions or local congestion, the available data cannot identify the cause.
R2.S5	Parameter sensitivity and spatial homogeneity F_L and JTP are derived from a domain-wide sensitivity analysis (Sect. 4.1). Are these	The constant F_L cases are described as spatially uniform in Section 4, and Section 4.1.2 states that a single JTP is applied to all road-type transitions. We clarify our assumption, and indicate we do not estimate separate inner- and outer-London values.

REVIEWER 2 COMMENTS

	parameters assumed spatially homogeneous, or do results (e.g. Fig. 6, Fig. S6) show differences between inner-London and suburban areas? A brief comment on this would be useful.	Clarification added to Section 4.1, lines 393-396: Within each configuration, effective local road length fraction F_L (Sect. 4.1.1) and junction time penalty (JTP, Sect. 4.1.2) are applied uniformly across the London domain. In the distance-dependent cases, F_L varies with O-D straight-line distance, but not location. Figure S6 shows the relation between F_L and grid-cell junction density and does not test differences between inner and outer London. Separate regional values are therefore not determined.
R2.S6	Section 4.2 signposting The cases in Table 3 would benefit from clearer signposting or subheadings.	The relevant Table 3 cases are identified in each comparison in Section 4.2. To make the sequence easier to follow, we have added one opening sentence grouping the cases by purpose; the analysis is unchanged. Opening signpost added to Section 4.2, lines 466-468: Using the independent evaluation dataset (Table 2), we compare four aspects of the Table 3 configurations: input speed (run #1 versus #2), road-network detail (#2 versus #3), parameter choices (#3 versus #4 and #3 versus #5) and unconstrained full-domain routing (#6 and #7).
R2.S7	Integration of supplementary material Key results from Supplementary Sections S1-S4 could be summarized in the main text.	The main outcomes from S1-S4 are reported and cross-referenced in Sections 2.2, 3.1, 3.2, and 4.1. We add a concise synthesis to make those links easier to identify without repeating the processing detail from the Supplement. Synthesis added to the main text, lines 621-625: In addition, Sections S1-S4 document the construction of the 20,273-node database for 9,416 grid-cells, calculation of origin-to-all-destination paths, processing and splitting of the GM route data, and the F_L sensitivity analyses. From this we determine that a junction-dependent F_L formulation will introduce substantial computational cost in the current Pathfinder, while a distance-dependent F_L does not materially improve model performance over the constant F_L (= 0.16 for London).