Theory and Practice of Science and Technology
2025, VOL. 6, NO. 7, 159-167
DOI: 10.47297/taposatWSP2633-456931.20250607

Modelling Urban Traffic Congestion using Agent-Based Model: A Case of York, UK

Qiushi Jin, Heshi Hon*

FJ Dynamics International Limited, Nanjing, Jiangsu, 210000, China

ABSTRACT
Understanding urban traffic is crucial for rapidly urbanising historic cities like York, UK. Traditional traffic models face challenges due to limited origin-destination (OD) data. This study uses census data on car commuters as a proxy for OD matrices and simulates morning peak traffic with agent-based modelling (ABM). Results show frequent congestion in the city centre, near the River Ouse, and on key routes like the Northern Distributor Road and Eastern Expressway, aligning with real-world observations. The approach demonstrates the value of combining non-traditional data with traffic simulations for urban planning.

KEYWORDS
Urban Traffic Congestion; Agent-Based Modelling (ABM); Traffic flow simulation; GIS-based spatial analysis

1 Introduction

1.1 Background

With the growth of the automobile industry, traffic has an increasing impact on cities, and traffic congestion leads to economic decline through its impact on regional productivity business activities (Sweet, 2014). The exhaust and noise from motor vehicles can seriously damage the urban environment (Kondrashov et al., 2002). This has led to an urgent need to optimise traffic flow management and improve urban traffic patterns. Traffic simulation techniques can be used to improve road usage and reduce environmental pollution through road network planning, road design and optimised traffic signal control.

1.2 Research objectives

Aim 1: To develop and validate an agent-based model to simulate urban traffic dynamics and provide evidence-based decision support for urban traffic management.

Objective：Review existing agent-based models for urban traffic simulation.

Design a new agent-based model for urban traffic dynamics.

Run the simulation programme focusing on traffic flow and congestion in the morning peak hour.

Aim 2: Explore the potential of incorporating census data into agent-based traffic simulation models to provide a new approach to urban traffic modelling.

Objectives：Integrate selected census data into an agent-based model.

Analyse the impact of incorporating census data on model results and insights.

Assess the advantages and challenges of using census data in agent-based traffic simulation.

1.3 Scope and rationale

Whilst the scope of the study covers urban traffic scenarios, it zooms in specifically to the morning peak hour, which is typically a period of high traffic congestion. This focussed approach was adopted to provide insights during key traffic hours and to assess the efficacy of the model in the most challenging scenarios. In addition, census data was selected with the aim of improving the accuracy of the model by utilising easily accessible granular demographic information.

2 Literature review

2.1 Traffic Simulation

In the multifaceted domain of transportation research, traffic simulation stands out as a pivotal tool, offering valuable insights into the intricacies of traffic dynamics. It is a technique that utilizes computer models or mathematical models to

* **Corresponding Author:** Heshi Hon, hsh970822@gmail.com

simulate and analyze traffic flows and the performance of a transport system. Through this technique, the behavior of an existing or future transport network can be predicted, thereby allowing researchers to propose rational traffic planning and management strategies (Barceló, 2010).

2.2 Agent-Based Modeling（ABM）

Although macroscopic simulation can provide a global view of the traffic system, such models usually ignore the detailed behaviour of the participants in the traffic flow. While traditional mathematical modelling approaches are difficult to describe the traffic system effectively, the agent-based modelling approach can simulate the "bottom-up" behaviour of a complex traffic system, and its autonomy and adaptivity are in line with the intelligent characteristics of the elements in the traffic system, which can better simulate the autonomy and anisotropy of the behaviour. In this paper, the choice will be made to use an Agent-based model, ABM can provide a deeper understanding of how the behaviour of individual participants affects the dynamics of the whole system, such as traffic flow patterns and congestion (Alghais et al., 2018). In addition, ABM can handle more complex traffic scenarios such as traffic signal control and vehicle routing, which are often difficult to implement in other models (Shaharuddin et al. 2023).

2.3 Challenges and innovations

ABM offers distinctive insights in traffic simulation, yet is fraught with challenges. Accurate simulations demand extensive data which is often elusive and expensive. Specifically, obtaining origin-destination (OD) matrices entails significant complexities. Once acquired, data from diverse sources necessitates meticulous cleaning and conversion for ABM applicability. Missing or inaccurate data entries can skew results. Moreover, ABM's computational requirements, especially for intricate transport systems, can strain resources. A final challenge is model validation and reproducibility; lacking comprehensive documentation hinders further research. This study introduces a novel approach, leveraging easily accessible census and OpenStreetMap data for urban traffic simulations. This strategy promises to streamline data acquisition while maintaining accuracy, setting the stage for transparent and replicable research.

3 Methodology

3.1 Study design

This study aims to explore and simulate the morning rush hour traffic conditions in York, especially for the travel patterns of commuters.

The overall research process is divided into several key steps. First, the road data of York area will be downloaded from OpenStreetMap and converted into .net.xml format that can be recognized by SUMO software. In addition, specific area labels were set for each road using York's zoning vector data. Then, based on the population distribution and commuters' statistics in York, vehicle models were randomly generated in the corresponding zones and the best routes were selected based on their distances to their workplaces to simulate the traffic flow during the morning rush hour.

3.2 Study area and data sources

York, located in North Yorkshire at the confluence of the Ouse and Foss rivers. As highlighted by Rees Jones (2013), it is a historic city with narrow medieval streets that now face modern traffic pressures. The morning peak is particularly congested, making it an ideal case study for testing urban traffic models.This study uses census data from the UK Data Service to estimate commuting demand, and OpenStreetMap (downloaded via BBBike) to construct York’s road network.

3.3 Simulation environment

SUMO (Simulation of Urban Mobility), an open-source, globally used transportation simulation software, has been shown to be highly effective in several research and application scenarios (Krajzewicz et al., 2002). Its flexibility, accuracy of detail, and ability to integrate with a variety of other tools and data sources made it the preferred choice for this study.

3.4 Data conversion and processing

The OSM data obtained via BBBike was processed into SUMO-compatible format. Non-motorised lanes were removed, only relevant road classes retained, and left-hand driving rules applied to match UK conditions, resulting in a validated network for simulation. The York boundary data were projected to UTM Zone 30 to ensure consistency with the SUMO network. Administrative attributes were integrated into the road file, producing a location-specific network for subsequent simulations.

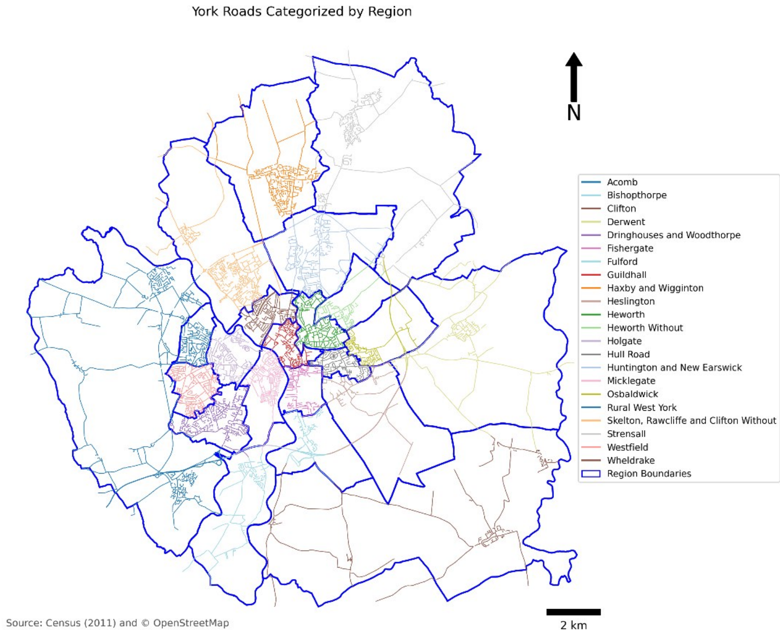

Figure 1 York roads categorized by region

3.5 Agent-Based Modeling（ABM）

In order to better understand how this paper uses ABM to simulate urban traffic, the next subsections delve into vehicle generation, path selection, and travel patterns in the network:

3.5.1 Vehicle generation model

One of the biggest challenges in traffic simulation is obtaining realistic origin and destination data. Cascetta (1984) proposed simulating origin and destination data by combining traffic counts with survey data. Cascetta's approach provides a solution for this paper, which will use census data from the "driving a car or van" commute to generate origin and destination data in the model.

By analyzing census data, it is possible to determine the number of people in each region who choose to drive a car or van to work. Based on these numbers, the number of vehicles that should be generated in each region can be estimated. After determining the number of vehicles in each region, the residential roads in the region are further selected and a specific road is randomly chosen as the starting point for the vehicle. Since information on the distance of each commuter from their workplace is already available, this data can be utilized to simulate the end point data of the generated vehicles. To ensure the realism of the simulation, the vehicle's departure time is an integral factor. In the model, the departure time range of the vehicles is from 7:00 am to 9:00 am, totaling 2 hours, and is divided into 8 sections, one every 15 minutes, with a group of vehicles departing from each section. And through multithreading allows multiple regions can generate vehicles at the same time.

3.5.2 Vehicle path selection model

Route selection is an important part of traffic flow simulation and involves deciding how a vehicle will reach its destination from its starting point. Based on the literature this paper uses a shortest path based approach to plan paths for all vehicles, selecting 30% of the intelligences that can avoid traffic congestion, while the other 70% always stay on their original path (DE PALMA et al., 1999).

(1) Static path selection based on shortest paths

Vehicle routes were assigned using Dijkstra's shortest-path algorithm (Dijkstra, 1959), with road-segment length as the edge weight. To improve computational efficiency, paths were pre-computed and stored for use during simulation.

(2) Dynamic path selection

In order to simulate real-world traffic flows, especially when vehicles need to dynamically change their routes during traffic congestion, this paper uses Traffic Control Interface (TraCI), a powerful tool of SUMO. TraCI allows the researcher to run the simulation with the SUMO in real-time while the simulation is running, allowing the researcher to dynamically change the vehicle's route so that it can respond to changes in traffic conditions.

TraCI is a software interface that allows external programs to interact with SUMO while the simulation is running. With TraCI, the status of various entities such as vehicles, road sections, traffic lights, etc. can be acquired at each simulation step and decisions can be made based on this information, e.g., changing the route of a vehicle.

Getting the current vehicle list: Using the method provided by TraCI, it is possible to get all the vehicles in the simulation environment in real time.

Selection of Selected Vehicles for Route Changes: In order to simulate that the navigation systems of some vehicles can dynamically respond to traffic congestion and provide them with alternative routes, 30% of all vehicles were randomly selected.

Monitoring the routes of selected vehicles: These selected vehicles are closely monitored during each simulation step. When a possible traffic congestion is detected on the vehicle's intended route, TraCI develops a new optimal path for the vehicle based on the predicted travel time. The congestion is determined based on the Speed Performance Index (SPI):

$$SPI = v_{avg}/v_{max} \times 100$$

Where: v_{avg} denotes the average traveling speed; v_{max} denotes the maximum allowable traveling speed.

The Speed Performance Index (SPI) is used to evaluate congestion on urban roads.The SPI value ranges from 0 to 100 and is determined by the ratio between the average travel speed and the maximum allowable speed (He et al., 2016). In order to use this index to measure the traffic condition on the road the speed performance index (SPI) can be divided into four intervals, details of which are shown in Table 1. In this paper, a threshold value of 50 was chosen as a way to define the congestion that changes the vehicle's travelling route.

Table 1 Traffic states corresponding to different speed performance indices (SPI)

SPI	Traffic Status
(0,25)	Heavy congestion
(25,50)	Medium congestion
(50,75)	Smooth
(75,100)	Very smooth

This feature of TraCI not only more accurately simulates real-world traffic flow patterns, but also reveals the driver's behavioral responses when encountering traffic congestion. When the navigation system predicts that the road ahead may be congested, it provides drivers with alternative routes designed to bypass the congestion, improve the overall efficiency of the road, and ensure a smoother driving experience.

3.5.3 Vehicle Driving Mode

The vehicle travel model is one of the most critical components of a simulation because it describes how a vehicle responds to other vehicles and road conditions in a road network. Choosing the right model will determine the realism and accuracy of the simulation.

(1) Basic Driving Model

In SUMO simulation, the basic driving model of the vehicle is defined by several core parameters:

Acceleration (accel): defines the maximum acceleration of the vehicle under optimal conditions. Most modern family cars accelerate from 0-60 mph in about 7-10 seconds, which corresponds to an acceleration range of 2.4-3.8 m/s^2. An intermediate value of 2.8 m/s^2 has been chosen here to reflect the performance of a typical family car on urban roads.

Deceleration (decel): represents the maximum deceleration of a vehicle in a hard stopping situation. Most family cars may have better brakes, but in real-world driving, drivers usually don't use the maximum brakes unless it's an emergency. 3.0 m/s^2 is a moderate value.

Maximum Speed (maxSpeed): The maximum speed a vehicle can reach under given road and traffic conditions. Considering urban traffic and possible traffic restrictions, 35 m/s was chosen as the maximum speed.

(2) Vehicle following model

Vehicle-following models, particularly the Intelligent Driver Model (IDM), provide insight into traffic flow simulations, allowing understanding and predicting how individual driver behavior affects overall traffic dynamics. IDMs have been widely adopted for their accuracy and reproduction of actual traffic phenomena (Treiber et al., 2000).

The model describes how drivers make decisions based on the traffic ahead through the following equation:

$$a = a_{max}\left(1 - \left(\frac{v}{v_0}\right)^{\delta} - \left(\frac{s^*(v,\Delta v)}{s}\right)^2\right)$$

where:

a is the acceleration of the vehicle, which is directly related to the response of the driver.

a_{max} is the maximum acceleration of the vehicle.

v is the current speed of the vehicle, reflecting the current driving state of the driver.

v_0 is the desired speed of the vehicle.

δ is the exponent of acceleration.

s is the current spacing from the vehicle in front, a key factor in decision making.

$s^*(v,\Delta v)$ is the desired spacing, and the formula is:

$$s^*(v,\Delta v) = s_0 + vT + \frac{v\Delta v}{2\sqrt{a_{max}b}}$$

Where:

s_0 is the minimum spacing between the vehicle and the vehicle in front of it at a stop.

T is the safety time header, which represents the time interval that the driver wishes to maintain with the vehicle in front, usually depending on the driver's driving style and the current traffic conditions.

Δv is the relative speed to the vehicle in front.

b is the comfortable deceleration.

Safety Time Head: In the IDM model, the safety time head is a key parameter that represents the time interval that the driver wishes to maintain. It represents the maximum reaction time required by the vehicle behind to avoid a collision when the vehicle in front stops suddenly. In this study, the safety time head was set to 1.5 seconds as recommended by Treiber et al. (2000). This is based on the observation of real driving data, in which most drivers maintain such a time head under normal driving conditions.

Minimum Spacing: In addition, this paper considers the minimum physical spacing between the vehicle and the vehicle in front of it at a complete stop. This is the minimum spacing that a driver wants to maintain with the vehicle in front of him to ensure a safe and comfortable drive. Referring to the study by Treiber and Kesting (2013), 2 metres was chosen as the minimum spacing. This value is not only based on previous studies, but also takes into account real driving situations where drivers usually maintain such a distance in order to avoid possible collisions and driving stress.

The advantage of the model is that it takes into account a variety of realistic driving factors, such as the driver's desired speed, the actual distance to the vehicle in front, and the relative speed, thus providing an accurate description of the real traffic flow (Treiber et al., 2013).

(3) Vehicle lane changing model

Lane changing behavior is a key factor that cannot be ignored in road traffic simulation. Correctly modeling lane change decisions is essential to reproduce realistic traffic flow characteristics. Model of Behaviour in Lane Changes (MOBIL) is a microscopic model that provides this functionality by basing decisions on vehicle safety and driving efficiency (Kesting et al., 2007).

The MOBIL model not only considers the safety and benefits of the current vehicle, but also evaluates the impact of lane changing on other vehicles. The underlying assumption is that drivers weigh the benefits and potential risks of lane changing in their decision making.

MOBIL's lane change decisions are based on the following guidelines:

$$\Delta a_{own} + p \times \Delta a_{fol} > \Delta a_{thr}$$

Where：

Δa_{own} is the acceleration change of the car after the lane change.

$\Delta a_{fol}$is the acceleration change of the vehicle after the lane change.

p is the weighting factor considered for the safety of the rear vehicle.

Δa_{thr} is the minimum benefit threshold required for lane changing.

When conducting a vehicle lane changing modeling study, the selection of appropriate parameters is crucial. In this study, the following parameter values were chosen to better simulate real-world driver behavior.

Switching Safety Interval (lcDelta): This document sets a value of 0.2 for lcDelta. This means that drivers would like to maintain an additional safety interval of at least 0.2 meters from other vehicles after completing a lane change. This setting takes into account the general preference of drivers in real-world driving, who tend to maintain a certain safety distance in order to cope with potentially unpredictable events, such as a sharp braking of the vehicle in front of them or other unexpected events. Furthermore, this value falls within a reasonable range compared to previous studies and provides the model with a way to balance safety and efficiency (Kesting et al., 2007).

Gain Speed Threshold (lcSpeedGain): For lcSpeedGain, a threshold value of 0.2 m/s was chosen in this paper. This means that drivers will only consider changing lanes if the predicted speed gain is at least 0.2 m/s. This choice is based on the economic behavior of drivers, i.e., they want to gain a significant speed advantage by changing lanes, not just for a small speed gain. In this context, a speed gain of 0.2 m/s is considered to be a real and significant speed gain that warrants driver action (Kesting et al., 2007).

The main advantage of MOBIL is its simplicity and efficiency. Since it requires only a small number of parameters, it is easy to estimate from real data. In addition, the MOBIL model considers both benefits and safety, allowing it to more realistically simulate real driving behavior.

4 Data analysis and results

4.1 Road network data

The generated road network is shown below：

Figure 2 York road network and traffic signals

There are a total of 19,912 roads with 641 traffic signals.The main roads in York's urban centre are mainly primary type roads. These roads are the backbone of the city's traffic and are usually wide and provide the main access to commercial and residential areas. Between the primary roads are a large number of residential roads, which serve mainly residential areas and are narrower, usually one or two lanes only, with speed limits, and the city centre is encircled by a .trunk road, which acts as a ring road and provides a fast route for traffic in and out of the city centre. Extending outwards from the ring road are secondary and tertiary roads. These roads connect York to other urban and rural areas. In addition most of the traffic signals are concentrated in the city centre area, which is where the highest volumes of traffic occur. This design ensures a smooth flow of traffic and reduces congestion.

4.2 Simulation Overview

The traffic simulation was conducted in SUMO between 7:00 and 9:00 to capture morning peak dynamics. The model incorporated realistic driver behaviour, lane configurations, and traffic control, providing a reliable basis for subsequent analysis.

4.3 Simulation results

Two files were output at the end of the simulation. One is a record of the basic information of the roads during the simulation contains the traffic flow, speed and density on each road in the simulation. The other one is about the travelling information of each vehicle, including the waiting time, travelling time and travelling distance of the vehicle. Further analyses are carried out based on these data.

In order to visually represent the degree of congestion of vehicles on the road, this paper adopts the speed performance index (SPI) as the evaluation index. Based on the simulation data, the road thematic maps of SPI are produced for different time periods, as shown in Table 1 in the "Methodology" section, the value of SPI is inversely proportional to the degree of congestion on the road: a low value of SPI means that the road is more congested, and a high value means that the traffic on the road is smooth. This is shown in the graphs below. These graphs show the traffic

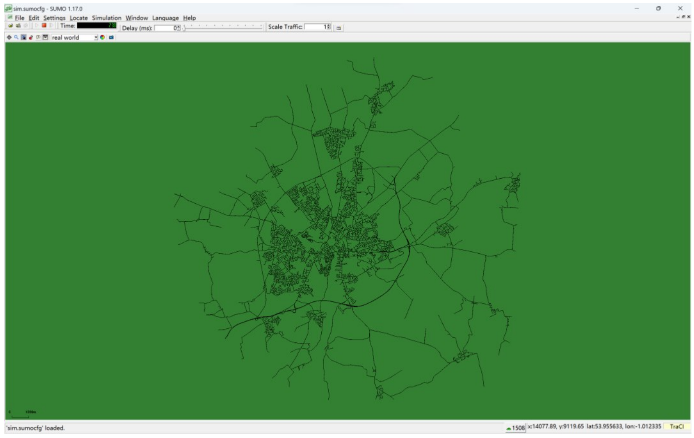

Figure 3 Simulation screen of SUMO

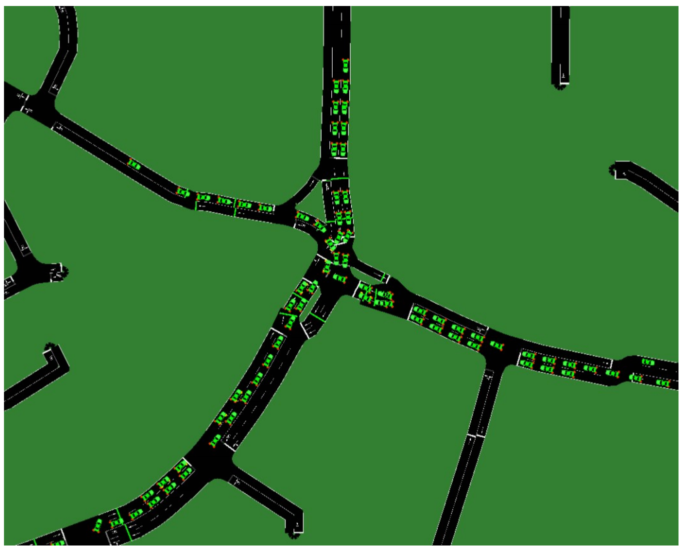

Figure 4 Congested junction

congestion in each area during the morning rush hour, providing a powerful visual support for urban traffic management.

Figure 5 Traffic congestion between 7am and 7.30pm

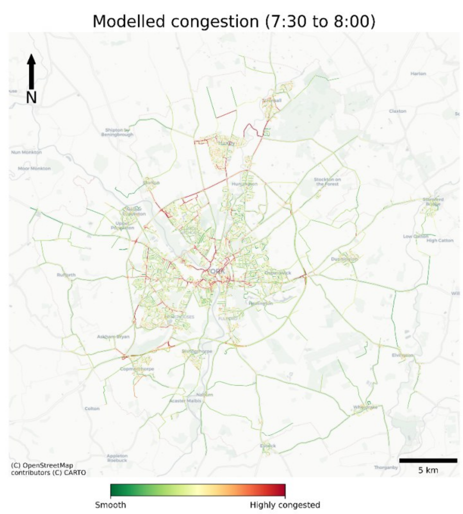

Figure 6 Traffic congestion between 7.30am and 8pm

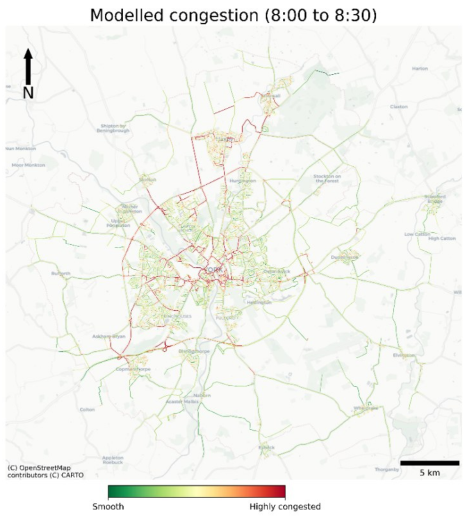

Figure 7 Traffic congestion between 8am and 8.30pm

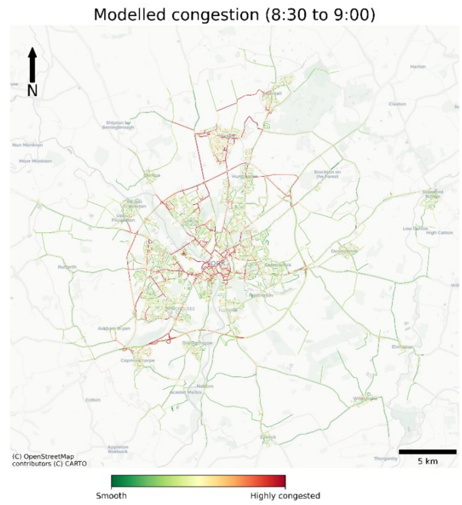

Figure 8 Traffic congestion between 8.30am and 8pm

In the thematic map, this paper uses red colour to denote the areas with traffic congestion (low SPI values) and green colour to denote the parts with smooth traffic (high SPI values). By analysing the thematic maps, it can be observed that the extent of road congestion gradually increases as the simulation time increases. This trend is mainly due to the vehicle generation mechanism: the number of vehicles in the road network gradually increases over time, which in turn exacerbates the level of traffic congestion. Obvious areas of congestion are concentrated in the city centre, on the main expressways in the north, and near the ring overpasses on the eastern expressways.

4.4 Model Validation

In this study, the traffic layer of Google Maps was captured for real-time traffic flow information during the time period of 7: 00 to 9: 00 pm on weekdays. This traffic information is represented based on colour coding, where green means smooth traffic, yellow indicates moderate congestion, and red represents severe traffic congestion. The model was validated by comparing the real-time traffic conditions from Google Maps with the previous thematic map of traffic congestion generated using SPI than.

As illustrated in the figure above, the real-time weekday traffic conditions show that the main distribution of congestion broadly matches that modelled in Figures 5~8, particularly in the city centre, the major expressways in the

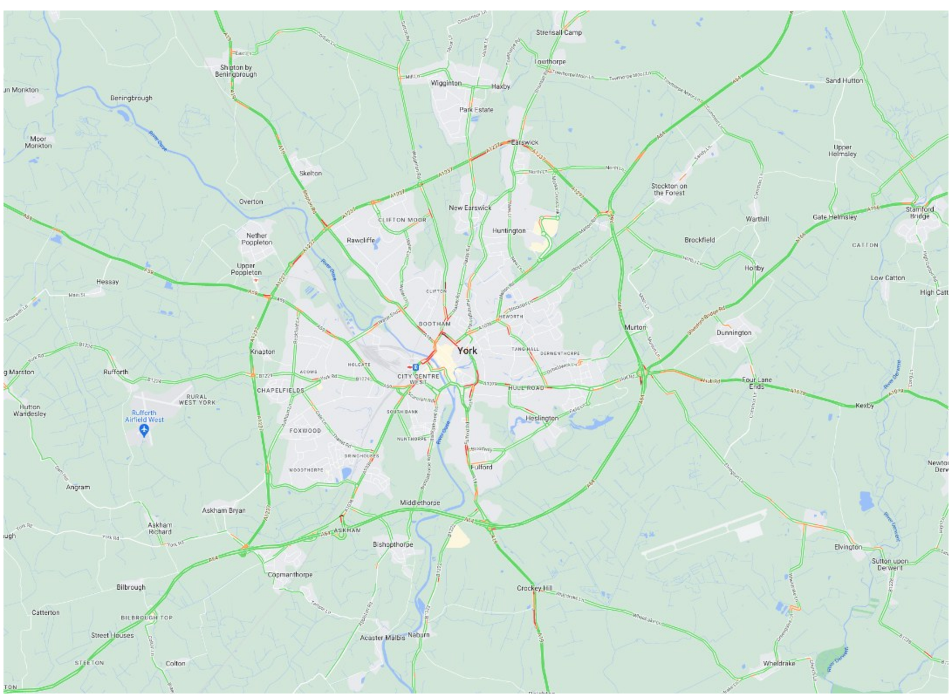

Figure 9 Real traffic congestion

north, and the ring flyover area in the east.

5 Conclusion

This study highlights both the potential and limitations of using census data in agent-based modelling (ABM) for urban traffic analysis. While census data provides valuable insights, it lacks real-time variability, and ABM itself requires further refinement and validation against established modelling techniques. Future research should integrate dynamic sources such as real-time traffic data, mobile GPS, or social media to enhance responsiveness and accuracy. Despite these limitations, the study demonstrates a promising approach to urban traffic modelling, aligning with its initial aims and offering important insights for data-driven urban traffic management.

References

[1] Alghais, N. and Pullar, D. 2018. Modelling future impacts of urban development in Kuwait with the use of ABM and GIS. *Transactions in GIS*[J]. *22(1)*, pp.20–42.

[2] Barceló, J. 2010. Fundamentals of traffic simulation[M]. *New York: Springer.*

[3] Cascetta, E. 1984. Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator[J]. *Transportation research. Part B: methodological*. 18(4), pp.289–299.

[4] DE PALMA, A. and ROCHAT, D. 1999. Understanding individual travel decisions : results from a commuters survey in Geneva[J].*Transportation (Dordrecht)*. 26(3), pp.263–281.

[5] Dijkstra, E.. 1959. A note on two problems in connexion with graphs[J].*Numerische Mathematik.* 1(1), pp.269–271.

[6] He, F., Yan, X., Liu, Y. et al., 2016. A traffic congestion assessment method for urban road networks based on speed performance index[J]. *Procedia Engineering*, 137, pp.425-433.

[7] Kagho, G.O., Balac, M. and Axhausen, K.W. 2020. Agent-Based Models in Transport Planning: Current State, Issues, and Expectations[J].*Procedia computer science*. 170, pp.726–732.

[8] Kesting, A., Treiber, M. and Helbing, D. 2007. General Lane-Changing Model MOBIL for Car-Following Models[J].*Transportation research record*. 1999(1), pp.86–94.

[9] Kondrashov, V.V., Reshetin, V.P., Regens, J.L. and Gunter, J.T. 2002. Modeling mobile source emissions during traffic jams in a micro urban environment[J].*Environmental science and pollution research international*. 9(5), pp.333–336.

[10] Rees Jones, S. 2013. York : the making of a city, 1068-1350[M] .*Oxford: Oxford University Press*.

[11] Schadschneider, A., Knospe, W., Santen, L. and Schreckenberg, M. 2005. Optimization of highway networks and traffic forecasting[J].*Physica A.* 346(1), pp.165–173.

[12] Sweet, M. 2014. Traffic Congestion's Economic Impacts: Evidence from US Metropolitan Regions[J].*Urban studies (Edinburgh, Scotland)*. 51 (10), pp.2088–2110.

[13] Treiber, M., Hennecke, A. and Helbing, D. 2000. Congested traffic states in empirical observations and microscopic simulations[J].*Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*. 62(2 Pt A), pp.1805–1824.

[14] Treiber, M. and Kesting, A. 2013. Traffic flow dynamics : data, models and simulation[M]. *Berlin* : Springer.