Transportmetrica A: Transport Science

ISSN: 2324-9935 (Print) 2324-9943 (Online) Journal homepage: www.tandfonline.com/journals/ttra21

Exploring Voronoi-based area-weighted network fundamental diagrams in shared urban spaces

Deng Pan, Pedram Beigi, Winnie Daamen, Victor Knoop, Alireza Talebpour & Samer H. Hamdar

To cite this article: Deng Pan, Pedram Beigi, Winnie Daamen, Victor Knoop, Alireza Talebpour & Samer H. Hamdar (25 Sep 2025): Exploring Voronoi-based area-weighted network fundamental diagrams in shared urban spaces, Transportmetrica A: Transport Science, DOI: 10.1080/23249935.2025.2558149

To link to this article: https://doi.org/10.1080/23249935.2025.2558149

TRANSPORTMETRICA A: TRANSPORT SCIENCE
https://doi.org/10.1080/23249935.2025.2558149

Exploring Voronoi-based area-weighted network fundamental diagrams in shared urban spaces

Deng Pan[a], Pedram Beigi [a], Winnie Daamen[b], Victor Knoop[c], Alireza Talebpour[d] and Samer H. Hamdar[a]

[a]Department of Civil and Environmental Engineering, The George Washington University, Washington, DC, United States; [b]Department of Transport and Planning, Delft University of Technology, Delft, The Netherlands; [c]Department of Civil Engineering of Technology, Delft University of Technology, Delft, The Netherlands; [d]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, United States

ABSTRACT

This research extracts trajectories of non-directional micromobility traffic (pedestrians, cyclists, and mopeds) in a shared right-of-way urban space, applying a Voronoi-based area-weighted framework to construct Network Fundamental Diagrams (NFDs). Using an aggregation technique that weights by link length, we explore two Voronoi-based approaches for generating NFDs from microscopic data: a standard mixed-mode approach and a novel mode-isolated approach. Results demonstrate that both methods accurately compute macroscopic traffic measures, and mode-isolated approach, in particular, reveals unique contribution of each mode to the NFD. Cyclists and mopeds drive performance dominance, shaping the NFD even when pedestrians dominate traffic composition (mode-share dominance). This study empirically validates existence of mixed two-dimensional traffic NFD including critical capacities and jam densities and highlights the aggregate impact of individual modes. The findings underscore the potential of area-weighted aggregation to account for heterogeneity in urban mixed traffic, offering insights into capacity and efficiency evaluations for non-motorized transport systems.

ARTICLE HISTORY
Received 27 January 2025
Accepted 2 September 2025

KEYWORDS
Network fundamental diagram; Voronoi cells; mixed traffic; video detection and tracking

1. Introduction

The continued advancements in artificial intelligence research are changing the technology industry productivity and the lifestyles of human beings (Cockburn, Henderson, and Stern 2018). In the area of computer vision, many scholars, and researchers from around the world have produced multiple tools that are being adopted in the engineering field (Simon 1983). Despite challenges such as visual occlusion in mixed urban environments, the improved video technology fosters their use in practical offline and online applications. There is a multitude of video files/data already produced in the transportation domain where object detection and tracking algorithms, based on neural networks, enable the identification of

CONTACT Pedram Beigi beigi@gwu.edu Department of Civil and Environmental Engineering, The George Washington University, Washington, DC 20052, United States

vehicles, pedestrians, and other obstacles on roads or in crowded cityscapes at a 'microscopic' level. These technologies contribute to enhancing road safety, optimising traffic flow, enabling the development of advanced driver-assistance systems (ADAS) and aiding in the implementation of efficient transportation logistics and predictive maintenance (Suvizi and Venkataramani 2025). While numerous commercial and academic systems detect various road users in real time, few solutions are specifically tailored and validated for the dense, multidirectional interactions of shared space environments (Bhat et al. 2021). Therefore, it is imperative to develop a system capable of detecting and tracking road users in a complex mixed environment for data extraction and analysis.

Once the detection and tracking algorithms provide real-world vehicle and pedestrian trajectory data, leveraging the Network Fundamental Diagram (NFD) enables an understanding of fundamental traffic flow attributes at a network level. Although we might not have a complete network-level dataset, the NFD enables us to derive insights from individual trajectories by aggregating them into network-wide metrics. This approach allows us to analyze relationships between traffic density, flow rate, and speed across different segments of the transportation system. NFD application allows traffic pattern analysis (Z. Zhang et al. 2024), congestion spot identification, and network optimisation. It aids in devising effective traffic management strategies, like synchronised traffic signals, congestion pricing, and dynamic routing guidance, while supporting the development of predictive models and simulations for assessing various scenarios and implementing traffic control measures to enhance network efficiency and safety. Concepts of density and flow originate from fluid mechanics and are typically defined for an infinite number of particles. Standard or directional methods of measuring these quantities locally suffer from large data scatter (Steffen and Seyfried 2010). In this case, a possible solution involves utilising Voronoi diagrams to allocate personal space to each object, minimising density scatter (Mullick et al. 2025; Vackova and Bukacek 2025). Similarly, calculating direction and speed from position differences between times with identical phases of movement gives low-scatter sequences for speed and direction.

In this study, we employ a detection and tracking approach to collect trajectories for measuring characteristics and relations between users in mixed environment. Then Voronoi-based space discretisation is used to bridge the gap between microscopic and macroscopic traffic analysis. The objectives are: (1) to enhance the accuracy of an automated tracking systems for identifying and analyzing various road users' movement in complex urban settings while validating such systems against manually extracted data; and (2) to leverage this unique data to investigate the existence of area-based NFD and the concept of mode dominance in two-directional mixed traffic.

2. Literature review

Accurate estimation and control of traffic flow remain pressing challenges for researchers and practitioners, especially in the context of increasingly dynamic and complex urban networks. To address this, it is essential to study macroscopic relationship between traffic density and flow across an entire network.

For traffic monitoring and control, the NFD has demonstrated successful traffic analysis and insightful performance measures (Aghamohammadi and Laval 2020; Du and Rakha

2019). The application of NFD is for estimating the traffic state of a neighbourhood in relation to its space mean speed (Mahmassani, Hou, and Saberi 2013). In addition, global traffic conditions of an urban road network in ordinary conditions may be measured with NFD based approximations (Musolino and Vitetta 2014). However, accurately constructing the NFD requires a complete and high-resolution dataset representing all road users. With advancements in detection and tracking algorithms, such detailed trajectory data is now attainable (Du and Rakha 2019). NFD-based traffic models are widely employed to support network planning and congestion mitigation strategies. These models help link aggregate flow and density measures, offering insights that inform signal control optimisation and overall traffic management (L. Zhang et al. 2020; Sleiman, Beigi, and Hamdar 2025). Many of the studies to analyze traffic flow have focused on models based on NFD that describe the relationship between aggregated flows and aggregated densities of transport networks (L. Zhang et al. 2020). Additional NFD studies aim to enhance signal control optimisation, incorporating the link fundamental diagram concept, and illustrating the link density and flow relationship, leading to recent advancements (Musolino and Vitetta 2014).

Saberi et al. (2014) evaluate measurement methods for traffic flow variables taken at the network level. Using three-dimensional vehicle trajectories, the method provides estimates of network flow, density, and speed. It extends the generalised Edie's definitions of fundamental traffic flow variables to the network level. Huang, Sun, and Zhang (2022) introduced a 3D-MFD (Macroscopic Fundamental Diagram) to analyze car – bicycle interactions using data from urban networks in Shanghai. Their results showed that higher densities and speed differences between modes negatively impact overall traffic performance, highlighting the need for improved multi-modal traffic management. Researchers have also explored pedestrian dynamics using video and modelling techniques. According to Plaue et al. (2011), data from crowds can be used for understanding different pedestrian dynamics. The authors suggested a method of extracting data from arbitrary angles which does not require other information such as the heights of the pedestrians. They adopted a Cellular Automata (CA) model (Ermentrout and Edelstein-Keshet 1993) and a Social Force (SF) model (Mehran, Oyama, and Shah 2009) to analyze the pedestrian trajectories and the associated flows and densities. The collected videos were analyzed using the width kernel density estimation. It was recognised that the unusual behaviours of the cyclists such as bending may affect the detection of their positions and thus the relative time and space headways with respect to other pedestrians and their heads' coordinates. While it is hard to measure the density of pedestrians or cyclists, the suggested method could use numerical simulations to validate the continuous locations of given objects. This was one of the first studies to utilise videos to analyze pedestrians' and crowds' dynamics.

Pedestrian and vehicular characteristics extracted and tracked from video allow for more elaborate and advanced traffic management applications. However, even though it is relatively simple to measure some target characteristics to differentiate between cyclists, pedestrians, and mopeds, it is harder to translate such characteristics to traffic measures (such as density, flow, and speed) accurately and efficiently. As an example, according to Steffen and Seyfried (2010), methods of measuring speed, density, and flow of traffic elements in a two-dimensional space, in general, and pedestrians, in particular, have been well studied; however, such methods often suffer from uncontrolled data scattering and lack of standards and common thresholds depending on the study area, the measurement method and the type of experiment/phenomena observed.

To address data scattering and measurement inconsistencies, researchers have increasingly adopted Voronoi-based spatial discretisation techniques. Steffen and Seyfried (2010) suggest the use of the Voronoi diagram (Erwig 2000) to reduce instances of data scattering. This approach also allows to calculate speeds from the positions of a relatively low number of objects and to translate such micro measures into basic macro quantities such as density and flow. However, even with the use of the Voronoi diagram, the measures found will depend on the boundaries of the study area(s) set: based on these boundaries, a density in an area A which accommodates N number of pedestrians will change significantly along with the reported flow-density data points. When dealing with non-constrained areas accommodating pedestrians or cyclists, one should be aware of the boundary setup. Once defined, the density in the area, the directional velocities (or the scalar speeds), and the flows along given virtual lines may be extracted. The authors further argue that the density can be measured per person per area increasing the possible data points gathered for added analysis. However, implementation has been done solely for pedestrian measurements.

In 2018, Nikolić and Bierlaire (2018) argued that pedestrian flow can be determined through a spatial–temporal discretisation model once trajectory data is provided. In particular, the authors claimed that it is challenging to understand and analyze continuous measures in space and time to understand different traffic variables. Accordingly, they developed Voronoi diagrams out of pedestrian trajectories. The discretisation parameters strongly influenced the outcome; in calculating pedestrian flows especially in complex traffic environments that may include motorised vehicles and bicycles, an analyzer should perform numerical explorations to identify the locations and the periods reflecting the general observed movement patterns (i.e. representability). In addition, the shape of the Voronoi cells plays a role in defining the traffic analysis outcome.

Other enhancements to Voronoi-based techniques include the use of phase and diffusion error corrections, which improve accuracy in areas of varying density (Wageningen-Kessels, Daamen, and Hoogendoorn 2014). The phase error corrects for estimated coordinates in high-density regions. The diffusion error corrects for pedestrians' locations in areas with medium to low densities. Using this approach is suitable for one- and two-dimensional pedestrian traffic flow measurements. These measurements can be extracted from live video feeds or be used for predicting future pedestrian behaviour. When tracking the pedestrians, the phase and diffusion errors are used to penalise any sudden drop or increase in speed or change in coordinates. The centres of mass of the Voronoi cells are used to define such changes.

Cao et al. (2017) uses Voronoi diagrams to calculate pedestrian flow, and Fundamental diagrams for uni-, bi- and multi-directional flows at corridors and crossings are investigated by a series of experiments. The study revealed that while measurement methods yield similar results under low densities, discrepancies emerge as density increases – highlighting the need for context-specific measurement techniques. However, this research focused solely on pedestrians and did not account for the added complexity of mixed traffic.

The literature reveals significant progress in both the modelling and measurement of traffic flow using NFD frameworks and trajectory-based data. However, most studies tend to focus either on motorised traffic or pedestrian-only environments, with limited attention to naturalistic, mixed urban settings that include cyclists, mopeds, and other vulnerable road

Figure 1. Security video snapshot of a transit station in Amsterdam, the Netherlands.

users. Moreover, while Voronoi-based methods effectively bridge micro- and macro-level analyses, their application in shared, multidirectional spaces remains underexplored.

This research aims to fill these gaps through two main objectives. The first objective is to develop an efficient trajectory tracking algorithm that is validated against manually extracted trajectory data of pedestrians, cyclists and mopeds in a mixed urban environment. The trajectory data extracted through the proposed detection and tracking system are plotted and analyzed in a quasi-real-time manner for extended durations of time. Using the outcomes from the first objective, the main objective is to analyze traffic with shared right of way in a naturalistic multidirectional setting by extracting aggregate macroscopic parameters through the Voronoi approach. The macroscopic parameters are then used to explore the existence of an empirically validated space-based network fundamental diagram that can be utilised for traffic management and control purposes.

3. Methodology

3.1. Data description

The video used in this study captures mixed traffic in a naturalistic multidirectional setting, provided by the Delft University of Technology (TU Delft), the Netherlands. The video is taken from a security camera located at a transit station in Amsterdam, the Netherlands, with a shared right of way. The video duration is 15 minutes and 6 seconds with a rate of 15 frames per second and a resolution of 704 × 480 pixels. To derive trajectories, detection and tracking were conducted. After that since the video does not offer a plan view of the study area, trajectory transformation and orthorectification were employed to project camera-observed trajectories onto an *x-y* top view using Homography matrix-based algorithms. Finally, smoothing is implemented to refine the trajectories. Figure 1 shows a snapshot of the video.

3.2. Detection and tracking algorithm

In this section, an enhanced version of DeepSORT is introduced to increase tracking accuracy. This is integrated within a detection algorithm, resulting in improved average precision, recognition speed, and overall tracking performance.

In the detection process, YOLO (Redmon et al. 2016) has been used since it is fast with real-time object detection capabilities, as it processes images in a single pass and is generally more efficient than some other methods. YOLOv5 has been used since it demonstrate that the Average Precision (AP) value and recognition speed of the improved YOLOv5s are both higher than those of YOLOv4 and YOLOv3 (Dwivedi 2021).

In the tracking process, several challenges arise, including the relatively low video resolution, which complicates the identification of distant objects from the camera. Additionally, the presence of pedestrians, mopeds, and cyclists in close proximity, combined with the camera's angled positioning, often leads to overlap, as illustrated in Figure 1. This overlap hinders the continuous tracking of target objects and gives rise to the 'clique' problem reported by Dehghan, Assari, and Shah (2015), where an object may be assigned different identity numbers due to interruptions in its detection.

In this study, an enhanced version of DeepSORT is used for multiple target tracking (MOT) (Andriyenko, Schindler, and Roth 2012). Using the DeepSORT algorithm's cosine distance, DeepSORT reduces the number of automatic identification switches in the SORT algorithm by calculating the similarity between tracked and detected objects. The squared Mahalanobis distance (McLachlan 1999) between Kalman predictions and actual detection is used to filter low likelihood matches. This study combines the Triplet Loss (Hermans, Beyer, and Leibe 2017) and centre loss function (He et al. 2018) with DeepSORT to improve the tracking accuracy. The Triplet loss function is currently a widely used loss function and commonly used in face recognition tasks which has made some progress in clustering features. Triplet loss tries to learn a feature space for the set triples (Anchor, Positive, Negative). The positive sample (Positive) is closer to the sample (Anchor), and the different types of Anchors are farther away from the negative sample (Negative). The concept is in line with the purpose of the image recognition task, which is to determine if two images belong to the same type of label given the training and test atlases. The function of the Triplet loss is:

$$L_T = \max(||f(A) - f(P)||^2 - ||f(A) - f(N)||^2 + a, 0) \tag{1}$$

where A is the input, P is the positive input of the same class as A and N is the negative input of a different class from A. a is the margin between the P and N.

Centre loss function L_{cl} (Pedregosa et al. 2011), includes the distance between the classes but also considers the reduction of intra-class differences. Applying the centre loss function can improve the feature resolution. The function of the Centre Loss is:

$$L_{cl} = \frac{1}{2}\sum_{i=1}^{N_b} ||x_i - c_{yi}||_2^2 \tag{2}$$

where N_b is the batch size, c_{y_i} is the centre value of all depth features of the y_i-th class and x_i is the features of the image. Combining these two loss functions improves deep feature discrimination and enhances feature discrimination while maintaining good intra-class compactness. We use the scalar γ to balance the centre loss functions, with a range of

Table 1. Comparison of the tracking methods by different performance measures.

Method	MOTA	MOTP	HOTA	IDF1	SFDA
DeepSORT	60.3	81.6	53.1	64.6	64.9
Enhanced method	61.4	82.2	53.6	64.7	65.4

$\gamma \in [0, 1.0]$: when γ is set to 0, the Centre Loss has no influence, and the total loss function is solely driven by the Triplet Loss. As γ increases, the influence of the Centre Loss becomes more significant. Also, keeping γ within the range of 0–1, we ensure that the Centre Loss contributes to the overall learning process without overwhelming the main task of distinguishing between different classes of objects. This adjustability allows for fine-tuning the model's training process to find an optimal balance that improves feature discrimination and intra-class compactness. The final function is:

$$L_f = \frac{\gamma}{2}\sum_{i=1}^{N_b} ||x_i - c_{yi}||_2^2 + \sum_{i=1}^{M} \max(||f(A) - f(P)||^2 - ||f(A) - f(N)||^2 + a, 0) \quad (3)$$

To test the proposed enhanced tracking method, MOT16 dataset (Milan et al. 2016) is used and its performance was evaluated using several metrics, including HOTA (Luiten et al. 2021), CLEAR (Bernardin and Stiefelhagen 2008), Identity, and VACE. As shown in Table 1, the method in this paper shows improvement for the MOTA and MOTP, HOTA, IDF1, and SFDA indicators; compared to the benchmark DeepSORT algorithm.

Figure 2 shows the process of detection and tracking adapted. The detection and tracking system requires the input videos and the identifiers assigned to each object are then specified. The algorithm can detect 13–15 frames per second with the GTX-2070 Graphics Processing Unit (GPU) using OpenCV. There may be a number of anomalies in the raw trajectory data, including sudden changes in movement and the detection of duplicates. To successfully eliminate these anomalies, smoothing and interpolation (Bezier-Curve fitting method (Han et al. 2010) – see the next section) is applied to account for appearing/disappearing objects and to extract realistic maneuvers. Figure 3 illustrates the detection and tracking results drawn on the video canvas.

3.3. Training and smoothing

Once the selection and tracking components of the hybrid system are specified and arranged, the 'label image master' tool is adopted to train the YOLOv5 and DeepSORT techniques (with modified loss function) to extract the pedestrians'/cyclists'/mopeds' trajectory data accurately. These algorithms are executed on the GPU to expedite the process and facilitate real-time applications (Suvizi et al. 2024).

To allow better evaluation, Bezier curve (Ketkar 2017) smoothing algorithm as shown in Figure 4 is used to smooth the trajectories. The target may be mistakenly recognised at a specific time step, which will cause serious errors when drawing the trajectories. Using the Bezier curve not only eliminates these kinds of errors but also allows for extracting traffic movement measures that are understandable and interpretable (An, Bae, and Talebpour 2023). Alternatives like Kalman filters or spline interpolation can also smooth trajectories. However, our preliminary tests showed that Bezier curves preserve natural turning maneuvers in highly multi-directional flows, without relying on explicit motion models (e.g.

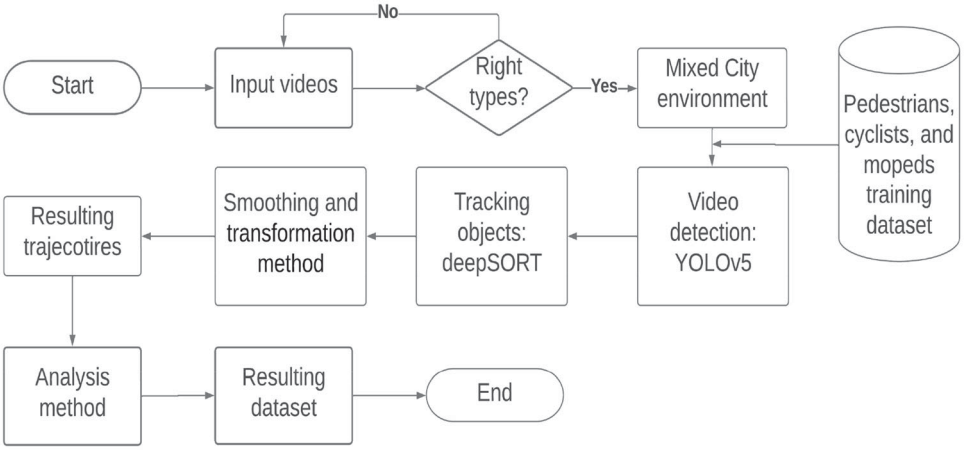


Figure 2. Detection and tracking process.

Figure 3. Sample illustration of the detection and tracking method adopted where the boxes bound the objects, and the lines retrace their trajectories.

constant velocity). When smoothing, the line between the first control point and the second control, the line between the last control point and the penultimate control point is the tangent line at the last control point. Written in a recursive manner, the function of the Bezier curve is:

$$B(t) = (1-t)p_0 + tp_1 \tag{4}$$

where the $p_0 and p_1$ are the first and second control points respectively, and t refers to the time difference between the detection/targeting timesteps associated with these points. The second-order function is:

$$B(t) = (1-t)^2 p_0 + 2t(1-t)p_1 + t^2 p_2 \tag{5}$$

The function of the n-order Bezier curve is:

$$P(t) = \sum_{i=0}^{n} P_i B_{i,n}(t) \tag{6}$$

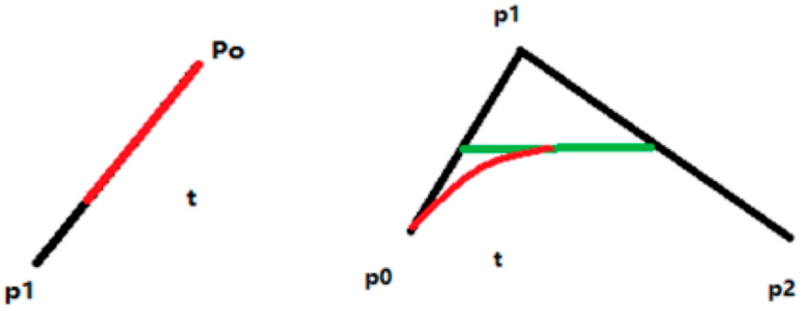


Figure 4. First order and second order Bezier Curve.

where:

$$B_{i,n}(t) = C_n^i t^i (1-t)^{n-i} = \frac{n!}{i!(n-i)!} t^i (1-t)^{n-i} \tag{7}$$

Once smoothed, the trajectories are transformed through orthorectification; the objective of the orthorectification process is to transform the trajectories from the camera view to the plan view. The wedge transformation method is used to achieve this goal. The ImageTracker (Knoppers, Van Lint, and Hoogendoorn 2012) is adopted in order to set the standard to determine the value of the relative parameters which include the scaling parameters.

3.4. Trajectory verification

For cross-validation of extracted trajectories, TU Delft has provided sample one-minute manual detection and tracking results. Following the steps mentioned earlier and to compute the speed and density measures and verify the feasibility of the results, the microscopic coordinate errors based on the trajectories of the bicyclists, pedestrians, and mopeds should be calculated. Having a set of Automated Extracted Data *(AED)* points extracted based on the steps mentioned earlier, at every real-time step and for every target, the measures of interest are expressed by c_x^{AED} and c_y^{AED} (where c can designate a coordinate or a velocity projection along a given coordinate axis). Having a manually extracted data set (*data*), each tracked target would have a matching verified measure designated by c_x^{data} and c_y^{data}. Accordingly, a relative error is calculated based on the following function:

$$F_{rel}[c^{AED}] = \left| \frac{\sqrt{c_x^{AED^2} + c_y^{AED^2}} - \sqrt{c_x^{data^2} + c_y^{data^2}}}{\sqrt{c_x^{data^2} + c_y^{data^2}}} \right| \tag{8}$$

To test for the error value between the recorded and the automatically detected trajectories, the average (across time and individuals) relative velocity error is used. The TU Delft's manually recorded trajectories were considered as the ground-truth based data. Accordingly, for the one-minute ground truth video, the errors for each period (10 seconds) may be computed as follows:

$$F_{rel}[V] = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\sqrt{v_{x_{i,j}}^{AED^2} + v_{y_{i,j}}^{AED^2}} - \sqrt{v_{x_{i,j}}^{data^2} + v_{y_{i,j}}^{data^2}}}{\sqrt{v_{x_{i,j}}^{data^2} + v_{y_{i,j}}^{data^2}}}}{n} \tag{9}$$

where*i* represents an object in the period of study,*j* is the time stamp (frame number) associated with one object in the period of study, *m* is the number of objects during the entire period of study based on the video detection dataset, *and n* is the total number of frames during which an object appears in the video detection dataset.

This error calculation does not require the identification of the processed video frames by frame number. It also does not necessitate the ability to map a specific frame number in the manually processed database (specific to different segments of the original recorded ∗15-minute video) to the frame number of the longer, more complete recording.

3.5. Fundamental diagram estimation

To estimate network flow, density, and speed, the method adopted in this study uses three-dimensional trajectory data (x, y, t) obtained from the automated video detection and tracking method discussed above. The space mean speed and density should be computed, as well as the associated errors, in order to estimate the fundamental diagram from the trajectory – data extracted from the mixed-city environment. The pedestrians, cyclists, and mopeds moving in the video have a variety of velocities. The following equation may be used to determine the scalar value of each velocity term:

$$V_s = \sqrt{V_x{}^2 + V_y{}^2} \tag{10}$$

where V_x is the scalar projection of the velocity vector in the x direction, and V_y is the scalar projection of the velocity in the y direction. These speeds can be then aggregated into space mean speed measures depending on the study area A under consideration and the time duration of interest (Knoop, Hoogendoorn, and van Zuylen 2009).

To calculate density using the Voronoi approach, every object is allocated a cell in area A_i. In the traditional approach, the area is fixed, such as the rectangle in the centre of the Figure 5. In contrast, for the Voronoi-based density, the area is defined as the collective sum of the Voronoi areas of the cells present within the designated analysis area. When a user is isolated, their larger Voronoi region accurately reflects low local interaction and thus a lower density. In contrast, in more crowded areas, the Voronoi cells become smaller, indicating higher local interaction. This dynamic weighting helps reduce scatter in fundamental diagrams (Steffen and Seyfried 2010; Wierbos 2021).

The steps to generate the Voronoi Diagram (Voronoi 1907) are: (1) locate the points that are distinct and create a Delaunay triangle network by automatically constructing a triangle network; (2) compute the number of the discrete points and triangles generated and keep track of the three discrete points that make up each triangle; (3) find and record the counts of all triangles next to each discrete point: to achieve this step, locating all triangles in the built triangle network that have the same vertex is needed; (4) produce a Voronoi polygon, sort the triangles adjacent to each discrete point in a clockwise or counterclockwise fashion; (5) calculate and record the centre of each triangle's circumscribed circle; (6) connect the centres of the circumscribed circles of these adjacent triangles according to the adjacent triangles of each discrete point to create the Voronoi polygon; a vertical bisector can be used to intersect the figure and produce a Voronoi polygon at the triangular mesh's edge.

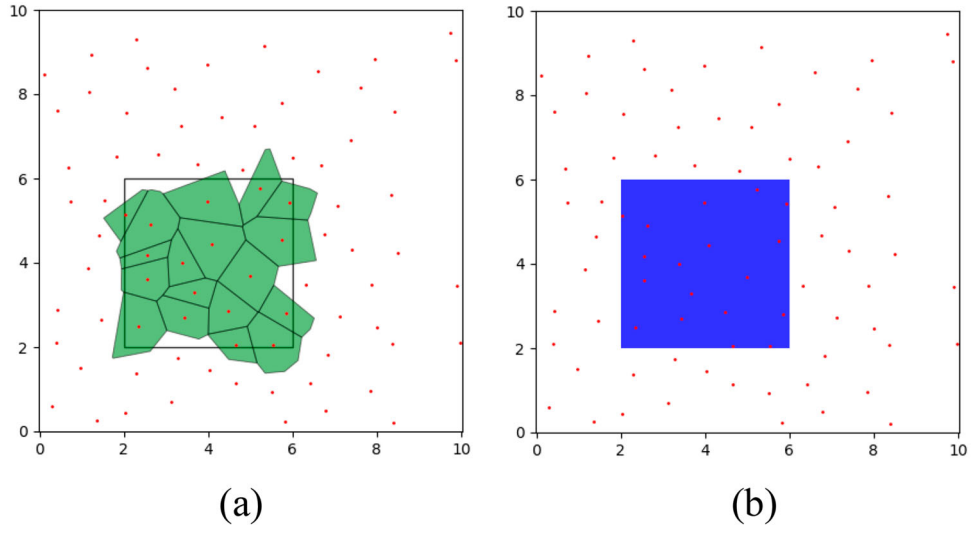


Figure 5. Comparison of density calculation methods: (a) Voronoi-based personal space approach (i.e. polygons associated with each object/point), and (b) standard fixed-area grid counting approach, accounting for partial grid areas within the study zone (i.e. centre square areas in both figures).

Flow q_i and density k_i for object type i during the time interval $[t, t + \Delta t$ is given by:

$$q_i = \sum_k \frac{d_{k,i}}{l_i \Delta t} \tag{11}$$

$$k_i = \sum_k \frac{t_{k,i}}{l_i \Delta t} \tag{12}$$

where $d_{k,i}$ and $t_{k,i}$ are the distance travelled, and the time spent by object k with type i. Also, l_i is the characteristic length associated with object type i and can be defined as the average linear dimension of the Voronoi cell for object i, $l_i = \sqrt{A_i}$. The q_i and k_i are calculated by averaging the values over the network. Using Mahmassani, Williams, and Herman (1984) equation to estimate networkwide link-based measurements, with modification to apply the Voronoi diagram method to account for space discretisation, the average network flow, average network density, and average network speed can be computed as follows:

$$Q = \frac{\sum_1^M A_i q_i}{\sum_1^M A_i} \tag{13}$$

$$K = \frac{\sum_1^M A_i k_i}{\sum_1^M A_i} \tag{14}$$

where Q is average network flow, K is average network density, and q_i, k_i are individual average flow and density, respectively, or each observation period. A_i is area of object i based on Voronoi diagram and M is total number of objects. These network-level equations are applied using two distinct approaches to assess their effectiveness. The first, standard aggregation approach, which generates Voronoi diagram including all road users present at a given moment to calculate metrics. The second, mode-isolated aggregation approach,

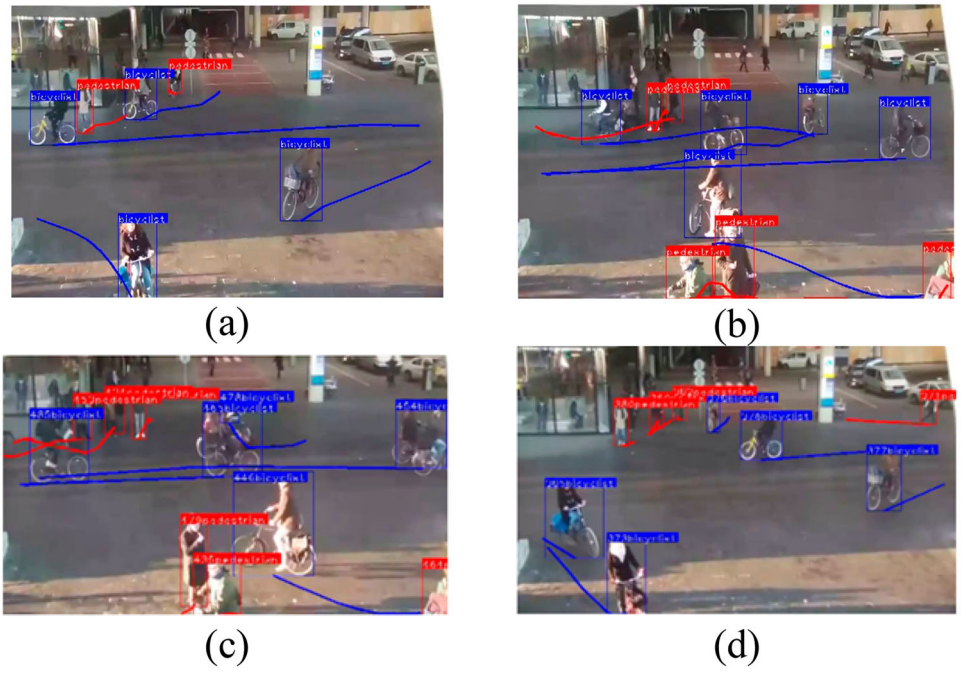


Figure 6. Video detection and tracking results in mixed city environment at different time instances.

which also begins by constructing Voronoi cells for all traffic units. Then, for each mode type (e.g. pedestrian, bicycle, mopeds), macroscopic quantities are computed using only the Voronoi areas associated with that specific mode, excluding areas from other modes. This enables mode-isolated analysis, resulting in separate density, velocity, and flow measures for each agent type. These measures support the derivation of fundamental diagrams that reveal the unique contribution and behaviour of each mode in the traffic stream, and bridge microscopic behaviour with macroscopic performance metrics, accounting for heterogeneity in both mode and behaviour.

4. Results and analysis

4.1. Trajectories

Figure 6 demonstrates the outcomes of the detection and tracking modules. Despite the challenges of relatively low video resolution, occlusion and overlap of targets, dynamic changes in target size and shape based on movement, and the similarity of target attributes, the fundamental modules exhibit accurate identification and trajectory reconstruction.

Figure 7 presents the trajectories of the total video duration after projecting the video field of vision into the plan view: the green lines denote cyclists' trajectories, the red lines denote pedestrians' trajectories, and the blue lines denote moped trajectories. Most cyclists move from left to right and vice versa – along a self-organised un-marked bike path. There is not a clear pattern for the pedestrians' and moped's trajectories. However, pedestrians are mostly concentrated at the lower end of the study area where there is a waiting area next to several bus stops (i.e. for alighting and boarding). These pedestrians seem to exit

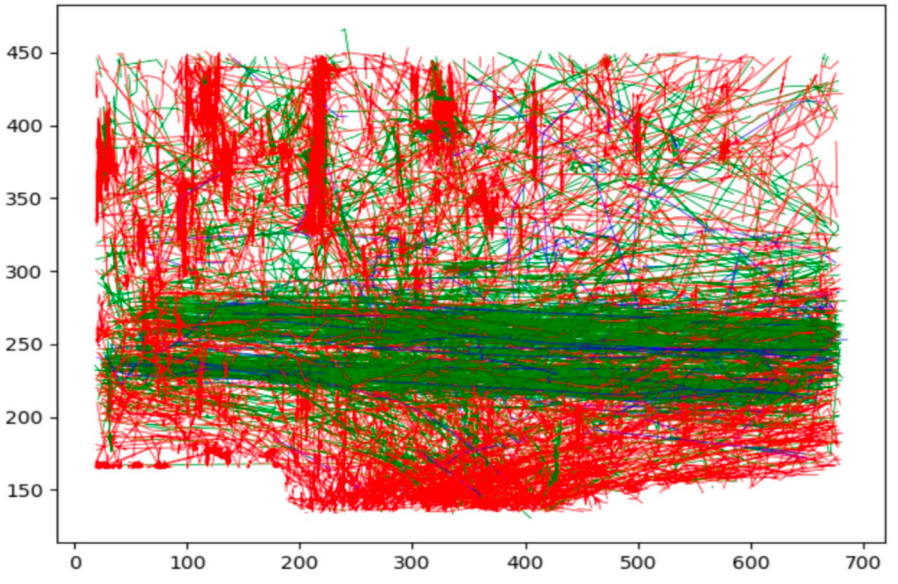


Figure 7. Trajectories plotted in a mixed city environment in a view based on a ∼15-minute complete video recording.

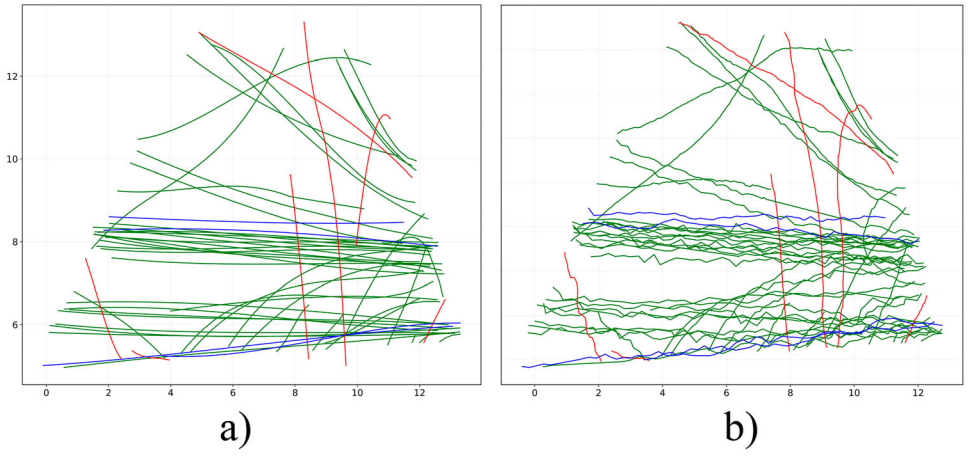


Figure 8. Extracted trajectories from a 1-minute video: (a) automatically extracted (smoothed), (b) manually extracted.

the train station (upper end of the study area) and move downward either to take a bus or be picked by other vehicles or walk toward the left or right directions.

Figure 8 shows the trajectories extracted manually and automatically from the sample video (1-minute video) after the transformation and orthorectification application. The (0,0) point is set to the top left side of the study area. Bezier curve fitting algorithm is used to reduce the coordinate error. The unit of each axis is metre; the red lines represent the trajectories of the pedestrians; the blue lines represent the trajectories of the mopeds; and the green lines represent the trajectories of the cyclists.

Table 2. Relative velocity error (cyclist and pedestrians).

Average Velocity of each period	0–10*s*	10–20*s*	20–30*s*	30–40*s*	40–50*s*	50–60*s*	0–60*s*
Manual (Cyclist)	2.80 *m/s*	2.73 *m/s*	3.03 *m/s*	2.64 *m/s*	3.05 *m/s*	4.15 *m/s*	2.97 *m/s*
Video Detection (Cyclist)	2.96 *m/s*	2.93 *m/s*	3.11 *m/s*	2.49 *m/s*	3.18 *m/s*	4.01 *m/s*	3.14 *m/s*
Mean relative error	0.05	0.07	0.02	0.08	0.06	0.03	0.05
Manual (Pedestrian)	1.20 *m/s*	–	0.94 *m/s*	0.68 *m/s*	0.70 *m/s*	0.85 *m/s*	0.90 *m/s*
Video Detection (Pedestrian)	1.17 *m/s*	–	0.89 *m/s*	0.68 *m/s*	0.74 *m/s*	0.81 *m/s*	0.93 *m/s*
Mean relative error	0.03	–	0.05	0.00	0.05	0.05	0.03

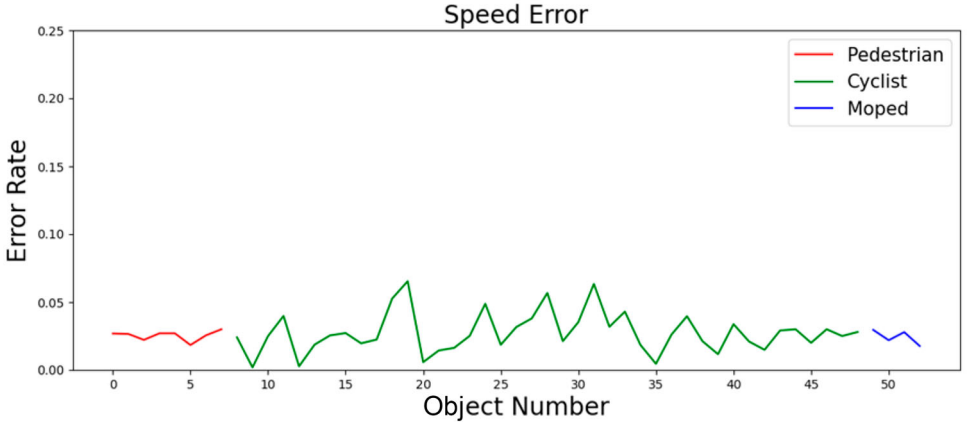


Figure 9. Per–object relative speed errors.

4.2. Errors and calibration results

According to TU Delft's manually extracted data, there are 7 pedestrians reported during the one-minute video, and 7 recorded by automated approach, indicating 100% accuracy in pedestrian detection. A total of 45 cyclists were reported by TU Delft, while 41 cyclists were recorded by automated approach, meaning 91% of the records were accurate. In total, TU Delft reports 4 mopeds and the automated approach records 4 mopeds, which represents an accuracy of 100 percent. Despite the complex environment, the adopted method detects almost all the objects (for example, cyclists, mopeds, and pedestrians) with an average of 93% accuracy – an encouraging result suggesting the method can be applied in a complex multi-directional shared environment.

Based on the directional values of the x and y vectors of the velocities, the relative speed error is calculated in Table 2. The mean relative error for cyclists ranged from 2% to 8% and for pedestrians from 0% to 5%. Mopeds were too few for reliable error metrics. Overall, the mean relative error was under 8%, demonstrating that despite the challenges of fast-moving objects, the method remains robust, establishing a strong foundation for macroscopic model validation.

Figure 9 shows the distribution of the speed error for each object. The results contain all the data which are compared frame by frame. The speed error during the manually analyzed time-period is around 0.05 which means that the manually recorded data and the detection data are highly matched.

In Table 3, the average coordinate error is displayed based on the project locations and timeframes. Approximately 12 segments of 5 seconds are used in the validation process in

Table 3. Relative coordinate error analysis (combining the x and y directional location).

Time period	Mean relative error (%)
Error from 0 *s*–5 *s*	0.01
Error from 5 *s*–10 *s*	0.03
Error from 10 *s*–15 *s*	0.01
Error from 15 *s*–20 *s*	0.03
Error from 20 *s*–25 *s*	0.04
Error from 25 *s*–30 *s*	0.05
Error from 30 *s*–35 *s*	0.05
Error from 35 *s*–40 *s*	0.08
Error from 40 *s*–45 *s*	0.01
Error from 45 *s*–50 *s*	0.01
Error from 50 *s*–55 *s*	0.02
Error from 55 *s*–60 *s*	0.03
Error from 0 *s*–60 *s*	0.03

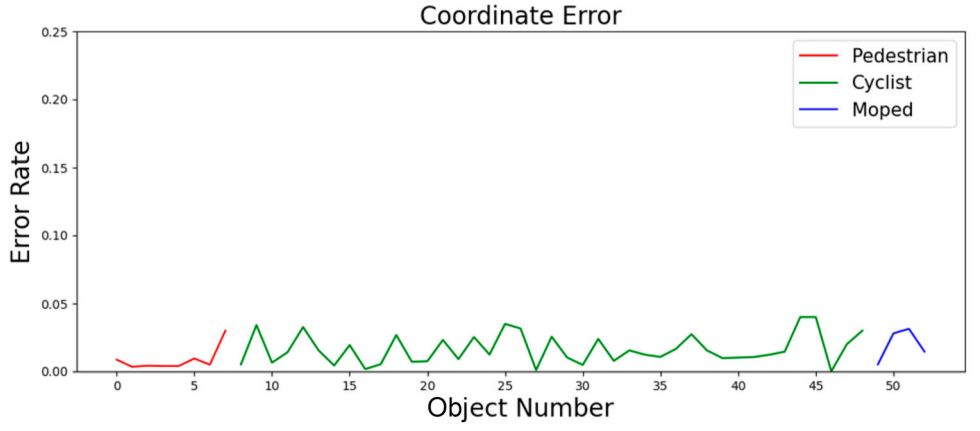


Figure 10. Per–object relative coordinate errors.

the one-minute video. The average relative coordinate error ranges from 1% to 8%. In this video, 3% is the average relative speed error.

Figure 10 shows the distribution of the coordinate error for each object. The error for each object is less than 0.04 which represents a high detection accuracy.

Therefore, the proposed system can convert video data efficiently into useful trajectories for pedestrians, cyclists, and mopeds in real time. Figures 11 and 12 illustrate the trend of speed and density over time, with colour indicating the speed and density values projected on a plan view.

4.3. Macroscopic flow measures

The benchmarked efficiently extracted trajectories for an extended period of time (around 15 minutes) are needed to quantify the macroscopic traffic performance measures used in analyzing the shape of different space-based network fundamental diagrams produced next. The outcomes presented in this section stem from two methodological approaches:

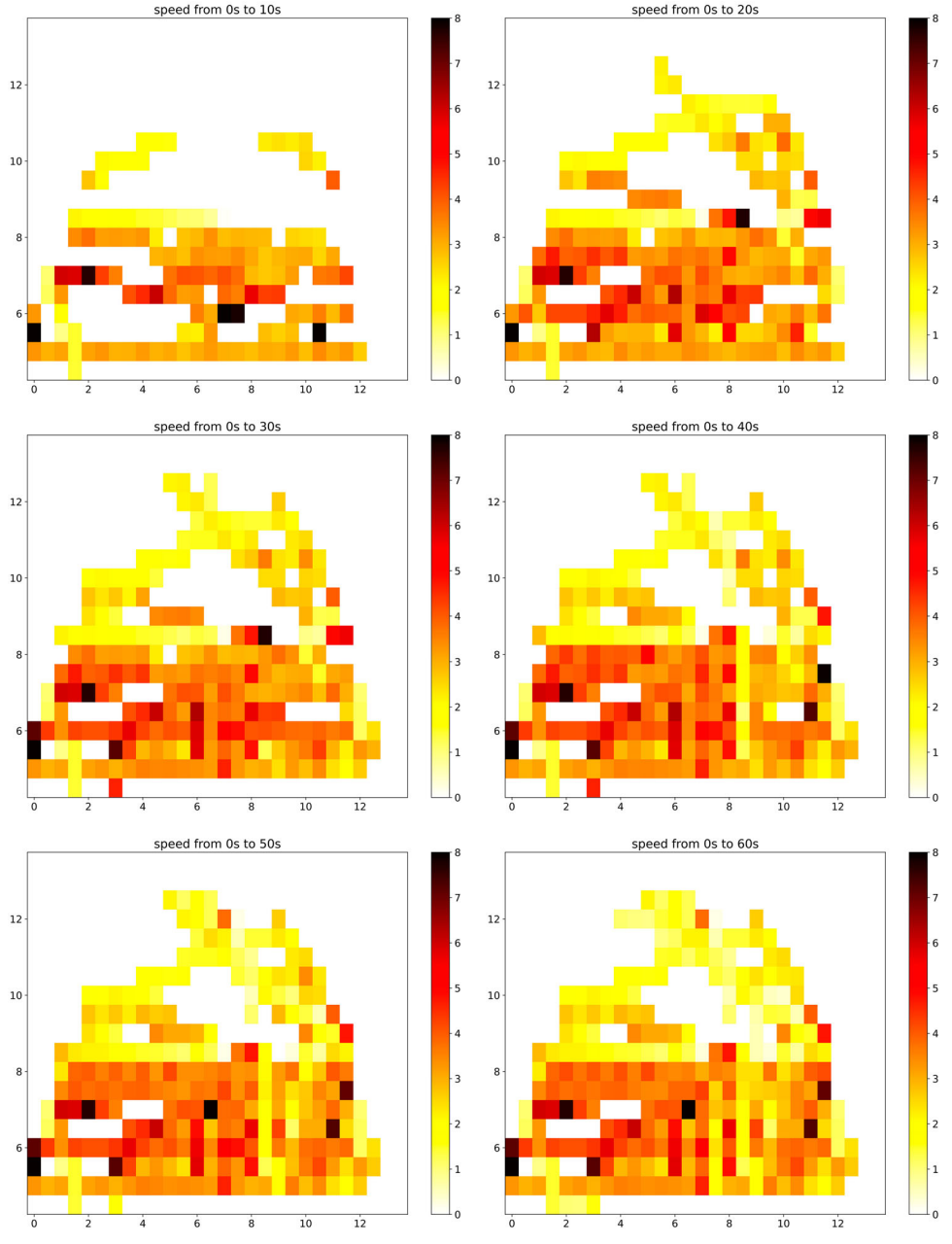


Figure 11. 2D spatial (*x-y* coordinates in *m*) speed distribution (colour indicates speeds in *m/s*).

(1) standard aggregation approach: construct a Voronoi diagram using all traffic units present in the area at a time step. Metrics are computed for the full population. The resulting values are then labelled based on the modal composition observed in each instance: if only a single mode (e.g. pedestrians, bicycles, or mopeds) is present, the point is labelled according to that mode; if more than one mode is present, the point is labelled as Mixed.

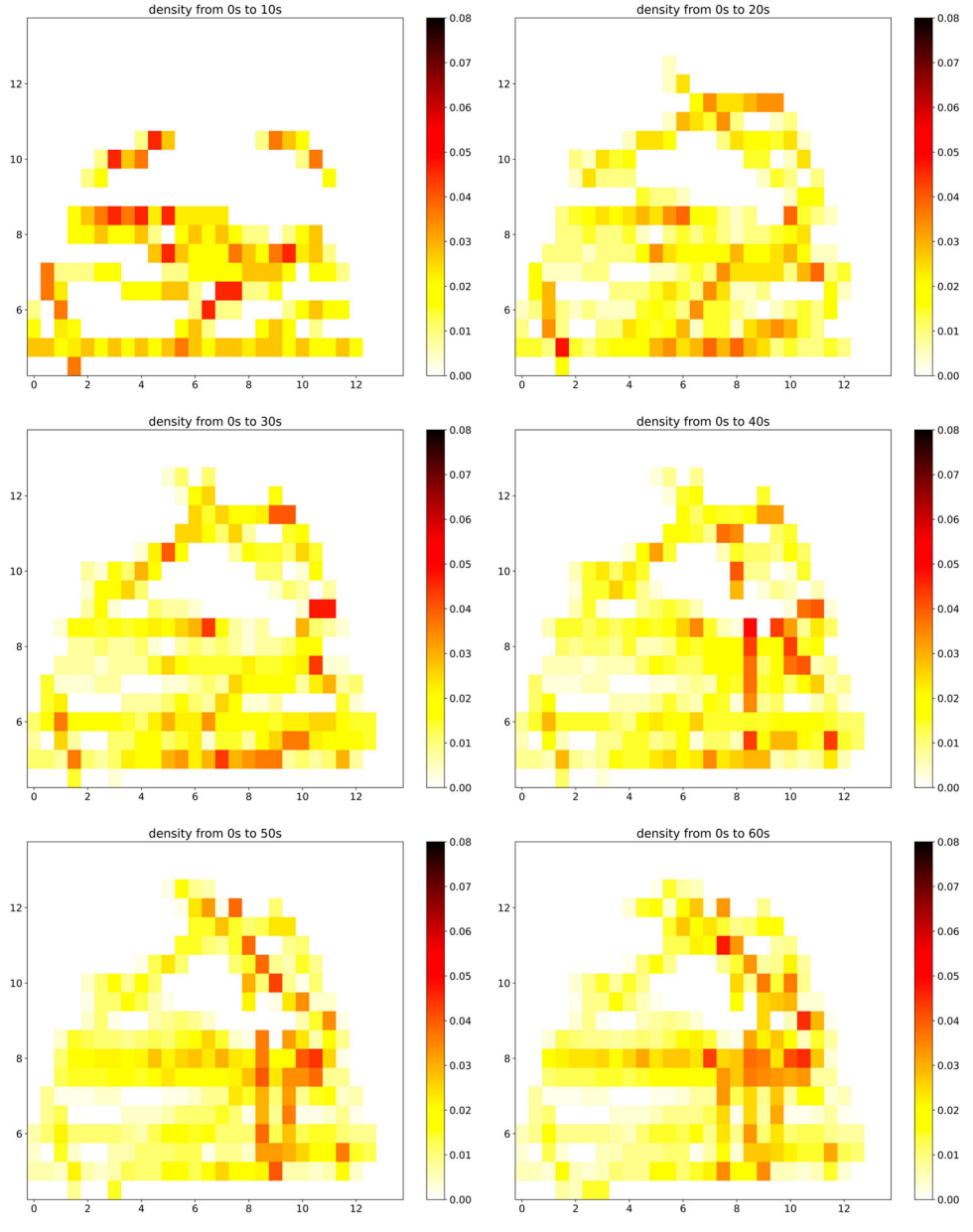


Figure 12. 2D spatial (*x-y* coordinates in *m*) density distribution (colour indicates densities in *objects/m*^{2}).

(2) mode-isolated aggregation approach: Voronoi diagram is constructed using all traffic units. Then, for each object type (e.g. pedestrian, bicycle, moped), metrics are computed using only the Voronoi areas associated with that specific mode, excluding others, allowing mode-isolated macroscopic analysis.

Figure 13 shows speeds, densities, and flows measurements using the standard aggregation method at each time step. In other words, densities are calculated by counting

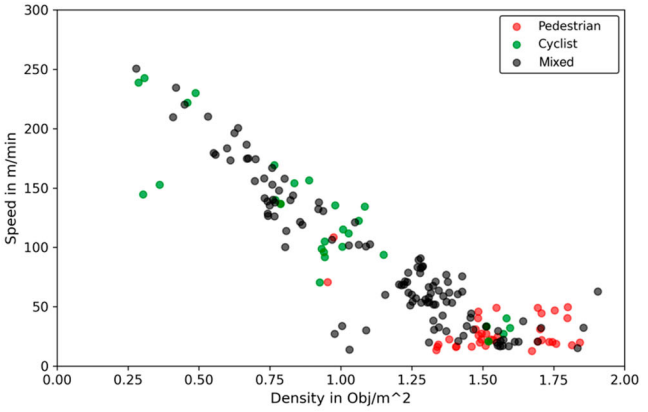
the number of objects within a general area at each time instance. The plots distinguish between different user types: 'Pedestrian' indicates that only pedestrians are present at that time, 'Cyclist' means only cyclists are in the area, and 'Mixed' refers to instances where a mix of users are present. The moped category is not in Figure 13 because there is no time instance where only mopeds are present in the area. The area we used to average for this number of objects is the area obtained by adding all Voronoi cell spaces for these objects (Figure 5(a)). The maximum flow recorded is 140 pedestrian-cyclist-moped/min (corresponding to a standardised value of 2.33 pedestrian-cyclist-moped/s-m); the maximum density is 1.87 *pedestrian-cyclist-moped/m*^{2} and the maximum free-flow speed is 250 *m/min* (corresponding to a value of 4.17 *m/s*.). The red points (pedestrian) indicate instances when only pedestrians are present in the area, and the same interpretation applies to cyclists. On the other hand, the black points (Mixed) denote situations where two or more types of objects coexist within the area (shared presence). Mopeds and cyclists tend to move at the highest speeds while pedestrians tend to stick around the origin (0-0 coordinates). Furthermore, there is little difference between the three modes of transportation in terms of jam density as high density is linked to overcrowding in waiting areas and other areas.

Using the standard aggregation method, we can see that the capacity (the maximum flow recorded in Figure 13(b): peak of the parabolic shape flow-speed function) and free-flow speed (maximum speed – left-hand side – of Figure 13(a)) aggregate measures are governed by the cycling mode (green points) while the jam-density (maximum density – right-hand-side of Figure 13(a)) measure is governed by both the pedestrian (red point) mode (to the greater extent) and the cycling (green points) mode (to the lesser extent). This is despite the fact that we have more pedestrians in the mix when recording the video. This finding is empirically supported with the entirety of the traffic fundamental regions covered starting with the congested region (left side of Figures 13(a,b)), going to the non-congestion break-down (moving from the ascending flow trend to the descending flow trend in the centre of Figure 13(b)) and ending with the non-congested region (right side of Figures 13(a,b)).

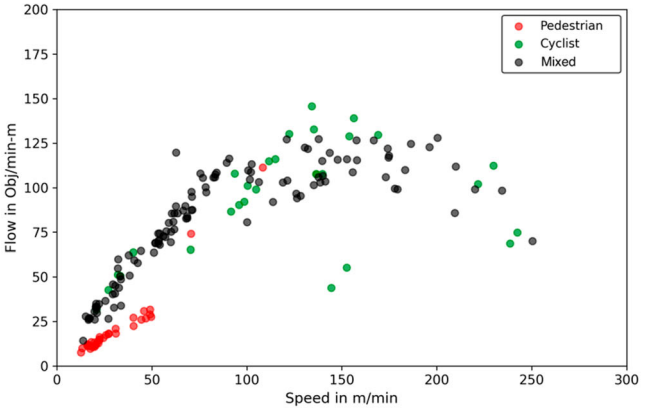
In the standard aggregation method, we construct a single Voronoi diagram over all objects present at each time step, regardless of mode. Flow, density, and speed metrics are computed using the Voronoi areas of all agents in the region, and the resulting measurements are tagged based on the modal composition (e.g. pedestrians only, cyclists only, or mixed).

In contrast, the mode-isolated aggregation method performs the macroscopic analysis separately for each mode after constructing a full Voronoi diagram that includes all agents. For each mode, only the Voronoi areas associated with agents of that mode are used in the calculations, while the influence of other modes is excluded. For example, if three pedestrians and one cyclist are present, the pedestrian metrics are computed using only the Voronoi cells of the pedestrians, ignoring the cyclist's cell. This process yields clean, per-mode estimates of flow, density, and speed at every time step – independent of the presence or absence of other modes – resulting in denser and more interpretable mode-isolated NFDs.

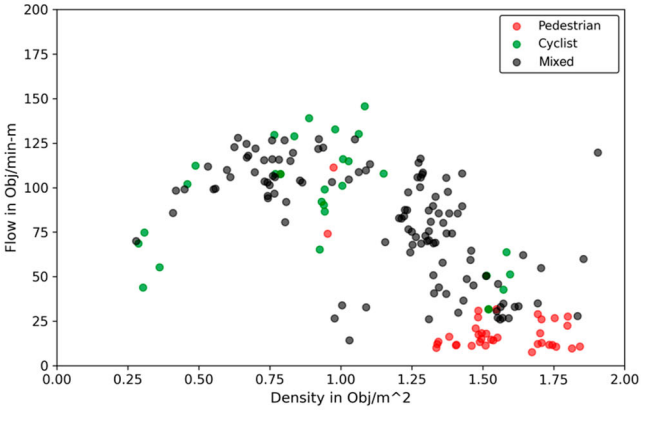
Results are shown in Figures 14–16 present results using the mode-isolated aggregation method, which are separated by user type – pedestrians, cyclists, mopeds, and the mixed mode – allowing clearer visualisation of the distinct flow-density-speed trends for



(a)



(b)



(c)

Figure 13. (a) Speed *(m/min)* – Density *(objects/ m^2)* diagram; (b) Flow *(objects/ min – m)* – Speed *(m/min)* diagram; (c) Flow *(objects/ min – m)* – Density *(objects/ m^2)* diagram from the ∼15-minute video using the standard aggregation method.

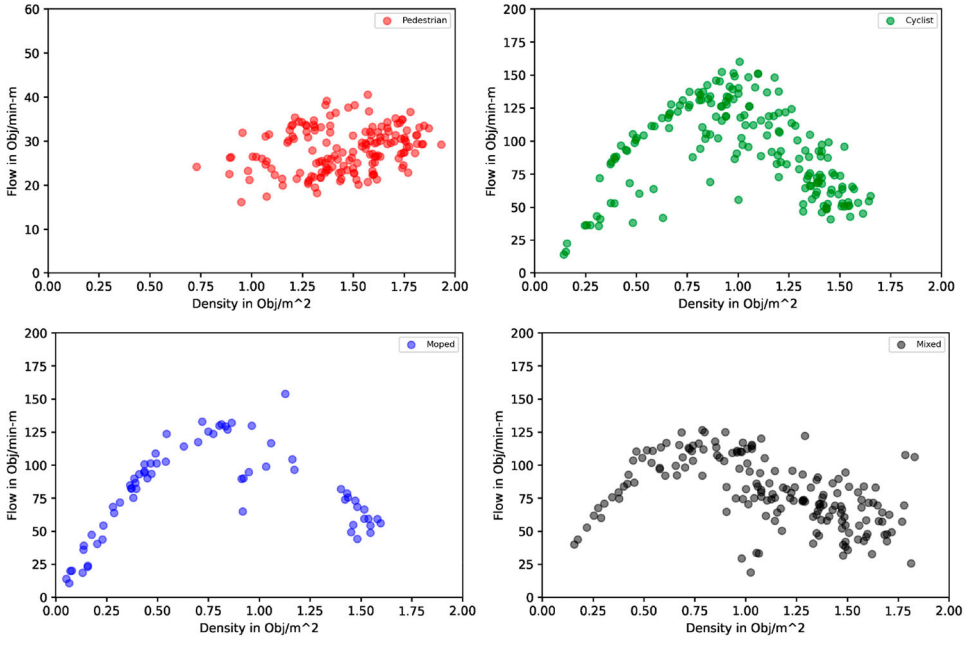


Figure 14. Relationship of flow and density for pedestrians, cyclists, mopeds and mixed.

each mode. Figure 17 combines all modes into a single plot and highlights how the different modes interact and collectively shape the macroscopic traffic dynamics in shared urban space.

After utilising the data filtering techniques to remove anomalies as described in Section 3C, the maximum flow recorded is 2.38 *pedestrian-cyclist-moped/s-m*, the maximum density is 1.93 *pedestrian-cyclist-moped/m*^{2} and the (maximum) free-flow speed is 4.9 *m/s*. The estimated parameters are context-specific and shaped by the local mode composition. The complete diagram covers a wider set of measurements based on the detailed trajectory specifically extracted for this purpose. The shape of the produced mixed space diagram is consistent with shapes reported in the literature and prove the existence of a space based mixed-environment network fundamental diagram for micro-mobility modes of transport. This diagram reported in Figure 17 remains one of the few empirically supported diagrams found in the literature. Since this study analyzes mixed traffic with shared right of way in a naturalistic multidirectional setting, the findings are unique as the authors are unaware of existing studies that produces validated results as the ones in Figure 13 for the standard aggregation approach, and Figure 17 for the mode-isolated aggregation approach.

Employing the Voronoi diagram for metric evaluation with an efficient trajectory data extraction from videos increases the number of data points produced. These data points are characterised by a reduced scattering and allow observing trends despite the fact that the number of objects detected and tracked is limited.

When using the mode-isolated aggregation technique to compute the mixed NFD within a given space, each instance can be transformed into three data point thus the ability to see more behavioural trends if examining Figure 17 if compared to Figure 13. It is obvious that three types of behaviours are seen in Figures 14–17 depending on the type of mode

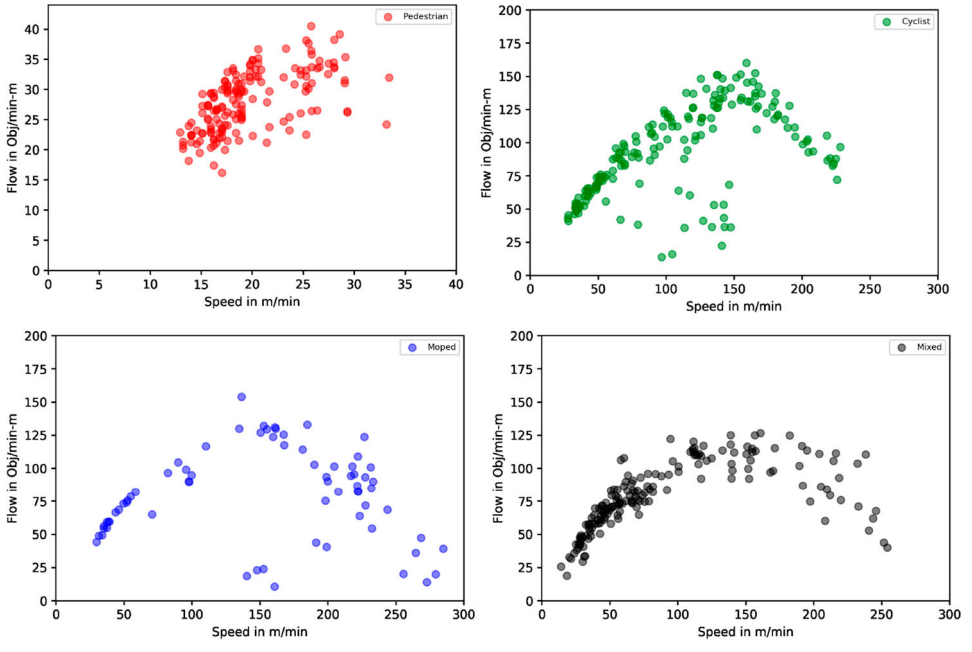


Figure 15. Relationship of flow and speed for pedestrians, cyclists, mopeds and mixed.

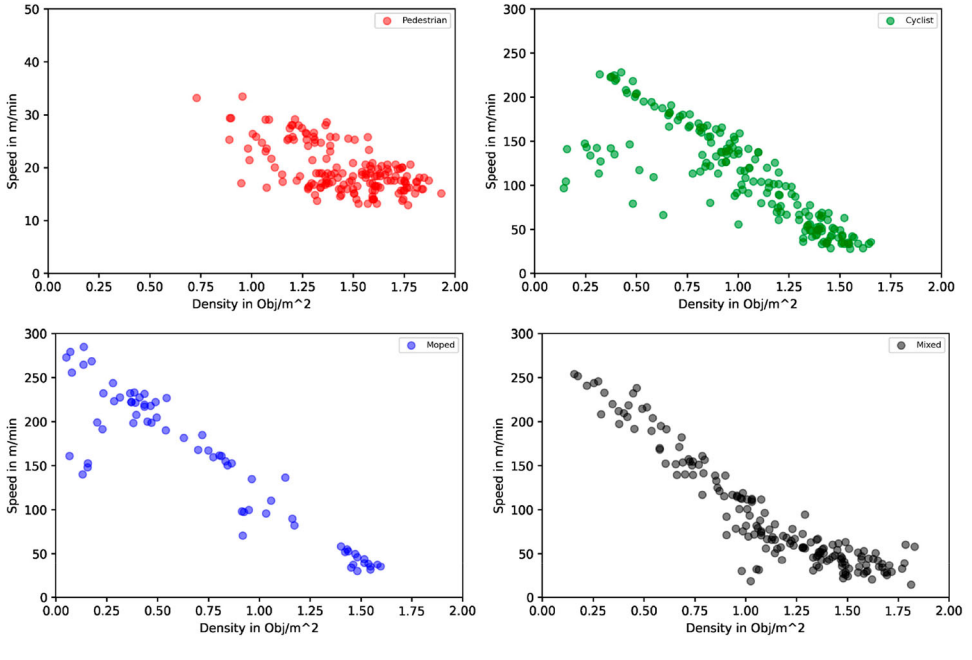
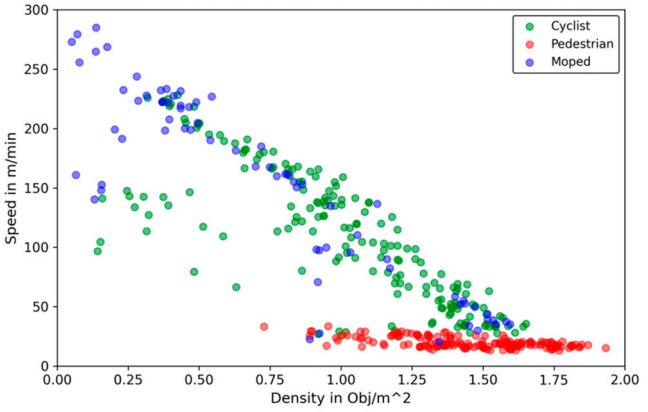
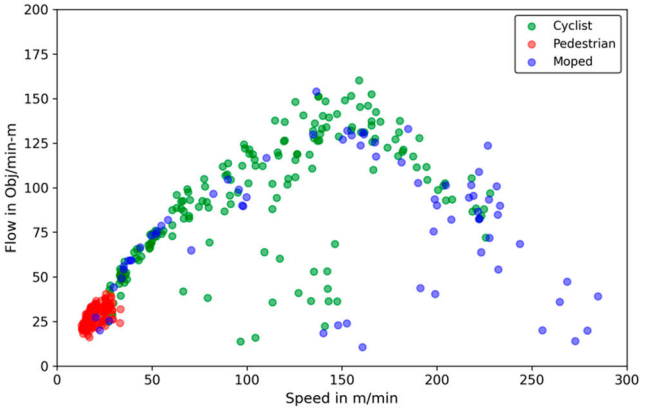


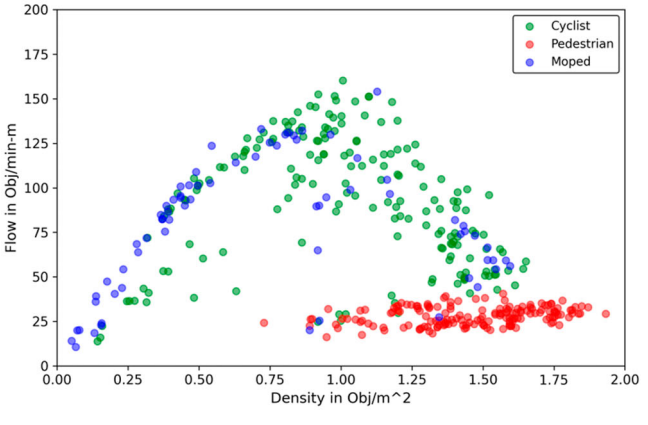
Figure 16. Relationship of speed and density for pedestrians, cyclists, mopeds and mixed.



(a)



(b)



(c)

Figure 17. (a) Speed *(m/min)* – Density *(objects/ m^2)* diagram; (b) Flow *(objects/ min – m)* – Speed *(m/min)* diagram; (c) Flow *(objects/ min – m)* – Density *(objects/ m^2)* diagram from the ∼15-minute video using mode-isolated aggregation approach.

being observed. As the jam density is dominated by the pedestrians' crowding phenomena, the free-flow speed is governed by the mopeds' speeding behaviour. The near–vertical cluster of points at high density but low flow – visible in Figure 17(c) – stems from pedestrians standing still while waiting at the bus station, producing the 'waiting–stem' effect that inflates density without adding throughput (Hoogendoorn and Daamen 2005). Finally, the highest capacity, jam density and free-flow speed (outer envelope of the fundamental diagram) are dictated by the movement of cyclists. In other words, despite having the pedestrians being the most observed in the study area – making them mode-share dominant, meaning they represent the largest proportion of traffic participants – the overall shape of the mixed network fundamental diagram is primarily governed by cyclists and, to a lesser extent, mopeds. These modes are therefore considered performance dominant, defined as those contributing most significantly to the maximum flow and free-flow speed in the NFD. This dominance is reflected by their presence along the upper envelope of the flow-speed and flow-density plots (see Figure 17), particularly in the regions representing peak performance. Accordingly, it is the cyclists that dominate the operation of this area outside of the studied Amsterdam Transit Station while the pedestrians set the jamming conditions forcing the bikes and the mopeds to be governed by the movement of pedestrians in the waiting areas. The newly produced NFDs (Figure 17) adhere to the fundamental shapes of other NFDs reported in the literature with different sub-dynamics possibly seen for each mode of transport: for example, the pedestrians display some stationary pedestrians waiting for transit thus the pedestrian data points can be clustered in one region rather than being spread across the entire flow-density spectrum. On the other hand, the cyclists show a triangular (sharp) transition in flow from the congested to the non-congested region. This transition is smooth for Mopeds. Additional analysis can be performed based on sub-regions of the study area or different time durations within the study recording/analysis period.

In addition to the above findings, some of the results are aggregated into 'per minute' measures to compare the results of this paper with the results reported by Rastogi and Chandra (2013) (as shown in Figure 18); in their study, Rastogi et al collected data at nineteen locations in five cities in India. We focus only on the flow-density and the speed-density diagrams. Most of the density points found in this paper are located between the 0 and the 2 *pedestrian-cyclist-moped/m*^{2} values, which is consistent with Rastogi et al.'s results. Rastogi et al. only focus on pedestrians, and Rastogi et al. speed is around 100 *m/minute*, flow range is between 0 and 55 *pedestrians/min-m*. Our pedestrian results show maximum speed is around 34 *m/minute* and flow ranges between 14 and 40 *pedestrians/min-m*.

It should be noted that the research team tried to regenerate all regions of Figure 18; however, the corresponding data sets were not explicitly provided. Moreover, a limitation of this comparative task is associated with the fact that (Rastogi and Chandra (2013) corresponds only to pedestrian detection. In future research, the authors will apply their detection and data aggregation methodologies on datasets collected in naturalistic environments and characterised by more complex pedestrians\bicycles\mopeds interactions. These data sets remain rare, and the research team is working on securing additional 2-D trajectories for such purposes (Ammourah et al. 2024).

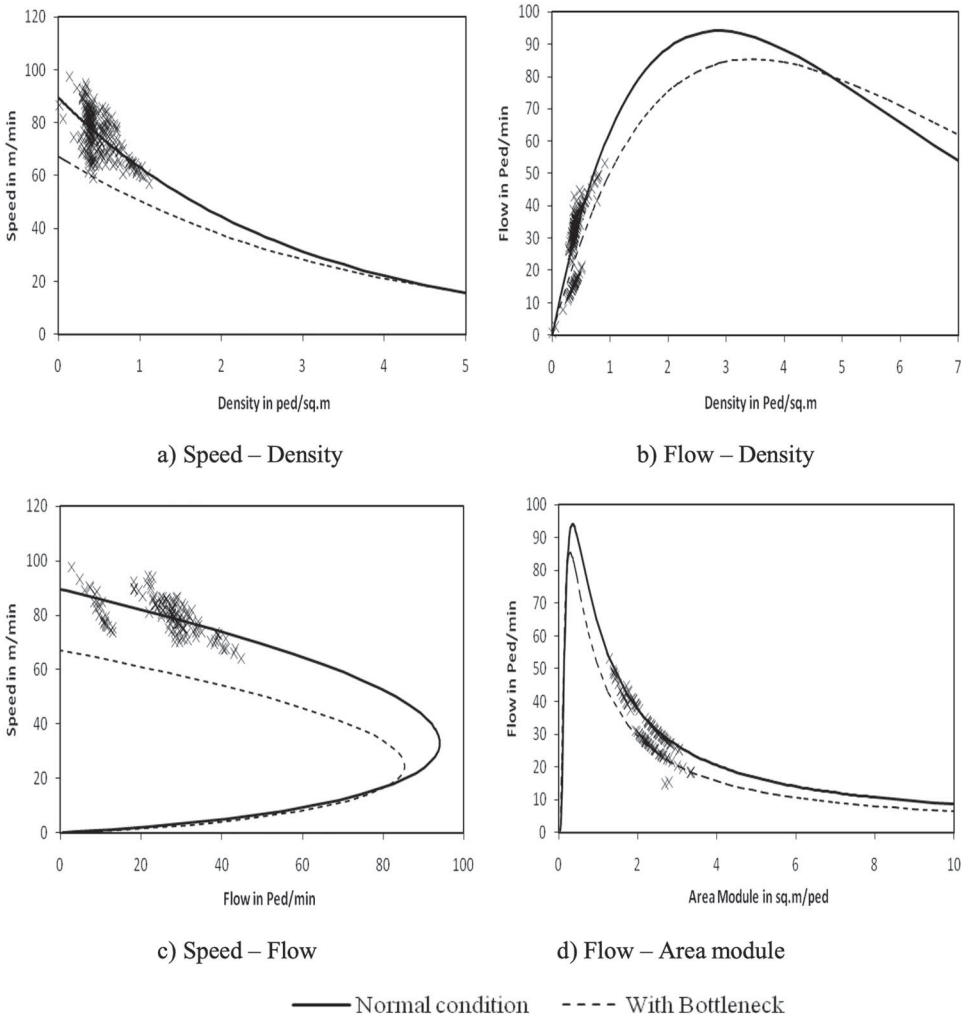


Figure 18. Flow-density-speed data points as reported by Rastogi and Chandra (2013).

5. Conclusion

This research introduces an enhanced approach for efficiently transforming video recordings of non-directional, non-bounded pedestrian, cyclist, and moped traffic in urban settings into valuable mobility and safety metrics. This method utilises YOLO, improved DeepSORT, and the Bezier technique. Cross-validation has been conducted to demonstrate the algorithm's performance, and the proposed method exhibits significant potential for application in detecting and controlling both non-motorised and motorised traffic.

This research develops and tests a novel, mode-isolated Voronoi aggregation method that empirically validates the existence of Network Fundamental Diagrams (NFDs) for mixed micromobility traffic in shared urban spaces. The study was designed around the implementation of two distinct approaches: a standard aggregation method and the novel

mode-isolated approach. It was this novel method that proved critical for disaggregating the complex traffic stream, which in turn enabled the identification of performance dominance versus mode-share dominance. By using the Voronoi cell to determine density at an individual level, this mode-isolated approach allows for the high-resolution analysis required to understand the nuanced dynamics of multidirectional traffic flows.

This Voronoi-based discretisation of space enables both mixed-mode and mode-isolated aggregation of macroscopic traffic metrics, revealing how different user types uniquely influence the shape of the Network Fundamental Diagram in shared, multidirectional urban traffic.

When assessing the NFD methodology, it becomes apparent that area-based NFDs exist for mixed two-directional traffic flow where some modes can dominate the traffic performance of the space (i.e. performance dominant) even if such modes do not constitute the majority of the road users (i.e. mode-share dominant). All empirically validated mixed naturalistic NFDs produced adhere to fundamental shapes reported in the literature.

Despite the aforementioned findings, the current analysis relies on a 15-minute dataset collected at a single site. While this duration sufficed for initial validation, the variety of observed states remains limited and the authors plan to collect large quantities of two-dimensional trajectories in complex urban settings. These settings include intersections (signalised and non-signalised) and mid-block crossing zones. Future work will extend our area-based NFD framework to include motorised vehicles and undertake a full hysteresis analysis, with the goal of informing mobility and safety assessments. In future studies, the following questions may be explored: (1) can the NFD shapes at different instances be linked to the microscopic time-to-collision measures? (2) Can the Level of Service (LOS) of a given intersection be calculated using the NFD technique per approach/movement? (3) At what market penetration rate, a given mode become performance dominant?

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Pedram Beigi http://orcid.org/0000-0002-3020-4085

References

Aghamohammadi, R., and J. A. Laval. 2020. "Dynamic Traffic Assignment Using the Macroscopic Fundamental Diagram: A Review of Vehicular and Pedestrian Flow Models." *Transportation Research Part B: Methodological* 137:99–118. https://doi.org/10.1016/j.trb.2018.10.017.

Ammourah, R., P. Beigi, B. Fan, S. H. Hamdar, J. Hourdos, C.-C. Hsiao, R. James, et al. 2024. "Introduction to the Third Generation Simulation Dataset: Data Collection and Trajectory Extraction." *Transportation Research Record* 2679: 1768–1784.

An, G., J. H. Bae, and A. Talebpour. 2023. "An Optimized Car-Following Behavior in Response to a Lane-Changing Vehicle: A Bezier Curve Based Approach." *IEEE Open Journal of Intelligent Transportation Systems* 4: 682–689.

Andriyenko, A., K. Schindler, and S. Roth. 2012. "Discrete-Continuous Optimization for Multi-Target Tracking." In *2012 IEEE Conference on Computer Vision and Pattern Recognition*.

Bernardin, K., and R. Stiefelhagen. 2008. "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics." *EURASIP Journal on Image and Video Processing* 2008:1–10. https://doi.org/10.1155/2008/246309.

Bhat, P. G., B. N. Subudhi, T. Veerakumar, and S. Esakkirajan. 2021. "A Fully Automatic Feature-Based Real-Time Traffic Surveillance System Using Data Association in the Probabilistic Framework." *IEEE Transactions on Intelligent Transportation Systems* 23 (8): 11088–11097. https://doi.org/10.1109/TITS.2021.3099729.

Cao, S., A. Seyfried, J. Zhang, S. Holl, and W. Song. 2017. "Fundamental Diagrams for Multidirectional Pedestrian Flows." *Journal of Statistical Mechanics: Theory and Experiment* 2017 (3): 033404. https://doi.org/10.1088/1742-5468/aa620d.

Cockburn, I. M., R. Henderson, and S. Stern. 2018. "The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis." In *The Economics of Artificial Intelligence: An Agenda*, edited by A. Agrawal, J. Gans, and A. Goldfarb, 115–146. University of Chicago Press. https://doi.org/10.7208/9780226613475-006.

Dehghan, A., S. M. Assari, and M. Shah. 2015. "GMMCP Tracker: Globally Optimal Generalized Maximum Multi Clique Problem for Multiple Object Tracking." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Du, J., and A. Rakha. 2019. "Constructing a Network Fundamental Diagram: Synthetic Origin – Destination Approach." *Transportation Research Record* 2673 (7): 479–488. https://doi.org/10.1177/0361198119851445.

Dwivedi, P. 2021. "YOLOv5 Compared to Faster RCNN. Who Wins?" Towards Data Science.com. Accessed 7 May 2021. https://towardsdatascience. om/yolov5compared-to-faster-rcnn-who-wins-a771cd6c9fb4.

Ermentrout, G. B., and L. Edelstein-Keshet. 1993. "Cellular Automata Approaches to Biological Modeling." *Journal of Theoretical Biology* 160 (1): 97–133. https://doi.org/10.1006/jtbi.1993.1007.

Erwig, M. 2000. "The graph Voronoi diagram with applications." *Networks: An International Journal* 36 (3): 156–163. https://doi.org/10.1002/1097-0037(200010)36:3 < 156::AID-NET2 > 3.0.CO;2-L.

Han, L., H. Yashiro, H. T. N. Nejad, Q. H. Do, and S. Mita. 2010. "Bézier Curve Based Path Planning for Autonomous Vehicle in Urban Environment." In *2010 IEEE Intelligent Vehicles Symposium*.

He, X., Y. Zhou, Z. Zhou, S. Bai, and X. Bai. 2018. "Triplet-Center Loss for Multi-View 3D Object Retrieval." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Hermans, A., L. Beyer, and B. Leibe. 2017. "In Defense of the Triplet Loss for Person Re-Identification." arXiv preprint arXiv:1703.07737.

Hoogendoorn, S. P., and W. Daamen. 2005. "Pedestrian Behavior at Bottlenecks." *Transportation Science* 39 (2): 147–159. https://doi.org/10.1287/trsc.1040.0102

Huang, Y., D. Sun, and S. Zhang. 2022. "Three-Dimensional Macroscopic Fundamental Diagram for Car and Bicycle Heterogeneous Traffic." *Transportmetrica B: Transport Dynamics* 10 (1): 312–339. https://doi.org/10.1080/21680566.2021.1994050

Ketkar, N. 2017. "Introduction to Keras." In *Deep Learning with Python*, 97–111. Berkeley, CA: Apress. https://doi.org/10.1007/978-1-4842-2766-4_7.

Knoop, V., S. P. Hoogendoorn, and H. van Zuylen. 2009. "Empirical Differences Between Time Mean Speed and Space Mean Speed." In *Traffic and Granular Flow'07*, 351–356. Springer.

Knoppers, P., H. Van Lint, and S. Hoogendoorn. 2012. "Automatic Stabilization of Aerial Traffic Images."

Luiten, J., A. Ošep, P. Dendorfer, P. Torr, A. Geiger, L. Leal-Taixé, and B. Leibe. 2021. "HOTA: A Higher Order Metric for Evaluating Multi-object Tracking." *International Journal of Computer Vision* 129 (2): 548–578. https://doi.org/10.1007/s11263-020-01375-2.

Mahmassani, H. S., T. Hou, and M. Saberi. 2013. "Connecting Networkwide Travel Time Reliability and the Network Fundamental Diagram of Traffic Flow." *Transportation Research Record* 2391 (1): 80–91. https://doi.org/10.3141/2391-08.

Mahmassani, H. S., J. C. Williams, and R. Herman. 1984. "Investigation of Network-Level Traffic Flow Relationships: Some Simulation Results." *Transportation Research Record* 971:121–130.

McLachlan, G. J. 1999. "Mahalanobis distance." *Resonance* 4 (6): 20–26. https://doi.org/10.1007/BF02834632

Mehran, R., A. Oyama, and M. Shah. 2009. "Abnormal Crowd Behavior Detection Using Social Force Model." In *2009 IEEE Conference on Computer Vision and Pattern Recognition*.

Milan, A., L. Leal-Taixe, I. Reid, S. Roth, and K. Schindler. 2016. "MOT16: A Benchmark for Multi-Object Tracking." arXiv preprint arXiv:1603.00831.

Mullick, P., C. Appert-Rolland, W. H. Warren, and J. Pettre. 2025. "Eliminating Bias in Pedestrian Density Estimation: A Voronoi Cell Perspective." *Physica A: Statistical Mechanics and its Applications* 657. https://doi.org/10.1016/j.physa.2024.130251

Musolino, G., and A. Vitetta. 2014. "Estimation of the Network Fundamental Diagram (NFD): An Urban Application in Emergency Conditions." *Transportation Research Procedia* 3:205–213. https://doi.org/10.1016/j.trpro.2014.10.106.

Nikolić, M., and M. Bierlaire. 2018. "Data-driven Spatio-temporal Discretization for Pedestrian Flow Characterization." *Transportation Research Part C: Emerging Technologies* 94:185–202. https://doi.org/10.1016/j.trc.2017.08.026.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg. 2011. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research* 12:2825–2830.

Plaue, M., M. Chen, G. Bärwolff, and H. Schwandt. 2011. "Trajectory Extraction and Density Analysis of Intersecting Pedestrian Flows from Video Recordings." In *Photogrammetric Image Analysis. PIA 2011. Lecture Notes in Computer Science.*

Rastogi, R., and S. Chandra. 2013. "Pedestrian Flow Characteristics for Different Pedestrian Facilities and Situations."

Redmon, J., S. Divvala, R. Girshick, and A. Farhadi. 2016. "You Only Look Once: Unified, Real-Time Object Detection." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Saberi, M., H. S. Mahmassani, T. Hou, and A. Zockaie. 2014. "Estimating Network Fundamental Diagram Using Three-Dimensional Vehicle Trajectories: Extending Edie's Definitions of Traffic Flow Variables to Networks." *Transportation Research Record* 2422 (1): 12–20. https://doi.org/10.3141/2422-02.

Simon, H. A. 1983. "Why Should Machines Learn?." In *Machine Learning*, 25–37. Elsevier.

Sleiman, W., P. Beigi, and S. Hamdar. 2025. "Percolation-Based Cloud Computing Framework for Network Reliability Assessment." In *2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI)*, MI, USA, 1–7. https://doi.org/10.1109/ICMI65310.2025.11139838.

Steffen, B., and A. Seyfried. 2010. "Methods for Measuring Pedestrian Density, Flow, Speed and Direction with Minimal Scatter." *Physica A: Statistical Mechanics and its Applications* 389 (9): 1902–1910. https://doi.org/10.1016/j.physa.2009.12.015.

Suvizi, A., S. Subramaniam, T. Lan, and G. Venkataramani. 2024. "Exploring In-Memory Accelerators and FPGAs for Latency-Sensitive DNN Inference on Edge Servers." *2024 IEEE Cloud Summit*, 1–6.

Suvizi, A., and G. Venkataramani. 2025. "Auto-Healer: Self-Healing Hardware for Perception Stage Faults in Autonomous Driving Systems." In *International Conference on Supercomputing (ICS), Salt Lake City, UT, USA*.

Vackova, J., and M Bukacek 2025. "Kernel Estimates as General Concept for the Measuring of Pedestrian Density." *Transportmetrica A: Transport Science* 21 (1): 2236236. https://doi.org/10.1080/23249935.2023.2236236.

Voronoi, G.. 1907. "Nouvelles applications des paramètres continus à la théorie des formes quadratiques." *Journal für die reine und angewandte Mathematik* 133: 97–178. http://eudml.org/doc/149276.

Wageningen-Kessels, F. V., W. Daamen, and S. Hoogendoorn. 2014. "Accuracy of Pedestrian and Traffic Flow Models." In *Symposium Celebrating 50 years of Traffic Flow Theory 2014 FTF Summer Meeting, Portland (USA), 11–13 August, 2014*. Authors version.

Wierbos, M. 2021. "Macroscopic Characteristics of Bicycle Traffic Flow." Doctoral diss., Ph. D. thesis, Delft University of Technology.

Zhang, L., Z. Yuan, L. Yang, and Z. Liu. 2020. "Recent Developments in Traffic Flow Modelling Using Macroscopic Fundamental Diagram." *Transport Reviews* 40 (6): 689–710. https://doi.org/10.1080/01441647.2020.1738588.

Zhang, Z., J Zhou, P Zhao, S Dong, and H Ding 2024. "Roles of Physical Collision Forces and Walking Characteristics on Pedestrian Traffic Flow Status." *Transportmetrica A: Transport Science* , in press. https://doi.org/10.1080/23249935.2024.2350645.