

Predicting traffic intensity in the urban area of Madrid: Integrating route network topology into a machine-learning model

Julián Moral-Carcedo

Universidad Autónoma de Madrid, Spain

ARTICLE INFO

JEL classification:

R4
R41
R58

Keywords:

Modeling traffic intensity
Urban mobility
Frequent routes
Neural network model
Anomaly detection

ABSTRACT

In addition to temporal effects, spatial effects must be considered in models of traffic intensity. Temporal effects (temporal autocorrelation) and spatial proximity shape observed traffic intensity, with repetitive patterns of peak and trough hours on weekdays and simultaneous peaks in traffic intensity picked up by traffic sensors along roadways of similar capacity. In this article we explore the ability of a neural network (NN) model to replicate both patterns, with the explicit inclusion of the route network's characteristics. Combining information on optimal routes between different origins and destinations in the Madrid metropolitan area with the localization of traffic sensors, we estimate the route use intensities that mimic the observed traffic patterns as those that can most accurately reproduce the traffic-intensity values recorded by the sensors throughout the urban area. We conclude that this modeling approach can in fact mimic the distribution of the main origins and destinations of displacements in the metropolitan area of Madrid, reproducing the spatial and temporal patterns observed in traffic intensity. In addition, we conduct an analysis of anomalies to test an eventual application of the proposed method to determine the effects on the traffic system of critical events, such as the implementation of the Madrid Central Low Emission Zone.

1. Introduction

The mobility of people living in urban areas, how they choose to travel and the routes they take have a direct impact on their quality of life. Mobility is also a fundamental challenge for urban planners, and managing the urban transport network is a priority for the authorities of large cities. Knowledge of traffic volume and intensity and their evolution over time are therefore fundamental for the correct dimensioning of the road network, the planning of public transport and the implementation of incentives for or restrictions on the use of private vehicles. In this article, we analyze traffic volume as a result of the multiple choices of route that the urban area's inhabitants make for their daily activities. Our objective is twofold: on the one hand, to provide an alternative approach to the modeling of traffic volume; on the other, and more interestingly, to provide a novel approach to characterizing critical urban routes as those that characterize the whole urban network's observed traffic.

Traffic volume/traffic intensity data (number of vehicles per unit of time) show a high correlation over time and across the urban road network. This correlation is due mainly to the temporal coherence of

circadian traffic patterns. Traffic intensity drops sharply at night and remains high during business hours. Because of this temporal correlation in traffic volume, as noted in (Manibardo et al., 2022), past traffic-intensity data, whether recorded on the same road or not, is generally a good short-term predictor of the traffic intensity measured at a given point.¹

High-performance traffic forecasting models ranging from ARIMA to complex neural-network models, as well as hybrid models, exploit this temporal correlation, and past values of traffic data are widely used to predict future values, as pointed out in (Tay et al., 2023), (Ma et al., 2020; Zhao et al., 2023), and (Boquet et al., 2020), among others.

When the focus is on modeling simultaneous data from multiple traffic sensors, spatial patterns become more relevant. These patterns appear in sensors located at nearby points, but in the same direction and/or on roads of similar capacity. Both spatial and temporal correlation can be exploited as well, as in (Deng and Qu, 2016; Yu H. W., 2017), (Kamarianakis and Prastacos, 2005), (Ma et al., 2020; Wang et al., 2019), among others. For instance, in the last case, critical paths from historical trajectories are first identified and each critical path is modeled and then stacked to incorporate temporal information.

E-mail address: julian.moral@uam.es.

¹ The traffic intensity at a given sensor on working days is correlated with the traffic intensity at any other sensor 24 h earlier.

Table 1

Overview (not exhaustive) of methodological approaches that exploit temporal and spatial dependence in traffic data.

Reference	Temporal dependence	Explicit analysis of Spatial dependence	Variable	Modelling
Wang et al. (2019)	yes	Critical Path	Speed	Path-LSTM. Critical paths derived from historical trajectories and modeled with LSTM NN model
Yu et al. (2020)	yes	Use of convolution layer on adjacency matrix.	Speed	Graph convolutional neural network
Kamarianakis and Prastacos (2005)	yes	Weighting matrices estimated on basis of distances between data-collection locations.	traffic intensity divided by road occupancy.	Space-time autoregressive integrated moving average
Peng and Zhang (2023)	yes	Graph convolutional network (GCN)	Traffic intensity	Memory Augmented Graph Convolutional Network
Huang et al. (2023)	yes	Sub-area division of road network	Traffic intensity	graph attention network (GAT) + long short-term memory (LSTM)
Zhao et al. (2023)	yes	Adjacency matrix + GCN	Traffic intensity	Multi-Spatiotemporal Fusion Graph Recurrent Network
Tay et al. (2023)	yes	Weighted adjacency matrix. Predefined weights, non-trainable.	Traffic intensity	Random Forest
Ghosh et al., (2009)	Yes	No	Intensity	Structural time-series model. Average travel time between two sites sharing same route or path within network is less than 15 min
Ma et al. (2020)	Yes	No	Intensity, Speed and Occupancy	Arima + Multilayer NN model
Polson and Sokolov (2016)	Yes	No	Speed	Multilayer NN model
Boquet et al. (2020)	No	No	Speed	variational autoencoder for missing data imputation

Other studies have used machine learning, as reviewed in (Tay et al., 2023), (Nguyen et al., 2018), and (Shaygan et al., 2022), to predict traffic intensity, congestion, speed, or accidents, because of their flexibility and non-parametric nature. These models exploit both the temporal correlation, usually with an RNN (recurrent neural network) or LSTM (long short-term memory network), and the spatial correlation, with a CNN (convolutional neural network), by considering traffic data at different points as if they were an image or a graph. Some use a combination of both (CNN and LSTM), as in (Peng and Zhang, 2023; Yu H. W., 2017; Bing Yu, 2018; Polson and Sokolov, 2016; Zhao et al., 2023; Zhao et al., 2020), where data from surrounding traffic sensors is taken into account or linked because they are on the same road or, as in (Yu et al., 2020), in wider areas. It is common to represent the sensor's network topology by adjacency matrices, with indications for when two sensors are connected by the same route. Network traffic-sensor topology captured in this way is also adopted in other approaches, such as that in (Huang et al., 2023), which combines a graph attention network with an LSTM model to capture spatio-temporal dependence in traffic data (see Table 1).

In general, NN models perform well in predicting the volume of traffic. However, as pointed out in (Manibardo et al., 2022), deep learning may not be the best modeling technique for every case. In some works, the sensors analyzed are selected on different criteria, such as their location on high-capacity routes, at accident hotspots, or on highly congested roads. However, in general, few studies have analyzed entire sensor networks (Manibardo et al., 2022). Compares the performance of different techniques, simple to complex, for different prediction horizons (1–4 periods). The authors conclude that there is no superior model for different prediction horizons. In any case, given the speed at which new developments are proposed, it is difficult to make a comparison of all techniques. Another limitation, especially for models that attempt to incorporate spatial dependence into their neural network, is the lack of direct interpretation of the estimated dependence between sensors. In some cases, the association estimated by the models indicates a high correlation of measurements between sensors, while in others, such as (Yu et al., 2020; Wang et al., 2019), it indicates complex behavior in propagating traffic patterns.

This paper, in contrast to other approaches, focuses mainly on the inclusion of the road's spatial characteristics, considering the most common itineraries in the urban area. The purpose is less to demonstrate this method's predictive accuracy than to explore the advantages of this type of modeling. The proposed approach provides a direct

interpretation of the results as well as important information on the intensity of origins and destinations of the urban area's "usual" displacements. Instead of representing the network topology by adjacency matrices between pairs of sensors,² our proposed method analyzes the use intensity of the routes that generate the observed traffic intensities, considering the entire sensor network. This approach allows us to estimate the use intensity of the most common routes, which is directly interpretable. The relative intensity of routes warns of changes in the conditions that affect mobility, providing valuable information for traffic management. To illustrate the potential uses of this approach, we propose an analysis of "anomaly detection" in the study of the effects of certain measures to restrict the circulation of traffic, as may be the case with the LEZ analyzed in (Moral-Carcedo, 2022) in the metropolitan area of Madrid.

The following section describes the proposed method. Section 3 describes the dataset. Finally, Section 4 describes the results, and Section 5 presents the main conclusions and limitations.

2. An NN model with embedded route network topology

Traffic in urban areas is usually monitored by a set of traffic sensors. Suppose there is a set of n traffic sensors in an urban area, which in the real world are typically placed along the routes most frequently used by residents in their daily travel. The distribution of the sensors in the road network can be sparse, with only a few sensors on selected routes, or dense, with many sensors throughout the urban road network, as in the Madrid metropolitan area. The traffic intensity (number of vehicles passing sensor i in an hour) at time t in the whole urban road network is collected in the vector $I_t = (i_{1,t} \dots i_{n,t})$ of size $(1 \times n)$. Typically, traffic intensity will be characterized by strong seasonal behavior, with similar values in the same time slots on the same days of the week, as can be seen in Fig. 1, with data from two traffic sensors in the Madrid metropolitan area.

Suppose the urban area consists of a set of L locations. Residents of the urban area continuously move between the L locations using a finite, but possibly very large, set Γ of m routes between each possible combination of origin and destination ($o-d$) of size $L \times L$. For rational agents who use only optimal routes (in terms of time) in their daily trips, we can

² By their relative proximity, if they are on the same road, or by the correlation of traffic recorded in two different locations.

assume that each $o-d$ pair has only one optimal route and that, by default, agents will use only that route. Then $m = |L| \times |L|$. Although the potential number of routes that each agent can take is huge, in the end there are only a limited number of routes that agents use in their daily lives, because of status quo bias (Samuelson and Zeckhauser, 1988). Let's assume that the set Γ contains only optimal routes. A route r_{ij} in Γ is composed of a set of points in \mathbb{R}^2 that connects origin $i, i \in L$, and destination $j, j \in L$, in the urban area on the available road network.

Each route can intersect the set of n traffic sensors placed in the urban area (see Fig. 2). Each route is associated with a sensor subset $T(r_{ij})$, which is composed only of sensors that are intersected by the route r_{ij} . It is worth noting that the number of routes is expected to be greater than the number of sensors $m > n$, and that any two subsets $T(r_{ij})$ and $T(r_{kl})$ can have common elements.

Assuming that the set of sensors is ordered in an arbitrary way in a row vector of length n (denoted id order), we can represent $T(r_{ij})$ as a row vector of length n , where a value of 1 at position k ($k = 1, 2, \dots, n$) means that the k sensor is intersected by route r_{ij} . Otherwise a value of zero is set at that position.

$$T(r_{ij}) = \begin{matrix} \text{n sensors} \\ (a_{i1} \quad \dots \quad a_{in}) \end{matrix} ; \quad a_{ik} = \begin{cases} 1 & \text{if } k \in T(r_{ij}) \\ 0 & \text{if } k \notin T(r_{ij}) \end{cases}, k = 1, \dots, n$$

We obtain the route-to-sensor intersection matrix T by stacking all the row vectors $T(r_{ij})$ corresponding to the m optimal routes:

$$T = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} ; \quad a_{ik} = \begin{cases} 1 & \text{if } k \in T(r_{ij}) \\ 0 & \text{if } k \notin T(r_{ij}) \end{cases} \quad k = 1, \dots, n; i = 1, \dots, m$$

Note that $\sum_j a_{ij}$ measures the adequacy of the sensor network to monitor traffic intensity along route r_{ij} , and $\sum_i a_{ij}$ measures the “centrality” of sensor j in the sensor network, i.e. the number of optimal routes that pass through sensor j .

Ideally, if we knew all the routes that agents living in the metropolitan area took at a given time, and their departure times, we could reproduce the observed traffic intensity across the entire network of traffic sensors. If the timing of the traffic data and the use of the routes coincided (the duration of the trip was no longer than the time interval for which traffic intensity was measured), it would be possible to reconcile the two. For example, knowing the number of people traveling in a given hour on a route that takes less than an hour to complete would provide an accurate estimate of the traffic intensities measured by the traffic sensors along that route in the same hour.

For the remainder of this paper, we will assume that route usage and sensor traffic records are coincident. In other words, we will assume that all routes take less than an hour to travel, and that all trips start and end between H:00 and H+1:00. We can expect this assumption not to hold perfectly in the real world and some randomness to end up in the data, because of the different timing of daily trips.

Let $R_{i,j,t}$ be the route usage intensity at time t — i.e., the number of people using route r_{ij} at time t , which is contained in the row vector of route usage intensities \mathcal{R}_t . We can obtain the $(1 \times n)$ vector of traffic intensities at time t , I_t , as the matrix product of the vector of route usage intensities (of size $1 \times n$) and the route-to-sensor matrix T (of size $m \times n$).

$$I_t = \mathcal{R}_t T \quad [\text{Eq. 1}]$$

where the element $i_{n,t}$ is the number of agents (vehicles) crossing sensor n at time t .

A simple example of the relationship between route usage intensity and traffic intensity through the node-to-route matrix is shown in Fig. 3. In it we consider two partially overlapping routes, EF and AB. The sensor network here consists of four sensors traversed by two routes, and the result in this particular case is the route-to-sensor matrix $T = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$.

In most cases, researchers have access to traffic intensity data, I_t , and the location of sensors in the urban area (which allow them to obtain the route-to-sensor matrix), but \mathcal{R}_t — i.e., intensity of use of the routes at hour t — is not available. To obtain an estimate of the intensity of route use, we can use the relationship shown in Equation (1), since it is possible to recover \mathcal{R}_t as a solution (James, 1978) of the linear system given by Eq. (1):

$$\mathcal{R}_t^* = T^+ \bar{I}_t \quad [\text{Eq. 2}]$$

where T^+ is the Moore–Penrose inverse of matrix T , and \bar{I}_t the vector $(1 \times n)$ of the mean hourly traffic intensity in hour t measured by the sensor network.⁴ This method allows us to obtain a numerical solution (because T does not have full column rank, the solution is not unique). However, there is no guarantee that this solution is meaningful. We will use the solution given by Equation (2) as a benchmark to highlight the advantages of the proposed method. Instead, we look for an optimal solution, \mathcal{R}_t' , that allows us to replicate the observed traffic intensities with the smallest possible MAE (mean absolute error) between the true data and the replicated data.

The proposed approach uses an autoencoder with a predefined, non-trainable decoder layer, the route-to-sensor matrix T , to determine route intensities. It also uses \mathcal{R}_t' as an intermediate layer, so as to output simulated traffic intensities. The autoencoder aims to replicate, with an MAE loss function, the true traffic intensities measured by the sensor network.

The autoencoder⁵ has the following structure (see Fig. 4), where I_t is the traffic intensity recorded at hour $t = (1, \dots, 24)$,

- Input layer (traffic intensity recorded by the n sensors at 15' intervals in the same hour, t , on working days), $I_t = \begin{bmatrix} i_{01:t:00}^1 & \dots & i_{01:t:00}^n \\ \dots & & \dots \\ i_{dd:t:45}^1 & \dots & i_{dd:t:45}^n \end{bmatrix}$.
- Hidden layer (Encoder): $\mathbf{h}_t = \text{ReLU}(\mathbf{W}_{1,t} \mathbf{I}_t + \mathbf{b}_{1,t})$.
- Layer for the determination of route intensities: $\mathcal{R}_t' = \text{ReLU}(\mathbf{W}_{2,t} \mathbf{h}_t + \mathbf{b}_{2,t})$.
- Decoder layer (non-trainable): $\hat{I}_{t,t} = \mathcal{R}_t' T$.

The NN will learn the non-negative route intensities, \mathcal{R}_t' , that mimic the observed traffic intensity measured by the entire network of sensors at hour t . That is,

$$\mathcal{R}_t' = \underset{i=1}{\text{argmin}} \frac{1}{N} \sum_{i=1}^N |I_{i,t} - \mathcal{R}_t' T|, i = 1, \dots, N \text{ number of available observations for hour } t.$$

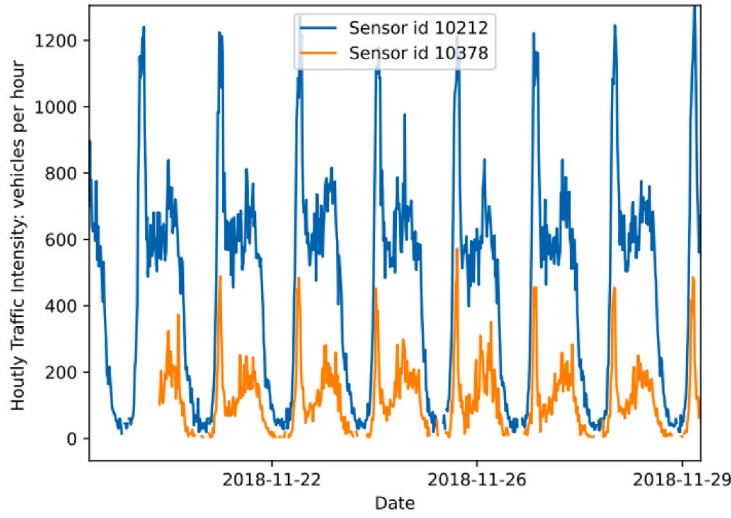
The NN inputs are the traffic intensities observed in hour t , and a different NN model is estimated for each hour $t = 0, 1, 2, \dots, 23$. Note that,

⁴ The complete set of solutions to Eq. (1) is given by: $\mathcal{R}_t^* = T^+ \bar{I}_t + (I - T^+ T)w$ Where I is the identity matrix of order m and w is an arbitrary m -dimensional vector.

⁵ Rectified Linear activation Unit (ReLU).

³ The subset can also have zero elements if no sensor is intersected by route r_{ij} .

Time Series



ACF plot

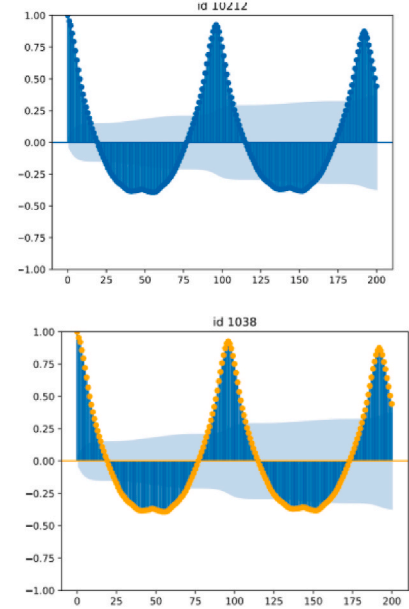


Fig. 1. Traffic intensity (vehicles per hour) recorded at 15' intervals by sensors ID 10212 and 10378. Working days, November 2018. Note: Hourly seasonality leads to a significant autocorrelation at lag 96 and its multiples.

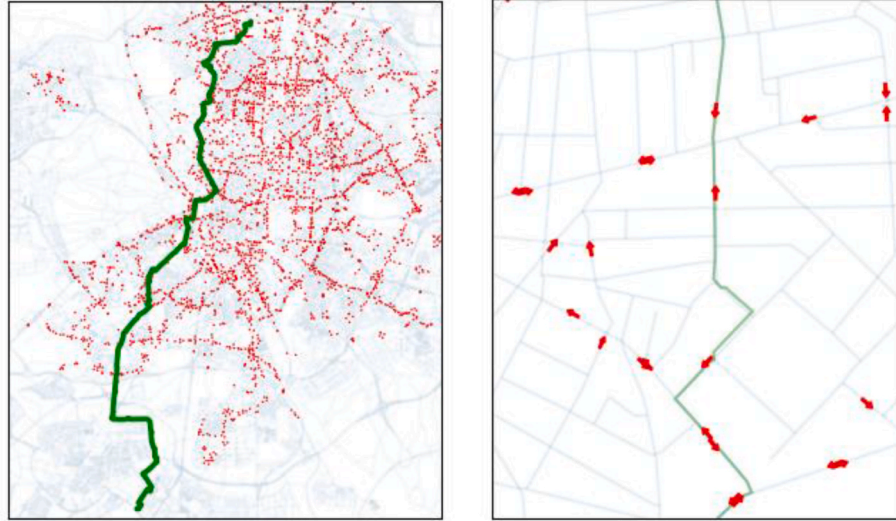


Fig. 2. Sample optimal route between two area centroids intersecting traffic sensors (red points, Madrid) and detail (right). Note: Arrows indicate direction of traffic.

although the traffic intensities measured by the sensors are very different in magnitude, the traffic intensity data enter the NN without any transformation. This is required by the final step in the NN architecture, since matrix T is applied to the non-negative route intensities estimated by the NN. Data transformation should therefore be avoided.

We have chosen a naïve modelling of the traffic intensity's time dependence, since the aim is to show that our procedure can reproduce the use intensity of the different routes. It is assumed that the traffic intensity at a given sensor depends only on the hour at which it is measured, $I_{it} = \bar{I}_i + e_{it}$, where e_{it} is a random error with zero mean. The determined intensity of use for a given route will therefore be the same for the same hours on different days, and will capture that route's hourly pattern of intensity of use. In the future, this simple approach could be modified by the inclusion of recurrent behavior, in either the use intensities or the traffic intensities of the routes.

3. Dataset description

3.1. Madrid traffic data

The Madrid region is densely populated, with 6.8 million inhabitants (Deloitte - Ipd, 2018) and about 3 million motor vehicles (Dirección General de Tráfico, www.dgt.es), of which 92.4% are cars. Close to the geographical center of the region is the municipality of Madrid (604 km², or 7% of the Madrid region's area), with about three million inhabitants. Madrid's traffic authorities use a network of sensors (around 3500 measuring points) to continuously monitor the intensity of traffic in the city. These sensors are located not only on the main roads, but also on other types of roads, providing a complete picture of traffic conditions in the metropolitan area. Traffic information is recorded by the sensors at 15-min intervals and provides data on traffic intensity

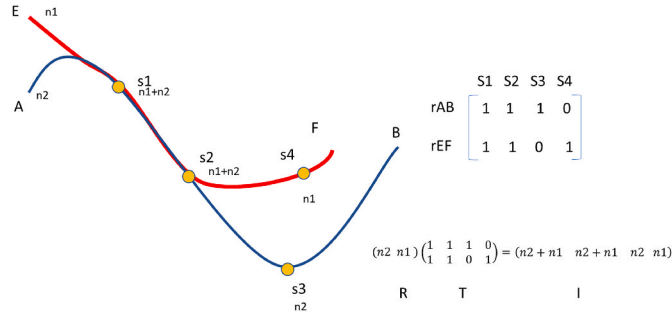


Fig. 3. Example. Simple route network with two optimal routes ($m = 2$) (AB and EF) and four traffic sensors ($n = 4$) (s_1, s_2, s_3, s_4). Note: $S = [s_1, s_2, s_3, s_4]$. Route AB intersects sensors s_1, s_2, s_3 . Route EF intersects sensors s_1, s_2 and s_4 . $T(r_{AB}) = [1 \ 1 \ 1 \ 0]$, $T(r_{EF}) = [1 \ 1 \ 0 \ 1]$. If n_1 drivers take route EF and n_2 take route AB, $\mathcal{R}_t = [n_1 \ n_2]$ then the sensors will measure intensities $I_t = \mathcal{R}_t T$, where I_t is the vector of traffic intensities measured by sensors S , and T is given by the stacked row vectors $T(r_{AB})$ and $T(r_{EF})$. Note that this setup can be extended to any number of sensors and routes.

(number of vehicles per hour) together with data on road occupancy and speed, but for these two indicators there is a significant lack of data in the period analyzed, so they are not included in this study.

Fig. 5 shows the location of the sensors, grouped into clusters by traffic intensity. It is important to note that the data used in this study do not cover the entire metropolitan area, as they are limited to the municipality of Madrid. As we will see later, mobility in the whole metropolitan area includes trips between areas outside the municipality of Madrid but along routes that pass through it. These too contribute to traffic intensity in the area where the sensors are located.

Fig. 6 shows a heatmap of the correlations of traffic intensities recorded by the Madrid City Council sensor network (3523 sensor points) in November 2018. The graph on the left includes all observations. The graph on the right includes only correlations calculated from traffic intensities recorded between 09:00 and 10:00 on weekdays in November 2018. In the latter case, if we select only those sensors that have a correlation greater than 0.8 with sensor ID 1001 (located on a

high-capacity entrance road to Madrid), we see (see Fig. 7) that they are located mainly on high-capacity roads, mainly at shorter distances, and in the same way and direction. Given these characteristics, the inclusion of spatiotemporal features may be useful in predicting traffic intensity, as in (Guo et al., 2019) (Fang et al., 2021; Xu et al., 2020; Peng et al., 2020), among others.

As shown in Fig. 5, there are clear differences in the intensity of traffic recorded by the sensors, in terms of both intensity levels and the hourly pattern of traffic intensity. In the case of the sensors included in the cluster with the highest intensity, the traffic peak is observed between 14:00 and 19:00, while in the remaining cases it occurs between 08:00 and 10:00. For the purposes of this article, we will consider all types, even though these differences might seem to recommend a separate analysis by sensors in the same group or an analysis limited to sensors located on the same type of road. We use traffic-intensity data recorded at 15-min intervals on working days in November 2018 and November 2019. Only sensors with the same location in both months are considered. The traffic data for each sensor is clipped at the mean plus two times its standard deviation to prevent outliers from influencing the estimates. The final dataset contains data from 2035 sensors after this cleaning.

3.2. Urban mobility data

Mobility characteristics in urban areas can be directly observed in mobile-phone positioning data throughout the day. These data, dated November 2019, were published by the Spanish National Statistics Institute (INE) during the COVID-19 outbreak (unfortunately, data publication has been discontinued). In this study, the entire region was divided into small mobility areas. These are basically areas where telephone companies can assign a position that preserves anonymity (each mobility area includes more than one census tract, with an average of 15,000 inhabitants). Mobility data provide the number of terminals moving from the “residence” of the phone to the “usual destination.” Each mobile phone’s “residence” is the area where the phone is located for the longest period between 00:01 and 06:00 on four selected working days (18–21 November 2019). The “daily destination” is the area

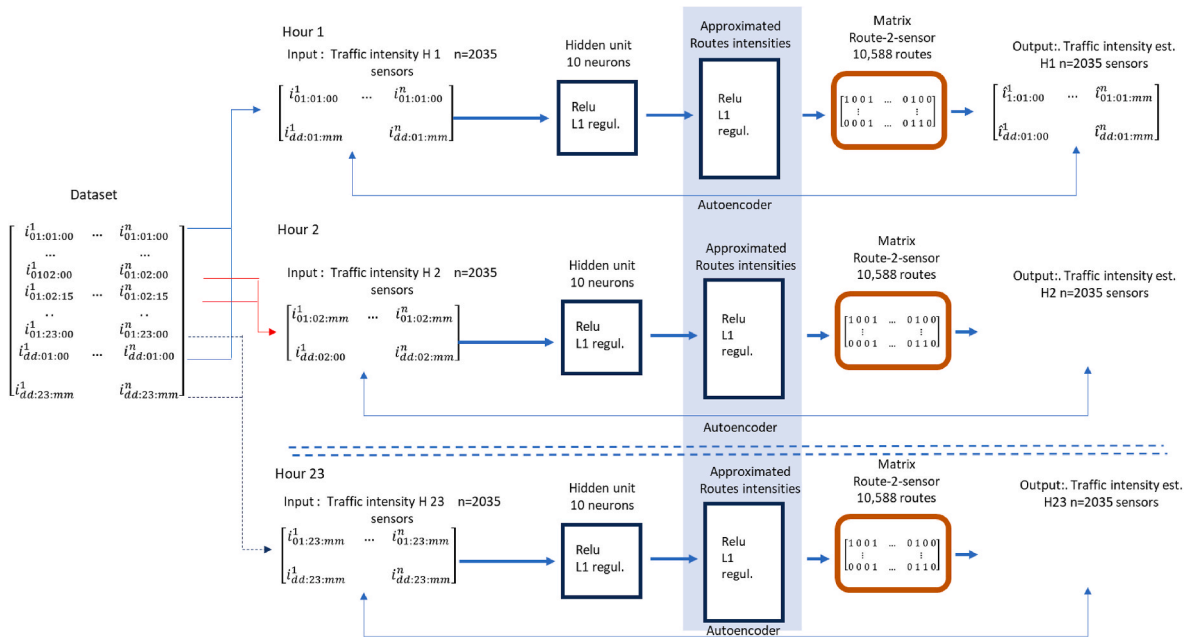
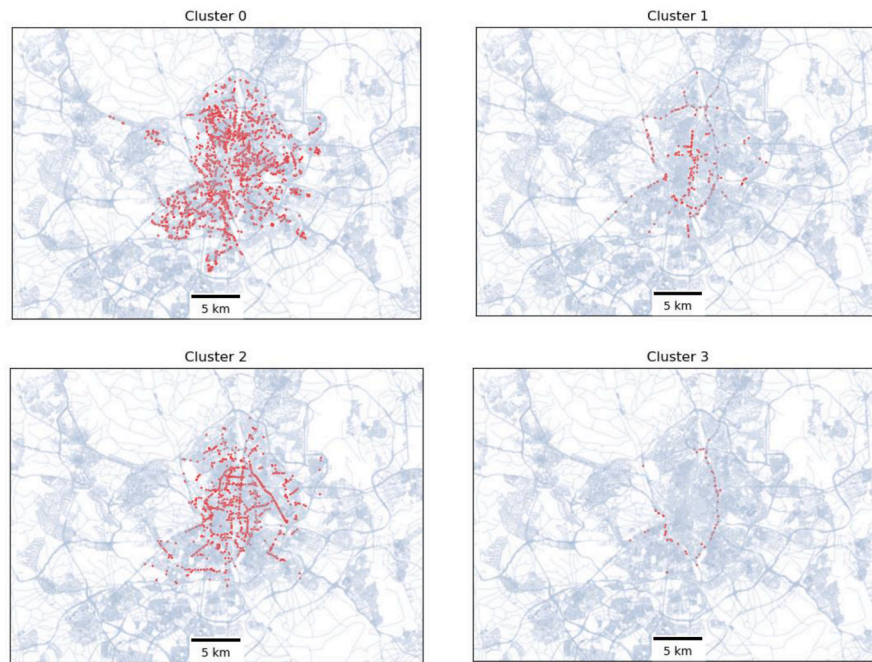


Fig. 4. Neural network architecture. Note: The intermediate dense layers include L1 regularization in the coefficients to reduce overfitting. The first intermediate dense layer (n inputs and 10 outputs) has a rectified Linear Unit (ReLU) activation function. The second layer (10 inputs and r outputs) also uses a ReLU function to ensure outputs values in the range $[0, \infty)$. The matrix “route-to-sensor” is fixed and coefficients are not updated, as it captures the topology of the traffic sensor network. As explained in the text, the input traffic intensity data is not transformed.

a. Location



b. Hourly traffic intensity by cluster. November 2018.

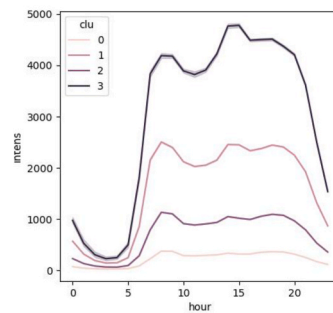


Fig. 5. Location of traffic sensors in Madrid. November 2018. Sensors grouped by traffic intensity. Note: Sensor grouped with the K-means algorithm with 4 groups. Traffic intensities recorded every 15 min on weekdays in November 2018.

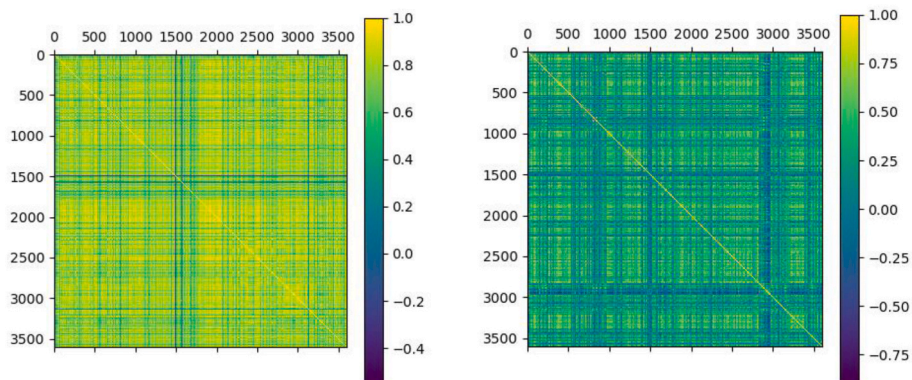


Fig. 6. Correlation in hourly traffic intensity measured by different sensors. Left, data recorded at all hours. Right, data recorded only at 09:00.

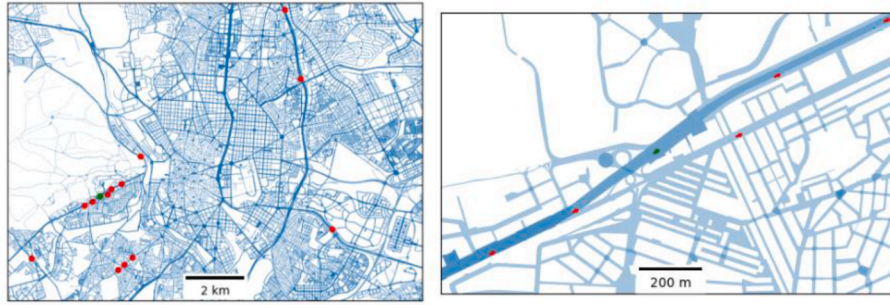


Fig. 7. Traffic sensors (red dots) that have a correlation greater than 0.8 in traffic intensity at 09:00 (15-min intervals) with sensor ID 1001 (green dot). Detail in the graph on the left. Data from November 2018. Source: Own elaboration based on Madrid City Council data (<https://datos.madrid.es/>).

(including residence) where the mobile phone is located between 10:00 and 18:00 (with a minimum stay of 4 h per day and on at least two of the four days observed). To preserve anonymity, no personal information is provided and only mobility flows involving more than 15 people are published. Despite these limitations, the study's information provides an overview of mobility in the Madrid region in sufficient spatial detail.⁶ Given that 39% of weekday trips in the region are made by car (Deloitte-Ipd, 2018), these mobility patterns should be reflected in the traffic conditions.

The mobility data (Fig. 8) show that the destinations with the highest intensity of displacement are relatively concentrated in the immediate surroundings of the municipality of Madrid. On the contrary, areas of origin with high outflows to other areas are distributed throughout the region and follow the layout of high-capacity roads. The greater spatial concentration of destination points, by comparison with the spatial distribution of origins, suggests a monocentric model of urban mobility, where people from the suburbs move to areas in the center of the metropolitan area for work or even leisure.

Using the information on the main origins and destinations of daily mobility, obtained from INE mobility data, and the urban network graph, we can construct optimal routes (minimum travel time) between the different origins and destinations. The INE mobility data do not provide explicit information on means of transport, but we will assume that residents use motorized vehicles to make their trips. In order to make this assumption more plausible, we consider only routes of at least 2 km along which at least 75 people reported the same origin and destination in their trips. Since we have traffic information only from sensors in the municipality of Madrid, we must consider only those optimal routes with origins and/or destinations in the Madrid metropolitan area. We have carried out this analysis with data from OpenStreetMap and Python libraries like NetworkX (Hagberg et al., 2008) and OSMNX (Boeing, 2017), restricting the road network graph to the bounding box delimited by latitude and longitude coordinates [40.665427, 40.277059, -3.3330721, -4.033846]. This procedure has resulted in a total of 10,588 routes for the same number of different origin-destination pairs.

4. Results

The main objective of the proposed approach is to estimate mobility patterns that mimic the spatial distribution of traffic intensity in the metropolitan area of Madrid and its changes throughout the day. As detailed above, we follow a naïve method to capture the time dependence of traffic intensity, focusing the analysis on the NN model's ability to reproduce the spatial pattern of traffic intensity. The dataset is

divided into hourly subsets (140,800 observations per hour with 80 observations per hour per sensor), and a different autoencoder, but with identical architecture, is estimated in each subset. In Table 2 and Fig. 9, we present different metrics for the performance of the tested model,⁷ such as MAPE, MedAPE, R2, and Kullback-Leibler divergence. In the same table, we present the results of different specifications, each varying the number of hidden layer neurons (the network architecture is represented in Fig. 4), and of the final specification, in which the number of hidden layer neurons is set to 10. The dimension of the last layer is kept fixed between specifications, since the number of routes in the metropolitan area is determined. The solution from Equation (2) fits the traffic data well, but the results are somewhat unsatisfactory, because of the data's high volatility. More importantly, the estimated use intensity of the routes as determined in Equation (2) is negative in 48.4% of the cases, which is unacceptable, whereas the NN method does not generate any negative intensities of use by virtue of its mathematical construction.

In Fig. 9, we compare the true hourly traffic intensity, as recorded by sensor, with the NN estimates, in terms of both their evolution and their comparative histograms. Fig. 10 shows the spatial distribution of true and estimated traffic intensity at different hours. These are highly similar, although the values of the traditional goodness of fit measures R^2 and PEAM are not fully satisfactory, because of a relative underestimation in sensors with medium and low traffic intensity. This is evident globally as well and is perhaps caused by high volatility in the dataset.

One of the advantages of the proposed method is that it can recover from the NN the intensity of use of the different routes, R'_t , insofar as it reproduces the traffic intensity as measured by the sensor network at hour t . With these estimations in hand, it is possible to calculate which areas are the greatest flow receptors and which are the greatest flow originators. Figs. 11–13 show both estimates, inflows to destinations and outflows from origins, and compares them with the INE mobility data. As explained above, the mobility data records only how many people living in each area are observed in another area for at least 2 h between 10:00 and 18:00, but it is silent on the route they take and their mode of transport. Data on people spending less than 2 h in the same place outside their residence is also missing. The proposed NN model uses traffic data to estimate hourly route use intensities that mimic the observed spatial distribution of recorded traffic intensities. Despite these differences, the correspondence between mobility data and estimated inflows (left graph in Fig. 13) by destination area (187 areas) and in terms of outflows by origin area (right graph in Fig. 13) appears clear. Unfortunately, there is no clear relationship between estimated trip intensities for each origin-destination pair and mobility flows for the same pairs as reported in the INE data.

⁶ The main limitations of this dataset are that it lacks information on mode of transport used, route taken, and trips that fail to meet the minimum stay criteria. Also, the 10:00–18:00 time window excludes leisure trips.

⁷ Tensorflow library for Python used for calculations (Martín Abadi, 2015).

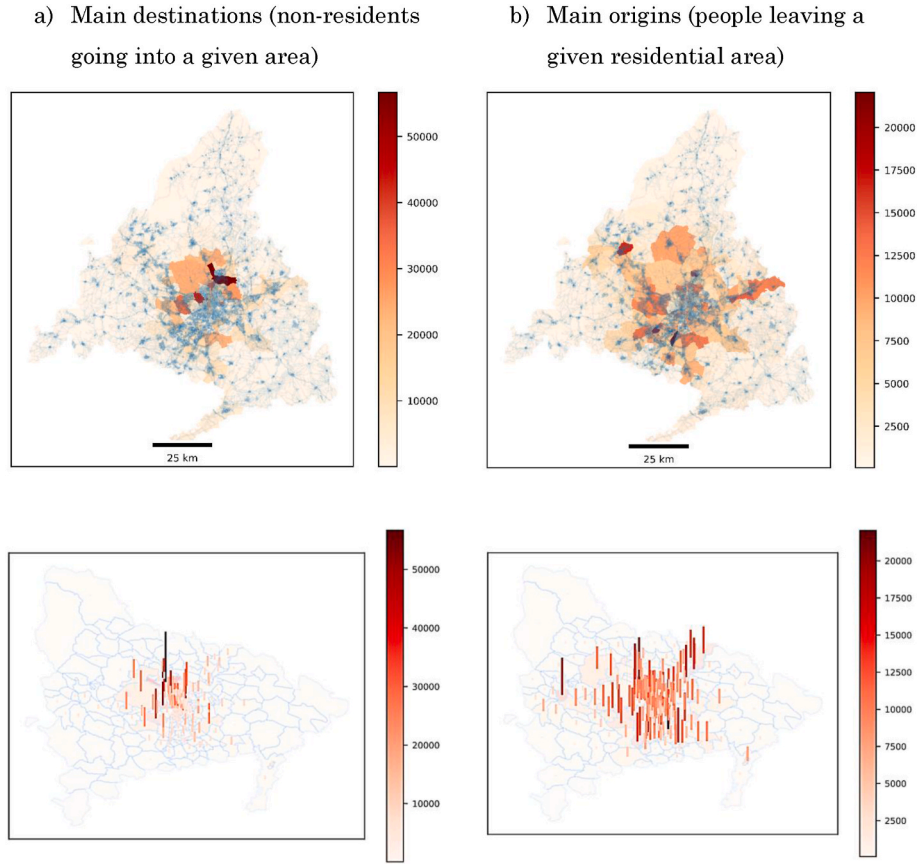


Fig. 8. Main origins and destinations of mobility flows. INE dataset. Note: Only people who go to a destination that is different from their residence. Source: Own elaboration with INE data.

Table 2
Model fit. Summary of results. Sample period: November 2018 (working days only).

	NN autoencoder (4 neurons in hidden layer)	NN autoencoder (10 neurons in hidden layer (*))	NN autoencoder (14 neurons in hidden layer)	Equation (2) Solution
•Mean absolute percentage error	0.4585	0.4501	0.4598	0.2477
$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{ I_{it} - \hat{I}_{it} }{I_{it}}$; I_{it} = traffic intensity recorded in sensor i at hour t , \hat{I}_{it} estimated traffic intensity in sensor i at hour t	0.2024	0.1828	0.2014	0.0755
•Median absolute percentage error	0.5329	0.5447	0.5318	0.7460
•Determination coefficient R^2	0.0176	0.0112	0.0192	0.0091
$R^2 = 1 - \frac{SS_e}{SS_t}$; $SS_e = \sum_{i=1}^N (I_{it} - \hat{I}_{it})^2$; $SS_t = \sum_{i=1}^N (I_{it} - \bar{I})^2$; $\bar{I} = \frac{1}{n} \sum_{i=1}^N I_{it}$				
•Kullback–Leibler divergence $KL(P, Q) = \sum P(x) \log \left(\frac{P(x)}{Q(x)} \right)$, where $P(x)$ is the probability of measuring a traffic intensity equal to x , and $Q(x)$ is the probability of estimating a traffic intensity equal to x . Both $P(x)$ and $Q(x)$ are estimated through Kernel Density Estimation using a Epanechnikov kernel.				

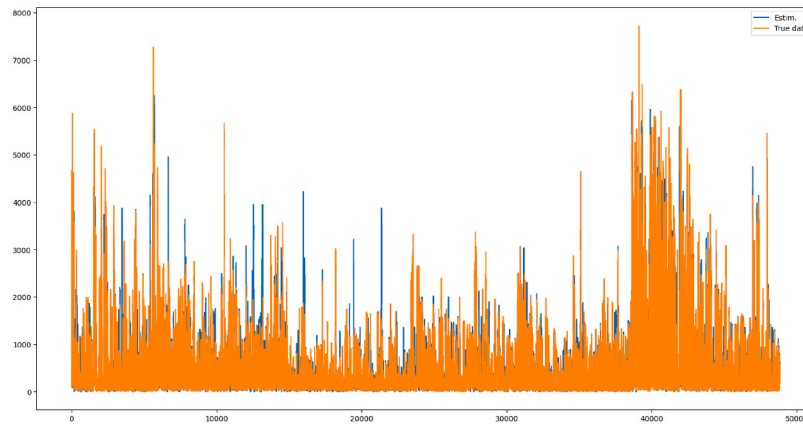
5. Madrid Central as a case study

Madrid Central was a Low Emission Zone (LEZ) from Friday, November 30, 2018, to July 2020, when the High Court of Justice of Madrid suspended the zone over formal deficiencies. The Madrid Central LEZ's ban on private cars above a certain pollution threshold led to a reduction in traffic intensity within the zone, as shown in (Moral-Carcedo, 2022). The LEZ amounts to an opportune test for our paper and its proposed method, because the policy presents an anomaly-detection problem of a type common in data analysis. To attempt to reproduce the resulting traffic intensities, we first estimate the NN model with data from November 2018 and apply it to November

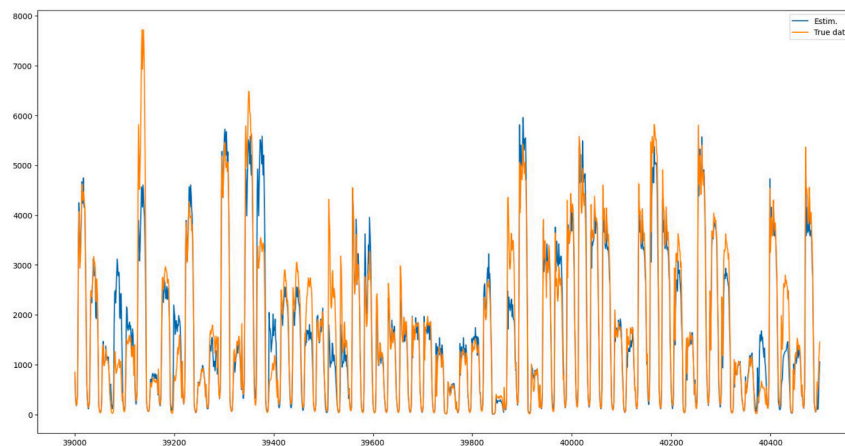
2019. We then compare the true and estimated data with the traffic behavior determined by the NN (trained only with data from November 2018).

Fig. 14a and b shows that although there are no clear differences in global traffic intensity between November 2019 and November 2018, we observe in both the observed and the estimated values clear divergences when we include only traffic sensors in the Madrid Central area. The NN model, trained on data from 2018 and fed data from November 2019 (after the Madrid Central LEZ became operational), provides estimates that are significantly higher than the actual values observed by sensors in the Madrid Central area (Fig. 14d). However, we do not observe any notable effect in the case of sensors elsewhere in the

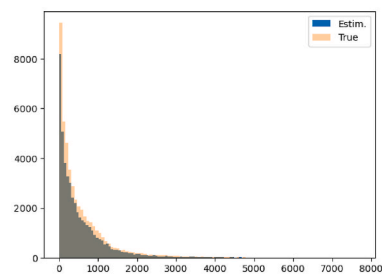
a. Full dataset, hourly means, 2035 sensors (2035 sensors x 24 hours = 48840 obs).



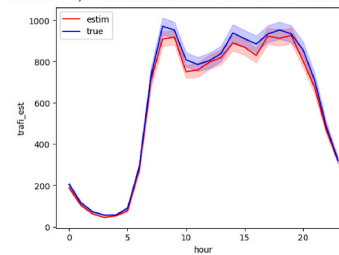
b.- Zoom in on observations (39500–40000). Data from 63 sensors.



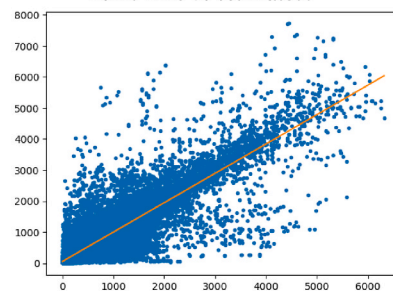
c. Traffic intensity histogram. Estimated and true values.



d. Mean hourly traffic intensity (all sensors). Estimated and true values.



e.-Mean traffic intensity by sensor and hour. True vs estimated.



f.-Mean traffic Intensity by ID. True vs estimated.

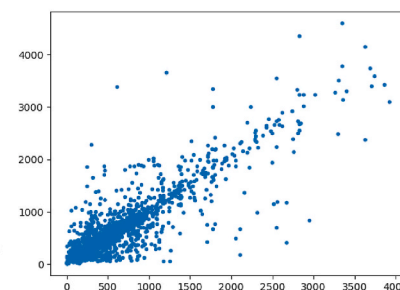


Fig. 9. Estimated and true traffic intensity comparison. November 2018 data.

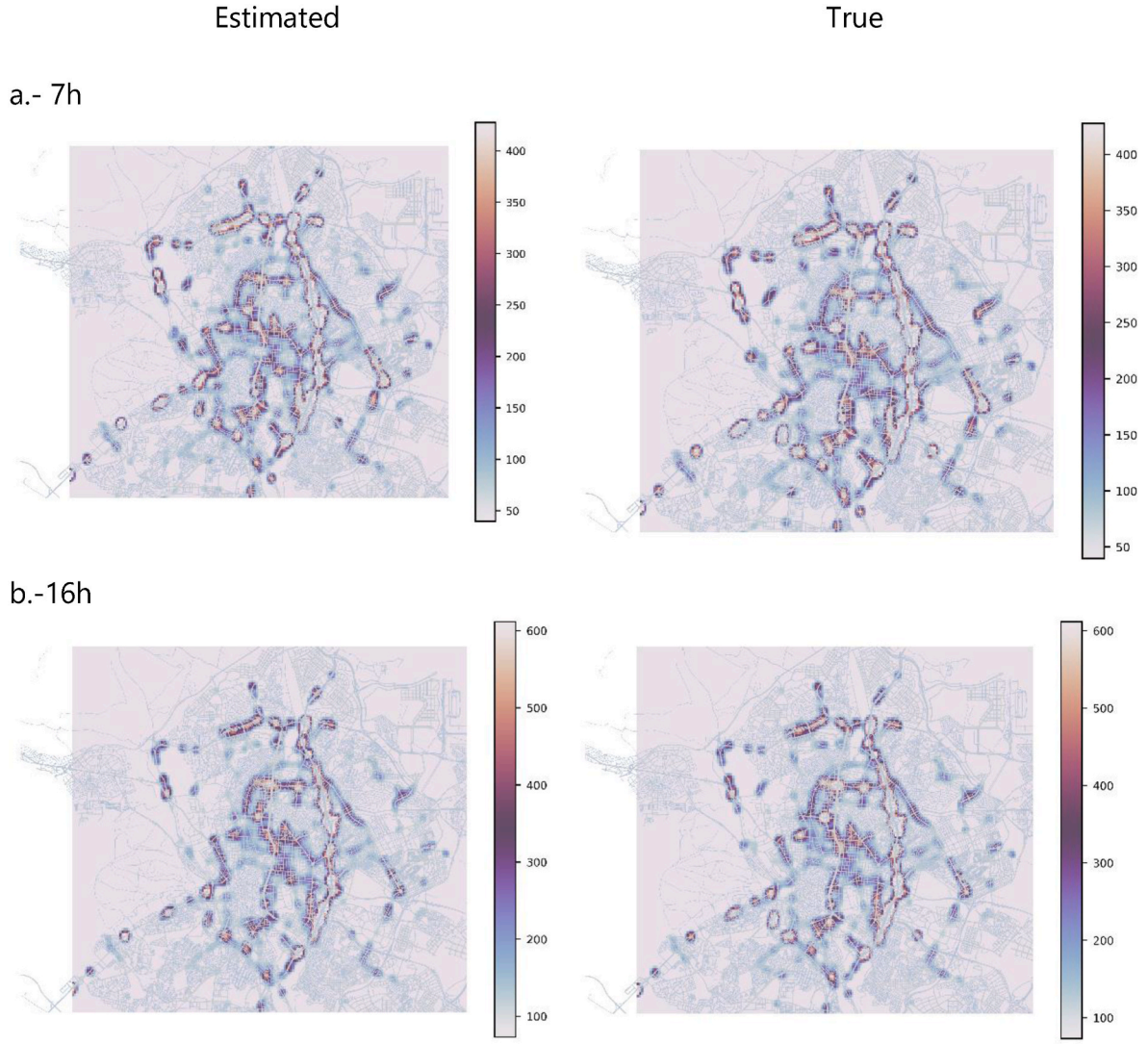


Fig. 10. Estimated and true traffic intensity. Traffic intensity heatmap at different hours. Note: The scale in the legend differs between graphs.

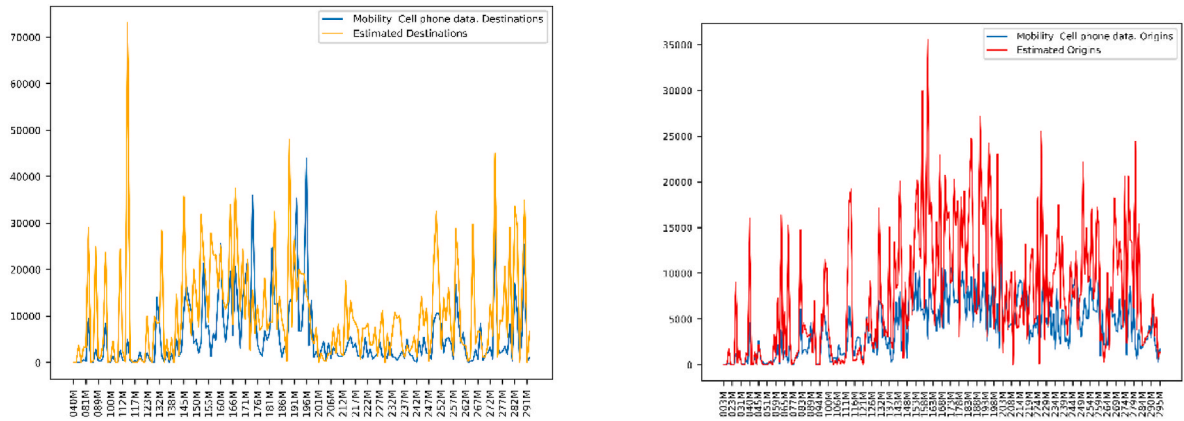


Fig. 11. Estimated flows by origin and destination compared with mobile-phone mobility data. Note: Mobile-phone mobility data are as of November 2019.

metropolitan area (Fig. 14f). This difference is consistent with the negative effect of the access restrictions on the LEZ's traffic intensity.

To provide some insight into this negative effect, Fig. 15 shows an anomaly indicator by traffic sensor, calculated as the difference between estimated and actual daily traffic intensity in November 2019. Using

Equation (1) we compute the estimated traffic intensity by hour t as $I_{t|2019}' = \mathcal{R}_{t|2019}'T$, where T is the route-to-sensor matrix and $\mathcal{R}_{t|2019}'$ is the vector with the estimated intensity of route use in hour t with November 2019 data but with the NN trained on November 2018. This step is an attempt to simulate what the volume of traffic would be

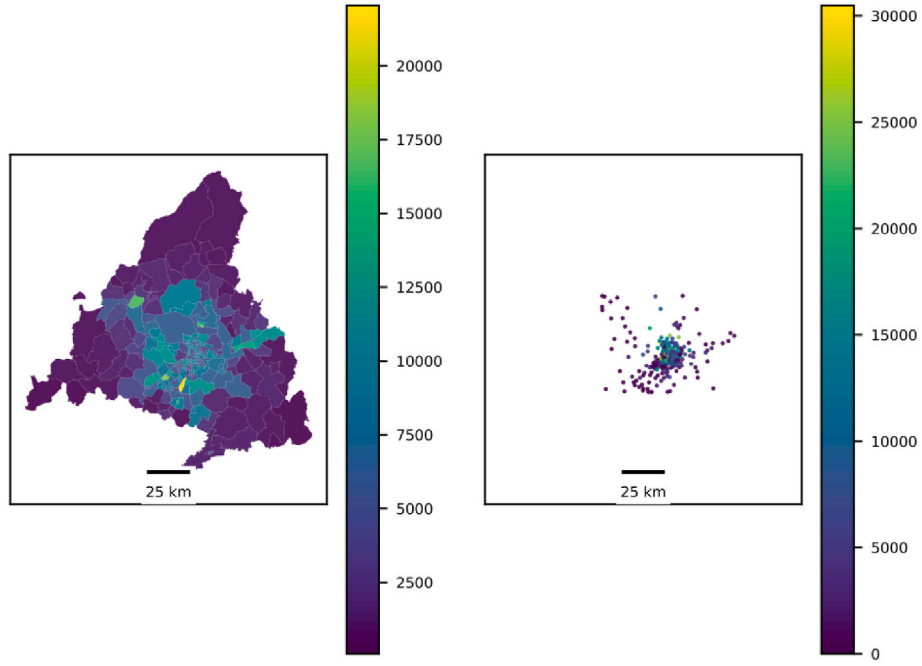


Fig. 12. Main origins of mobility flows. Mobile-phone mobility data (left) and origin of routes with high estimated usage intensity (right).

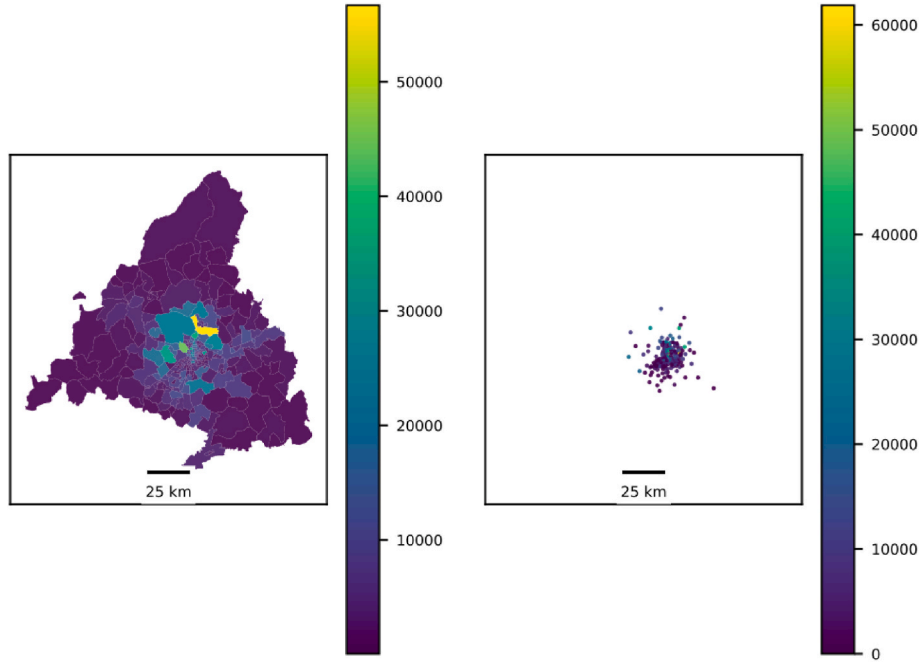


Fig. 13. Main destinations of mobility flows. Mobile-phone mobility data (left) and destinations of routes with high estimated usage intensity (right).

without traffic restrictions. We compute the daily differences as $V^{2019} = \sum_{t=0}^{23} I_{t|2019}' - \sum_{t=0}^{23} I_t^{2019}$, where I_t^{2019} is the vector of traffic intensities recorded by the sensor network at hour t . Once V^{2019} is computed we plot this value by sensor location but only for those sensors where this difference (in absolute terms) is 5 times higher than the difference computed in 2018, $V^{2018} = \sum_{t=0}^{23} I_{t|2018}' - \sum_{t=0}^{23} I_t^{2018}$.

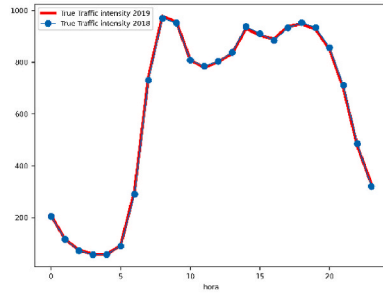
As mathematically constructed, V^{2019} will generate positive values when the NN overestimates actual traffic intensity. We therefore expect positive values in areas where traffic intensity is lower than expected. Fig. 15 shows how positive anomalies, such as t as defined above, occur in the Madrid Central area but also in other, more dispersed areas.

Interestingly, negative anomalies occur in areas close to Madrid Central, indicating higher than expected traffic intensity, which is consistent with the border effect analyzed in the literature.

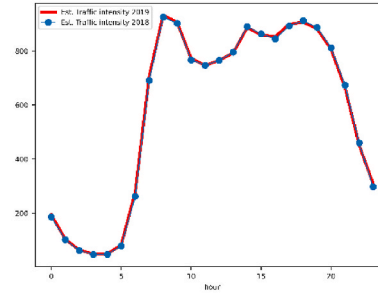
6. Limitations

Mobility patterns in urban areas are reflected in the data collected by the network of traffic sensors. How people move in the urban area, their main origins and destinations, and their preferred routes are reflected in the differences in recorded traffic intensity. This mobility pattern, insofar as it reflects the concentration of residents, economic activities, and jobs, can to some extent be considered a structural characteristic of

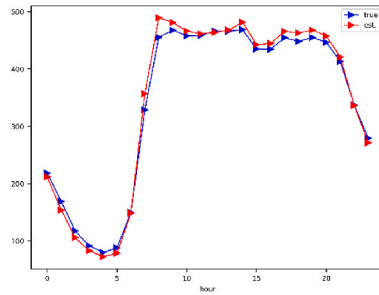
a. True data. Nov 2018 vs Nov 2019. All sensors, hourly mean.



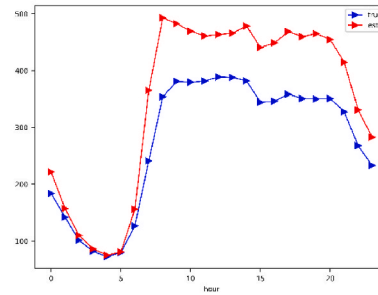
b. Estimated data (NN trained with 2018 data only). Nov 2018 vs Nov 2019. All sensors, hourly mean.



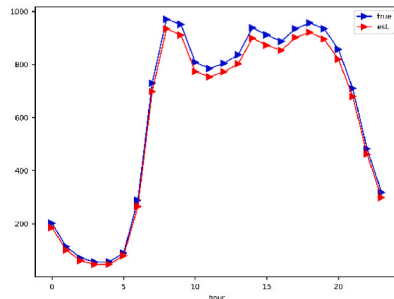
c. Sensors in Madrid Central area only (before restrictions). Nov 2018 true data vs Nov 2018 estimated data. Hourly mean.



d. Sensors in Madrid Central area only (after restrictions). Nov 2019 true data vs Nov 2019 estimated data (NN trained with 2018 data only). Hourly mean.



e. Sensors outside Madrid Central area (before restrictions). Nov 2018 true data vs Nov 2018 estimated data. Hourly mean.



f. Sensors outside Madrid Central area (after restrictions). Nov 2019 true data vs Nov 2019 estimated data (NN trained with 2018 data only). Hourly mean.

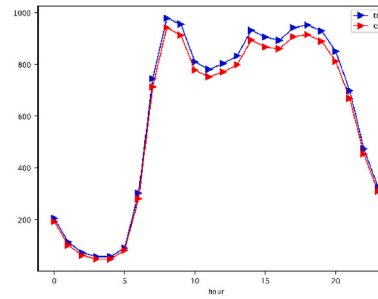


Fig. 14. Anomalies in average hourly traffic intensity in the metropolitan area of Madrid and the Madrid Central area. Comparison of actual traffic data and traffic data estimated with the NN model trained on November 2018 data.

the urban area. In this paper, we have proposed to use people-mobility data to model traffic intensity in the urban sensor network. The presented NN model can reproduce the traffic intensity with an emphasis on the link between route use intensities and recorded flows of vehicular traffic, but it could be improved. For instance, it could be used to predict route use intensities in relation to urban mobility patterns. It would be possible to predict normal traffic volumes and even simulate the effects of changes in the use intensity of certain routes, provided these patterns were relatively stable. Our study's main limitations relate to the characteristics of the datasets used. Mobile-phone positioning data has its advantages, but it does not supply us with routes taken or with all journeys made during the day; it therefore provides only two locations: the place of residence and the location of the longest stay. It would be possible to test other assumptions if other comprehensive data sources were available.

Another limitation relates to the traffic sensor network. The data

used in this study include only the sensor network of the municipality of Madrid, but the traffic volume is due not only to the inhabitants of the city but also to the surrounding municipalities, which have a similar population to the municipality of Madrid's. Traffic sensors set up in areas outside the municipality of Madrid would improve the model's ability to estimate route-use intensities and distinguish between overlapping routes whose origin/destination is outside the analyzed area.

7. Conclusions

NN-based models, proliferating in recent years, have shown remarkable success in predicting traffic intensity, with models that exploit spatial and temporal correlations in traffic data. When the traffic datasets, which show a strong seasonality in their temporal patterns, are obtained from sensors with similar characteristics, most applications perform excellently. Despite this apparent success, however, few studies

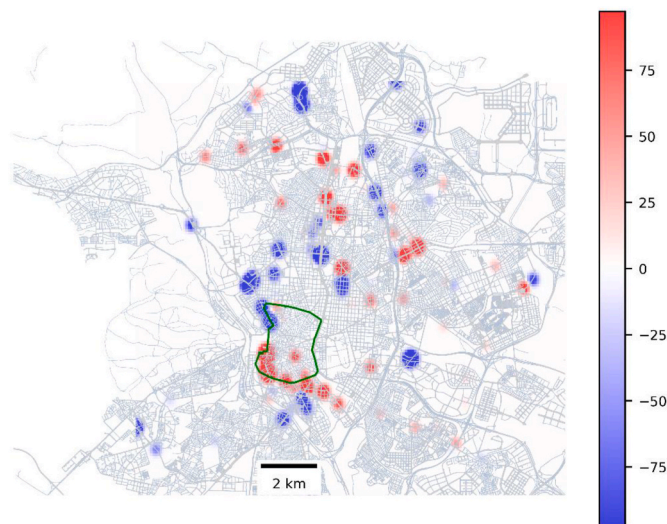


Fig. 15. Location of sensors with the highest anomaly values. Madrid Central area border plotted in green.

have attempted to apply this method to analyze entire sensor networks in an urban area. Even rarer are attempts to relate observed traffic data to the mobility patterns of a given urban area's inhabitants. This study makes both attempts.

Drawing our data from the entire network of traffic sensors in Madrid's urban area, we analyze the differences in traffic intensity observed between different areas and try to determine whether these differences are related. Our starting assumption is that relationships between traffic intensities must result from the characteristics of an urban area's mobility patterns. It is possible, albeit imperfectly, to use mobile-phone positioning data to determine the most frequent daily origins and destinations on weekdays between different areas of the city. With this mobility pattern, it is possible to construct a complete set of the urban area's normally used routes that lie behind the normally observed traffic intensity records. That is, we would exclude exceptional events that affect the traffic data.

In this work, we have estimated the use intensities of the usual routes as those that most accurately reproduce the values of traffic intensity in the entire urban area, including both sensors on high-capacity roads and sensors on main roads that register less traffic. This method makes it possible not only to estimate the entire sensor network's traffic-intensity levels with a significant degree of similarity, especially in terms of spatial distribution, but also to reproduce the mobility-flow values by destination and origin areas as measured by mobility studies that use mobile-phone positioning. In addition to these contributions, this method is useful for observing possible anomalies in traffic behavior. For example, this article has shown how the implementation of the Madrid Central LEZ has caused anomalies not only in the sensors within the restricted area itself, with estimated intensity values higher than those actually observed, but also in the sensors of nearby areas.

Availability of data and material

Datasets are publicly available.

- Traffic data: Madrid City Council data. <https://datos.madrid.es/>
- Mobility in the Madrid region INE. https://www.ine.es/experimenta/movilidad/experimental_em1.htm
- Routes: Open Street Map and OSMNx library for Python.

CRediT authorship contribution statement

Julián Moral-Carcedo: Writing – original draft, Visualization,

Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data is publicly available. Links to data are provided in the text

References

- Bing Yu, H.Y., 2018. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*.
- Boeing, G., 2017. OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex Street networks. *Comput. Environ. Urban Syst.* 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>.
- Boquet, G., Morell, A., Serrano, J., Vicario, J.L., 2020. A variational autoencoder solution for road traffic forecasting systems: missing data imputation, dimension reduction, model selection and anomaly detection. *Transport. Res. Part C* 115, 102622.
- Deloitte - Ipd, 2018. Encuesta de Movilidad de la Comunidad de Madrid 2018. Consorc. Reg. de Transp. de Madrid.
- Deng, M., Qu, S., 2016. Road short-term travel time prediction method based on flow spatial distribution and the relations. *Math. Probl Eng.* <https://doi.org/10.1155/2016/7626875>.
- Fang, Z., Long, Q., Song, G., Xie, K., 2021. Spatial-temporal graph ODE networks for traffic flow forecasting. In: *KDD '21: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. <https://doi.org/10.1145/3447548.3467430>.
- Ghosh, B., Basu, B., O'Mahony, M., 2009. Multivariate Short-Term Traffic Flow Forecasting Using Time-Series Analysis. In: *IEEE Transactions on Intelligent Transportation Systems* 10 (2), 246–254. <https://doi.org/10.1109/TITS.2009.2021448>.
- Guo, S., Lin, Y., Feng, N., Song, C., Wan, H., 2019. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. *Proc. AAAI Conf. Artif. Intell.* 33 (1), 922–929. <https://doi.org/10.1609/aaai.v33i01.3301922>.
- Hagberg, A.A., Schult, D.A., Swart, P.J., 2008. Exploring network structure, dynamics, and function using NetworkX. In: *Proceedings of the 7th Python in Science Conference (SciPy2008)*.
- Huang, H., Hu, Z., Wang, Y., Lu, Z., Wen, X., Fu, B., 2023. Train a central traffic prediction model using local data: a spatio-temporal network based on federated learning. *Eng. Appl. Artif. Intell.* 125, 1–12.
- James, M., 1978. The generalised inverse. *Math. Gaz.* 62 (420), 109–114.
- Kamarianakis, Y., Prastacos, P., 2005. Space-time modeling of traffic flow. *Comput. Geosci.* 31, 119–133.
- Ma, T., Antoniou, C., Toledo, T., 2020. Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. *Transport. Res. Part C* 111, 352–372.
- Manibardo, E.L., Laña, I., Ser, J.D., 2022. Deep learning for road traffic forecasting: does it make a difference? *IEEE Trans. Intell. Transport. Syst.* 23 (7).
- Martín Abadi, A.A., 2015. TensorFlow: large-scale machine learning on heterogeneous systems. <https://doi.org/10.5281/zenodo.4724125>.
- Moral-Carcedo, J., 2022. Dissuasive effect of low emission zones on traffic: the case of Madrid Central. *Transportation*. <https://doi.org/10.1007/s11116-022-10318-4>.
- Nguyen, H., Kieu, L.-M., Wen, T., Cai, C., 2018. Deep learning methods in transportation domain: a review. In: *IET Intelligent Transport Systems, Special Issue: Selected Papers from the 25th ITS World Congress*. <https://doi.org/10.1049/iet-its.2018.0064>.
- Peng, D., Zhang, Y., 2023. MA-GCN: a memory augmented graph convolutional network for traffic prediction. *Eng. Appl. Artif. Intell.* 121, 1–12.
- Peng, H., Wang, H., Du, B., Bhuiyan, M.Z., Ma, H., Liu, J., Yu, P.S., 2020. Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting. *Inf. Sci.* 521, 277–290. <https://doi.org/10.1016/j.ins.2020.01.043>.
- Polson, N., Sokolov, V., 2016. Deep learning predictors for traffic flows. *Transport. Res. Part C Emerging* 79, 1–17.
- Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. *J. Risk Uncertain.* 1, 7–59.
- Shaygan, M., Meese, C., Li, W., Zhao, X., Nejad, M., 2022. Traffic prediction using artificial intelligence: review of recent advances and emerging opportunities. *Transport. Res. Part C* 145, 103921.
- Tay, L., Lim, J.M.-Y., Liang, S.-N., Keong, C.K., Tay, Y.H., 2023. Urban traffic volume estimation using intelligent transportation system crowdsourced data. *Eng. Appl. Artif. Intell.* 126, 1–12.
- Wang, J., Chen, R., He, Z., 2019. Traffic speed prediction for urban transportation network: a path based deep learning approach. *Transport. Res. Part C* 100, 372–385.
- Xu, M., Dai, W., Liu, C., Gao, X., Lin, W., Qi, G.-J., Xiong, H., 2020. Spatial-temporal transformer networks for traffic flow forecasting. *Arxiv*. <https://doi.org/10.48550/arxiv.2001.02908>.

- Yu, B., Lee, Y., Sohn, K., 2020. Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network (GCN). *Transport. Res. Part C* 114, 189–204, 114,189-2024.
- Yu, H.W., 2017. Spatiotemporal recurrent convolutional networks for traffic prediction in transportation. *Sensors*. <https://doi.org/10.3390/s17071501>.
- Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Li, H., 2020. T-GCN: a temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transport. Syst.* 21–9, 3848–3858.
- Zhao, W., Zhang, S., Zhou, B., Wang, B., 2023. Multi-spatio-temporal fusion graph recurrent network for traffic forecasting. *Eng. Appl. Artif. Intell.* 124, 1–10.