USING POISSON PROXIMITY-BASED WEIGHTS FOR TRAFFIC FLOW STATE PREDICTION

E. Uglickich[*], *I. Nagy*[*†]

Abstract: The development of traffic state prediction algorithms embedded in intelligent transportation systems is of great importance for improving traffic conditions for drivers and pedestrians. Despite the large number of prediction methods, existing limitations still confirm the need to find a systematic solution and its adaptation to specific traffic data. This paper focuses on the relationship between traffic flow states in different urban locations, where these states are identified as clusters of traffic counts. Extending the recursive Bayesian mixture estimation theory to the Poisson mixtures, the paper uses the mixture pointers to construct the traffic state prediction model. Using the predictive model, the cluster at the target urban location is predicted based on the traffic counts measured in real time at the explanatory urban location. The main contributions of this study are: (i) recursive identification and prediction of the traffic state at each time instant, (ii) straightforward Poisson mixture initialization, and (iii) systematic theoretical background of the prediction approach. Results of testing the prediction algorithm on real traffic counts are presented.

Key words: *traffic counts, traffic flow state, cluster prediction, Poisson mixture, recursive mixture estimation*

Received: April 27, 2023 **DOI:** 10.14311/NNW.2023.33.017
Revised and accepted: August 30, 2023

1. Introduction

A traffic count is defined as the number of vehicles that pass through a given section of road during a period of time. The multimodal behavior of traffic counts (e.g., peak/off-peak, night, congestion due to accidents, restricted traffic flow due to sports or social city events, etc.) naturally creates clusters that express some of the traffic flow states. The definitions of the traffic flow states in the literature vary from free flow to congested flow depending on the specific road characteristics and flow density [56]. The classification into three traffic flow states is commonly

[*]Evženie Uglickich – Corresponding author; Ivan Nagy; Department of Signal Processing, Institute of Information Theory and Automation of the CAS, Pod Vodárenskou věží 4, 18208 Prague, Czech Republic, E-mail: uglickich@utia.cas.cz, nagy@utia.cas.cz

[†]Ivan Nagy; Faculty of Transportation Sciences, Czech Technical University, Na Florenci 25, 11000 Prague, Czech Republic

used [9, 70], but some sources mention up to six states [66]. The development of algorithms for identification and early prediction of traffic flow states embedded in intelligent transportation systems is of great importance for improving traffic conditions for drivers and pedestrians.

Detailed reviews of existing methods provided by [32, 35, 44] divide traffic flow prediction methods into (i) statistical approaches, (ii) machine learning algorithms, and (iii) deep learning techniques.

The most commonly used statistical approaches to traffic flow prediction are the Kalman filter [17, 61], autoregressive integrated moving average (ARIMA) [63, 64], nonparametric regression [4], and physics of traffic flow [34]. As noted by [66], statistical approaches often require prior assumptions about the data to be analyzed, such as normality, due to the large number of models that have been extensively developed for normally distributed data. This introduces a number of limitations, especially for non-negative and asymmetrically distributed data such as traffic counts. The papers [32, 52] mention the lower prediction accuracy of statistical approaches compared to other state-of-the-art methods.

Another extensive group of methods that have become popular in predicting traffic flow data due to its non-linear nature are machine learning algorithms [65]. In particular, these include artificial neural networks [31, 33, 43, 55], support vector machines (SVM) [13, 23, 30, 66, 68], k-nearest neighbors [3, 19, 23, 30, 47], and random forests [8, 10, 47]. In contrast to statistical approaches, they provide higher accuracy and flexibility of traffic prediction. However, [30] note the sensitivity of artificial network structure settings to traffic characteristics and the lack of theoretical background for its formulation. SVM-based methods have better potential, but [30] mention the high computational cost of these algorithms, which limits their application.

In recent years, deep learning techniques, which are a powerful extension of machine learning tools and provide more accurate prediction results, have received increasing attention [29, 30, 32, 56]. In traffic flow prediction, convolutional neural networks [7], short-term memory networks [69], encoders [62], graphical convolutional networks [67], and hybrid models that combine several techniques [60] are the most common. Despite the higher prediction accuracy provided by deep learning techniques, they have a number of limitations in practical application. As listed by [30], their shortcomings include the need for a huge amount of data to find an appropriate model structure, high computational complexity, and poor interpretability of results from a theoretical point of view.

Despite the large number of studies (not limited to those mentioned above), the limitations of existing approaches discussed by [32, 35, 44] still confirm the necessity of finding a systematic solution to traffic flow prediction as well as its adaptation to specific traffic data.

Automatic recorders at various locations in the urban area provide traffic counts in real time. The dependence between traffic flow at the target and explanatory urban locations was discussed by [46], who claimed that information from the target road section is insufficient for predicting traffic flow. [30] assumed the non-linear relationship between traffic flows on the target and explanatory road sections and used the maximum information coefficient instead of the correlation coefficient. The present paper focuses on the relationship between traffic flow states in different urban locations, where these states are identified as clusters of traffic counts.

Traffic counts are non-negative discrete count data. In a probabilistic approach, this type of random variable is well suited for description by the Poisson distribution. Given the multimodality of traffic counts, a mixture of Poisson distributions is an appropriate tool for detecting and predicting clusters of count data. The Poisson distribution is known for not having a general conditional form, which makes it difficult to model a relationship between target and explanatory variables. Traditional approaches in this case is the use of Poisson or negative binomial regressions [2, 12, 20] as well as the mixtures of them [1, 6]. The latter are well suited for continuous explanatory variables, but may not be suitable for specific explanatory multimodal counts.

The presented paper proposes to predict traffic count clusters based on the recursive Bayesian mixture estimation theory [27, 28, 36]. Unlike the commonly used model-based clustering with mixture estimators using the iterative expectation-maximization (EM) algorithm [18], the recursive algorithms of the adopted theory have two key features: (i) they identify and predict clusters at each time instant and update them with the new data, and (ii) provide the fixed computational time not limited by the algorithm convergence typical for iterative methods. The recursive Bayesian mixture estimation algorithms are extensively developed for normal and categorical mixtures [27, 28, 37], etc. [36] generalized the approach for different types of distributions, which was later developed for exponential [39, 54], uniform [53], and binomial [26] mixtures. [42, 58] studied recursive clustering algorithms for predicting Poisson multimodal counts based on discretized normal explanatory data.

The presented paper extends the aforementioned theory to Poisson mixtures to model traffic counts at the target and explanatory urban locations. For this purpose, the conditional model of the mixture pointers is used to construct the traffic state prediction model. The Poisson proximity [36] computed using the traffic counts at the explanatory location is incorporated into the weights of the traffic states. The number of traffic states is allowed to vary between locations. The predictive model predicts the cluster at the target urban location based on real-time traffic counts at the explanatory urban location. The models of individual clusters can also be used to generate traffic count predictions.

Using the categorization of prediction methods in [66], the presented algorithm combines solutions to traffic flow prediction as a classification problem based on traffic state along with a regression problem based on specific traffic variables (here, traffic counts). The main contributions of the study are: (i) recursive identification and online prediction of traffic count clusters at each time instant, (ii) straightforward Poisson mixture initialization, and (iii) systematic theoretical background of the prediction approach.

The paper is organized as follows: Section 2 provides the general theoretical background of traffic flow state prediction, including the recursive algorithm in Section 2.3.1. Section 3 demonstrates its application to traffic count data, including the data set description in Section 3.1, mixture initialization in Section 3.2, results in Section 3.3, and discussion in Section 3.4. Conclusions in Section 4 conclude the work.

2. Theoretical background

2.1 Problem formulation

Let us consider the traffic counts y and x measured at selected urban locations per unit of time, generally denoted by $t = 1, 2, \ldots, T$. The variable y is the target count measured at the major urban location to be modeled, and x is the explanatory count observed at secondary locations. Their observations produce the data sets $\{y_t\}_{t=1}^T$ and $\{x_t\}_{t=1}^T$ to be analyzed. Due to changes in traffic flow behavior (e.g., during the day, hour, season, depending on the selected time unit), the data values have a multimodal character. This means that clusters are created in the traffic data space that express traffic flow states. If the clusters express only rush hours, they may naturally be common for both traffic counts y and x. However, due to special circumstances at the locations where the counts are taken (e.g., a marathon at one location and an accident at the other) or due to the distance between them, the clusters may not be common. Having the data of both the variables, the clusters can be recognized. Estimating the models of the variables, the clusters are identified and under the condition that the counts are actually measured, they are classified into the clusters that can also be predicted.

However, the problem is how to predict the occurrence of clusters of the target traffic count y when its values can no longer be measured. One of the possibilities is to use its estimated model, but in this case sudden changes in traffic flow behavior in the current time may not be taken into account. The model describing the relationship between y and x can be used for cluster prediction, assuming that the observations of x are still measured. The problem is that the traffic counts y and x are naturally the non-negative discrete variables. As mentioned above, mixtures of Poisson distributions are used to model both individually, but not their relationship due to the lack of the conditional form of the Poisson model. This paper proposes to use the clusters of the actually measured data of the explanatory variable x, detected in real time, to predict the clusters of the target variable y.

Verbally, the problem to be solved is formulated as follows:

(i) construct a model describing the dependence of clusters of traffic counts y on clusters of x

(ii) and predict the clusters of y in real time based on knowledge of x.

The theoretical background for solving this problem is given in the next section.

2.2 Poisson mixture models

Due to the possible different character of the multimodality of the traffic counts y and x, each of them is individually described by a mixture of the Poisson components. This means that each cluster expressing a traffic state generated by the data of y and x is modeled by a single component belonging to their mixtures. The individual mixture models considered consist of the Poisson components, i.e., for y and x, respectively, they are

$$\mathcal{P}oi_y\left(y_t, \lambda, c\right) = e^{-\lambda_c} \frac{\lambda_c^{y_t}}{y_t!}, \quad c = \{1, 2, \ldots, N\},$$

$$\mathcal{P}oi_x\left(x_t, \mu, s\right) = e^{-\mu_s} \frac{\mu_s^{x_t}}{x_t!}, \quad s = \{1, 2, \ldots, M\}, \tag{1}$$

where $\lambda \equiv \{\lambda_c\}_{c=1}^N$ and $\mu \equiv \{\mu_s\}_{s=1}^M$ are the collections of all unknown Poisson component parameters, where N and M denote fixed numbers of the components. The random discrete variables c and s are the so-called *pointers* [28], whose values indicate the active components that generate the current realizations y_t and x_t at the time instant t respectively. This means that each traffic count has its own pointer. Each pointer is generally modeled by the parameterized categorical distribution with the probabilities $\beta \equiv \{\beta_c\}_{c=1}^N$ and $\gamma = \{\gamma_s\}_{s=1}^M$, which correspond to the values of the pointers c and s, respectively.

Note that the conditional form of the Poisson component (1), which could be used to describe the dependence of y on x (or vice versa), is not feasible for this type of distribution. This paper proposes to model the relationship of these variables through the dependence of their pointers. Therefore, the joint pointer model in the form of the probability function denoted by $f(\cdot, \cdot|\cdot)$ with the unknown parameter $\alpha \equiv \{\alpha_{sc}\}_{s,c=1}^{M,N}$ is used instead of the marginal pointer models. It is

$$f(c, s|\alpha) = \begin{array}{c|cccc} {}_{s}\diagdown{}^{c} & 1 & 2 & \ldots & N \\ \hline 1 & \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1N} \\ 2 & \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2N} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ M & \alpha_{M1} & \alpha_{M2} & \ldots & \alpha_{MN} \end{array}, \tag{2}$$

from which the conditional distributions of the pointers can be derived (either $f(c|s, \alpha)$ or vice versa $f(s|c, \alpha)$). In the case that the measurements of the traffic count y are no longer available and the pointer s is estimated in real time using the data of x, the conditional distribution $f(c|s, \alpha)$ indicating the occurrence of clusters of y is used to predict the values of the pointer c for each time point. The recursive cluster prediction algorithm is introduced in the next section.

2.3 Recursive cluster prediction algorithm

The derivation of the cluster prediction algorithm is based on basic principles of the recursive Bayesian mixture estimation theory [27, 28, 36]. Here they are elaborated for Poisson components and conditional pointer prediction.

The *learning* phase of the algorithm is executed on the data sets $\{y_t\}_{t=1}^T$ and $\{x_t\}_{t=1}^T$ of both variables. Inspired by the adopted methodology [27, 28, 36], the estimation of the Poisson components (1) and the joint pointer model (2) is based on the Bayes rule and the chain rule [14, 41]

$$\begin{aligned} & f\left(c, s, \lambda, \mu, \alpha|y(t), x(t)\right) \propto f\left(y_t, x_t, c, s, \lambda, \mu, \alpha|y(t-1), x(t-1)\right) \\ = \; & \underbrace{\mathcal{P}oi_y\left(y_t, \lambda, c\right)}_{\text{component}} \underbrace{\Gamma\left(\lambda|y(t-1)\right)}_{\text{gamma prior}} \times \underbrace{\mathcal{P}oi_x\left(x_t, \mu, s\right)}_{\text{component}} \underbrace{\Gamma\left(\mu|x(t-1)\right)}_{\text{gamma prior}} \\ \times \; & \underbrace{f\left(c, s|\alpha\right)}_{\text{joint model (2)}} \underbrace{\mathcal{D}ir\left(\alpha|y(t-1), x(t-1)\right)}_{\text{Dirichlet prior}}, \end{aligned} \tag{3}$$

where $y(t)$ and $x(t)$ denote the collections of measurements up to time t, including the prior knowledge needed to recursively update the statistics. The conjugate prior gamma distributions [21] are used for the recursive Bayesian estimation of the Poisson component parameters λ and μ. For the pointer model estimation, the conjugate prior Dirichlet distribution [27] is used. Note that the parallel computations can optionally include the estimation of the marginal pointer models based on the similar principle. The factorized right-hand side of the relation (3) is marginalized to obtain the posterior joint pointer distribution

$$\begin{aligned} f\left(c,s|y(t),x(t)\right) &= \int_{\lambda*} \underbrace{\mathcal{P}oi_y\left(y_t,\lambda,c\right)}_{\text{component}} \underbrace{\Gamma\left(\lambda|y(t-1)\right)}_{\text{gamma prior}} \mathrm{d}\lambda \\ &\times \int_{\mu*} \underbrace{\mathcal{P}oi_x\left(x_t,\mu,s\right)}_{\text{component}} \underbrace{\Gamma\left(\mu|x(t-1)\right)}_{\text{gamma prior}} \mathrm{d}\mu \\ &\times \int_{\alpha*} \underbrace{f\left(c,s|\alpha\right)}_{(2)} \underbrace{\mathcal{D}ir\left(\alpha|y(t-1),x(t-1)\right)}_{\text{Dirichlet prior}} \mathrm{d}\alpha \\ &\approx \underbrace{\mathcal{P}oi_y\left(y_t,\hat{\lambda}_{c;t-1}\right)}_{\text{proximity } m_{y;c}} \times \underbrace{\mathcal{P}oi_x\left(x_t,\hat{\mu}_{s;t-1}\right)}_{\text{proximity } m_{x;s}} \times \underbrace{\hat{\alpha}_{cs;t-1},}_{\text{point estimate}} \\ & \quad \forall c=\{1,2,\dots,N\},\ \forall s=\{1,2,\dots,M\}. \end{aligned} \tag{4}$$

As defined in (4), the concept of the *proximity* is the value of the Poisson component calculated from the actual observation (i.e., y_t or x_t) and the point estimate of this component parameter ($\hat{\lambda}_{c;t-1}$ or $\hat{\mu}_{s;t-1}$) computed on the basis of the past data up to the time $t-1$ [36], i.e.,

$$\begin{aligned} m_{y;c} &= \mathcal{P}oi_y\left(y_t,\hat{\lambda}_{c;t-1}\right),\forall c=\{1,2,\dots,N\}, \\ m_{x;s} &= \mathcal{P}oi_x\left(x_t,\hat{\mu}_{s;t-1}\right),\forall s=\{1,2,\dots,M\}. \end{aligned} \tag{5}$$

The proximity gives the approximate "closeness" of the current realization to each component (more details can be found in [25, 36, 38]). The normalized proximities are used as *weights*

$$w_{y;c;t} = \frac{m_{y;c}}{\sum_{i=1}^{N} m_{y;i}}, \quad w_{x;s;t} = \frac{m_{x;s}}{\sum_{j=1}^{M} m_{x;j}}, \tag{6}$$

of the components of each mixture at time t.

It is clear that to obtain the estimate of the joint pointer distribution (4), the point estimates of the component parameters $\hat{\lambda}_{c;t-1}$, $\hat{\mu}_{s;t-1}$ and the pointer model parameters $\hat{\alpha}_{cs;t-1}$ should be computed, first for the past data and then recomputed. As for the component parameters, they are obtained after the *recursive update of statistics* of the Poisson components of the mixtures of the variables y and x, denoted by

$$S_{y;c;t} = S_{y;c;t-1} + w_{y;c;t}y_t, \quad \kappa_{y;c;t} = \kappa_{y;c;t-1} + w_{y;c;t}, \tag{7}$$

$$S_{x;s;t} = S_{x;s;t-1} + w_{x;s;t}x_t, \quad \kappa_{x;s;t} = \kappa_{x;s;t-1} + w_{x;s;t}, \tag{8}$$

where the recursions start with the initial statistics of the components. The formulas (7)–(8) are derived based on the Bayesian mixture estimation theory [27, 28], see the detailed derivations in [57]. The point estimates of the component parameters based on the current data at time t are then

$$\hat{\lambda}_{c;t} = \frac{S_{y;c;t}}{\kappa_{y;c;t}} \text{ and } \hat{\mu}_{s;t} = \frac{S_{x;s;t}}{\kappa_{x;s;t}}, \tag{9}$$

which is the recursive recomputation of the Poisson expectation identical to its maximum likelihood estimate.

The statistic of the joint pointer model (2) is a matrix denoted by $\nu_t \equiv \{\nu_{sc;t}\}_{s,c=1}^{M,N}$. Its recursive update has a form similar to the case of the dynamic pointer [37], where the weighting matrix is composed as the product of the vectors $w_{x;t}$ and $w_{y;t}$ of all normalized proximities

$$\nu_t = \nu_{t-1} + w_{x;t} w'_{y;t}. \tag{10}$$

The result is used to recalculate the point estimate [27, 28] via the normalization of the statistic (10)

$$\hat{\alpha}_{sc;t} = \frac{\nu_{sc;t}}{\sum_{i=1}^{M} \nu_{ic;t}}, \tag{11}$$

which represents the point estimate of the joint pointer distribution (4) based on the actual measured data. Now, both the component and the pointer models are estimated. The recursive phase of the estimation of all models is finished when all values from the data sets $\{y_t\}_{t=1}^{T}$ and $\{x_t\}_{t=1}^{T}$ are used.

The *prediction* phase is based on the learned joint pointer distribution (4), which takes into account the relationship between the pointers c and s. The proximity (5) is computed only for the measured count x using its current value x_t. The proximity normalization (6) gives the weights $w_{x;s;t}$ $\forall s = \{1, 2, \ldots, M\}$. The conditional distribution of the pointer c depending on the pointer s is constructed from (4) by normalizing over columns

$$f\left(c|s, y(T), x(t)\right) = \frac{\hat{\alpha}_{sc;T}}{\sum_{i=1}^{N} \hat{\alpha}_{si;T}}, \tag{12}$$

where the x values are taken online at the current time t as opposed to the y data that was last available up to time $t = T$. Then

$$f\left(c|s, y(T), x(t)\right) w_{x;s;t} = w_{y;c;t}, \quad \forall c = \{1, 2, \ldots, N\}, \ \forall s = \{1, 2, \ldots, M\}, \tag{13}$$

returns the prediction of the actual weight of the cth component of the target count y based on the data x_t. The prediction for the pointer c is

$$\hat{c}_t = \arg\max_{c \in \{1, \ldots, N\}} \left[w_{y;1;t}, w_{y;2;t}, \ldots, w_{y;N;t}\right]', \tag{14}$$

which points to the active Poisson component $\mathcal{P}oi_y\left(y_t, \hat{\lambda}_{\hat{c}_t;T}\right)$ describing the cluster of traffic count data y. Either the active distribution or the weighted average of the components can be used to predict data in clusters. The presented theory is summarized in the algorithm below.

2.3.1 Algorithm

The presented algorithm has been tested in Scilab (www.scilab.org), a free and open-source programming environment for engineering and scientific computations.

Algorithm 1 Presented algorithm.

{Mixture initialization (for $t = 1$)}
Set the numbers of components M, N.
for all $s \in \{1, \ldots, M\}$, $c \in \{1, \ldots, N\}$ **do**
  Set the initial statistics $S_{x;s;t}$, $\kappa_{x;s;t}$, $S_{y;c;t}$, $\kappa_{y;c;t}$ and $\nu_{sc;t}$.
  Compute the point estimates $\hat{\mu}_{s;t}$, $\hat{\lambda}_{c;t}$ and $\hat{\alpha}_{sc;t}$ according to (9) and (11).
end for
{Offline learning phase}
for $t = 2, 3, \ldots, T$ **do**
  Measure the current realization x_t, y_t.
  for all $s \in \{1, \ldots, M\}$, $c \in \{1, \ldots, N\}$ **do**
    Compute the proximities $m_{x;s}, m_{y;c}$ via (5).
    Obtain the weights $w_{x;s;t}, w_{y;c;t}$, according to (6).
    Update the statistics $S_{x;s;t}$, $\kappa_{x;s;t}$, $S_{y;c;t}$, $\kappa_{y;c;t}$ and $\nu_{sc;t}$ according to (7)–(8) and (10).
    Recompute the point estimates $\hat{\mu}_{s;t}$, $\hat{\lambda}_{c;t}$ and $\hat{\alpha}_{sc;t}$ according to (9) and (11).
  end for
end for
{Online prediction phase}
for $t = T + 1, T + 2, \ldots,$ **do**
  Measure the current realization x_t.
  for all $s \in \{1, \ldots, M\}$ **do**
    Compute the proximities $m_{x;s}$ via (5).
    Compute the weights $w_{x;s;t}$ according to (6).
  end for
  Predict the weights $w_{y;c;t}$ according to (12) and (13).
  Predict the pointer c according to (14).
end for

---

3. Traffic flow state prediction

This section demonstrates the application of the above algorithm to the prediction of the traffic flow state in the form of clusters based on the data from the explanatory urban location.

3.1 Data

The data set contains hourly traffic counts measured at 5 traffic stations in Trondheim, Norway from December 2018 to January 2020. The data set from [48] described in [49] and related to the paper [50] is used. 9 758 measurements are available from each of the 5 locations.

Fig. 1(a) shows the traffic count histograms. For better visibility, the histograms from urban locations 1, 2, and 5, where values from 0 to 5 500 vehicles were observed, are shown in the top plot, while locations 3 and 4, with values up to 1 000 vehicles, are given in the bottom plot. The figure clearly shows the multimodal nature of the traffic counts due to the peaks in the distributions. Moreover, the number of these peaks differs from location to location: locations 1 and 5 show 4 peaks, while only 2 peaks can be guessed from the histograms of the remaining locations. This confirms that the multimodality of the traffic counts is different and each of them has the specific number of traffic states. Therefore, they should be modeled by individual mixtures with the specific number of components, i.e., 4 for locations 1 and 5, and 2 for locations 2, 3, and 4.

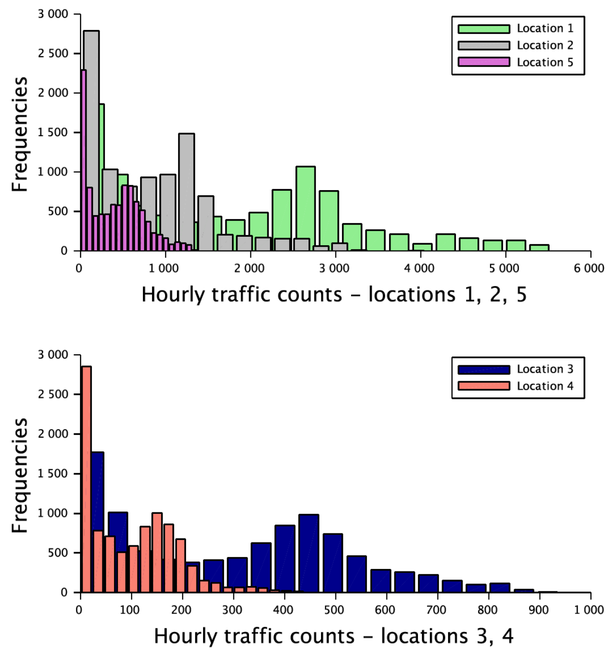
The Poisson components with a higher mean (which is natural for hourly traffic counts) may look symmetric in histograms, but boxplots of the counts in Fig. 1(b) show their skewness.

3.1.1 Choice of traffic counts for prediction

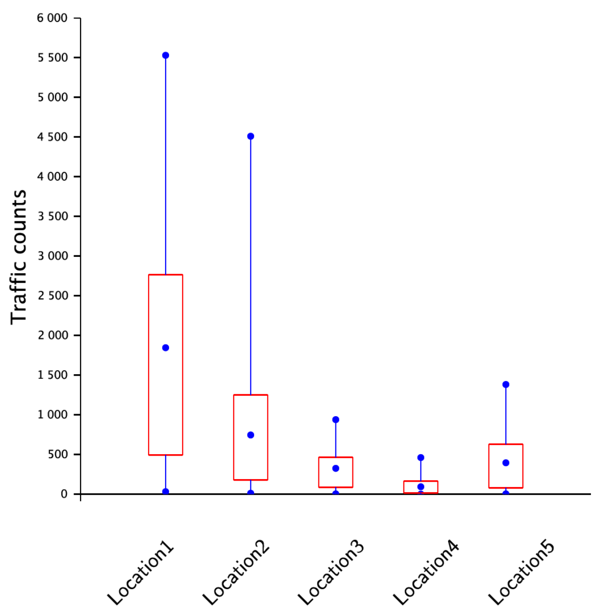
To test the algorithm, traffic counts from 5 locations were divided into 20 pairs x-y, so that each of them can be the target variable y, whose clusters are to be predicted based on the clusters of x. The pairs from all locations show a high correlation, with Spearman's rank correlation coefficient ranging from 0.92 to 0.97. However, in order to be suitable for the proposed algorithm, pairs of counts must have the mutual influence of their traffic states determined by their pointers. To find suitable pairs, parallel clustering of individual traffic counts was performed during the offline learning phase of the algorithm to obtain the point estimates of the marginal pointers, which indicate the clusters of counts detected during learning. The contingency tables of the marginal pointers were analyzed using Goodman & Kruskal's lambda [16] to show the agreement of the traffic flow states. The pairs with lambda values $\Lambda \leq 0.5$ listed in Tab. I were not taken for the prediction because of their weak agreement. The threshold 0.5 for the value of Λ was chosen to get a representative sample of pairs where both locations can have both 2 and 4 traffic states. The 14 pairs were used for traffic state prediction.

Location x	Location y				
	1	2	3	4	5
1	–	0.87	0.78	0.836	0.725
2	**0.488**	–	0.863	0.87	**0.376**
3	**0.492**	0.88	–	0.822	**0.377**
4	**0.445**	0.86	0.784	–	**0.373**
5	0.709	0.752	0.846	0.83	–

Tab. I *Goodman & Kruskal's* Λ *of the contingency tables of the marginal pointers.*



(a)



(b)

Fig. 1 *Histograms (a) and boxplots (b) of traffic counts from 5 urban locations.*

3.2 Mixture initialization

Mixture initialization is a critical task that strongly affects the start of the recursive clustering algorithm. The use of individual mixtures [26] has the important advantage that the mixtures are initialized using histograms of the traffic counts used for the offline learning phase. For Poisson mixtures, this initialization approach is simplified due to a single parameter of the distribution. The algorithm in Section 2.3.1 requires that the number of components of each count mixture and their initial statistics be specified. The number of components of each mixture corresponds to the number of traffic flow states at the location. They are set according to the peaks of the histograms in Figs. 1 and 2, i.e., the 4 traffic states are initialized for locations 1 and 5 and the 2 traffic states for the rest. The key point is to correctly initialize the Poisson component statistics $S_{x;s;t}$ and $S_{y;c;t}$, which is done based on the centers of the histogram peaks. Looking at the enlarged view of the histograms in Fig. 2, it can be seen that the initial statistics of the components can be set to the vector $[168.5 \;\; 1\,543.5 \;\; 2\,643.5 \;\; 4\,293.5]'$ for location 1 and $[37.475 \;\; 382.255 \;\; 520.125 \;\; 1\,140.675]'$ for location 5. The values of the initialized statistics are placed in the red boxes on the x axis in Fig. 2. The rest of the histograms are not zoomed here, but the initial statistics for them were set similarly according to the centers of their peaks. The component statistics $\kappa_{x;s;t}$ and $\kappa_{y;c;t}$ start from 1. The initial statistic of the pointer model ν_t is set uniformly.

3.3 Results

For each of the 14 traffic count pairs, 5 758 measurements were used for the offline learning phase and 4 000 for the prediction phase. The correctness of the predicted clusters cannot be verified directly from the data because they are not measurable. To verify the predictions, observations of each target count were clustered in parallel and then the clusters obtained were compared with the predicted ones. This evaluation was done in two ways.

Pointer comparison First, the pointer values indicating traffic flow states were compared using (i) predictions from the conditional pointer model of the proposed algorithm and (ii) estimates from the marginal pointer model. Fig. 3(a) demonstrates fragments of the pointer comparison obtained for the target location 5 based on the counts from the explanatory location 1 (pair 1-5). Fig. 3(b) compares the pointers of the target location 2 (pair 3-2). The graphical results of the two target traffic counts that have the different number of clusters are shown. To save space, the rest of the graphical evaluation of the pointers is not shown. Fragments for the last 100 hours are displayed for better visibility. In Fig. 3, the target traffic count at location 5 has 4 clusters, while there are only 2 clusters in the case of location 2. Due to the different multimodality at the locations obtained during the initialization, the predicted clusters reflect different traffic flow states. In practice, they should be expertly assigned according to the specific road characteristics at the selected locations. The characteristics of the roads are not known to the authors of this paper, so for these experiments the traffic states can be interpreted verbally according to common practice (see, e.g., [66]) and the number of pointer values as follows:

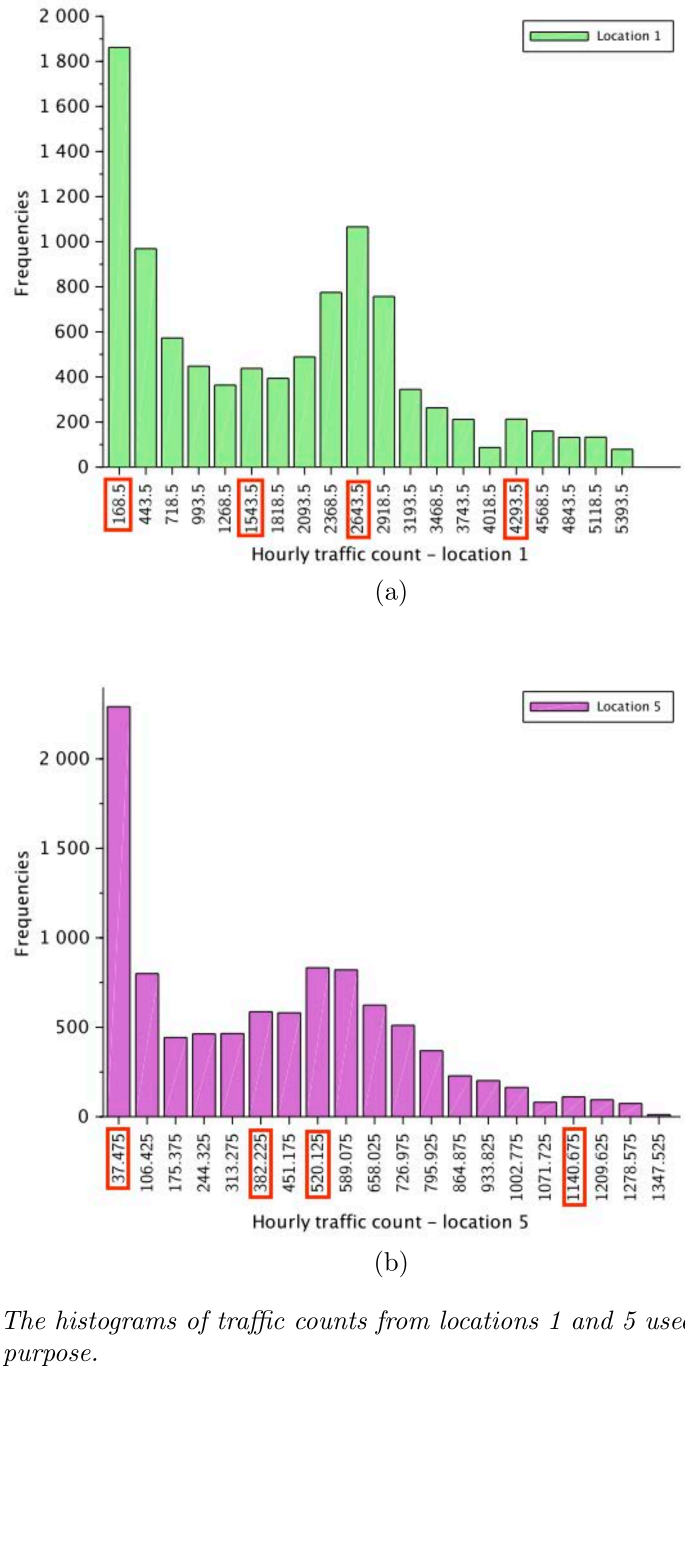
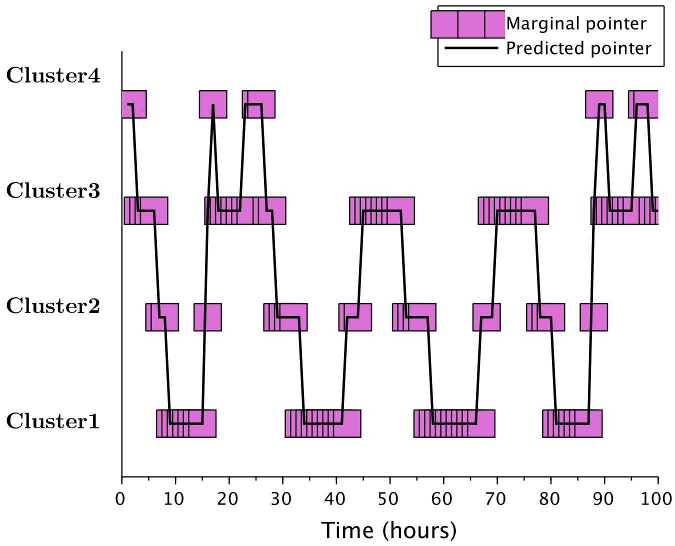
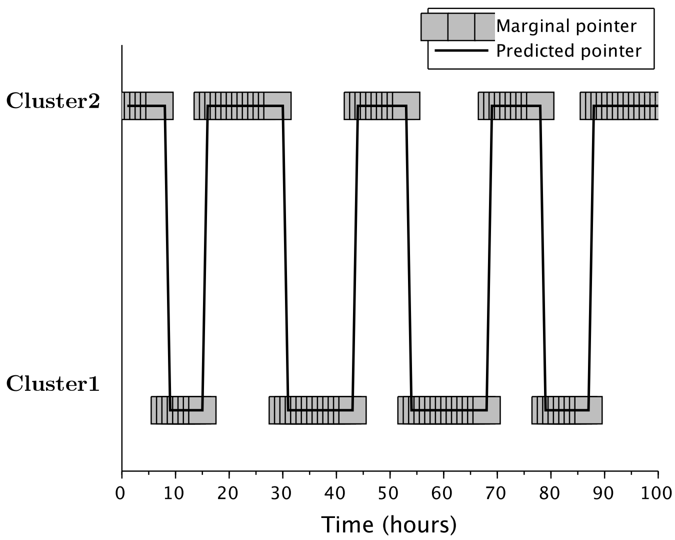


Fig. 2 *The histograms of traffic counts from locations 1 and 5 used for the initialization purpose.*



(a)



(b)

Fig. 3 *The fragments of the pointer prediction for the target locations 5 (a) and 2 (b).*

$$\text{Locations 1, 5:} \qquad \{\underbrace{\text{“free flow”}}_{\text{cluster 1}}, \underbrace{\text{“stable flow”}}_{\text{cluster 2}}, \underbrace{\text{“approaching unstable flow”}}_{\text{cluster 3}}, \underbrace{\text{“unstable flow”}}_{\text{cluster 4}}\},$$

$$\text{Locations 2, 3, 4:} \qquad \{\underbrace{\text{“stable flow”}}_{\text{cluster 1}}, \underbrace{\text{“unstable flow”}}_{\text{cluster 2}}\}.$$

In Fig. 3, the pointers show how the active components describing the traffic count clusters were switching during this time period. The predictions in both plots follow the trend of the marginal pointers, although after the switch (see e.g., the intervals around 35 and 55 hours) it takes some time for the weights to adjust by obtaining new Poisson proximities from the counts at the explanatory locations (here 1 and 3). A higher number of these switches naturally increases the number of errors when evaluating the prediction accuracy, which is calculated as the percentage of errors (PE) obtained when comparing the predicted pointer values and the estimates of the marginal pointers. The evaluation of the PE was calculated over the 4 000 hour prediction phase, averaged over 100 runs of the algorithm. The average PE of the 14 traffic count pairs used at all locations is 7.35%. This means that the data from the explanatory urban locations can serve well enough to predict the occurrence of traffic flow clusters at the target locations.

However, Fig. 4 reports that the PEs of the pairs where both locations have 4 traffic states (i.e., 5-1 and 1-5) are higher than the others. They are 23% and 16%, respectively. The PEs of the pairs where either both locations have 2 traffic states (1-2, 3-2, 4-2, 1-3, 2-3, 4-3, 1-4, 2-4, 3-4) or the locations have a mixed number of traffic states (5-2, 5-3, 5-4) reach the maximum of 10.4%.

Clustering comparison Second, the predicted clustering was compared with well-known existing algorithms applied to each target traffic count. Five iterative methods were chosen: (i) the iterative Poisson mixture estimation based on the EM algorithm [18], (ii) DBSCAN [11], (iii) k-means [22], (iv) fuzzy c-means [15, 40],

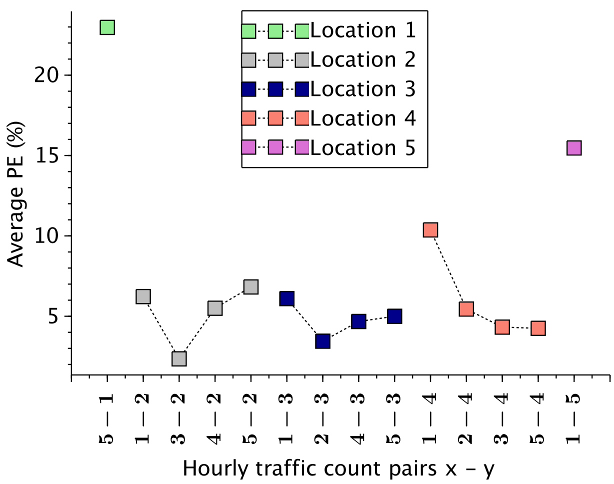


Fig. 4 *The average PE of the pointer prediction of the 14 traffic count pairs x-y.*

and (v) k-medoids [24]. The clustering obtained with the marginal pointer model was also used for the comparison.

The quantitative evaluation of the clustering agreement was done with the help of clustering validity indices [59]. Three indices based on different approaches were calculated to evaluate the similarity between the proposed algorithm and existing methods. They are: (i) the $\kappa_{\max}$ statistic [45], which is an extension of Cohen's kappa coefficient [5], (ii) the normalized mutual information (NMI) [51], and (iii) the Goodman & Kruskal lambda Λ [16]. Their values, averaged over 14 target urban locations, are given in Tab. II. According to them, the highest degree of similarity on average is between the predicted clustering and EM. It was rated as almost perfect agreement by $\kappa_{\max}$ together with Λ and as substantial agreement by NMI. The lowest agreement is with DBSCAN, ranging from fair by Λ to moderate by $\kappa_{\max}$ and NMI. The agreement between the predicted clustering and the other methods ranges from moderate (NMI) to substantial ($\kappa_{\max}$ and Λ).

	EM	DBSCAN	Marginal pointer	k-means	fuzzy c-means	k-medoids
$\kappa_{\max}$	**0.847**	0.456	0.837	0.755	0.76	0.755
NMI	**0.695**	0.408	0.691	0.592	0.597	0.592
Λ	**0.831**	0.346	0.823	0.746	0.751	0.746

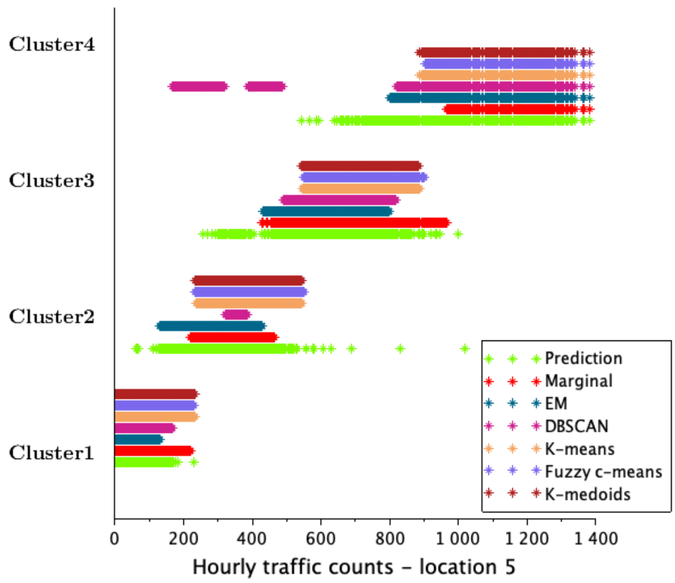
Tab. II *Average clustering validity indices.*

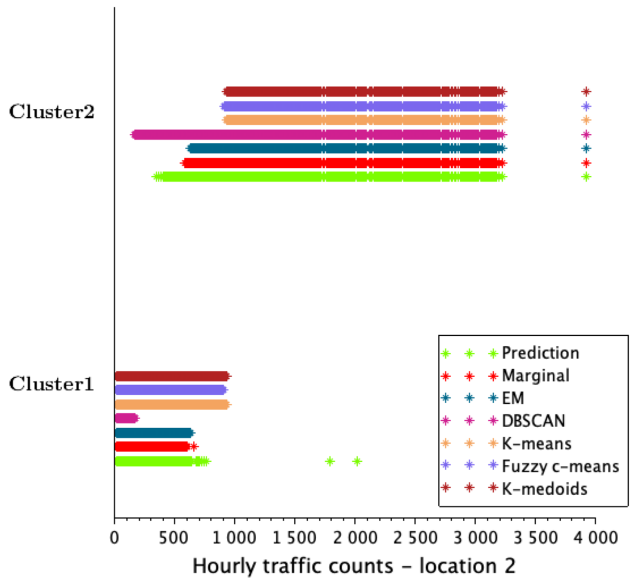
The graphical comparison of the clusterings is shown in Fig. 5. The figure shows the results obtained for the target locations 5 and 2 with 4 and 2 clusters, respectively. In this figure, the traffic counts are plotted against the clustering indices of the compared methods. The clusters obtained using the marginal pointer model are also shown. In Fig. 5(a), the predicted clusters are more overlapping when compared to the "true" clusters obtained by the iterative methods directly from the counts at the target location, which are clearly separated except for DBSCAN. This is because the proposed probabilistic approach needs some time to adjust the mixture weights using the proximity from the explanatory location. The overlap is more evident in clusters 2 and 3 in the Fig. 5(a), where the switch between them was not precisely determined by the proximities. In Fig. 5(b), the overlap is not significant in the case of switching between 2 clusters.

Data prediction from active components The above clusters express the predicted states of the traffic flow. In addition, their models in the form of predicted active Poisson components can be used to generate traffic count predictions. For each of the 14 traffic count pairs, the normalized root mean square error (NRMSE) was computed

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{\sum_{t=1}^{T_{\mathrm{p}}}(y_t-\hat{y}_t)^2}{T_{\mathrm{p}}}}}{y_{\max}-y_{\min}}, \tag{15}$$

where $T_{\mathrm{p}} = 4\,000$ hours denotes the time of the prediction phase, y_t are the actual traffic count values, and $\hat{y}_t$ are the data predictions. The average NRMSE for the





(b)

Fig. 5 *The cluster comparison of the target locations with 4 (a) and 2 (b) traffic flow states.*

14 pairs is given in Tab. III, where the traffic count at target location 2 has the lowest average prediction error and location 4 has the highest.

	Location 1	Location 2	Location 3	Location 4	Location 5
NRMSE	0.163	**0.136**	0.199	0.2	0.154

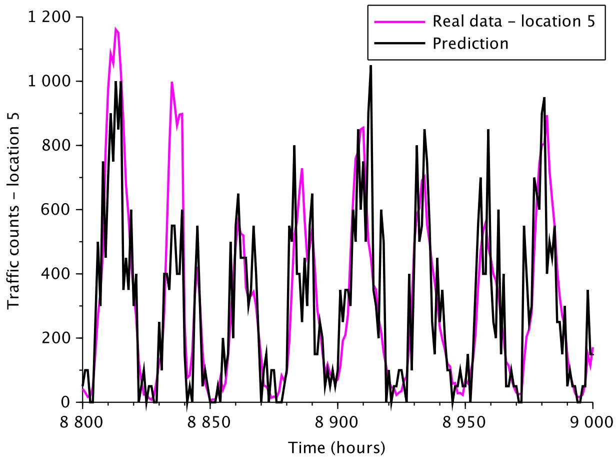
Tab. III *Average NRMSE.*

Graphical results can be found in Fig. 6, which shows prediction fragments for 200 hours for better visualization. Fig. 6 (a) shows the traffic prediction for target location 5 (pair 1-5) with 4 predicted traffic states. The prediction for target location 3 (pair 2-3) with 2 traffic states is shown in Fig. 6(b). It can be seen that the predicted values follow the actual traffic counts.

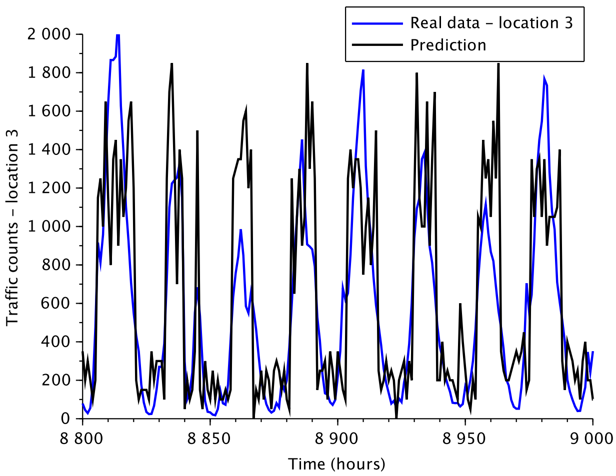
3.4 Discussion

The main objective of the study was to verify the cluster prediction algorithm on traffic count data to predict clusters expressing traffic flow states. The clusters at the target urban locations were predicted using the traffic counts measured at each time instant at the explanatory locations. The proposed prediction algorithm allows the traffic states to be updated with each new measured traffic count. The objective was successfully achieved: the experiments conducted provide the adequate correspondence between the predictions and the clusters actually detected at the target locations.

As reported in Tab. II, the results of the proposed algorithm are closest to the clusters obtained by the EM algorithm. This is logical given the probabilistic model-based clustering philosophy to which both EM and the proposed algorithm belong. However, EM is an iterative method that depends on the convergence of the algorithm, which is avoided in the presented approach. Worse agreement of predicted traffic states was achieved with clusters detected by the DBSCAN method. It is the only algorithm among those compared that does not require a pre-specified number of clusters. However, it is extremely sensitive to the setting of the neighborhood radius and the minimum number of points to detect the required number of clusters in the data. This explains the larger differences in clusters. On the other hand, DBSCAN is more suitable for finding clusters of non-spherical shape (which is the case of non-Gaussian traffic counts), so it was chosen for comparison with the proposed algorithm. In contrast to DBSCAN, centroid-based k-means, fuzzy c-means, and k-medoids detected practically identical clusters without overlap. They also differ slightly from the clusters provided by the proposed algorithm, which predicts overlapping clusters by adaptively computing component weights online, i.e., taking into account the probabilities with which each traffic flow state is active at the current time. In contrast to the presented recursive clustering, the algorithms selected for comparison perform clustering in offline mode, i.e., they process all data at once. They cannot be used to update the traffic state prediction with new counts actually measured at the explanatory location, which is one of the



(a)



(b)

Fig. 6 *Fragments of the traffic count prediction for location 5 (a) and location 3 (b).*

main contributions of the proposed algorithm, but only to verify the results. For this reason, the computation time comparison was not performed.

Another significant contribution of the study is a straightforward Poisson mixture initialization procedure. The use of individual Poisson mixtures leads to 1D clustering of traffic counts. In this case, initialization based on histograms of counts from different urban locations greatly facilitates the search for initial cluster centers. As reported in Figs. 1 and 2, the proposed approach takes into account the different multimodality of the counts, which allows us to distinguish different numbers of traffic states in urban locations. In this paper, the learning data set was used for initialization, but a relatively small prior or expert knowledge data set can also be used. It is important to have prior data that includes all traffic states that may occur in the testing data.

Among the drawbacks of otherwise very successful deep learning prediction methods, which are known to produce more accurate prediction results, [30] mention the lack of a systematic theoretical approach to model training and the use of empirical intuitions instead. The additional strong advantage of the proposed method is the systematic solution through the well-developed Bayesian mixture estimation theory, potentially extendable to other types of data (e.g., pedestrian flow), and the generalized recursive clustering approach. Moreover, although [66] classify the traffic state prediction based on hourly data as a medium-term prediction, the presented theory is not limited by the time unit used.

A practical application of online traffic flow state prediction is mainly seen in intelligent transportation systems, starting with informing drivers about traffic forecasts and more efficient routes to reduce travel time, and ending with supporting traffic control centers in managing the urban road network, depending on the level of automation.

In terms of limitations, the presented theory obviously requires multimodal data sets to distinguish individual traffic flow states. However, it is not sensitive to overestimation in the case of smaller data sets, which is the problem of deep learning techniques [30]. The correlation between traffic states at the target and explanatory locations should also be present in the data.

4. Conclusion

The paper is devoted to the prediction of a traffic flow state based on the conditional model of clusters of traffic counts identified at target and explanatory urban locations. The recursive Bayesian mixture estimation theory has been extended to Poisson mixtures in order to use the pointers of the mixtures that indicate the actual traffic flow state. Four and two traffic states related to traffic count clusters were predicted for target locations using the online prediction algorithm using various multimodal traffic counts measured at explanatory locations. The algorithm was also used to predict traffic counts. The recursive clustering based on individual mixtures also allowed us to use the specific initialization based on histogram analysis of the available data, which is crucial for the successful execution of the algorithm.

In addition, the general solution brings the advantage of universality of the presented algorithm with respect to the data to be predicted. As mentioned in [66], the given statistical assumptions can be limiting for many applications. Here, the success of the prediction depends on the amount of useful information contained in the data sets. The algorithm is not limited by specific types of variables or by the application domain. The often used normality assumption can be limiting for many applications. Here, the Poisson distribution was used as the most appropriate tool for traffic counts. However, alternative types of data distributions with reproducible statistics and a different application domain can be chosen. Data analysis using discrete mixtures with respect to model dimensionality, possible overdispersion of data in clusters, and automation of mixture initialization are open problems to be investigated in this research project.

In addition, [35] noted that current approaches to intelligent transportation systems lack a universal system capable of handling traffic, pedestrian, and bicycle data simultaneously. Research in this area still requires new adaptive solutions that can help to adjust urban traffic plans in time, avoid potential congestion on urban roads, and improve traffic conditions.

Acknowledgement

This work was supported by the StorAIge project and the corresponding Czech institutional support project No. 8A21009. This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No. 101007321. The JU receives support from the European Union's Horizon 2020 research and innovation program and France, Belgium, Czech Republic, Germany, Italy, Sweden, Switzerland, Turkey.

The authors also thank to Olga Cinková for her brilliant help and fruitful discussions during preparing the manuscript.

References

[1] AGRESTI A. Categorical Data Analysis. 3rd Ed. John Wiley & Sons, 2012.

[2] AGRESTI A. An Introduction to Categorical Data Analysis. 3rd Ed. Wiley, 2018.

[3] CAI P., WANG Y., LU G., CHEN P., DING C., SUN J. A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting, Transp. Res. C, 2016, 62, pp. 21–34.

[4] CLARK S. Traffic prediction using multivariate nonparametric regression, J. Transp. Eng., 2003, 129, pp. 161–168.

[5] COHEN J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, pp. 196037-196046.

[6] CONGDON P. Bayesian Models for Categorical Data. John Wiley & Sons, 2005.

[7] DJENOURI Y., BELHADI A., CHEN H.-C., LIN J. C.-W. Intelligent deep fusion network for urban traffic flow anomaly identification. Computer communications, 2022, 189, pp. 175–181.

[8] DOGRU N., SUBASI A. Traffic accident detection using random forest classifier. In: 2018 15th Learning and Technology Conference (L & T), IEEE, 2018, pp. 40–45.

[9] DONG C., SHAO C., XIONG Z. Identification of traffic states with optimized SVM method on urban expressway network. J. Beijing Jiaotong Univ., 2011, 35(6), pp. 13–16.

[10] ELLAHYANI A., EL ANSARI M., EL JAAFARI I. Traffic sign detection and recognition based on random forests, Appl. Soft Comput., 2016, 46 pp. 805–815.

[11] ESTER M., KRIEGEL H.-P., SANDER J., XU X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Simoudis, E., Han, J.; Fayyad, U. M. (eds.) Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press, 1996, pp. 226–231.

[12] FALISSARD B. Analysis of Questionnaire Data with R. Chapman & Hall/CRC, Boca Raton, 2012.

[13] FENG X., LING X., ZHENG H., CHEN Z., XU Y. Adaptive multi-kernel SVM with spatial temporal correlation for short-term traffic flow prediction, IEEE Trans. Intell. Transp. Syst, 2018, 20, pp. 2001–2013.

[14] GELMAN A., CARLIN J.B., STERN H.S., DUNSON D.B., VEHTARI A., RUBIN D.B. Bayesian Data Analysis (Chapman & Hall/CRC Texts in Statistical Science). 3rd edn. Chapman and Hall/CRC, 2013.

[15] GHOSH S., DUBEY S.K. Comparative analysis of k-means and fuzzy c-means algorithms. International Journal of Advanced Computer Science and Applications, 2013, 3, pp. 45–47.

[16] GOODMAN L.A., KRUSKAL W.H. Measures of association for cross classifications III: Approximate sampling theory. Journal of the American Statistical Association, 1963, 58(302), pp. 310–364.

[17] GUO J., HUANG W., WILLIAMS B.M. Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification, Transp. Res. C, 2014, 43, pp. 50–64.

[18] GUPTA M.R., CHEN Y. Theory and Use of the EM Method. (Foundations and Trends(r) in Signal Processing). Now Publishers Inc., 2011.

[19] HABTEMICHAEL F.G., CETIN M. Short-term traffic flow rate forecasting based on identifying similar traffic patterns, Transp. Res. C, 2016, 66, pp. 61–78.

[20] HEERINGA S.G., WEST B.T., BERGLUNG P.A. Applied Survey Data Analysis. Chapman & Hall/CRC, 2010.

[21] HYVPÖNEN V., TOLONEN T. Bayesian Inference 2019. https://vioshyvo.github.io/Bayesian_inference/. Last accessed March 2019.

[22] JAIN A. K. Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 2010, 31(8), pp. 651–666.

[23] JI K.K., LI Z.Z., CHEN J., LIU K.L., WANG G.Y. Freeway accident duration prediction based on social network information, Neural Network World, 2022, 32(2), pp. 93–112.

[24] JIN X., HAN J. K-Medoids Clustering. In: Sammut, C., Webb, G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA, 2011, doi: 10.1007/978-0-387-30164-8_426.

[25] JOZOVÁ Š., UGLICKICH E., NAGY I. Bayesian mixture estimation without tears. In: Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics - ICINCO, ISBN 978-989-758-522-7, 2021, pp. 641–648.

[26] JOZOVÁ Š., UGLICKICH E., NAGY I., LIKHONINA R. Modeling of discrete questionnaire data with dimension reduction. Neural Network World, 2022, 32(1), pp. 15–41.

[27] KÁRNÝ M., BÖHM J., GUY T.V., JIRSA L., NAGY I., NEDOMA P., TESAŘ L. Optimized Bayesian Dynamic Advising: Theory and Algorithms. Springer, London, 2006.

[28] KÁRNÝ M., KADLEC J., SUTANTO E.L. Quasi-Bayes estimation applied to normal mixture. In: Rojíček, J., Valečková, M., Kárný, M., Warwick, K. (eds.). Preprints of the 3rd European IEEE WORKSHOP ON COMPUTER-INTENSIVE METHODS IN CONTROL AND DATA PROCESSING, Prague, CZ, 1998, pp. 77–82.

[29] KONG F., LI J., JIANG B., SONG H. Short-term traffic flow prediction in smart multimedia system for internet of vehicles based on deep belief network, Future Gener. Comput. Syst., 2019, 93, pp. 460–472.

[30] LIN G., LI A., GU D. Using support vector regression and K-nearest neighbors for short-term traffic flow prediction based on maximal information coefficient. Information Sciences, 2022, 608, pp. 517–531.

[31] MALINOVSKÝ V. Comparative analysis of freight transport prognoses results provided by transport system model and neural network, Neural Network World, 2021, 31(4), pp. 239–259.

[32] MEDINA-SALGADO B., SÁNCHEZ-DELACRUZ E., POZOS-PARRA P., SIERRA J.-E. Urban traffic flow prediction techniques: a review. Sustainable Computing: Informatics and Systems, 2022, 35, 100739.

[33] MORE R., MUGAL A., RAJGURE S., ADHAO R.B., PACHGHARE V.K. Road traffic prediction and congestion control using artificial neural networks. In: 2016 International Conference on Computing, Analytics and Security Trends, CAST, IEEE, 2016, pp. 52–57.

[34] NAGATANI T. The physics of traffic jams, Rep. Progr. Phys., 2002, 65, pp. 1331–1386.

[35] NAGY A.M., SIMON V. Survey on traffic prediction in smart cities. Pervasive and Mobile Computing, 2018, 50, pp. 148–163.

[36] NAGY I., SUZDALEVA E. Algorithms and Programs of Dynamic Mixture Estimation. Unified Approach to Different Types of Components, SpringerBriefs in Statistics. Springer International Publishing, Heidelberg, 2017.

[37] NAGY I., SUZDALEVA E., KÁRNÝ M., MLYNÁŘOVÁ T. Bayesian estimation of dynamic finite mixtures. International Journal of Adaptive Control and Signal Processing, 2011, 25(9), pp. 765–787.

[38] NAGY I., SUZDALEVA E., PECHERKOVÁ P. Comparison of various definitions of proximity in mixture estimation. In: Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics – Volume 1: ICINCO, ISBN 978-989-758-198-4, 2016, pp. 527–534.

[39] NAGY I., SUZDALEVA E., PETROUŠ M. Clustering with a model of submixtures of different distributions. In: Proceedings of IEEE 15th International Symposium on Intelligent Systems and Informatics SISY, 2017, pp. 315–320.

[40] PAL N.R., PAL K., KELLER J.M., BEZDEK J.C. A possibilistic fuzzy c-means clustering algorithm. IEEE Transactions on Fuzzy Systems, 2005, 13(4), pp. 517–530.

[41] PETERKA V. Bayesian system identification, in Eykhoff, P. (Ed.), Trends and Progress in System Identification. Oxford, Pergamon Press, 1981, pp. 239–304.

[42] PETROUŠ M., UGLICKICH E. Modeling of mixed data for Poisson prediction. In: Proceedings of IEEE 14th International Symposium on Applied Computational Intelligence and Informatics (SACI 2020), 2020, pp. 77–82.

[43] PURKRÁBKOVÁ Z., RŮŽIČKA J., BĚLINOVÁ Z., KOREC V. Traffic accident risk classification using neural networks. Neural Network World, 2021, 31(5), pp. 343–353.

[44] RASKAR C., NEMA S. Metaheuristic enabled modified hidden Markov model for traffic flow prediction. Computer Networks, 2022, 206, 108780.

[45] REILLY C., WANG CH., RUTHERFORD M. A rapid method for the comparison of cluster analyses. Statistica Sinica, 2005, 15, pp. 19–33.

[46] RYU U., WANG J., KIM T., KWAK S.J.U. Construction of traffic state vector using mutual information for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies, 2018, 96, pp. 55–71.

[47] SOORYA V.B., SHRINIWAS S.A., LELITHA V. Exploratory study on approaches for traffic count prediction using toll-way traffic count. Transportation Research Procedia, 202048, pp. 3880–3895.

[48] SØRENSEN Å.L. Residential electric vehicle charging datasets from apartment buildings, Mendeley Data, V1, doi: 10.17632/jbks2rcwyj.1.

[49] SØRENSEN Å.L., LINDBERG K.B., SARTORI I., ANDRESEN I. Residential electric vehicle charging datasets from apartment buildings, Data in Brief, 2021, 36, 107105, ISSN 2352-3409, doi: 10.1016/j.dib.2021.107105.

[50] SØRENSEN Å.L., LINDBERG K.B., SARTORI I., ANDRESEN I. Analysis of residential EV energy flexibility potential based on real-world charging reports and smart meter data, Energy and Buildings, 2021, 241, 110923.

[51] STREHL A., GHOSH J. Cluster ensembles – a knowledge reuse framework for combining multiple partitions. Journal of Machine Learning Research, 2002, 3, pp. 583–617.

[52] SUN T., YANG C., HAN K., MA W., ZHANG F. Bidirectional spatial-temporal network for traffic prediction with multisource data, Transp. Res. Rec., 2020, 2674(8), pp. 78–89.

[53] SUZDALEVA E., NAGY I., PECHERKOVÁ P., LIKHONINA R. Initialization of recursive mixture-based clustering with uniform components. In: Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2017), 2017, pp. 449–458.

[54] SUZDALEVA E., NAGY I., PETROUŠ M. Recursive clustering hematological data using mixture of exponential components. In: Proceedings of International Conference on Intelligent Informatics and BioMedical Sciences ICIIBMS, 2017, pp. 63–70.

[55] TANG J., LIU F., ZOU Y., ZHANG W., WANG Y. An improved fuzzy neural network for traffic speed prediction considering periodic characteristic, IEEE Trans. Intell. Transp. Syst., 2017, 18, pp. 2340–2350.

[56] TU W., XIAO F., LI L., FU L. Estimating traffic flow states with smart phone sensor data. Transportation Research Part C, 2021, 126, 103062.

[57] UGLICKICH E., NAGY I. Recursive mixture estimation with univariate multimodal Poisson variable. Research Report 2394. ÚTIA AV ČR, v.v.i., Prague, 2022, http://library.utia.cas.cz/separaty/2022/ZS/uglickich-0557467.pdf

[58] UGLICKICH E., NAGY I., PETROUŠ M. Prediction of Multimodal Poisson Variable using Discretization of Gaussian Data. In: Gusikhin, O., Nijmeijer, H., Madani, K. (eds.) Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2021), online, 2021, pp. 600–608.

[59] VENDRAMIN L., CAMPELLO R.J.G.B., HRUSCHKA E.R. Relative clustering validity criteria: a comparative overview. Statistical Analysis and Data Mining, 2010, 3(4), pp. 209–235.

[60] VIJAYALAKSHMI B., RAMAR K., JHANJHI N., VERMA S., KALIAPPAN M., VIJAYALAKSHMI K., VIMAL S., GHOSH U. An attention-based deep learning model for traffic flow prediction using spatiotemporal features towards sustainable smart city, Int. J. Commun. Syst., 2020, e4609.

[61] WANG Y., PAPAGEORGIOU M. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach, Transp. Res. B, 2005, 39, pp. 141–167.

[62] WEI W., WU H., MA H. An autoencoder and LSTM-based traffic flow prediction method, Sensors, 2019, 19(13), 2946.

[63] WILLIAMS B.M. Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling, Transp. Res. Rec., 2001, 1776, pp. 194–200.

[64] WILLIAMS B.M., HOEL L.A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results, J. Transp. Eng.,2003, 129, pp. 664–672.

[65] XIE P., LI T., LIU J., DU S., YANG X., ZHANG J. Urban flow prediction from spatiotemporal data using machine learning: A survey, Inf. Fusion, 2020, 59, pp. 1–12.

[66] YAN H., FU L., QI Y., CHENG L., YE Q., YU D.-J. Learning a robust classifier for short-term traffic state prediction. Knowledge-based systems, 2022, 242, 108368.

[67] ZHANG D., KABUKA M.R. Combining weather condition data to predict traffic flow: a GRU-based deep learning approach, IET Intell. Transp. Syst., 2018, 12(7), pp. 578–585.

[68] ZHANG Y., XIE Y. Forecasting of short-term freeway volume with v-support vector machines, Transp. Res. Rec., 2007, 2024, pp. 92–99.

[69] ZHENG H., LIN F., FENG X., CHEN Y. A hybrid deep learning model with attention-based conv-lstm networks for short-term traffic flow prediction, IEEE Trans. Intell. Transp. Syst., 2020, 22, pp. 6910–6920.

[70] ZHU G., LIU K., QIAO L. Modeling and simulation for traffic states with ANFIS based on T-S model. J. Beijing Jiaotong Univ., 2012, 36(6), pp. 96–101.