MODELLING OCCUPANCY-QUEUE RELATION USING GAUSSIAN PROCESS

Jan Přikryl, *Juš Kocijan*†

Abstract: One of the key indicators of the quality of service for urban transportation control systems is the queue length. Even in unsaturated conditions, longer queues indicate longer travel delays and higher fuel consumption. With the exception of some expensive surveillance equipment, the queue length itself cannot be measured automatically, and manual measurement is both impractical and costly in a long term scenario. Hence, many mathematical models that express the queue length as a function of detector measurements are used in engineering practice, ranging from simple to elaborate ones. The method proposed in this paper makes use of detector time-occupancy, a complementary quantity to vehicle count, provided by most of the traffic detectors at no cost and disregarded by majority of existing approaches for various reasons. Our model is designed as a complement to existing methods. It is based on Gaussian-process model of the occupancy-queue relationship, it can handle data uncertainties, and it provides more information about the quality of the queue length prediction.

1. Introduction

Queue length has been regarded as one of the key parameters in the process of signal plan design, as estimates of queue length may be used as a part of a criterion that is minimised by urban traffic control systems that provide coordinated control of signalised intersections.

The main difficulty of using the queue-length as a part of performance criterion is the fact that the length is difficult to measure automatically – the automatic systems are based exclusively on image processing [46, 47].

Numerous studies discuss the problem of modelling the queue development, see for example [9, 27, 38, 41]. Typical queueing models are for example those of Akçelik [1], Hensher [12], or Mück [25]. The American Highway Capacity Manual 2000 uses a modified version of Akçelik's model [37]. These models are derived from

---

*Jan Přikryl – Corresponding author, Czech Technical University in Prague, Faculty of Transportation Sciences, and Department of Adaptive Systems, Institute of Information Theory and Automation of the ASCR, Prague, Czech Republic. E-mail: `prikryl@fd.cvut.cz`

†Juš Kocijan, Jozef Stefan Institute, Ljubljana and University of Nova Gorica, Slovenia

underlying physical principles of the queue formation and dissipation processes and include some ad-hoc corrections accounting for the stochastic nature of the queuing process.

Stochastic properties of queue development are directly taken into account by Markov chain models [38, 39, 42, 43]. This class of models describes queueing as a stochastic process with probabilities of queue change being given by probability distributions.

The third class of models found in literature are black-box models trying to predict the queue length based on known "training" data. These include autoregressive models [13], neural networks [4, 20], combination of neural networks and fuzzy logic [31], genetic algorithms [10], or neural network constructed with the help of genetic algorithms [44].

All the models mentioned above compute queue length from vehicle count provided by upstream detectors. With the exception of [6] and publications of [29] and [40] (which all concentrate on estimating the total number of vehicles) and [24] (which estimates the queue tail position by dense sampling the output signal of a detector loop, reading out the lengths of active pulses), detector time-occupancy is not used to provide additional information about the queue formation process, although this quantity is usually provided together with the vehicle count by an intersection controller.

In our opinion, the reason for disregarding time-occupancy information can be twofold: First, pure time-occupancy gives us a reasonable measure of queue length only for limited range of queue tails. If the distance from the detector to the downstream queue tail is high, the time-occupancy stays low regardless of the queue length. If the queue tail reaches upstream far behind the detector, the time-occupancy will be high regardless of the queue length. Second, the occupancy readings from the detector may be influenced by other parameters of the traffic flow, as inter-vehicle gaps, vehicle speeds, and their length.

Several papers appeared that try to make use of the time-occupancy measurements for providing additional information about the traffic state at an approach to a signalised intersection. Authors of [5] combine vehicle count and time-occupancy measurements with known signal state and a model of vehicle dynamics to estimate the queue length. Another approach has been taken in [29], where the authors start with thorough theoretical analysis of the time-occupancy measurements and follow with a Kalman filter implementation of vehicle count estimator in the follow-up paper [40]. In our publication [30] we presented an empirical time-occupancy based queue-length model which combines two separate Gaussian-process (GP) models for the low and high occupancy ranges.

In this paper we present an empirical approach to queue length estimation from time-occupancy data, that is meant as a complement of the existing methods and that is able to provide estimation of queue length in the vicinity of a loop detector based on sparse occupancy measurements.

The paper is further organised as follows: in Section 2 we will briefly discuss the practical observations that led to the method proposed in the Section 3 of this article. Section 4 contains a short introduction to GP models that will be used as a base of the queue length estimation method. This Section also lists the parameters of the GP model used in the paper. Results of the queue length estimation using the model and their discussion are presented in Section 5.

2.   Motivation

In the last few years the research group at the Institute of Information Theory and Automation (UTIA, for its acronym in Czech) developed and tested a state-space queue length model for urban arteries based on ideas of [14]. This model is an extension of the models developed by [28] and [6]. All the mentioned models are based on macroscopic vehicle conservation principles.

The state equation of the UTIA model consists of two components – a queue-length part, modelling the queue length using vehicle count readings from upstream traffic detectors, and a linear detector occupancy model that is used to correct the queue length in cases of near saturation. The parameters of the occupancy model are being jointly estimated together with the model state, requiring the use of non-linear Kalman filtering. The current version of the model uses information from strategic upstream traffic detectors for modelling the queue length. Information from downstream strategic detectors is used to provide feedback measurements for the non-linear Kalman filter.

When adopting this model to a smaller urban road network with shorter distances between traffic signals, the observed queue length estimates provided by the model were quite unsatisfactory. However, this type of networks is often equipped also with dilemma-zone detectors, used to extend the green length for an approach in case of vehicle-actuated control. The measurements provided by these detectors will be ignored by the model, although they could reasonably improve the queue length estimate. Hence, we would like to find a mechanism for a dilemma-zone detector to provide additional information about the queue tail development, and to identify if and when the dilemma-zone detector does not contribute to a reasonable estimate due to saturation.

Various types of traffic sensors are used to measure traffic flow [16]. Probably the most common one is an inductive loop detector. This detector is built into the pavement and consists of one or several wire loops, forming a very simple coil, and an electronic unit that transmits energy at chosen frequency (typically between 10 kHz to 200 kHz) into the wire loops. Once a metallic object moves over the loop detector, the inductance of the loop decreases, and the detector signal processing unit records a change in occupancy of the detector.

An intersection controller then provides detector measurements in discrete time intervals of length Δt. This period is typically ranging from $60\,\mathrm{s}$ to $300\,\mathrm{s}$ dependent on hardware, in our case $\Delta t = 90\,\mathrm{s}$. At the end of every measurement interval, the detector signal processing unit provides typically two measured quantities, *vehicle flow* (expressed usually as vehicle count over Δt) and *time-occupancy* (usually a number of "detector active" pulses recorded in Δt or, more conveniently, a percentage of detector activation time in the measurement period Δt). While the vehicle count provided by a dilemma-zone detector may give us an estimate of vehicle distribution over the two parts of the link divided by the detector, the saturation of the link up to the detector is only observable from the occupancy measurements – low vehicle counts may be observed for both light traffic and oversaturated traffic conditions.

Given the nature of time-occupancy measurements, the position of the queue tail at some distance l from the stop-bar will have different influence on detectors
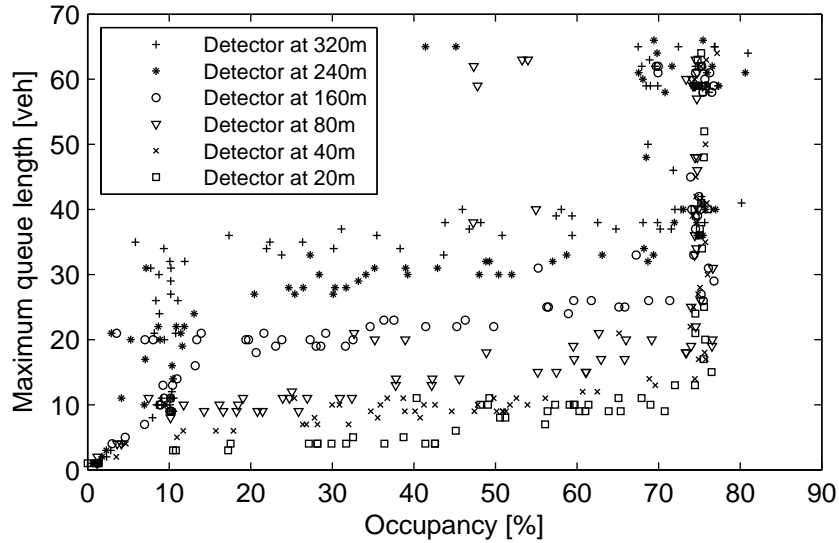
Fig. 1 *Occupancy-queue length dependency for different detector distances from the stop-bar and different effective green lengths for fixed green length (32 s out of 80 s cycle) and different detector distances. Simulated in Aimsun [36].*

placed at different distances from the stop-bar. As we outlined already, the further is the detector installed from the stop-bar, the less reliable it will be for estimating the tail of a short queue – the occupancy changes caused by short queues will become neglectable with the growing distance of the detector from the stop-bar. On the other hand, if the queue tail reaches far behind the detector and stays there over the whole measurement period Δt, the queue length cannot be measured anymore for the same reason.

This phenomenon can be easily demonstrated using a traffic micro-simulator. Results of such a demonstration are depicted in Figs. 1 and 2. As we can see from the graph in Fig. 1, in the case of a fixed-length green signal, the distance of the detector from the stop-bar causes vertical shifts in the queue-length vs. time-occupancy graphs.

However, if we take into account variable lengths of the green signal as a result of vehicle actuated control mechanisms, the resulting measurements provided by a detector at a fixed distance from the stop-bar become quite noisy – the maximum queue length is definitely influenced by the length of the downstream green signal, and the readings of the detector are influenced by the amount of free flow of vehicles during the green phase. This is demonstrated in Fig. 2.

There are two other phenomena that can adversely influence the time-occupancy measurement: The first is the case when queueing vehicles do not activate the detector at all, due to its small size or too large inter-vehicle gap. In order to suppress this "zero speed-zero occupancy phenomenon" [29], we will further suppose that the sizes of the detector loops exceed the typical maximum gap between vehicles. This, according to our observations, is a realistic assumption for urban traffic net-
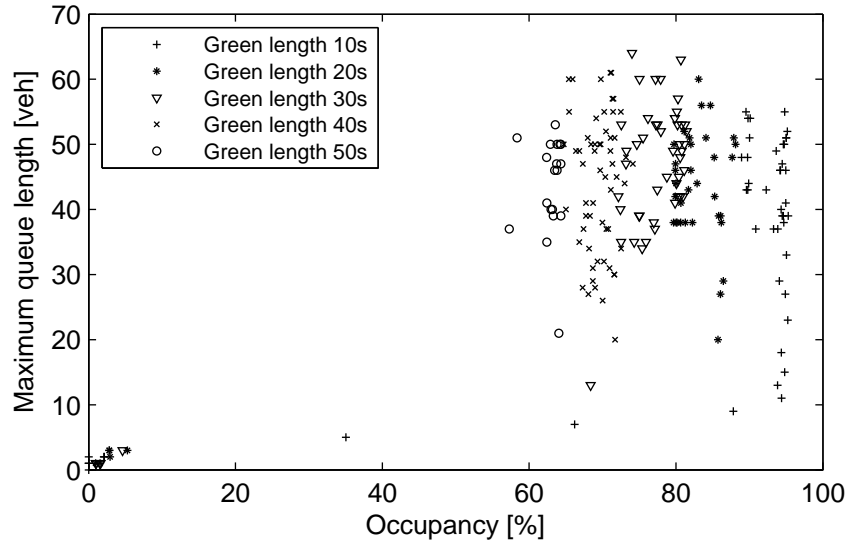
Fig. 2 *Occupancy-queue length dependency for different detector distances from the stop-bar and different effective green lengths for fixed detector distance* 30 m *and different green lengths. Simulated in Aimsun [36].*

works. The second phenomenon is the result of different vehicle length. For the purpose of this paper we will suppose that the vehicle stream consist solely of passenger vehicles. Should this not be the case, the internal logic of the detector signal processing unit can help by providing also vehicle counts separated according to vehicle size into several vehicle groups.

3.   Proposed model

Our queue model is inspired by [6] who used log-exponential S-curve fit to experimental data to provide point estimates of vehicle count within an urban traffic network link. The chosen approximation, though, did not work well for us in cases approaching saturation.

The queue length model of [14] does not behave well for high-occupancy situations as well and it also provides a point estimate only. While it uses a linear dynamic first-order formula for queue-occupancy relationship, our experiments indicate that the additional information provided by the dynamic part of the model does not bring significant improvements to the results obtained and the model can be expressed as static. This is consistent with findings of [6] or [24].

We could use other models that can be found in literature, but we would like to have a mechanism for assessment of the quality of the model prediction – the model is meant as a complement of a physically-based model of [14]. One of the possibilities of such assessment would be to evaluate the interval estimates of queue length rather than point estimates, provided by existing models.

We will therefore express the general model of queue length ξ and its standard deviation σ_ξ, measured by a detector in some distance Δ from the stop-bar, as a static relationship

$$[\xi, \sigma_\xi] = f_\Delta(o, g), \tag{1}$$

where $o \in \langle 0, 1 \rangle$ denotes the measured time-occupancy and g is the green length in seconds. Fixing the distance from the stop-bar means that we will have to provide separate models f_Δ for different detector distances from the stop-bar.

The relationship between the queue-length, time-occupancy and green signal length described by Eq. (1) is quite elaborate. Hence, we will build our model as a black-box model, with parameters trained from a set of training data. As one of our requirement is that the model has to provide interval estimates of the modelled variable, our choice is the Gaussian-process model. Modelling with GP models is introduced in more detail in the next Section.

4. Gaussian-process models

GP model is an example of a flexible, probabilistic, nonparametric model with uncertainty predictions. It fits naturally in the Bayesian modelling framework in which instead of parameterising a model of a non-linear input-output mapping function $f(\mathbf{x})$, a *prior* belief in the form of prior probability distribution is placed directly on the space of possible functions $f(\mathbf{x})$ which could represent the mapping from input vector $\mathbf{x}$ to output y. Its use and properties for modelling are reviewed in [3, 18, 32, 33]. In our particular case we will build a static model where the input vector $\mathbf{x}$ will be composed of time-occupancy and green length and the output y will correspond to the estimated queue length.

The model of the nonlinear input-output mapping is called the *GP model* as the output of the GP model is by prior belief considered to be GP.

GP is a stochastic process and is a generalisation of the Gaussian probability distribution. It can be viewed as a collection of random variables $f(\mathbf{x}_i)$ with joint multivariate Gaussian distribution: $p(f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)) = \mathcal{N}(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f)$, where $\mathcal{N}(\cdot, \cdot)$ stands for Gaussian distribution determined with mean value $\boldsymbol{\mu}_f$ and, as we deal with a multivariate distribution, covariance matrix $\boldsymbol{\Sigma}_f$.

A single element of the covariance matrix

$$\Sigma_{f_{pq}} = \mathrm{Cov}(y_p, y_q) = C(\mathbf{x}_p, \mathbf{x}_q)$$

holds the covariance between values of functions $y_p = f(\mathbf{x}_p)$ and $y_q = f(\mathbf{x}_q)$ which are expressed as values of some covariance function C that works on the input data $\mathbf{x}_p$ and $\mathbf{x}_q$ and has the property of generating a positive definite covariance matrix [32]. Thus, to completely specify a GP, the mean $\boldsymbol{\mu}_f$ (usually assumed to be zero) and the covariance function $C(\mathbf{x}_p, \mathbf{x}_q)$ have to be given.

A common choice for C is

$$C(\mathbf{x}_p, \mathbf{x}_q) = v_1 \exp\left[-\frac{1}{2}\sum_{d=1}^{D} w_d(x_{dp} - x_{dq})^2\right] + \delta_{pq}v_0, \tag{2}$$

where $\boldsymbol{\theta} = [w_1, \ldots, w_D, v_1, v_0]^{\mathrm{T}}$ are the parameters of the covariance functions, called *hyperparameters*, v_0 is estimated noise variance, v_1 is the estimate of the

vertical scale of variation, D is the input dimension, x_{dp} and x_{dq} are d-th components of input vectors $\mathbf{x}_p, \mathbf{x}_q$ and $\delta_{pq} = 1$ if $p = q$ and 0 otherwise. The covariance function given by Eq. (2) is composed of two parts: the Gaussian or squared exponential covariance function for the modelling of system function, that is the first summation term of two in Eq. (2), and the covariance function for the modelling of noise, that is the second summation term in Eq. (2). The noise, in our case, is presumed to be white and therefore it can be represented with constant variances v_0 on the diagonal of covariance matrix, which is achieved with δ_{pq}. Other forms of covariance functions C suitable for different applications can be found in [32].

4.1 Identification of a GP model

For a given problem, the hyperparameters in $\boldsymbol{\theta}$ are learned using the data at hand. Consider a matrix composed of N D-dimensional input vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ and a vector of output data $\mathbf{y} = [y_1, y_2, \ldots, y_N]^{\mathrm{T}}$. Based on the data $(\mathbf{X}, \mathbf{y})$, and given a new input vector $\mathbf{x}^*$, we wish to find the predictive distribution of the corresponding output y^*. Unlike with other models, e.g. parametric models or models formed by combinations of basis functions, there is no model parameter determination as such, within a fixed model structure. With this model, most of the effort consists in determining the parameters of the covariance function. This is done in the Bayesian context by maximisation of the logarithm of marginal likelihood $p(\mathbf{y}|\mathbf{X})$ which for the Gaussian prior probability distribution is

$$
\begin{aligned}
\mathcal{L} &= \log(p(\mathbf{y}|\mathbf{X})) \\
&= \log\left(\int_{\boldsymbol{\theta}} p(\mathbf{y}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|\mathbf{X})\, \mathrm{d}\boldsymbol{\theta}\right) \\
&= -\frac{1}{2}\log(|\,\mathbf{K}\,|) - \frac{1}{2}\mathbf{y}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{y} - \frac{N}{2}\log(2\pi)
\end{aligned}
\tag{3}
$$

where $\mathbf{K}$ is the $N \times N$ training covariance matrix. Each element of $\mathbf{K}$ is calculated with the covariance function (2), i.e. $\mathbf{K} = [C(\mathbf{x}_p, \mathbf{x}_q)]$ and $p, q = 1, \ldots, N$. See [32] for more details about GP model selection in the Bayesian context and the computational implementation.

The computation of the logarithm of marginal likelihood and, depending on used optimisation algorithm, its derivatives necessary for the optimisation involves the computation of the inverse of the $N \times N$ covariance matrix $\mathbf{K}$ at every iteration, which can become computationally demanding for large N. In each iteration of the optimisation, elements of $\mathbf{K}$, that is $C(\mathbf{x}_p, \mathbf{x}_q)$ change with the change of hyperparameters $\boldsymbol{\theta}$.

Nevertheless, the number of parameters to be optimised is small ($D + 2$, see Eq. (2)), which means that optimisation convergence towards optimal values might be faster as it would be in the case of large number of parameters. Small number of parameters that is proportional to the number of regressors means also that the "curse of dimensionality" [15, 26] is decreased (the curse of dimensionality occurs when an increase of model dimensions causes nonlinear increase in the number of parameters and is commonly met in the application of various machine learning methods).

4.2 Regression calculation

The described approach can be easily utilised for regression calculation. Based on training inputs collected in $\mathbf{X}$ a covariance matrix $\mathbf{K}$ of size $N \times N$ is determined. As already mentioned, the aim is to find the distribution of the corresponding output y^* at some new input vector $\mathbf{x}^*$.

For a new test input $\mathbf{x}^*$, the predictive distribution of the corresponding output y^*, over cases in the training set $(\mathbf{X}, \mathbf{y})$, is $p(y^*|(\mathbf{X}, \mathbf{y}), \mathbf{x}^*)$ and is Gaussian, with mean and variance

$$\mu(y^*) = \mathbf{k}(\mathbf{x}^*)^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{y}, \tag{4}$$
$$\sigma^2(y^*) = \kappa(\mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*)^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*),$$

where $\mathbf{k}(\mathbf{x}^*) = [C(\mathbf{x}_1, \mathbf{x}^*), \ldots, C(\mathbf{x}_N, \mathbf{x}^*)]^{\mathrm{T}}$ is the $N \times 1$ vector of covariances between the test and training cases, and $\kappa(\mathbf{x}^*) = C(\mathbf{x}^*, \mathbf{x}^*)$ is the covariance between the test input elements themselves. Vector $\mathbf{k}(\mathbf{x}^*)^{\mathrm{T}} \mathbf{K}^{-1}$ in (4) can be interpreted as a vector of *smoothing* terms which weights the training outputs collected in $\mathbf{y}$ to make a prediction at the test point $\mathbf{x}^*$.

In our particular case $\mathbf{x}_i = [o_i, g_i]$, where o_i is the detector occupancy and g_i is the effective green length of the signal. The corresponding component of the output vector $\mathbf{y}$ is the queue length ξ_i measured for the given o_i and g_i.

4.3 Utility of GP models

The reasons to select modelling with GP models are small amounts of data relative to the number of selected regressors, data corrupted with noise and measurement errors and the need for prediction confidence intervals. If there is not enough data or it is heavily corrupted with noise, even the GP model cannot perform well, but in that case the inadequacy of the model and the training data is indicated through higher variance of the predictions.

The utility to provide the prediction confidence intervals made GP models attractive for modelling case studies in various domains like chemical engineering [17] and process control [23], biomedical engineering [8], biological systems [2], environmental systems [11], power systems [21] and engineering [22], motion recognition [45], etc., to list just a few. We believe that this utility makes it interesting also for use in the domain of traffic modelling and its applications.

We should note that *there is no assumption that complete signals measured in a traffic process will have Gaussian distribution.* The Gaussian prior probability distribution is put over the space of input-output mapping functions meaning that every predicted output data point that is made by GP model has Gaussian distribution. This implies that every prediction of queue length that is made has some most likely value and the less likely values are equally possible on both sides of the most likely value. This is, by our convenience, not unrealistic.

Another assumption of GP modelling as we presented it is that noise that corrupts data is equal for all input data and has Gaussian distribution. When this is not the case, the data can be transformed in the form that will be better modelled by the GPs. More details about these transformations, called also GP warping, can be found in [19, 34, 35], where both the case of heteroscedastic data, i.e., data with

variable variance, and the case of non-Gaussian noise are dealt with. The basic principle is that input and output data are transformed via a monotonic function to match the assumption of Gaussian noise. The form of both the covariance matrix and the nonlinear transformation are learnt simultaneously under the same probabilistic framework. The obtained model predictions need to be transformed back to be comparable with measured data.
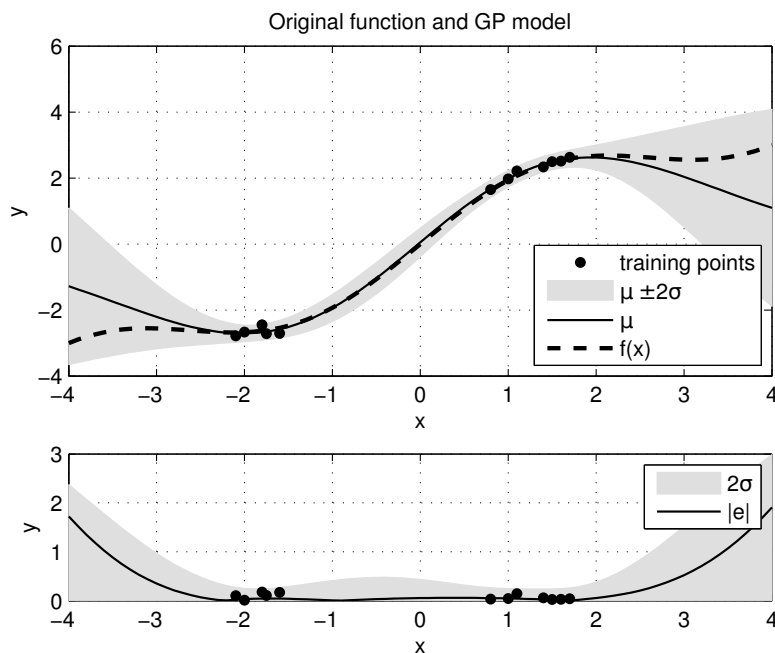

Fig. 3 *Modelling example with GP model. Upper figure shows original function $f(x)$, training points and predictions of GP model shown as their mean value μ with 95 % predictive confidence band $\mu \pm 2\sigma$, lower figure shows absolute value of difference between model predictions and real function e and 95 % tolerance band (2σ) of GP model predictions.*

4.4   Illustrative example

This example is aimed to demonstrate the utility of modelling with GP models on a simple example. Twelve samples are taken from the function

$$y = f(x) = x + \sin 1.2x \tag{5}$$

to which a white noise with mean value 0 and variance 0.01 was added. Input-output data points from these samples were used as estimation data for a GP model. We presume that the modelled function is smooth and stationary so that the covariance function (2) mentioned above can be used.

After optimisation of hyperparameters, the predictions were calculated from obtained GP model. The original function (5), training data and predictions in the chosen range are depicted in Fig. 3.

It can be seen from 95 % tolerance band around the mean values of model predictions on Fig. 3 that variance of the model output is narrow and close to the variance of estimation data noise in the region populated with estimation data and it increases with the distance from this region. It can be clearly seen from the example that GP model does not only give the prediction of values, but also the corresponding confidence intervals. Therefore, where there was not enough data for modelling the GP model does not perform well, but this is indicated with higher variance of model predictions which are not to be trusted in this region.
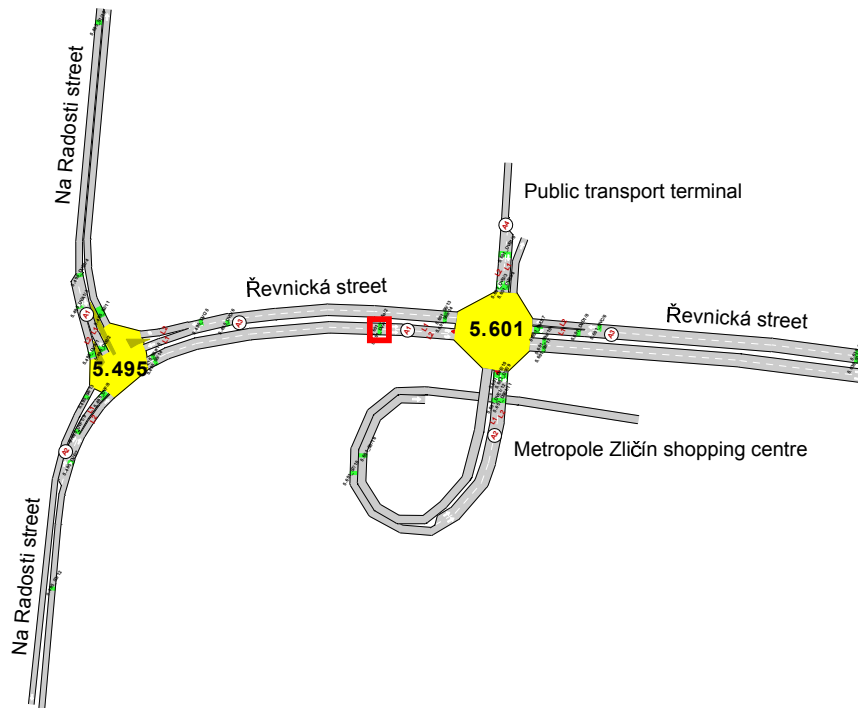

Fig. 4 *Test network model. The marked detector at intersection 5.601 has been used for measurements.*

5. Results

As it is not possible for us to conduct deliberate experiments in real-world traffic, in order to demonstrate the behaviour of the proposed GP model, one day of traffic was simulated on a simple network (part of our Zličín test site, see Fig. 4) using the Aimsun micro-simulator [36] that, once calibrated, should be able to closely mimic the real traffic [7]. The results shown in this section were obtained using data from a single detector placed at the distance of 32 metres (approximately 5-6 vehicles) from

the stop-bar. This distance was selected as an average distance of dilemma-zone detectors in Prague. The traffic demand data used the simulation were real-world measurements from our test site at Prague-Zličín. In order to generate enough training data even for less likely green lengths, we decided to mimic the behaviour of the vehicle-actuated traffic signal control by varying the duration of the green signal from 10 to 55 seconds in a 90 second cycle. Every 90 seconds of the simulation the maximum queue length, the relative green length, and detector measurements were stored, resulting in 960 samples of data consisting of green length, occupancy, and maximum queue length. This approach corresponds to the common traffic engineering practice in Prague where measurements are collected every 90 seconds.

The simulated measurements were then divided into two groups for model cross-validation. From the original 960 measurements, 254 training and 706 validation data samples were selected, using stratified sampling of a 2D histogram of queue lengths versus measured occupancy. The size of the queue length bin was 2 vehicles, the size of occupancy bins was set to 4 %. As the process of GP model learning (3) requires a repetitive inverse of the training covariance matrix, the number of 254 training samples was selected by hand as a limit providing the tuned covariance function in an acceptable time on our hardware — in this case, at most 4 measurements were randomly selected as training data from every non-empty histogram bin.

The variance of data in our case is far from being uniform in the range where we would like to model the relation between occupancy, green length and queue length: Measurements become more noisy as the lane reaches the point of saturation. Consequently, the estimation data needed transformation — warping — as mentioned in Section 4.3. In our case the selected monotonic function was

$$f(t) = t + a_1 \tanh(b_1(t + c_1)) \quad a_1, b_1 \geq 0, \tag{6}$$

where t denotes target data and a_1, b_1, c_1 are constants.

Figs. 5 and 6 show the performance of the model for its own training data. We can see that the model captures well the relationship between the occupancy, green length and the queue length, expanding the prediction confidence interval towards the higher occupancy regions where the measured data exhibit higher variance. The visible cluster of measured values near the peak occupancy at Fig. 5 can be explained by the operation mode of our demonstration – the intersection was repetitively driven into the saturated state and out of it, providing the most measurements in the high occupancy-long queue region.

The behaviour of the model for the validation data set is shown in Figs. 7 and 8 and two cuts – Fig. 9 for low green lengths and in Fig. 10 for high green lengths. We can see that in this case the observed variance in the training data (due to detector saturation and uncertainty in measurements, where certain occupancy value may correspond to different queue lengths, see Figs. 5 and 6) results also in predictions with high variance. This is the correct result as the quality of measured data does not permit better queue length prediction.

Looking at the confidence interval for queue length in Figs. 9 and 10 we can also observe that if the maximum queue length estimates are close to zero, the confidence interval predicted by the GP model may contain negative values. This is the nature of the estimation process as GP models can not use truncated Gaussian
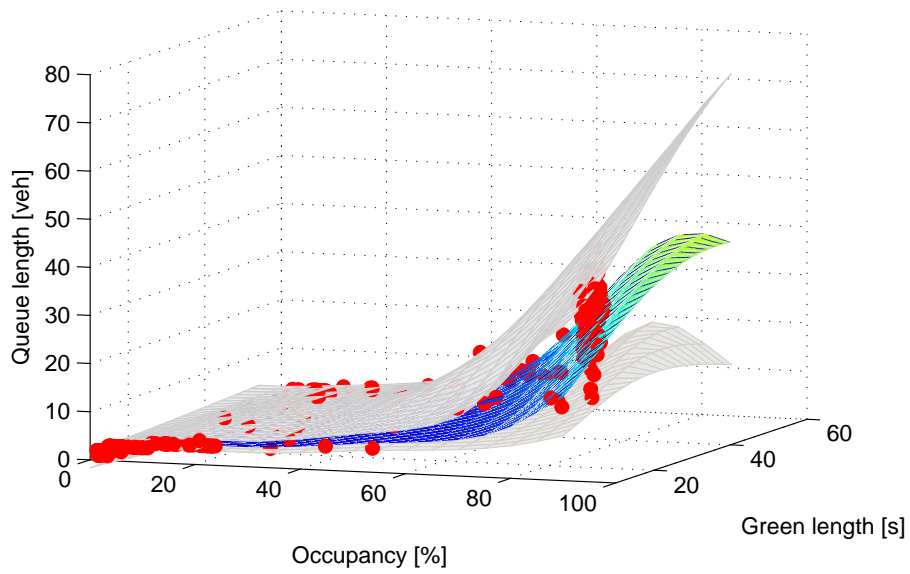
Fig. 5 *Model prediction on training data. Circles denote the training data, the middle surface is the mean prediction and the upper and lower gray surfaces limit the 95% confidence interval.*
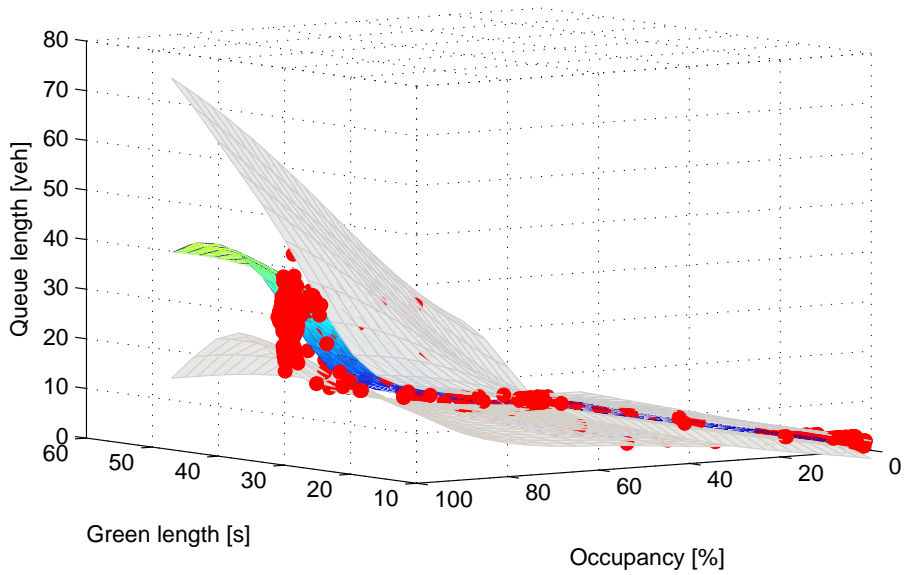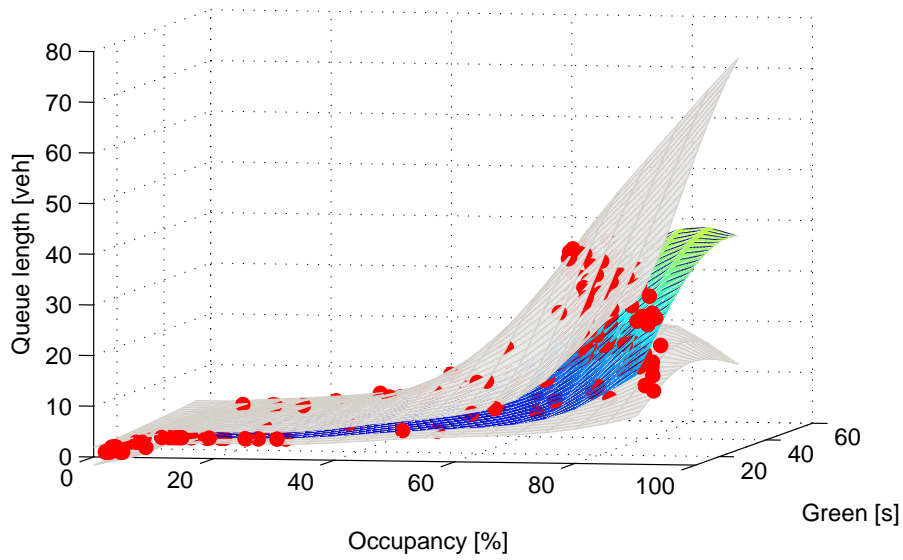

Fig. 6 *Another view of the model prediction on training data. Circles denote the training data, the middle surface is the mean prediction and the upper and lower gray surfaces limit the 95% confidence interval.*

Fig. 7 *Model prediction on validation data. Circles denote the validation data, the middle surface is the mean prediction of the model and the upper and lower gray surfaces limit the 95% confidence interval.*
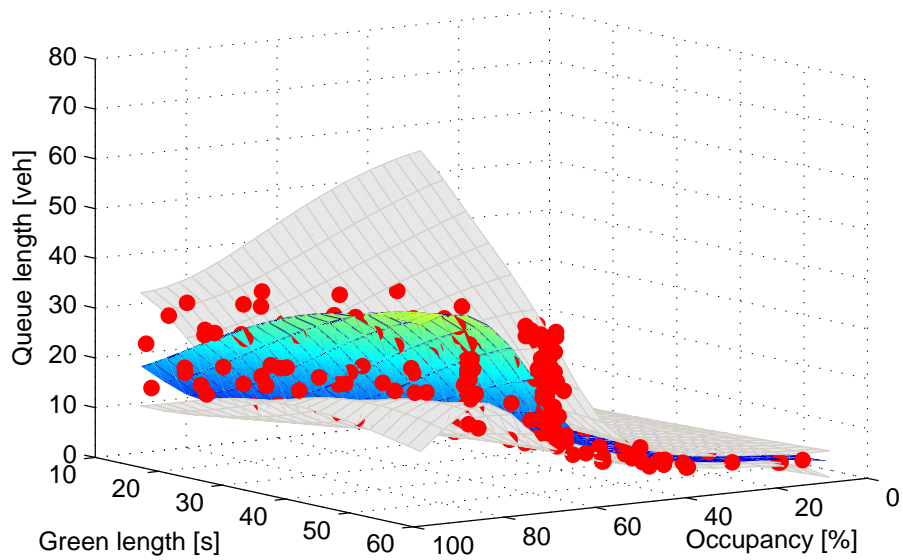

Fig. 8 *Another view of model predictions on validation data. Circles denote the validation data, the middle surface is the mean prediction of the model and the upper and lower gray surfaces limit the 95% confidence interval.*
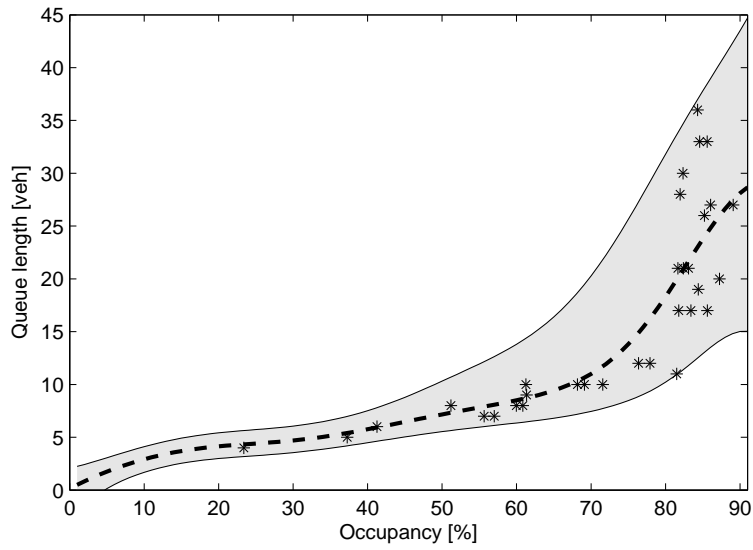
Fig. 9 *Cut along the line $g = 25$ s through the graph on Figs. 7 and 8. Stars denote selected validation occupancy-queue length pairs for which the green length falls in the range $g \in \langle 20, 30 \rangle$ s, dashed line denotes the predicted mean value, and the gray filled area limits the 95 % confidence interval.*
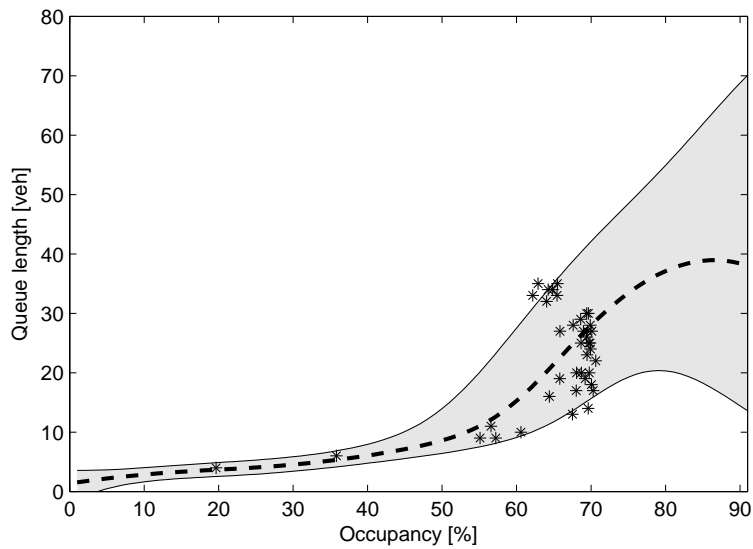

Fig. 10 *Cut along the line $g = 50$ s through the graph on Figs. 7 and 8. Stars denote selected validation occupancy-queue length pairs for which $g \in \langle 45, 55 \rangle$ s, dashed line denotes the predicted mean value, and the gray filled area limits the 95% confidence interval.*

distributions and in such a case the confidence interval lower bound can be clamped at zero. This clamping corresponds in fact to adding prior knowledge "queue length can not be negative" to the model.

6. Conclusions

We have presented an empirical static model for determining queue length from sparse time-occupancy data and green signal length measurements, based on GP modelling framework. Due to the large differences in the variance of the observed data, the presented version of the model uses a warping technique to transform the data into a form suitable for GP. This allows us to use a single static model for the relevant range of parameters.

The presented model is based on simulated data as we do not have access to equipment that would allow us to measure the queue directly.

Experimental results confirm that occupancy measurements may provide reasonable queue length estimates for unsaturated conditions on the detector and that the GP-based model can identify situations where the occupancy measurements will not provide reasonable estimate due to detector saturation. The effect of rising queue length on standard deviation of its estimate is similar to the observations of [39] for an intensity-based model. As our model is meant to complement an existing intensity-based estimator, we will in the latter case ignore the information provided by the GP model, although we may use the knowledge of detector saturation e.g. to set a lower bound on the queue length provided by an intensity-based queue length estimator.

The main advantage of using a GP model over other similar modelling approaches is that it provides the confidence interval of the model predictions, which is valuable additional information – results provided by different models may or may not prove good in practice, but the confidence interval reflects the uncertainty of the predictions due to, for example, lack of representative data, or due to uncertainty in the original training data set. Prediction confidence intervals can be calculated also when using other regression based methods, but with predictions of GP models this information is already inherently available.

Many possible extensions to this model can be developed: The model is unimodal, relying on the fact that the observed traffic stream is fairly homogeneous. Should this not be the case, more elaborate version could be developed, taking into account also different vehicle classes and their lengths. Also, the model is static, meaning that it does not make use of information about traffic dynamics, as some of the intensity based approaches do [6]. If a higher sampling rate could be achieved (for example, 5 or 10 seconds instead of current 90 seconds) the dynamic behaviour of the queueing process could be possibly captured by a dynamic model. Such a model would provide better estimates and improve the reliability of the estimates in the region of detector saturation. Finally, some of the ideas provided by [29] and [24] may be used to combine the sparse time-occupancy measurements with other information about the traffic on an approach to a signalised intersection.

Acknowledgement

References

[1] AKÇELIK R. Time-Dependent Expressions for Delay, Stop Rate and Queue Length at Traffic Signals. Vermont South: Australian Road Research Board, 1980. Technical report AIR 367-1.

[2] AŽMAN K., KOCIJAN J. Application of Gaussian processes for black-box modelling of biosystems. *ISA Transactions*. 2007, 46(4), pp. 443–457, doi: 10.1016/j.isatra.2007.04.001.

[3] BISHOP C.M. *Pattern recognition and machine learning*. New York: Springer Science + Business Media, 2006. Information Science and Statistics series.

[4] CHANG G.-L., SU C.-C. Predicting Intersection Queue with Neural Network Models. *Transportation Research Part C: Emerging Technologies*. 1995, 3(3), pp. 175–191, doi: 10.1016/0968-090X(95)00005-4.

[5] CHANG J., LIEBERMANN E.B., SHENK PRASSAS E. Queue Estimation Algorithm for Real-Time Control Policy using Detector Data [online]. Huntington Station, NY: KLD Associates, 2000 [viewed 2015-01-19]. Technical report. Available from: http://www.kldcompanies.com/cmsCSoft/uploads/realtime2000.pdf.

[6] DIAKAKI C. Integrated Control of Traffic Flow in Corridor Networks. Chania, 1999. PhD thesis, Technical University of Crete.

[7] FANG F.C., ELEFTERIADOU L. Some Guidelines for Selecting Microsimulation Models for Interchange Traffic Operational Analysis. *Journal of Transportation Engineering*. 2005, 131(7), pp. 535–543, doi: 10.1061/(ASCE)0733-947X(2005)131:7(535).

[8] FAUL S., et al. Gaussian process modelling of EEG for the detection of neonatal seizure. IEEE Transactions on *Biomedical Engineering*. 2007, 54(12), pp. 2151–2162, doi: 10.1109/TBME.2007.895745.

[9] FRIEDRICH B., et al. Data Fusion Techniques for Adaptive Traffic Signal Control. In: S. TSUGAWA, M. AOKI, eds. *Control in Transportation Systems 2003 (CTS '03)*. A proceedings volume from the 10th IFAC Symposium, Tokyo, Japan. Amsterdam: Elsevier, 2003, pp. 86–91.

[10] GIRIANNA M., BENEKOHAL R.F. Using genetic algorithms to design signal coordination for oversaturated networks. *Journal of Intelligent Transportation Systems*. 2004, 8(2), pp. 117–129, doi: 10.1080/15472450490435340.

[11] GRAŠIČ B., MLAKAR P., BOŽNAR M.Z. Ozone prediction based on neural networks and Gaussian processes. *Nuovo Cimento della Societa Italiana di Fisica, Sect. C*. 2006, 29(6), pp. 651–662, doi: 10.1393/ncc/i2006-10011-5.

[12] D.A. HENSHER, K.J. BUTTON, eds. *Handbook of Transport Modelling*. Oxford: Pergamon Press, 2000.

[13] HO C.-H., HWANG T.-L. Modeling Real-Time Dynamic Queue Length for Urban Traffic Control Systems. In: *Proceedings of the Intelligent Vehicles '94 Symposium*, Paris, France. Piscataway, NJ: IEEE, 1994, pp. 438–442, doi: 10.1109/IVS.1994.639558.

[14] HOMOLOVÁ J., NAGY I. Traffic model of a microregion. In: P. HORÁČEK, M. ŠIMANDL, P. ZÍTEK, eds. *Preprints of the 16th World Congress of the International Federation of Automatic Control*, Prague, Czech Republic. Prague: IFAC, 2005, pp. 1–6.

[15] JOHANSEN T.A., SHORTEN R., MURRAY-SMITH R. On the Interpretation and Identification of Dynamic Takagi-Sugeno Fuzzy Models. *IEEE Transactions on Fuzzy Systems*. 2000, 8(3), pp. 297–313, doi: 10.1109/91.855918.

[16] KLEIN L.A., MILLS M.K., GIBSON D.R. Traffic Detector Handbook: Third Edition – Volume I. McLean, VA: Federal Highway Administration, Turner-Fairbank Highway Research Center, 2006. Technical report FHWA-HRT-06-108.

[17] KOCIJAN J., LIKAR B. Gas-Liquid Separator Modelling and Simulation with Gaussian Process Models. In: *Proceedings of the 6th EUROSIM Congress on Modelling and Simulation – EUROSIM 2007*, Ljubljana, Slovenia. Ljubljana: SLOSIM and University of Ljubljana, 2007.

[18] KOCIJAN J., et al. Dynamic systems identification with Gaussian processes. *Mathematical and Computer Modelling of Dynamic Systems*. 2005, 11(4), pp. 411–424, doi: 10.1080/13873950500068567.

[19] LAZARO-GREDILLA M. Bayesian warped Gaussian processes. In: F. PEREIRA, et al., eds. *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012*, Vancouver, Canada. Red Hook, NY: Curran Associates, Inc., 2013, pp. 1619–1627. Advances in Neural Information Processing Systems series. Available also from: http://papers.nips.cc/paper/4494-bayesian-warped-gaussian-processes.pdf.

[20] LEDOUX C. An Urban Traffic Flow Model Integrating Neural Networks. *Transportation Research Part C: Emerging Technologies*. 1997, 5(5), pp. 287–300, doi: 10.1016/S0968-090X(97)00015-6.

[21] LEITH D.J., HEIDL M., RINGWOOD J. Gaussian process prior models for electrical load forecasting. In: *Proceedings of 2004 International Conference on Probabilistic Methods Applied to Power Systems*, Ames, IA. Piscataway, NJ: IEEE, 2004, pp. 112–117.

[22] LEITHEAD W.E., ZHANG Y., NEO K.S. Wind turbine rotor acceleration: Identification using Gaussian regression. In: *Proceedings of 2nd International conference on informatics in control automation and robotics (ICINCO 2005)*, Barcelona, Spain. Setúbal: INSTICC, 2005, pp. 84–91.

[23] LIKAR B., KOCIJAN J. Predictive control of a gas-liquid separation plant based on a Gaussian process model. *Computers and Chemical Engineering*. 2007, 31(3), pp. 142–152, doi: 10.1016/j.compchemeng.2006.05.011.

[24] MA D., et al. A Method for Queue Length Estimation in an Urban Street Network Based on Roll Time Occupancy Data. *Mathematical Problems in Engineering*. 2012, 2012, Article ID 892575, 12 pp., doi: 10.1155/2012/892575.

[25] MÜCK J. Estimation Methods for the State of Traffic at Traffic Signals using Detectors near the Stop-Line. *Traffic Engineering and Control*. 2002, 43(11), pp. 429.

[26] R. MURRAY-SMITH, T.A. JOHANSEN, eds. *Multiple model approaches to modelling and control*. London: Taylor and Francis, 1997.

[27] MYSTKOWSKI C., KHAN S. Estimating Queue Lengths Using SIGNAL94, SYNCHRO3, TRANSYT-7F, PASSER II-90, and CORSIM. In: *Proceedings of 78th Transportation Research Board Annual Meeting*, Washington, DC. Washington: Transportation Research Board of the National Academies of Science, 1998.

[28] PAPAGEORGIOU M. An integrated control approach for traffic corridors. *Transportation Research Part C: Emerging Technologies*. 1995, 3(1), pp. 19–30, doi: 10.1016/0968-090X(94)00012-T.

[29] PAPAGEORGIOU M., VIGOS G. Relating time-occupancy to space-occupancy and link vehicle-count. *Transportation Research Part C: Emerging Technologies*. 2008, 16(1), pp. 1–17, doi: 10.1016/j.trc.2007.06.001.

[30] PŘIKRYL J., KOCIJAN J. An Empirical Model of Occupancy-Queue Relation. In: A. G. CHASSIAKOS, ed. *Proceedings of 12th IFAC Symposium on Transportation Systems*, Redondo Beach, CA, USA. Laxenburg, Austria: IFAC, 2009, pp. 456–461, doi: 10.3182/20090902-3-US-2007.00068.

[31] QUEK C., PASQUIER M., LIM B.B.S. POP-TRAFFIC: A Novel Fuzzy Neural Approach to Road Traffic Analysis and Prediction. *IEEE Transactions on Intelligent Transportation Systems*. 2006, 7(2), pp. 133–146, doi: 10.1109/TITS.2006.874712.

[32] RASMUSSEN C.E., WILLIAMS C.K. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[33] SHI J.Q., CHOI T. *Gaussian process regression analysis for functional data*. Boca Raton, FL: Chapman and Hall/CRC, Taylor & Francis group, 2011.

[34] SNELSON E. Flexible and efficient Gaussian process models for machine learning. London, 2007. PhD thesis, University College London.

[35] SNELSON E., RASMUSSEN C.E., GHAHRAMANI Z. Warped Gaussian processes. In: S. THRUN, L. K. SAUL, B. SCHÖLKOPF, eds. *Advances in Neural Information Processing Systems 16*. Proceedings of NIPS 2003, Vancouver, Canada. Cambridge, MA: MIT Press, 2004, pp. 337–344. Available also from: http://papers.nips.cc/paper/2481-warped-gaussian-processes.pdf.

[36] TRANSPORT SIMULATION SYSTEMS. AIMSUN v. 5.0.1. Advanced Interactive Microscopic Simulator for Urban and non-Urban Networks [software]. 2008-05-12.

[37] TRANSPORTATION RESEARCH BOARD. *Highway Capacity Manual 2000*. Washington, DC: Transportation Research Board of the National Academies of Science, 2000.

[38] Van ZUYLEN H. J., VITI F. Delay at controlled intersections: the old theory revised. In: *Proceedigns of the 2006 IEEE Intelligent Transportation Systems Conference (IEEE ITSC 2006)*, Toronto, Canada. Piscataway, NJ: IEEE, 2006, pp. 68–73, doi: 10.1109/ITSC.2006.1706720.

[39] Van ZUYLEN H. J., VITI F. Uncertainty and the Dynamics of Queues at Controlled Intersections. In: S. TSUGAWA, M. AOKI, eds. *Control in Transportation Systems 2003 (CTS '03)*. A proceedings volume from the 10th IFAC Symposium, Tokyo, Japan. Amsterdam: Elsevier, 2003, pp. 43–48. Available also from: http://www.kuleuven.be/traffic/dwn/P2003I.

[40] VIGOS G., PAPAGEORGIOU M.,WANG Y. Real-time estimation of vehicle-count within signalized links. *Transportation Research Part C: Emerging Technologies*. 2008, 16(1), pp. 18–35, doi: 10.1016/j.trc.2007.06.002.

[41] VILORIA F., COURAGE K., AVERY D. Comparison of Queue-Length Models at Signalized Intersections. *Transportation Research Record*. 2000, 1710, pp. 222–230, doi: 10.3141/1710-26.

[42] VITI F., van ZUYLEN H.J. Probabilistic models for queues at fixed control signals. *Transportation Research Part B: Methodological*. 2009, 44(1), pp. 120–135, doi: 10.1016/j.trb.2009.05.001.

[43] VITI F., van ZUYLEN H.J. The dynamics and the uncertainty of queues at fixed and actuated controls: A probabilistic approach. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*. 2009, 13(1), pp. 39–51, doi: 10.1080/15472450802644470.

[44] VLAHOGIANNI E.I., KARLAFTIS M.G., GOLIAS J.C. Optimized and metaoptimized neural networks for short-term traffic flow prediction: A genetic approach. *Transportation Research Part C*. 2005, 13(3), pp. 211–234, doi: 10.1016/j.trc.2005.04.007.

[45] WANG J.M., FLEET D.J., HERTZMANN A. Gaussian Process Dynamical Models for Human Motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2008, 30(2), pp. 283–298, doi: 10.1109/TPAMI.2007.1167.

[46] YANG D., et al. A robust vehicle queuing and dissipation detection method based on two cameras. In: *Proceedings of 14th International IEEE Conference on Intelligent Transportation Systems (IEEE ITSC 2011)*, Washington, DC. Piscataway, NJ: IEEE, 2011, pp. 301–307, doi: 10.1109/ITSC.2011.6082828.

[47] ZHENG J., et al. Measuring Signalized Intersection Performance in Real-Time with Traffic Sensors. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*. 2013, 17(4), pp. 304–316, doi: 10.1080/15472450.2013.771105.