Validating state-dependent queueing models for uninterrupted traffic flows using simulation

B. Wuyts, T. Van Woensel, N. Vandaele
University of Antwerp, Belgium

January 26, 2004

Abstract

The main purpose of this paper is twofold. First, the relationship between flow and speed is assessed by means of simulation. Second, the effectiveness of some queueing based traffic models with state dependency is tested by comparing their outcomes to the recorded simulation-based results. It appears that the M/G/1 queueing model with gaussian state dependency outperforms all other models.

1 Introduction

In macroscopic models (as opposed to microscopic models, which are outside the scope of this paper) all individual vehicles are aggregated and described as flows. Traditionally, uninterrupted traffic flows are modeled by determining the speed-flow relationship empirically (Li [12]). Two methodologies are described in literature (Daganzo [2]). First, the speed-flow relationship can be estimated econometrically using data on observed traffic flows and speeds, and fitted into a speed-flow specification (Daganzo [2]). Obviously, this method suffers from limited feasibility, since although traffic data concerning volumes or counts is easily obtainable, the corresponding average speed of the traffic flow is not available. The second method to assess the speed-flow relationship consists in postulating a functional form, followed by calibrating its parameters based on a few observations (De Borger and Proost [13]). The main disadvantage of the latter method is the incompetence of the model to predict speeds that significantly differ from the speeds used to calibrate the parameters.

In this paper the relationship between flow and speed is simulated in order to test the effectiveness of some queueing based traffic models with new state dependent functions. Jain and Smith [5] show that state dependent queueing models where the service rate is a function of the flow already on the road, can be used to obtain a better fit to the speed-flow-density diagrams than the non-state dependent queueing models. This result is confirmed by Van Woensel and Vandaele [17]. This paper is organized as follows. First the method to simulate a speed-flow relationship is explained in detail. Second, we elaborate

on previously reported and new different state dependency functions. Third, these functions are evaluated against the simulated speed-flow data. This paper ends with a summary of conclusions.

2 Speed-flow simulation

In order to test the effectiveness of some queueing models by comparing their outcomes with real life situations, a speed-flow profile is simulated. No empirical speed-flow data is used since (i) no extensive dataset is available and (ii) empirical speed-flow data typically contain a significant amount of outliers. Therefore, we opt to use simulation to generate a speed-flow profile.

2.1 Method

Taken into account the typical bimodal character of flow data (morning and evening rush), a straightforward application of a stationary distribution for the arrival process is not appropriate. Therefore, several authors have put forward procedures where the underlying distribution of the arrival process is nonstationary (Leemis [11], Kao and Chang [7] and Johnson et al. [6]). A recently developed methodology (Preston White [15]) concerns a Nonstationary Poisson process using Bivariate Thinning (NPBT). For this model, arrivals are assumed to be nonstationary Poisson, with a piecewise-constant arrival rate independently regulated by hour and by day. This methodology will be applied to a dataset of observed flow. First, the data are converted to standard values (i.e. the mean is substracted from the dataset values and subsequently divided by their standard deviation). Observations with a standard value exceeding 2.5 will be removed (Hair [1]). Let x_{ij} be the observed flow data for hour i ($i = 1..24$) and day j ($j = 1..7$). Since the pattern of weekdays and weekenddays is typically different (weekenddays do not have an outspoken morning and evening rush, considerably more traffic is present during weekdays, etc.), the NPBT method will be applied separately on both types of days. Below, the procedure for the weekdays is described - weekenddays are treated in a similar way. The mean $\overline{x}_{ij}$ of each hour ($i = 1..24$) per day ($j = 1..5$) is calculated from the dataset. For each hour of the day and each day of the week we have

$$\begin{aligned} \overline{x}_i &= \frac{1}{5}\sum_{j=1}^{5}\overline{x}_{ij}, \qquad \forall i = 1..24 \\ \overline{x}_j &= \frac{1}{24}\sum_{i=1}^{24}\overline{x}_{ij}, ...\forall j = 1..5 \end{aligned}$$

Next, hourly and daily thinning factors are computed:

$$\eta_i = \frac{\overline{x}_i}{\max_i(\overline{x}_i)}, \qquad \forall i = 1..24$$
$$\delta_j = \frac{\overline{x}_j}{\max_j(\overline{x}_j)}, ...\forall j = 1..5$$

The hour block with the highest mean is combined with the day block with the highest mean:

$$\lambda_{\max} = \overline{x}_{ij} \; for \; (i,j) \; such \; that \; \eta_i \delta_j = 1$$

The arrivals are generated using an exponential interarrival rate with mean $\frac{1}{\lambda_{\max}}$ and consecutively thinned with the bivariate acceptance probability

$$p_{ij} = \eta_i \delta_j \qquad \forall i, j$$

Preston White [15] shows that the nonstationary poisson process with bivariate thinning is an adequate method to simulate arrivals at a consumer store. In section 2.3 it is shown that the NPBT method is appropriate for the simulation of arrivals in a traffic context as well.

With the aim to assess the ratio of service, the empirically recorded maximum number of vehicles is assumed to be a proxy for the service rate. However, since internal as well as external elements impact upon the ratio of service (e.g. accidents and rain), a traffic service rate can not be considered as a constant quantity. The service rate is defined as the product of free flow speed v_f with the maximum traffic density k_j (Heidemann [4]). And because internal and external elements can influence free flow speed, the service rate will be changed as well due to the latter variability.

With the intention of constructing a distribution of the maximum number of vehicles, nonparametric bootstrap samples are taken. The key idea is to resample from the original data to create replicate datasets, from which the variability of the quantities of interest (in our case the maximum number of vehicles) can be assessed without long-winded and error-prone analytical calculation (Davison and Hinkley [3]). Let Y be a vector which contains all the flow observations. Let then $Y_1^*, ..., Y_{1000}^*$ be 1000 bootstrap samples (i.e. each Y^* independently sampled at random with replacement from Y).

$$Y_{\max}^* = \{\max(Y_1^*), ..., \max(Y_{1000}^*)\}$$

The empirical discrete distribution vector $Y_{\max}^*$ will now be used as a proxy for the distribution of the service ratio.

2.2 Data

The data are collected by the Department of Environment and Infrastructure of the Ministry of the Flemish Community. This department is responsible for collecting and reporting the counting data for all Flemish roads. The dataset

consists of observations of the number of vehicles per hour for 273 days of the A1/E19 highway, section Mechelen North - Rumst, headed for Brussels (i.e. counter 19012).

2.3 Simulation

An example of the Nonstationary Poisson process using Bivariate Thinning methodology applied to the described dataset is presented in figure 1. In order to test the adequacy of the NPBT method, the Kruskal-Wallis test is applied, verifying the hypothesis H_0 whether all samples are drawn from the same population. This test (Kruskal and Wallis [9]) is the nonparametric version of one way ANOVA and is a straightforward generalization of the Wilcoxon test for two independent samples. If there are K independent samples of sizes $n_1, n_2, ..., n_K$, all the samples are combined into one large sample of size $n = \sum n_i$, the result is sorted from smallest to largest and ranks are assigned (the average rank is assigned to any observation in a group of tied observations), and then the average $\overline{R}_i$ of the ranks of the observations in the ith sample is found. The test statistic is then

$$H = \frac{12}{n(n+1)} \sum \frac{\overline{R}_i^2}{n_i} - 3(n+1)$$

and the null hypothesis that all K distributions are the same is rejected if $H > \chi^2_{K-1}$. The p-value indicates the probability that one would obtain a test statistic which is more extreme than the observed one when the H_0 is true. The rule is that we reject H_0 if $p < \alpha$.

A sufficiently large amount of 30 simulations versus 30 empirical observations of flow data of one week of the A1/E19 highway is compared by means of this test. The realization of this procedure reveals a p-value of 0.9922. Two substantial conclusions can be drawn from the former statistic: (i) the empirical observations are generated by one underlying process and (ii) the hypothesis that both the observed and simulated data originate from the same underlying distribution can not be rejected at the 99 procent significance level.

To cope with the problem of the initial transient phase inherent to simulation, the general technique of Welch (Law and Kelton [10]) is applied. The number of vehicles on Monday between 0 and 1 a.m. is taken as the random variable of interest, and 10 replications of the simulation -each of length 10- are made (Law and Kelton [10] propose a minimum number of replications of 5). It appears that no warmup period is necessary. No significant number of cars in the system on Monday between 0 and 1 a.m. is emanating from the preceeding hour. This is due to the fact that (i) the unit of time is one hour and (ii) the simulation starts at 0 a.m. (if the simulation would start at 7 a.m. (i.e. morning rush) an initial transient phase is more likely to occur). The simulation can therefore be considered as a terminating system, which implies the independence of the different weeks.

The simulation results in a vector of total times, and since the effective speed v is calculated as the division of the length of the road segment $\frac{1}{k_j}$ by the total

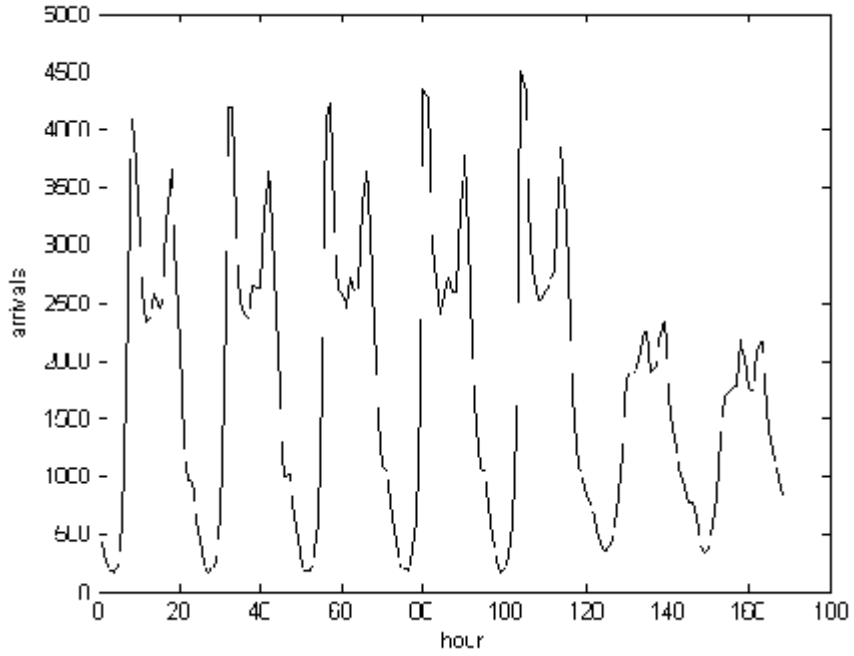

Figure 1: Simulation example of a nonstationary poisson process using bivariate thinning

time in the system W, (Vandaele, Van Woensel and Verbruggen [14]) or

$$v = \frac{1/k_j}{W}$$

with k_j (the maximum traffic density) estimated by the empirically observed maximum flow of 6034 vehicles, in total 16800 (i.e. 10 replications of a simulation of 10 weeks, with each week containing 168 hours) coordinates in the speed-flow space are obtained. Figure 2 shows the simulated speed-flow relationship, which has the form of a quadratic function $(v_f)_q = \alpha q^2 + \beta q + \gamma$.

Any quadratic function can be estimated by use of the OLS procedure. Table 1 gives the estimated coefficients and their respective significance levels. The $R^2_{adjusted}$ is very satisfying: 99.5 procent of the variation of speed is explained by the variation of the flow specification. From the presented p-values (between brackets) the coefficients appear to be significantly different from zero.

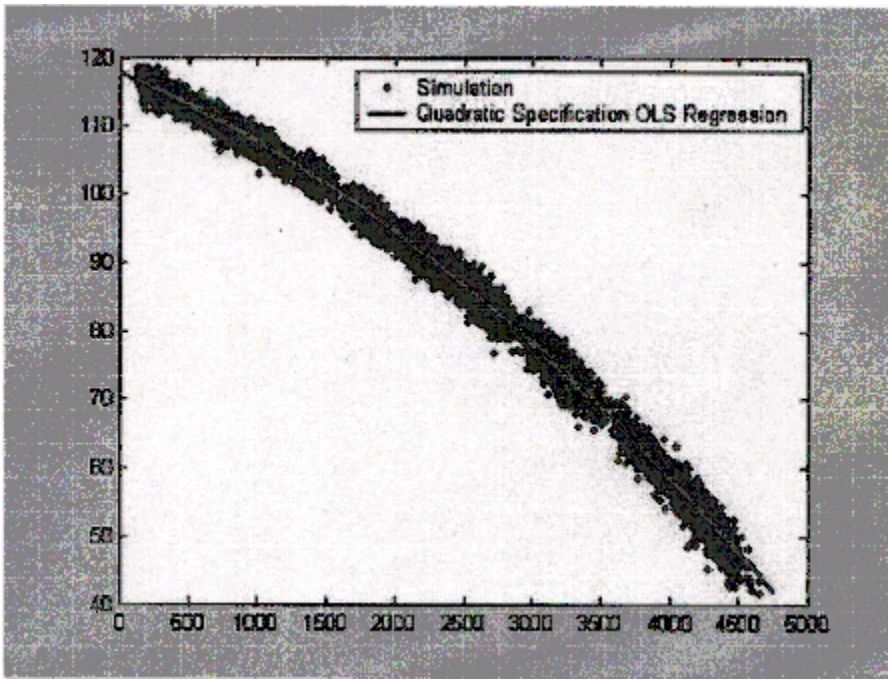

Figure 2: The simulated speed-flow diagram

3 Comparison of queueing models

The comparison shows to what extent the outcome of different queueing models is similar to the simulated results. Several queueing models are taken into consideration: $M/M/1$, $M/G/1$ and $GI/G/1$. Furthermore, for each of the mentioned queueing models different previously reported and new state dependent functions are implemented. First, the methodology of comparison is explained. Then, important properties of the queueing parameters are discussed, consecutively some existing and new state dependent functions are presented. Finally an overview of the results with some conclusions will be given.

3.1 Method of comparison

The speeds are compared by Theil inequality coefficients. The Theil coefficient of the queueing speed v_q and the earlier simulated speed v_s is defined as

Coefficient	Estimation
α	117.751
	(0.000)
β	-0.00844
	(0.000)
γ	-0.00000158
	(0.000)
$R^2_{adjusted}$	0.995

Table 1: OLS Regression of the quadratic specification of the speed-flow relationship $v = \alpha + \beta q + \gamma q^2$

$$
\begin{aligned}
Theil_{vq,vs} &= \frac{\sqrt{\frac{\sum (v_q - v_s)^2}{n}}}{\sqrt{\frac{\sum v_q^2}{n}} + \sqrt{\frac{\sum v_s^2}{n}}} \\
v_s &: \textit{simulated speed} \\
v_q &: \textit{queueing speed}
\end{aligned}
$$

The Theil inequality coefficient can be decomposed into the following three inequality proportions. In this way the error can be broken up into three different characteristic sources.

1. The bias proportion:

$$Theil_{bias} = \frac{(\overline{v_q} - \overline{v_s})^2}{\frac{1}{n}\sum (v_q - v_s)^2} \tag{1}$$

   This proportion reveals information about how far the mean of the series of queueing speeds v_q is from the mean of the series of simulated speeds $v_{s.}$

2. The variance proportion:

$$Theil_{var} = \frac{(\sigma_q - \sigma_s)^2}{\frac{1}{n}\sum (v_q - v_s)^2} \tag{2}$$

   This proportion gives information about how far the variation of the series of queueing speeds v_q is from the mean of the variation of the series of simulated speeds.

3. The covariance proportion:

$$Theil_{cov} = \frac{2 \times (1 - \rho) \times \sigma_q \times \sigma_s}{\frac{1}{n}\sum (v_q - v_s)^2} \tag{3}$$

The covariance proportion measures (with use of the correlation coeffcient ρ) the remaining unsystematic errors.

The bias, variance and covariance proportions all add up to one.

For each queueing model, the speed-flow diagram is constructed. Consequently, for every flow, the queueing speed is compared with the simulated speed for the same flow. The latter speed will be approximated by the next equation derived from table 1:

$$v_s = 117,751 - 844,012 \times 10^{-5} q - 158,002 \times 10^{-8} q^2 \tag{4}$$

Flow is hereby allowed to range to a maximum of 4647 vehicles per hour (i.e. the maximum observed flow in the simulation).

It is important to notice that although for every queueing model the maximum traffic density k_j is strictly determined, we still allow this parameter to fluctuate within certain predefined bounds in order to obtain the most appropriate (i.e. the lowest) Theil coefficient. This decision is relatively easy to motivate since for every considered queueing model the calculation of k_j involves both the free flow speed v_f and the maximum observed flow $q_{\max}$. Both parameters are merely estimates of the true underlying values, and therefore k_j can differ from its value that is appropriate for every queueing model (see paragraph 3.2). If the acquired parameters contrast too much with the queueing formulae, the practical use concerning that parameter setting is supposed to be rather limited. Next to an adjustment of the queueing parameters to become a smaller Theil coefficient, an implementation of a state dependent function (see paragraph 3.3) as well can improve the Theil inequality coefficient. We will examine which state dependent function will result in the most favourable Theil coefficient. Figure 3 presents an overview of the methodology: for three different queueing models, the parameters and the state dependent function are allowed to be modified simultaneously. The obtained queueing models can be represented by a speed-flow relationship, which is each time compared with the simulated speed flow equation (4). The state dependent queueing model with the lowest Theil coefficient will be preferred.

An inherent feature of the queueing models we use, concerns the limited flow range: the models are confronted with a maximum flow for which an exact speed can be calculated (due to the restricted setting of the parameters). It is therefore feasible that although a certain flow occured in the simulation, this flow can not generate a speed-flow profile in some queueing models due to the limiting behaviour of the parameters. Consequently, in the former situation the Theil inequality coefficient can not be calculated, and such limited flow models will accordingly be considered as irrelevant for purposes of our study.

A final remark about the methodology of comparison. Since the simulation described in this paper is straightforward macroscopic: all individual vehicles are aggregated and described as flows (Van Woensel [16]), the slope of the speed-flow relationship will be typically negative. Therefore, only the descending part of the queueing models will be used to be compared with the simulation

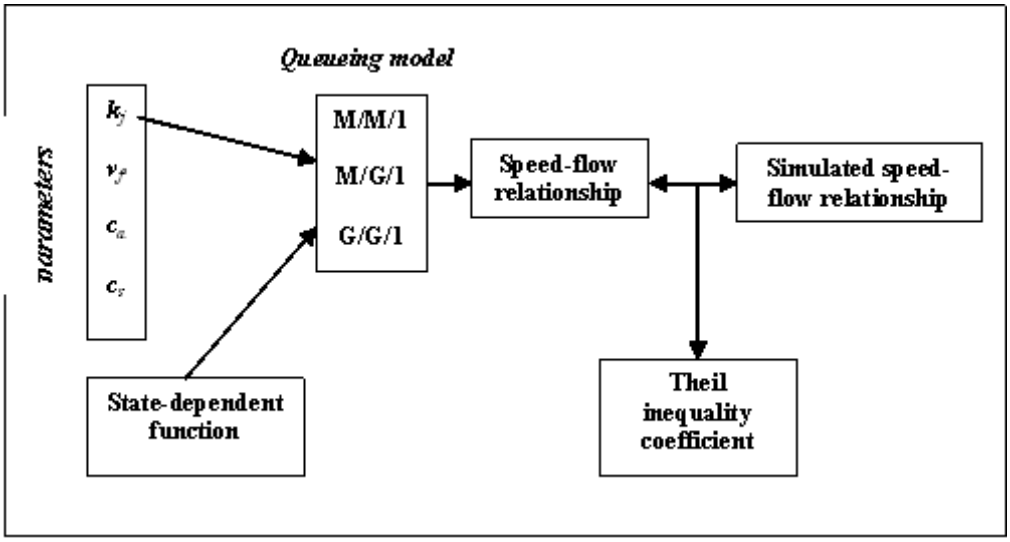

Figure 3: Methodology to validate state dependent queueing models

results. For each queueing model the parameter setting that minimizes the Theil coefficient is chosen.

3.2 Queueing parameters

In this paragraph, the appropriate value for the parameters of the different queueing models will be discussed. Furthermore, a necessary elaboration on the $GI/G/1$ queueing model will be explained.

3.2.1 The M/M/1 queueing model

The value for the maximum traffic density k_j can be calculated by (Van Woensel [14])

$$k_j = \frac{4q_{\max}}{v_f} \tag{5}$$

with a maximum speed allowed in Belgium of 120 km/hr and a maximum historical flow of 6034 vehicles per hour, the maximum traffic density k_j equals close to 201 vehicles per kilometer for the $M/M/1$ queueing model.

3.2.2 The M/G/1 queueing model

The maximum traffic density k_j can be expressed by (Vandaele, Van Woensel and Verbruggen [16])

$$k_j = \frac{q_{\max}}{2v_f \left[\frac{\sqrt{c_s^2+1}-\sqrt{2}}{c_s^2-1}\right]^2} \tag{6}$$

A measure for the variability of the service times is the coefficient of variation of the service times; it is calculated as $\frac{\sigma_{servicetimes}}{\mu_{servicetimes}}$ and is denoted as c_s. The observed maximum flow of 6034 per hour is assumed to have taken place in the best possible conditions. Therefore the coefficient of variation of the service time c_s is set to zero (Van Woensel [16] argues that the variability under ideal circumstances -or the natural variability- is equal to zero). The utopic idea of a continuous cruise control modus for every vehicle would for instance lead to zero variability. Consequently k_j equals approximately 146 vehicles per kilometer for the $M/G/1$ queueing model.

3.2.3 The GI/G/1 queueing model

Under the assumption that the maximum historical flow of 6034 vehicles per hour occured in the best possible conditions, the coefficients of variation of the interarrival time and the service time (i.e. c_a and c_s) are both set equal to their natural variability level of zero. The former leads to the $D/D/1$ queueing model with a speed v equal to v_f (Van Woensel [16]):

$$k_j = \frac{q_{\max}}{v_f}$$

Therefore k_j equals ca. 50 vehicles per kilometer for the $GI/G/1$ queueing model.

For the $GI/G/1$ queueing model, there does not exist an explicit form for speed as a function of merely flow. Therefore an iterative method (Van Woensel [16]) is used which for every speed calculates the corresponding flow (see appendix). Since for the calculation of the Theil inequality coefficient an equal amount of observations for both the simulated and queueing results is required, two possible methodologies exist.

First, the OLS estimated function in equation (4) can be inverted. Consequently, for every given speed a simulated flow can be calculated, and the Theil inequality coefficient can be determined. However, although the calculated Theil coefficient is accurate, its interpretation differs essentially from that of the $M/M/1$ and $M/G/1$ Theil inequality coefficient. This is a direct consequence of the invertion of the OLS estimated function. The $M/M/1$ and $M/G/1$ Theil inequality coefficients are a measure for the speed variation (the speed of the respective queueing model compared with the simulated speed) that occur over the entire flow horizon. In contrast, the formerly mentioned $GI/G/1$ Theil coefficient represents the flow variation (the flow of the respective queueing model compared with the simulated flow) that appears over the speed horizon. Figures 4 and 5 show graphically the difference in calculating the Theil inequality coefficient between respectively the $M/M/1$ queueing model (for the $M/G/1$ queueing model, the Theil coefficient is calculated similarly) and the

$GI/G/1$ queueing model. For reasons of completeness this methodology will be elaborated in appendix.

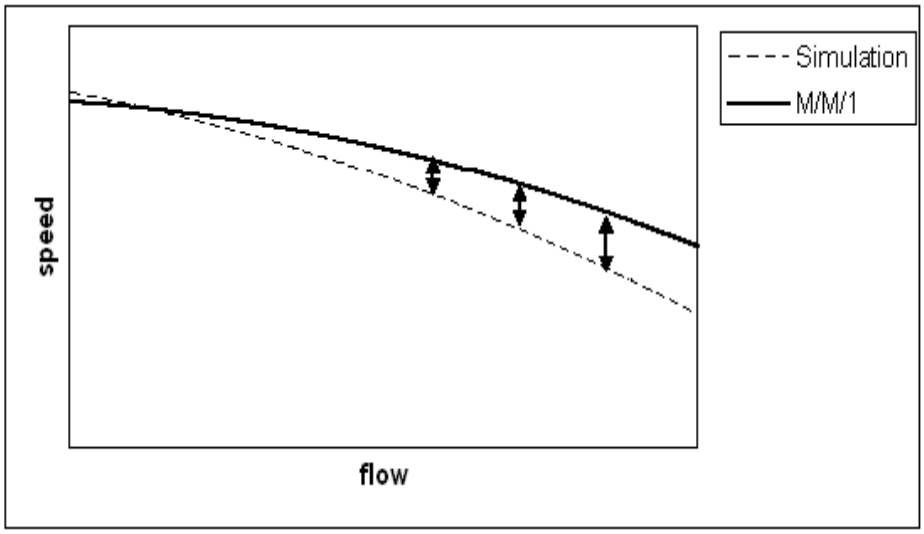

Figure 4: Calculation Theil coefficient for the *M/M/1* queueing model

A second methodology to cope with the previously mentioned difficulty -the method that will be used in this paper- consists of the implementation of a piecewise cubic interpolation of the speed-flow relationship obtained from the $GI/G/1$ iteration (i.e. the interpolation of the entire curve with a succession of cubics, one for each interval). The outcome of the interpolation will be a $(n \times 2)$-matrix (with n e.g. equal to 12.000) with in the first column the different speeds, and in the second column the correspondent flows. Now it can be verified what speed a certain flow provokes. Evidently, two problems may arise: (i) not necessarily every flow will be available since the slope of the curve enables the possibility that an increase in the speed of 1 km/hr provokes an increase in the flow of more than 2 vehicles and (ii) a flow may have two or more correspondent speeds. The former problem will be dealt with by taking that flow which is closest to the intendend flow, the latter difficulty will be dealt with by taking the arithmetic mean of the correspondent speeds.

3.3 State dependent functions

Based on the work of Jain and Smith [5], Vandaele, Van Woensel and Verbruggen [14] developed a state dependent $GI/G/1$ model. Instead of using a fixed service rate, the service rate is a function of the traffic flow. Vehicles are served at a certain rate dependent upon the number of vehicles in the system. In the state

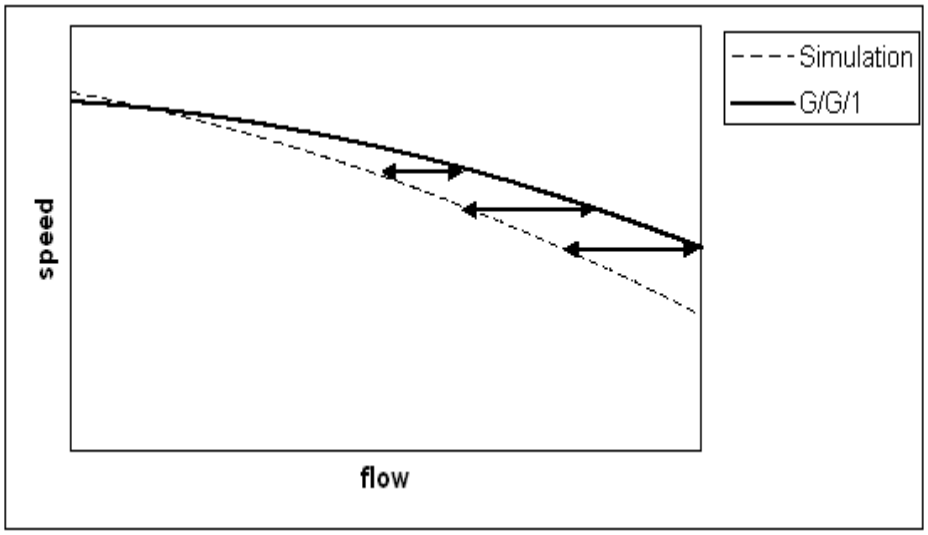

Figure 5: Calculation Theil coefficient for the *G/G/1* queueing model

dependent $GI/G/1$ model the nominal speed v_f is assumed to be a function of the traffic flow or $v_f = f(q)$. Concerning the state dependent functions, Jain and Smith [5] presented two possible specifications: the linear and the exponential function. In this paragraph, the application of state dependent functions will be extended to $M/M/1$ and $M/G/1$ queueing models. Since the linear and exponential functions might be too simple as a representation of the state dependent property of the nominal speed, the state dependent specification will be extended to other alternatives.

State dependent function	Equation
Linear	$(v_f)_q = v_f \left[\frac{q_{\max}+1-q}{q_{\max}}\right]$
Exponential	$(v_f)_q = v_f e^{-a\frac{q}{q_{\max}}}$, $with\ \alpha > 0$
Elliptic	$(v_f)_q = \sqrt{\left(1-\frac{q^2}{q_{\max}^2}\right)v_f^2}$
Gaussian	$(v_f)_q = v_f e^{\left(-\alpha q^2/q_{\max}^2\right)}, with\ \alpha > 0$
Quadratic	$(v_f)_q = v_f + \alpha q^2 + \beta q$

Table 2: The different state dependent functions

A state dependent function is subject to three self-explanatory conditions. First, the function should be descending since an increase in the number of vehicles provokes a decreasing free flow speed. Second, a saturated traffic situation

leads to a free flow speed equal to zero. Third, if a single vehicle is present, the free flow speed equals the free flow speed used in the non- state dependent queueing models. The former conditions are written as follows:

$$\begin{aligned} \forall q : \frac{\partial f(q)}{\partial q} < 0 \\ f(q_{\max}) \cong 0 \\ f(0) \cong v_f \end{aligned} \tag{7}$$

Since it is a priori quasi impossible to determine analytically a correct specification, we put forward functions which can easily be modified by a change of the parameter(s): the quadratic state dependent function and the gaussian state dependent function. Since the exponential function (Jain and Smith [5]) has an outspoken convex nature, an elliptic state dependent function (which is concave) will be subject to our study as well. The state dependent functions that are subject to evaluation in this paper are presented in table 2. Figure 6 plots the different state dependent functions (with arbitrarily chosen parameter values for the exponential, gaussian and quadratic state dependent functions).

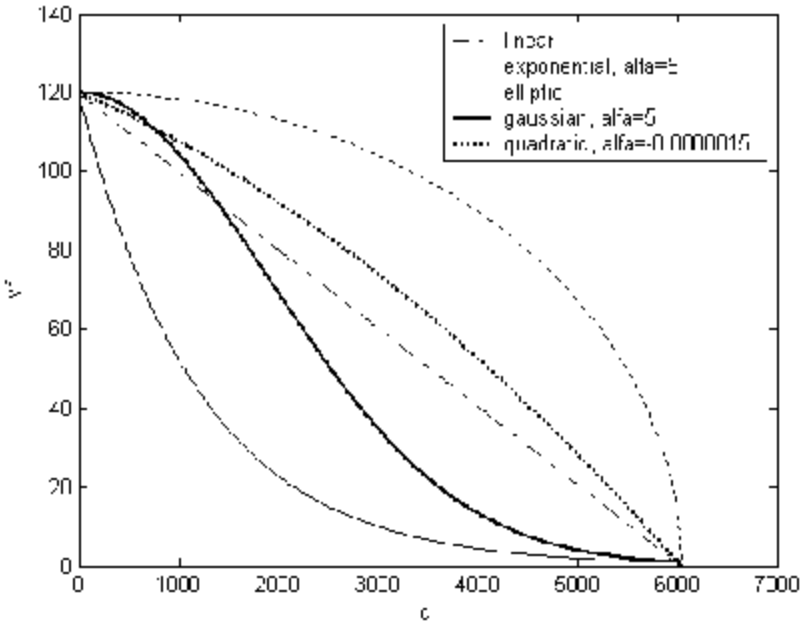

Figure 6: The different state dependent functions

The functions mentioned in table 2 are -except for the quadratic function- obviously in accordance with the state dependent conditions described in equa-

tion 7. In order to set the quadratic function to a useful state dependent function, additional parameter restrictions should be imposed.

$$\begin{aligned} \alpha &\in \left[-\frac{v_f}{q_{\max}^2}, \frac{v_f}{q_{\max}^2}\right] \\ \beta &= \frac{-\alpha q_{\max}^2 - v_f}{q_{\max}} \end{aligned} \tag{8}$$

A proof voor the former conditions can be found in appendix C.

3.4 Results

The best Theil inequality coefficient for every queueing model with matching parameter setting is shown in table 3. The Theil decomposition in respectively the bias proportion, the variance proportion and the covariance proportion is given in table 4.

Queueing Model	*Theil*	c_a	c_s	k_j	v_f	α	$k_{j,t.v.}^{***}$	$\frac{k_j - k_{j,t.v.}}{k_{j,t.v.}}$
$M/M/1$	0.0517	-	-	174	107	-	226	-23.0%
$M/M/1, linear$	0.0556	-	-	∞	137	-	176	$+\infty$
$M/M/1, exponential$	0.0112	-	-	228	120	0.50	201	+13.4%
$M/M/1, elliptic$	0.0395	-	-	247	118	-	205	+20.5%
$M/M/1, gaussian$	0.0169	-	-	248	116	0.73	208	+19.2%
$M/M/1, quadratic$	0.0273	-	-	394	116	$-0.32\times10^{-5**}$	208	+89.4%
$M/G/1$	0.0510	-	1.10	178	110	-	161	+10.6%
$M/G/1, linear$	0.0556	-	NI^*	∞	137	-	129	$+\infty$
$M/G/1, exponential$	0.0130	-	0.40	194	121	0.63	155	+25.2%
$M/G/1, elliptic$	0.0231	-	0.90	250	111	-	207	+20.8%
$M/G/1, gaussian$	0.0097	-	0.15	220	113	1.00	157	+40.1%
$M/G/1, quadratic$	0.0315	-	0.05	280	119	$-0.33\times10^{-5**}$	148	+89.2%
$GI/G/1$	0.0693	0.90	0.50	135	105	-	57	+136.8%
$GI/G/1, linear$	0.0592	0.10	0.10	163	140	-	43	+279.1%
$GI/G/1, exponential$	0.0347	0.10	0.10	80	123	0.82	49	+63.3%
$GI/G/1, elliptic$	0.0630	0.13	0.13	82	102	-	59	+39.0%
$GI/G/1, gaussian$	0.0359	0.10	0.10	97	120	1.30	50	+94.0%
$GI/G/1, quadratic$	0.0432	0.10	0.10	106	120	$-0.33\times10^{-5**}$	50	+112.0%
*The variable has no influence on the value of the Theil inequality coefficient								
**The minimum allowed value for α								
***$k_{j,t.v.}$ is the theoretical value for k_j for a certain queueing model								

Table 3: The best Theil coefficients with corresponding parameter settings

A first obvious observation concerns the relatively good performance of the state dependent queueing models: in general the state dependent queueing models outperform the non-state dependent models (only in the linear state dependent case, worse results are to be notified). Second, all the $GI/G/1$ models (whether state dependent or not) are of no real practical use since (i) the Theil inequality coefficients are relatively high and (ii) the ideal value of k_j is mostly beyond an acceptable value. A remark should be made upon the coefficient of variation of the service time in the $M/G/1$ case with linear state dependency: it appears that the mentioned coefficient does not influence the Theil coefficient for this specific parameter setting (i.e. $k_j \to \infty$). Furthermore, the Theil coefficient is similar for both the $M/G/1$ model with linear state dependency and the $M/M/1$ model with linear state dependency. A proof for both observations can be found in appendix D.

	$Theil$	$Bias$	Var	Cov
$M/M/1$	0.0517	0.0703	0.8969	0.0329
$M/M/1, linear$	0.0556	0.0633	08430	0.0937
$M/M/1, exponential$	0.0112	0.0214	0.3447	0.6339
$M/M/1, elliptic$	0.0395	0.8077	0.0193	0.1730
$M/M/1, gaussian$	0.0169	0.6831	0.0370	0.2799
$M/M/1, quadratic$	0.0273	0.0197	0.6480	0.3323
$M/G/1$	0.0510	0.2128	0.7555	0.0317
$M/G/1, linear$	0.0556	0.0633	0.8430	0.0937
$M/G/1, exponential$	0.0130	0.0235	0.2505	0.7260
$M/G/1, elliptic$	0.0231	0.1248	0.3549	0.5203
$M/G/1, gaussian$	0.0097	0.0163	0.0613	0.9224
$M/G/1, quadratic$	0.0315	0.3651	0.3619	0.2732
$GI/G/1$	0.0693	0.1031	0.8744	0.0225
$GI/G/1, linear$	0.0592	0.0622	0.8537	0.0841
$GI/G/1, exponential$	0.0347	0.0530	0.5830	0.3640
$GI/G/1, elliptic$	0.0630	0.0295	0.9071	0.0634
$GI/G/1, gaussian$	0.0359	0.9272	0.0284	0.0444
$GI/G/1, quadratic$	0.0432	0.8710	0.0003	0.1287

Table 4: Theil decomposition for the best parameter values

In order to verify which model performs best, the set of queueing models is reduced to those with an acceptable difference for the k_j values (each time, the k_j value that results in the best Theil coefficient is compared with $k_{j,t.v.}$, i.e. the value for the maximum density that is given by the respective queueing model). The maximum allowed difference is set to 40 %. For both the $M/M/1$ and the $M/G/1$ queueing model the maximum allowed difference of 40 % can be caused by e.g. an underestimation of the maximum flow $q_{\max}$ by 16.67 procent (i.e. $\frac{1}{6}$) and an overestimation of the free flow speed v_f by 16.67 %. This is easy to understand, because the occurence of the previously mentioned under- and

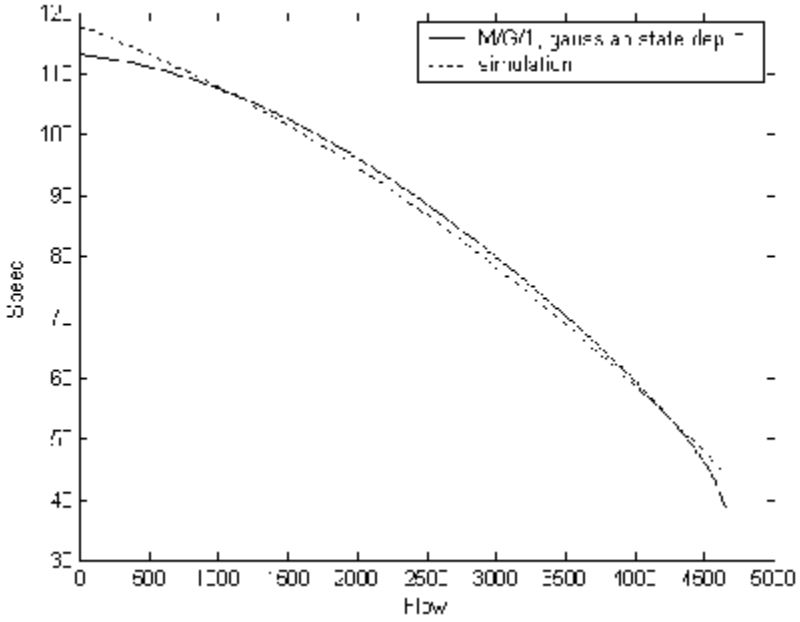

Figure 7: The M/G/1 queueing model with gaussian state dependency vs the simulation

overestimation in e.g. the formula for k_j in the $M/M/1$ queueing model (see equation 5) would lead to

$$k_j = \frac{4 \times (\frac{q_{\max}}{0.83})}{\frac{v_f}{1.17}} = \frac{1.17}{0.83} \times \frac{4q_{\max}}{v_f} = 1.401 \times k_{j,t.v.}$$

Self evidently, for the $M/G/1$ queueing model, the same methodology can be applied.

In this way only 9 queueing models remain: $M/M/1$, $M/M/1$ (exponential), $M/M/1$ (elliptic), $M/M/1$ (gaussian), $M/G/1$, $M/G/1$ (exponential), $M/G/1$ (elliptic), $M/G/1$ (gaussian) and $GI/G/1$ (elliptic).

A next step concerns the Theil inequality coefficient and its decomposition: ideally the covariance proportion (i.e. the unsystematic errors) should be 1. If those models are deleted that have a covariance proportion significantly different from 1, only one model remains: the $M/G/1$ with gaussian state dependency. Since the $M/G/1$ queueing model with gaussian state dependency has both the lowest Theil inequality coefficient and the highest covariance proportion, this model can be considered as the most accurate approximation of real-life situation. The conclusion that queueing models with a Poisson distributed arrival process perform better than models where arrival times follow a general distri-

bution, is most likely the direct result of the assumed exponential interarrival process of the simulation. However, as the Kruskal Wallis test has shown the effectivity of the exponential distribution to approximate the real life interarrival process (see paragraph 2.3), the choice for this distribution can not be considered as arbitrary.

4 Conclusion

In this paper we tried to verify which combination of queueing model and state dependent function best fits a reality based traffic-flow simulation. Testing different queueing models and implementing previously reported and new state dependent functions revealed the poor performance of the $GI/G/1$ model and showed the inaccuracy of the linear state dependent function. One queueing model appeared to outperform the other models: the $M/G/1$ queueing model with gaussian state dependency.

Different future research topics can be examined based on the used methodology. Is it useful to distinguish between the different days (Monday, Tuesday, etc.) and the congested and non-congested traffic situation? How would the addition of on-ramps and off-ramps influence the results described in this paper?

References

[1] R.L. Tatham J.F. Hair E. Anderson and W.C. Black. *Multivariate Data Analysis.* Prentice Hall International, New Jersey, 1998.

[2] C.F. Daganzo. *Fundamentals of Transportation and Traffic Operations.* Elsevier Science, 1997.

[3] A.C. Davison and D.V. Hinkley. *Bootstrap Methods and their Application.* Cambridge University Press, Cambridge, 1998.

[4] D. Heidemann. A queueing theory approach to speed-flow-density relationships. In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France, 1996. Transporation and traffic theory.

[5] R. Jain and J. MacGregor Smith. Modeling vehicular traffic flow using M/G/C/C state dependent queueing models. *Transportation Science*, 31:324–336, 1997.

[6] S. Lee M.A. Johnson and J.R. Wilson. Experimental evaluation of a procedure for estimating nonhomogeneous poisson processes having cyclic behavior. *ORSA J. Comput.*, pages 356–368, 1994.

[7] E.P.C. Kao and S.L. Chang. Modeling time-dependent arrivals to service systems: A case in using a piecewise-polynomial rate function in an nonhomogeneous poisson process. *Management Science*, 34:1367–1379, 1988.

[8] W. Kraemer and M. Lagenbach-Belz. Approximate formulae for the delay in the queueing system GI/GI/1. In *Congressbook of the Eight International Teletraffic Congress*, pages 235–1/8, Melbourne, 1976.

[9] W.H. Kruskal and W.A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47:583–634, 1952.

[10] A.M. Law and W.D. Kelton. *Simulation Modeling and Analysis*. McGraw Hill, 2000.

[11] L. Leemis. Nonparametric estimation of the cumulative intensity function for a nonhomogeneous poisson process. *Management Science*, 37:886–900, 1991.

[12] Michael Z.F. Li. The role of speed-flow relationship in congestion pricing implementation with an application to singapore. *Transportation Research Part B*, 36:731–754, 2002.

[13] B. De Borger S. Proost. *Reforming Transport Pricing in the European Union: A Modelling Approach*. Edward Elgar, 2001.

[14] N. Vandaele T. Van Woensel A. Verbruggen. A queueing based traffic flow model. *Transportation Research D*, 5(2):121–135, February 2000.

[15] K. Preston White. Simulating a nonstationary poisson process using bivariate thinning: The case of "typical weekday" arrivals at a consumer electronics store. *Proc. 199 Winter Simulation Conference*, pages 458–461, 1999.

[16] T. Van Woensel. *Models for Uninterrupted Traffic Flows A Queueing Approach*. Phd Thesis, Department of Applied Economics, UFSIA, Antwerp, 2003.

[17] T. Van Woensel and N. Vandaele. Empirical validation of a queueing approach to uninterrupted traffic flows. 2002.

A Iterative method: the $GI/G/1$ speed-flow relationship

Van Woensel [16] proposed an iterative method for the $GI/G/1$ queueing model in order to obtain a matching speed for every flow.

For each effective speed v^* going from 1 km/hr to v_f km/hr:

Step 1: Initialization:

Set q equal to an initial value: $q = q_0$

Step 2: Calculation

Calculate the speed v_{calc} with e.g. the Kraemer-Lagenbach-Belz approximations (Kraemer and Lagenbach-Belz [8])

Step 3: Evaluation
If $v_{calc} - v^* > d$ *(with* d *a predefined maximum error)*
Then
Set $q_0 = q_0 + 1$ *and return to Step 2*
Else
Continue the loop for the next speed s^*.

B Inversion of the simulated speed-flow equation

Proof. From equation 4 it is known that

$$\begin{aligned} v &= \alpha + \beta q + \gamma q^2 \quad \text{with } \alpha > 0, \beta < 0, \gamma < 0 \\ \Rightarrow\ & v = -\left(\sqrt{-\gamma}q - \frac{\beta}{2\sqrt{-\gamma}}\right)^2 - \frac{\beta^2}{4\gamma} + \alpha \\ \Rightarrow\ & -\left(v + \frac{\beta^2}{4\gamma} - \alpha\right) = \left(\sqrt{-\gamma}q - \frac{\beta}{2\sqrt{-\gamma}}\right)^2 \\ \Rightarrow\ & \frac{\sqrt{-\left(v + \frac{\beta^2}{4\gamma} - \alpha\right)} + \frac{\beta}{2\sqrt{-\gamma}}}{\sqrt{-\gamma}} = q \quad \text{for } v \leq \alpha - \frac{\beta^2}{4\gamma} \end{aligned}$$

As a result, equation 4 can only be inverted for v smaller than or equal to 129,022 km/hr. ■

C Parameter restrictions quadratic function

Proof.

$$(v_f)_q = v_f + \alpha q^2 + \beta q$$

Since $(v_f)_{q_{\max}}$ should tend to zero, we have

$$\begin{aligned} \frac{-\beta - \sqrt{\beta^2 - 4\alpha v_f}}{2\alpha} &= q_{\max} \\ \Rightarrow \beta &= \frac{-\alpha q_{\max}^2 - v_f}{q_{\max}} \end{aligned} \tag{9}$$

A state dependent function should be decreasing, and therefore

$$\frac{\partial (v_f)_q}{\partial q} < 0$$
$$\Rightarrow 2\alpha q + \beta < 0$$

For a quadratic function, we then have

$$\Rightarrow 2\alpha q_{\max} + \beta < 0 \text{ with } \beta < 0 \tag{10}$$

Formula 9 is subject to the non-negativity condition:

$$\frac{\beta^2}{4v_f} \geq \alpha$$
$$\Rightarrow \beta \leq -\sqrt{4\alpha v_f} \ \left(\text{and } \beta \geq \sqrt{4\alpha v_f}\right) \text{ for } \alpha \geq 0^1 \tag{11}$$

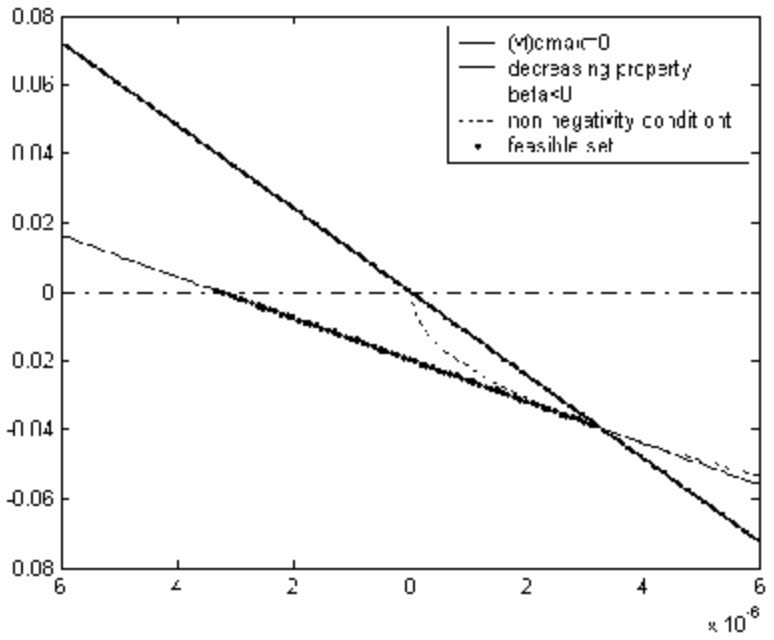

Figure 8: Feasible region for the quadratic state dependent function parameters

Figure 8 shows the discussed restrictions for the parameters of the quadratic state dependent function.The first coordinate (i.e. the coordinate with the lowest α) of the feasible region is the intersection of the $(v_f)\, q_{\max} = 0$ curve and the $\beta = 0$ curve.

$$\frac{-\alpha q_{\max}^2 - v_f}{q_{\max}} = 0$$
$$\iff \alpha = \frac{-v_f}{q_{\max}^2}$$

The last coordinate (i.e. the coordinate with the highest α) of the feasible region is the intersection of the $(v_f)_{q_{\max}} = 0$ curve, the $2\alpha q + \beta = 0$ curve and the $\beta = -\sqrt{4\alpha v_f}$ curve.

$$\begin{aligned} -2\alpha q_{\max} &= \frac{-\alpha q_{\max}^2 - v_f}{q_{\max}} = \sqrt{4\alpha v_f} \\ \iff \alpha &= \frac{v_f}{q_{\max}^2} \end{aligned}$$

Therefore, the original three conditions are now reduced to one condition:

$$\beta = \frac{-\alpha q_{\max}^2 - v_f}{q_{\max}} \; for \; \alpha \in \left[-\frac{v_f}{q_{\max}^2}, \frac{v_f}{q_{\max}^2} \right]$$

■

D The case for $k_j \to \infty$

This section will prove that for $k_j \to \infty$ the $M/M/1$ and $M/G/1$ queueing models with linear state dependency are similar. Moreover, the latter queueing model will be independent of the coeffcient of service variation.

Proof. For the $M/M/1$ queueing model, the dscending part of the speed v is explicitely written as a function of the flow q as follows (Van Woensel [16]):

$$v = \frac{k_j \times v_f + \sqrt{k_j \times v_f \left(k_j \times v_f - 4q\right)}}{2k_j} \tag{12}$$

Implementing the linear state dependent function (2) in equation 12 results in the $M/M/1$ queueing model with linear state dependency:

$$v = \frac{k_j \frac{v_f(q_{\max}+1-q)}{q_{\max}} + \sqrt{k_j \frac{v_f(q_{\max}+1-q)}{q_{\max}} \times \left(k_j \frac{v_f(q_{\max}+1-q)}{q_{\max}} - 4q\right)}}{2k_j}$$

Application of l'Hopital's rule gives:

$$\overset{H}{\Rightarrow} \lim_{k_j \to \infty} v = \lim_{k_j \to \infty} \frac{\frac{\partial \left[k_j \frac{v_f (q_{\max}+1-q)}{q_{\max}} + \sqrt{k_j \frac{v_f (q_{\max}+1-q)}{q_{\max}} \times \left(k_j \frac{v_f (q_{\max}+1-q)}{q_{\max}} - 4q \right)} \right]}{\partial k_j}}{\frac{\partial 2k_j}{\partial k_j}}$$

$$\Rightarrow \quad v = \frac{v_f \left(q_{\max} + 1 - q \right)}{q_{\max}} \quad \text{for } k_j \to \infty \tag{13}$$

For the $M/G/1$ queueing model, the descending part of the speed v is explicitely written as a function of the flow q as follows (Van Woensel [16]):

$$v = \frac{q(c_s^2 - 1) + 2k_j \times v_f + \sqrt{\left[q \left(c_s^2 - 1 \right) - 2k_j \times v_f \right]^2 - 16k_j \times v_f \times q}}{4 \times k_j} \tag{14}$$

Implementing the linear state dependent function (2) in equation (14) results in the $M/G/1$ queueing model with linear state dependency:

$$v = \frac{q(c_s^2 - 1) + 2k_j \times \frac{v_f (q_{\max}+1-q)}{q_{\max}}}{4 \times k_j} + \frac{\sqrt{\left[q \left(c_s^2 - 1 \right) - 2k_j \times \frac{v_f (q_{\max}+1-q)}{q_{\max}} \right]^2 - 16k_j \times \frac{v_f (q_{\max}+1-q)}{q_{\max}} \times q}}{4 \times k_j}$$

Application of l'Hopital's rule gives:

$$\overset{H}{\Rightarrow} \lim_{k_j \to \infty} v = \lim_{k_j \to \infty} \frac{\frac{\partial \left[q(c_s^2-1) + 2k_j \times \frac{v_f (q_{\max}+1-q)}{q_{\max}} + \sqrt{\left[q(c_s^2-1) - 2k_j \times \frac{v_f (q_{\max}+1-q)}{q_{\max}} \right]^2 - 16k_j \times \frac{v_f (q_{\max}+1-q)}{q_{\max}} \times q} \right]}{\partial k_j}}{\frac{\partial 4k_j}{\partial k_j}}$$

$$\Rightarrow v = \frac{v_f \left(q_{\max} + 1 - q \right)}{q_{\max}} \quad for\ k_j \to \infty \tag{15}$$

Obviously, equations 13 and 15 are identical, and apparantly both the $M/M/1$ and $M/G/1$ queueing models with linear state dependency are equal to the linear state dependency function itself. Moreover, for the $M/G/1$ queueing model with linear state dependency, the coefficient of service variation is eliminated in the speed/flow equation. ■