

NISS

Calibrating and Validating Deterministic Traffic Models: Application to the HCM Control Delay at Signalized Intersections

Rui Miguel Batista Paulo, Jiong Lin,
Nagui M. Roupail, and Jerome Sacks

Technical Report Number 144
July 2004

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

Acknowledgment and Disclaimer

This material is based upon work supported by the National Science Foundation under Grant No. DMS-00739520

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Calibrating and Validating Deterministic Traffic Models: Application to the HCM Control Delay at Signalized Intersections

Rui Miguel Batista Paulo¹, Jiong Lin², Nagui M. Rouphail², Jerome Sacks³

¹University of Bristol, ²North Carolina State University, ³National Institute of Statistical Sciences

Submitted for Consideration for Publication and Presentation at the 84th Annual Meeting of the
Transportation Research Board, Washington, DC

July 29, 2004

ABSTRACT

Calibrating and validating traffic models is a process that depends on field data that are often limited, but essential for determining inputs to the model and for assessing its reliability. A quantification and systemization of the calibration/validation process exposes statistical issues inherent in the use of such data. Formalization of the calibration/validation process leads naturally to the use of Bayesian methodology for assessing uncertainties in model predictions arising from a multiplicity of sources especially statistical variability in estimating and calibrating input parameters and model discrepancy. In an earlier paper the general problem was elucidated; in this paper we will carry out the full calibration/validation process in the context of a widely used deterministic traffic model, namely the Highway Capacity Manual (HCM) model for control delay at signalized intersection approaches. In particular we are able to assess the reliability of the model through quantification of the uncertainty in estimation of model parameters, predictions of model delay and predictions using data-adjustments to the model. While the methods are described in a specific context they can be used generally, inhibited at times by computational burdens that must be overcome.

Key words and phrases: Bayesian Analysis; Posterior Distribution; Model Calibration; Model Validation; Highway Capacity Manual; Control Delay.

0. Introduction

Calibrating and validating traffic models are inherently complex processes that are commonly treated informally and through a mix of ad hoc methodologies. Explicit ingredients of such a process are field data that are often limited and expensive to acquire but essential for determining inputs to the model and for assessing the reliability of the model. A quantification and systemization of the calibration/validation process exposes statistical issues inherent in the use of such data for assessing the validity of a model. In previous work by Bayarri et al (1) these issues were elucidated and a methodology advanced to address them.

As described in (1), a clear statement of what “validation” means is rarely set forward. Usually, the question is put as “does the model faithfully represent reality?” But, the answer to this question is simple: no, models are not perfect. But models can make useful, reliable predictions in particular settings; they may be useful for some purposes, useless for others. We can state this more formally as

$$\Pr[| \text{"reality"} - \text{prediction} | < \delta] > \alpha . \quad \text{Eq(1)}$$

Here we must specify δ = tolerable difference (how close) and α = level of assurance (how certain), say what is meant by “reality”, define prediction and, ultimately, compute the probabilities involved.

What we mean by “reality” is, operationally, a feasible measure of actual performance of a particular network. For the cases we examine here control delay will be the key measure. To compare actual performance with prediction from the model will require access to field data and model output that relate to the performance measure. As noted in (1) little attention has been

previously paid to the characterization of the *uncertainty* in model *inputs* and resulting effect on calibration and validation.

In (1) a framework that focuses on analyses that combine calibration and validation, accounting for a multiplicity of uncertainties, was only partially fleshed out in the context of the traffic simulation model, CORSIM. The focus of this paper is to fully treat specific algebraic mathematical models used in the Highway Capacity Manual (HCM). Our comments and methods are general but we will utilize as a test bed, a traffic network in Chicago and the specific control delay models that appear in the signalized intersection analyses of the HCM2000 (2). A detailed account of the network, the data and their use in a signal study is in Sacks et al. (3).

Inputs (including model parameters) to models, such as control delay in the HCM2000, come in various forms. Some, such as geometric inputs (lane widths, bus stops, parking, etc.) are readily supplied by accurate measurement or through documented sources. Others can be obtained by calibration that is, through the use of field data. These latter can be classified as

- (A) parameters that can be directly estimated, perhaps with error, from field data (e.g., demand volumes)
- (B) parameters not directly measurable (e.g., driver type)
- (C) tuning parameters (e.g., ideal saturation flow rate) that are not “real” but are required by the model.

Some parameters in (B) and (C) may be set to default values and thus removed from the calibration process. The remaining parameters in (B) and (C) are typically treated by adjusting them until the model output seems to ‘fit’ the field data, data from the real traffic system being

modeled. Sometimes the adjustment is informal; sometimes more formal by optimizing the parameters, using least squares or other optimization tools to minimize the difference between features of the field data and predictions from the model (see for example Jha et al. (4) and Hourdakis et al. (5)). Such approaches to calibration, as pointed out in (1), may have unpleasant consequences. For example, model imperfections may be masked by “over-tuning” which can then lead to highly inaccurate predictions in later usage of the model.

As discussed in (1) Bayesian analysis (described there and below as well) provides an attractive path to simultaneously calibrate the parameters of type (A), (B) and (C) and deal with the possible presence of model bias. *Calibration and assessment of validity can be done in one combined analysis.*

Bayesian analysis determines the posterior (or summary) distribution of model parameters and inputs, given the observed field data. The resulting distribution will then reflect the actual uncertainty in the parameters and inputs, adjudicate between the tuning parameters and the possible model bias thereby providing resistance to over-tuning, and quantify and assess the model bias. In short, a Bayesian analysis can provide a complete answer to Equation (1).

1. Test Bed Models, Network and Data

We consider two forms of the HCM control delay model at the signalized intersections that were investigated. In the first Model (M1), and based on the empirical data for the first intersection, there were no queues that spilled over from one cycle to the next. Therefore, only the first two

control delay components (d_1 and d_2) in the HCM2000 Chapter 16 (2) were considered. That restriction was removed in the second model (M2) (for the second intersection) where the initial queue delay term d_3 was added. This was due to the presence of downstream blockage in some cycles, resulting in an apparent drop in the saturation flow rate and the occurrence of an initial queue at the end of the green phase. In both cases, the unit of analysis for model calibration and validation was a single cycle, as opposed to a peak 15 minute period or a full hour.

The intersections where the models are applied are sketched in Figure 1. They come from a Chicago traffic network connecting an important freeway and major arterials with the central business district. The intersections we focus on are denoted as SB Wells-Grand during the morning peak period (M1) and NB LaSalle-Ontario during the evening peak period (M2).

[FIGURE 1 GOES HERE]

2. Model M1 at SB Wells-Grand (2 through lanes)

Model Inputs

T = duration of analysis period, CL = cycle length, V = volume, P = proportion of vehicles arriving on green, g = effective green time ($g = G + E - SL$ = displayed green time + used yellow and all red time – start up lost time), s = adjusted lane group saturation flow rate per lane (i.e. after accounting for all factors such as lane width, parking, grade, etc), k = incremental delay factor (normally defaulted to 0.50 for fixed time signals), I = upstream metering/filtering factor (normally defaulted to 1.0 but considered variable due to the short block length at the site), f_{PA} = supplemental delay adjustment factor for platoon arrivals.

Let $c = 2s \cdot g/CL =$ lane group capacity; $X=V/c$, $PF = \frac{(1-P)f_{PA}}{1-g/CL}$. Then there are two

components to model delay, y^M :

$$d_1(V, P, s) = \frac{0.5CL(1-g/CL)^2}{1 - [\min(1, X)g/CL]} \cdot PF$$

$$d_2(V, P, s, kI) = 900T \left[(X-1) + \sqrt{(X-1)^2 + \frac{8kIX}{cT}} \right] \quad \text{Eq (2)}$$

$$y^M(V, P, s, kI) = d_1(V, P, s) + d_2(V, P, s, kI)$$

Several of these inputs are known and fixed: $CL = 75$ sec, $T = CL/3600$, while $g = 31$ sec is obtained from the field. Because k and I appear only as the product kI , we will only work with this product. Both kI and s are to be taken as tuning parameters. From cycle to cycle the only inputs that change are V , P and f_{PA} (as a consequence X and PF also change every cycle).

Data

Video cameras perched atop nearby buildings were placed that recorded all vehicles passing through the intersections over one-hour periods during the morning and evening rush hours. The SB Wells-Grand movement was analyzed for the morning rush hour; the NB LaSalle-Ontario movement was analyzed with the evening rush-hour data.

There are 48 cycles each of 75 seconds (under fixed-time signal control). The video data were examined and for each cycle i , ($1 \leq i \leq 48$), V_i , P_i were extracted and (field) delays, $y^F(V_i, P_i)$, measured in accordance with the procedure in Appendix A of HCM Chapter 16 (2). For the model we can produce $y^M(V, P, s, kI)$ for any values of the four inputs. The field data, of course,

do not depend on s, kI – nature has no need for tuning parameters. A summary of the field measurements for the two approaches is provided in Tables 1 and 2, respectively.

[TABLES 1 and 2 GO HERE]

Analysis of Model M1

The statistical structure connecting the field data, the model, randomness in field delay measurements, the inputs and tuning parameters, and possible model bias is given as follows:

Let x_i be the vector of counts (V_i, P_i) ; let u = the vector of tuning parameters (s, kI) . Then for each cycle i

$$y^F(x_i) = y^M(x_i, u) + b_u(x_i) + \varepsilon_i \quad \text{Eq (3)}$$

where: y^F, y^M denote field and model delays, b the model bias, ε the variability in measured delays. The interpretation of Eq(3) is that “reality” = $E[y^F]$ = model + bias.

The ε_i 's are assumed to be independent normal random variables with mean 0 and precision (=1/variance) λ^F . From Eq (3) it is clear that the tuning parameters u and the bias b are intertwined (confounded) and the way we treat the bias later reflects that condition.

The bias is an unknown function and treating it requires non-standard techniques. There is no rationale leading to specific forms for b and we therefore adopt a non-parametric Bayes approach. The approach, taken (and discussed) in Bayarri et al. (6) (see also Santner et al. (7)), is described in Appendix I and introduces a parameterized family of (prior) distributions for b .

The parameters of the family, described in the appendix and denoted by $\lambda^b, \theta, \beta_1, \dots, \beta_J$, introduce a level of complexity to the Bayes analysis we now describe.

The basics of Bayesian analysis consist of taking the data density, given the unknown parameters, including the unknown b , and combining it, via Bayes theorem, with a prior distribution of the parameters. The parameters are $s, kI, \lambda^F, b, \lambda^b, \theta, \beta_1, \dots, \beta_J$. Denote the collection of these parameters by γ . Then given γ , the likelihood of the data, $y^{Data} = \{y^F(V_i, P_i); i=1, \dots, 48\}$, is multivariate normal with mean and covariance depending on γ . Denote this likelihood as $f(y^{Data} | \gamma)$. If $\pi(\gamma)$ denotes the prior distribution of the parameters then what we seek is (Bayes formula) the posterior distribution of the parameters given the data

$$\pi^{post}(\gamma | y^{Data}) = \frac{f(y^{Data} | \gamma) \pi(\gamma)}{\int f(y^{Data} | \gamma) \pi(\gamma) d\gamma} \quad \text{Eq(4)}$$

The posterior distribution captures the uncertainty present about γ after the data are collected. Uncertainty resulting from model predictions can then be assessed by treating $\pi^{post}(\gamma | y^{Data})$ as the "random input distribution" for the model, and making repeated evaluations of the model, initialized by draws from this distribution.

Choice of the Prior Distribution of γ

s is uniform on an interval $[s_0, s_1]$; kI is uniform on an interval $[kI_0, kI_1]$; λ^F has an exponential density with mean 0.172, λ^b has an exponential density with mean 0.394. The priors for λ^F, λ^b are constructed using the calculations described in Appendix II. For $\theta, \beta_1, \dots, \beta_J$ we adopt a computational shortcut by setting $\theta = 0$ and fixing β_1, \dots, β_J in a manner described in Appendix II. In effect, these parameters are removed from the Bayesian analysis.

For fixed $\lambda^b, \theta, \beta_1, \dots, \beta_J, b$ has a prior distribution as described in Appendix I. The prior for γ is taken as the product of these five prior densities:

$$\pi(\gamma) = \pi^{prior}(s) \times \pi^{prior}(kI) \times \pi^{prior}(\lambda^F) \times \pi^{prior}(\lambda^b) \times \pi^{prior}(b | \lambda^b)$$

To treat $\pi^{post}(\gamma | y^{Data})$ we first use laws of conditional probability and write

$$\pi^{post}(\gamma | y^{data}) = \pi^{post}(b | y^{data}, s, kI, \lambda^F, \lambda^b, \theta, \beta_1, \dots, \beta_J) \times \pi^{post}(s, kI, \lambda^F, \lambda^b, \theta, \beta_1, \dots, \beta_J | y^{data}) \quad \text{Eq(5)}$$

As described in Appendix I the first term on the right side of Eq(5) is readily obtained through properties of the multivariate normal distribution (MVN). But, even with fixed values for $\theta, \beta_1, \dots, \beta_J$, closed form expressions for the needed posterior distribution of the second term are not available. Consequently, posterior probabilities must be obtained by numerical methods or by simulation. The most commonly used Bayesian computational technique is called Markov Chain Monte Carlo (MCMC); see Robert and Casella (8) for a thorough description of this simulation method. Fixing β_1, \dots, β_J makes the MCMC computations more feasible.

The MCMC process will produce a sample of size, say N , from the posterior distribution

$$\pi^{post}(s, kI, \lambda^F, \lambda^b | y^{Data}) \quad (\text{we have removed } \theta, \beta_1, \dots, \beta_J \text{ because they are now fixed}). \quad \text{For the } n^{th}$$

element of the sample we can make a draw from $\pi^{post}(b | s_n, kI_n, \lambda_n^F, \lambda_n^b, y^{Data})$ (described in

Appendix I) to get a value for b_n . Adjoining this to the n^{th} element of the MCMC sample will

produce a sample of size N from $\pi^{post}(\gamma | y^{Data})$. How we use this sample is explored below in

the Results section. In our examples we have $N=10,000$.

Results

For the choice of prior distributions for s and kI set $[s_0, s_1] = [1400, 2000]$ and $[kI_0, kI_1] = [0.1,$

0.5]. The sample of size N from the posterior distribution of γ described above produces the

histograms in Figure 2 for the (marginal) posterior distributions of the two tuning parameters. The histograms for the precision parameters may also be obtained but they are of less interest.

[FIGURE 2 GOES HERE]

The histogram for s (recall that s is the *adjusted* not the *ideal* saturation flow rate) indicates that the median is about 1750 vph/lane and that [1650, 1920] is a 90% uncertainty interval for s . The histogram for kI indicates that nothing is learned about kI from the data. This suggests that virtually any value in [0.1, 0.5] is an acceptable choice for kI . We choose the posterior means $s_{mean} = 1786$, $kI_{mean} = 0.305$.

These values for the tuning parameters are very similar to those estimated using conventional calibration methods. For example, Figure 3 shows the relationship of the Mean Square Error (MSE) between the cycle-by-cycle field and model delays against various values of s and kI . The best (smallest MSE) results are achieved in the range of 1650-1850 vph/lane for s and, in that regime, the errors are unaffected by the value of kI .

[FIGURE 3 GOES HERE]

We turn next to the question of predicting reality $= y^M + b$ as noted following Eq(3) and of estimating the bias, both of which are connected to the issue of validation. Start with the MCMC generated N draws, $\{\gamma_n\}$, from $\pi^{post}(\gamma | y^{Data})$, evaluate $y^M(V, P, s_n, kI_n)$ and obtain draws $b_n(V, P)$ from $\pi^{post}(b | s_n, kI_n, \lambda_n^F, \lambda_n^b, y^{Data})$. Then, $\{y^M(V, P, s_n, kI_n) + b_n(V, P)\}$ is a sample of the posterior distribution of reality at V, P given the data and the prior distributions. We then take as the so-termed *bias-corrected prediction of reality* at (V, P)

$$\hat{y}^R(V, P) = \frac{1}{N} \sum_{n=1}^N [y^M(V, P, s_n, kI_n) + b_n(V, P)]. \quad \text{Eq(6)}$$

If s_{mean}, kI_{mean} are the posterior means of the tuning parameters we define the *pure model prediction of reality* by evaluating the model with these values of the tuning parameters plugged in: $y^M(V, P, s_{mean}, kI_{mean})$. The bias is then estimated by subtraction:

$$\widehat{b}(V, P) = \widehat{y}^R(V, P) - y^M(V, P, s_{mean}, kI_{mean}) \quad \text{Eq(7)}$$

Uncertainty bounds for $\widehat{b}(V, P)$ are obtained as follows: $\{b_n(V, P)\}$ is a sample from the posterior distribution of bias given the data and the prior distributions on γ . If we choose δ_b so that 90% of $\{b_n(V, P) - \widehat{b}(V, P)\}$ are between $+\delta_b$ and $-\delta_b$ then

$$\Pr\left[|bias(V, P) - \widehat{b}(V, P)| < \delta_b\right] = .90 \text{ for each } V, P \quad \text{Eq(8)}$$

In Figure 4a we plot $\widehat{b}(V, P)$ with uncertainty bounds for all V, P ; $\delta_b \sim 1.2$. We do so for a single $P = .20$ since $\widehat{b}(V, P)$ changes little, if at all, for different values of P in this example.

[FIGURE 4 GOES HERE]

Uncertainty bounds for predictions of reality are obtained similarly: For uncertainty bounds on pure model predictions find δ_M such that 90% of the MCMC sample satisfies

$$\left|y^M(V, P, s_n, kI_n) + b_n(V, P) - y^M(V, P, s_{mean}, kI_{mean})\right| \leq \delta_M \quad \text{Eq(9)}$$

We then get a precise statement of Eq(1):

$$\Pr\left[|reality - pure\ model\ prediction| < \delta_M\right] = .90 \quad \text{Eq(10)}$$

For the example at hand Figure 4c plots pure model predictions; $\delta_M \sim 1.5$.

If we use the bias-corrected prediction $\hat{y}^R(V, P)$ instead of $y^M(V, P, s_{mean}, kI_{mean})$ and find δ_R so that 90% of the MCMC sample satisfies

$$|y^M(V, P, s_n, kI_n) + b_n(V, P) - \hat{y}^R(V, P)| \leq \delta_R \quad \text{Eq(11)}$$

we get the counterpart of Eq(10):

$$\Pr[|\text{reality} - \text{bias-corrected prediction}| < \delta_R] = .90 \quad \text{Eq(12)}$$

For the example at hand Figure 4b plots bias-corrected predictions; $\delta_R \sim 1.3$.

Because the bias is near zero it is unsurprising that δ_M and δ_R are the same. Because the bias-corrected predictions are best predictions we will always have $\delta_R < \delta_M$. If $\delta_M \sim 1.5$ is small enough and 0.90 is the desired level of assurance then Eq(10) can be interpreted as stating that model M1 is valid for the Wells-Grand study lane group. Thus, we have in a single analysis treated the calibration/tuning (Figure 2) and the assessment of model validity (Figure 4).

We could have proceeded in a more “traditional” way of finding values of the tuning parameters by minimizing the squared error differences between model and field data, and then estimating the bias by fitting a function to the differences between field and model (for example, as in Figure 3 and Appendix II). Because the bias is nearly zero we would be led to similar estimates as found by the Bayes analysis but with no assessment of uncertainty. If the bias is not zero such an analysis could be misleading and, moreover, we would be unable to assess the utility of the model as we can via Eq(10) and Eq(12).

3. Model M2 at NB LaSalle-Ontario (3 through lanes)

The model inputs include those for M1 and, in addition, the input Q = initial queue length at the beginning of the cycles. Based on the HCM model with initial queue,

let $c = 3 s \cdot g/CL$ = capacity; $X=V/c$, $PF = \frac{(1-P)f_{PA}}{1-g/CL}$. Let $t = \min \left\{ T, \frac{Q}{c[1-\min(1, X)]} \right\}$;

$w = 0$ if $t < T$, else $w = 1 - \frac{cT[1-\min(1, X)]}{Q}$. There are three terms to model delay, y^M :

$$d_1(V, P, Q, s) = 0.5CL(1-g/CL) \frac{t}{T} + \frac{0.5CL(1-g/CL)^2}{1-\min(1, X)g/CL} \cdot PF \cdot \frac{T-t}{T}$$

$$d_2(V, P, Q, s, kI) = 900T \left[(X-1) + \sqrt{(X-1)^2 + \frac{8kIX}{cT}} \right] \quad \text{Eq(13)}$$

$$d_3(V, P, Q, s) = \frac{1800Q(1+w)t}{cT}$$

$$y^M(V, P, Q, s, kI) = d_1 + d_2 + d_3$$

The data here are for one-hour during the evening rush hour. The results are as follows:

The histograms in Figure 5 indicate that, unlike M1, the data are informative about kI .

The posterior means are $s_{mean} = 1616$ vph/lane, $kI_{mean} = 0.521$. The mean-square-error (MSE) analysis in this case yielded $s_{mean} = 1550$ vph/lane and $kI_{mean} = 0.50$. The plots in Figure 6 are of the bias (with uncertainty bounds) as a function of volume for four different fixed pairs of values for P, Q . The plots show that bias is present especially at $P = .50$. Figure 7 shows the effect of the bias on how well the model predicts reality for each of these 4 pairs of P, Q . In particular,

$$\Pr \left[\left| \text{reality} - \text{pure model prediction} \right| < 4.0 + \right] = .90$$

when $P=.50$ and $Q = 0$ or 4 . If 4 seconds is an acceptable error then, despite the presence of bias, the model provides usable estimates of delay. But 4 seconds is considerable, based on the range of observed delays in Table 2, and suggests that the model is somewhat unreliable. The reason: possibly because s may be very different when $Q=0$ than when $Q > 0$.

[FIGURES 5, 6 and 7 GO HERE]

The data-adjustment of the model by use of the bias-corrected predictions (Figure 8) produces discrepancies within 2 seconds of reality, a clear improvement. (The dashed line in Figure 8 is the pure-model prediction; its distance from the bias-corrected prediction is noticeable.) This 2 second level of uncertainty cannot be improved upon without more data or different assumptions. Thus, the bias-corrected predictions provide the best answer for the example at hand and overcome the model's shortcomings.

[FIGURE 8 GOES HERE]

4. Comments

The analysis used relied on the capability of producing model output quickly and cheaply. This is the case for the algebraic mathematical models found in the HCM. For more complicated situations where the model may require extensive computation to produce output for a single set of inputs (as is the case for solving differential equations in other engineering contexts) we would need an additional step of approximation of the model through a computer experiment (6).

An open issue and one that demands further research is how to determine the amount of data that must be collected to allow a given assurance for a specific discrepancy in (1).

5. Summary

We have shown how a Bayesian analysis of commonly used mathematical models can produce, in a single analysis, complete answers to the issues of calibration/validation. Though computationally demanding the process described is transferable to any readily computed model for which field data are available. The applications to control delay in the HCM show the possible range of assessments that can be made: quantifying the uncertainty in the model parameters, asserting validity and quantifying the utility of the model and the data-adjusted (bias-corrected) modification. In the first case presented in this paper we were able to assert the validity of the HCM model (predictions within 1.5 seconds of reality with 90% assurance). In the second case a tolerable difference of 2 seconds is possible only if the HCM model is bias-corrected, the model without such adjustment is not reliable enough.

Appendix I: Statistical Structure of b

The values of b at any finite set of points are initially unknown. Assume that these values have a multivariate normal distribution with some mean vector and covariance matrix (a prior distribution in Bayesian parlance). The dependence between values, reflected in the covariance matrix, will also reflect the fact that the values of a “tame” function at nearby points are close (“correlated”). By choosing the mean vectors and covariance matrices appropriately this can be done consistently for every finite set of points on which b is defined. The result is to give the whole function, b , a prior distribution. If we can evaluate b at several points we can learn about the behavior of b at new places through the conditional distribution of the values of b at the new set given the values on the set where the evaluation is made. (The conditional distribution is

again Gaussian but with a mean and covariance updated and dependent on the values observed.) A choice for the means and covariances for this process that has proved to be highly effective in a variety of settings is to take the mean vector to be constant (say θ) and the covariance between two points x and w (this determines all the covariance matrices for the multivariate normal) to be given by

$$Cov^b(w, z) = \frac{1}{\lambda^b} \prod_1^J \exp\left(-\beta_j |w^{(j)} - z^{(j)}|^2\right) \quad \text{Eq(A1)}$$

(J is the dimension of the vectors w, z). When the β 's are small the correlation (the product term in Eq (A1)) will get closer to 1 indicating a high degree of relationship between values of b . Large values of the β 's indicate a low degree of relationship and allow b to oscillate more rapidly. The size of the oscillations is greater if λ^b is small (low precision or high variability). The exponent 2 on the distance between coordinates of w and z in Eq(A1) assures that the function will be very smooth with infinitely many derivatives – we impose that assumption on b . Also, we do not let the distribution of b depend on u , thereby allowing a degree of decoupling of the bias from the tuning.

This approach is typical of so-termed “kriging” methods in common use in spatial and geostatistics literature (Cressie, 1993) (9). In essence, we have described a correlation structure between values of the function at the locations where it is defined; our choice of correlation structure is meant to capture the idea that bias should be smoothly varying

Appendix II. Selection of Parameters for the Prior on b

First fix values of the calibration parameters (say at the midpoint of their prior intervals). Form the differences $y^F - y^M$ for the input data, with the fixed values of the calibration parameters.

Then compute a maximum likelihood estimate of $\lambda^F, \lambda^b, \beta_1, \dots, \beta_J$ based on the MVN of the differences. The MVN will have a mean assumed to be 0 and a covariance matrix that depends on λ^F, λ^b as well as β_1, \dots, β_J . The maximum likelihood estimates of λ^F, λ^b are used to construct priors for λ^F, λ^b ; the estimates for β_1, \dots, β_J are retained for the analysis of b . Any errors in these estimates will have small effect on the ultimate use: prediction of b and prediction of reality.

This characteristic was observed in previous work where the tactic was used (6). The reduction in computational effort is considerable and makes the Bayesian analysis more feasible.

5. Acknowledgments

This research was supported by the U.S. National Science Foundation, Grant DMS-0073952.

REFERENCES

1. Bayarri, M.J., Berger, J.O., Molina, G., Roupail, N.M., Sacks, J. (2003). Assessing uncertainties in traffic simulation: a key component in model calibration and validation. (to appear in *Transportation Record*?) *Technical Report 137, NISS*, www.niss.org/downloadabletechreports.html
2. Transportation Research Board (2000). *Highway Capacity Manual*. National Research Council, Washington, D.C.
3. Sacks, J., Roupail, N., Park, B., Thakuria, P. (2002). Statistically based validation of computer simulation models in traffic operations and management. *J. Transp. Stat.* **5** 1-24
4. Jha, M., Darda, D., Tomer, T. and M. Ben-Akiva (2003), Joint Estimation of Time Dependent O-D and Calibration of Microscopic Simulation Model, Presented at the Workshop on Traffic Modeling: Trends and Challenges, Sitges, Spain, June 2-4.
5. Hourdakis, J. and P. Michalopoulos (2003), Practical Methodology for Microscopic Simulation Calibration, Presented at the Workshop on Traffic Modeling: Trends and Challenges, Sitges, Spain, June 2-4.
6. Bayarri, M.J., Berger, J.O., Higdon, D., Kennedy, M., Kottas, A., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J.C., Lin, C-H., Tu, J. (2002) "A Framework for Validation of Computer Models". *Technical Report 128, NISS*, www.niss.org/downloadabletechreports.html
7. Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments*. Springer Series in Statistics, Springer, New York.
8. Robert, C.P. and Casella, G. (1999), *Monte Carlo Statistical Methods*, Springer-Verlag, New York.
9. Cressie, N. A. (1993). *Statistics for Spatial Data*. J. Wiley, New York.

LIST OF TABLES

Table 1. Summary of Field Measurements at SB Wells-Grand Intersection

Table 2. Summary of Field Measurements at NB LaSalle- Ontario Intersection

Table1. Summary of Field Measurements at SB* Wells-Grand Intersection

Cycle Number	Count per Cycle (V)	Proportion Arriving on Green (P)	Measured Control Delay (y^F)
1	22	0.30	21.76
2	24	0.08	28.02
3	22	0.26	26.67
4	25	0.20	26.82
5	21	0.19	30.61
6	24	0.25	20.85
7	21	0.24	22.29
8	24	0.21	21.83
9	18	0.11	23.38
10	19	0.21	22.15
11	21	0.27	23.31
12	14	0.29	23.21
13	22	0.17	27.75
14	25	0.24	23.28
15	23	0.22	27.86
16	21	0.14	27.10
17	16	0.00	33.73
18	17	0.33	23.52
19	17	0.21	27.29
20	22	0.26	21.27
21	16	0.12	26.85
22	21	0.19	24.34
23	15	0.20	25.60
24	18	0.16	25.78
25	21	0.30	20.74
26	22	0.08	30.87
27	21	0.13	26.69
28	17	0.05	29.84
29	21	0.23	24.34
30	20	0.20	26.18
31	20	0.18	21.58
32	18	0.11	30.80
33	18	0.16	25.67
34	17	0.17	26.54
35	18	0.28	19.81
36	19	0.21	22.72
37	18	0.22	27.69
38	14	0.07	28.56
39	25	0.18	33.22
40	20	0.10	25.66
41	21	0.14	22.80
42	20	0.23	20.26
43	16	0.13	26.85
44	21	0.14	28.74
45	20	0.14	29.88
46	20	0.24	19.86
47	13	0.07	26.60
48	19	0.16	25.09

*through lane group only

Table2. Summary of Field Measurements at NB* LaSalle- Ontario Intersection

Cycle Number	Count per Cycle (V)	Proportion Arriving on Green (P)	Initial Queue (Q)	Measured Control Delay (\bar{y})
1	35	0.63	0	12.11
2	30	0.80	0	4.96
3	34	0.76	0	12.32
4	29	0.55	0	16.54
5	34	0.68	0	16.63
6	36	0.72	0	12.86
7	41	0.71	0	14.75
8	32	0.91	0	1.82
9	34	0.68	0	14.06
10	34	0.88	1	3.42
11	34	0.82	0	5.77
12	28	0.82	0	5.49
13	35	0.80	0	6.70
14	32	0.78	3	8.31
15	31	0.81	0	7.19
16	31	0.71	0	19.66
17	37	0.73	6	12.68
18	37	0.78	0	10.04
19	32	0.88	2	4.00
20	33	0.79	0	7.23
21	29	0.66	0	16.28
22	35	0.71	1	12.75
23	25	0.52	5	27.81
24	38	0.50	7	22.48
25	34	0.62	3	22.35
26	29	0.59	3	20.01
27	31	0.68	5	15.47
28	33	0.70	0	14.48
29	38	0.58	0	24.74
30	30	0.57	7	24.25
31	22	0.50	3	26.45
32	33	0.52	3	26.04
33	32	0.75	3	16.01
34	25	0.44	7	39.62
35	34	0.32	17	39.46
36	33	0.64	6	24.68
37	32	0.56	9	27.71
38	27	0.67	3	26.43
39	34	0.74	1	9.22
40	37	0.73	0	11.19
41	37	0.65	6	16.85
42	31	0.65	2	18.40
43	30	0.73	4	17.54
44	27	0.67	3	19.05
45	34	0.65	5	17.35
46	35	0.77	0	12.37
47	31	0.71	5	15.63
48	29	0.72	0	17.60

*through lane group only

LIST OF FIGURES

Figure 1. Intersection Configurations

Figure 2. Histograms of the posterior densities of the tuning parameters s and kI . (Model M1; SB Wells-Grand)

Figure 3. Control Delay MSE* vs. s and kI for Wells-Grand SB Lane Group

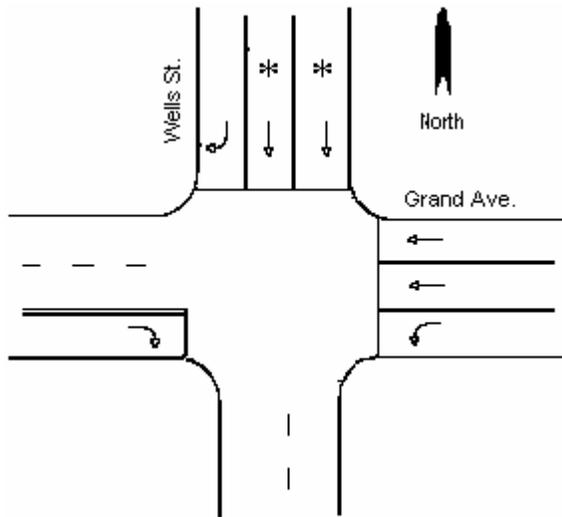
Figure 4. Estimates of bias, bias-corrected predictions and pure-model for Model M1 and SB Wells-Grand. For (a) the solid curve is the estimate of b and the dotted curves form 90% uncertainty bounds. For (b) the solid curve is the bias-corrected prediction of reality; the dotted curves form its 90% uncertainty bounds. and the dashed curve is the pure-model prediction of reality. For (c) the solid curve is the pure-model prediction of reality and the dotted curves form its 90% uncertainty bounds.

Figure 5. Histograms for posterior densities of the tuning parameters s and kI . (Model M2; NB LaSalle-Ontario)

Figure 6. Estimates of bias (Model M2; NB LaSalle-Ontario). The solid curves are estimates of b as a function of volume for four different combinations of P, Q ; the dotted curves form the 90% uncertainty bounds

Figure 7. Pure-model predictions of reality (Model M2; NB LaSalle-Ontario). The solid curves are predictions as a function of volume for four different combinations of P, Q ; the dotted curves form the 90% uncertainty bounds.

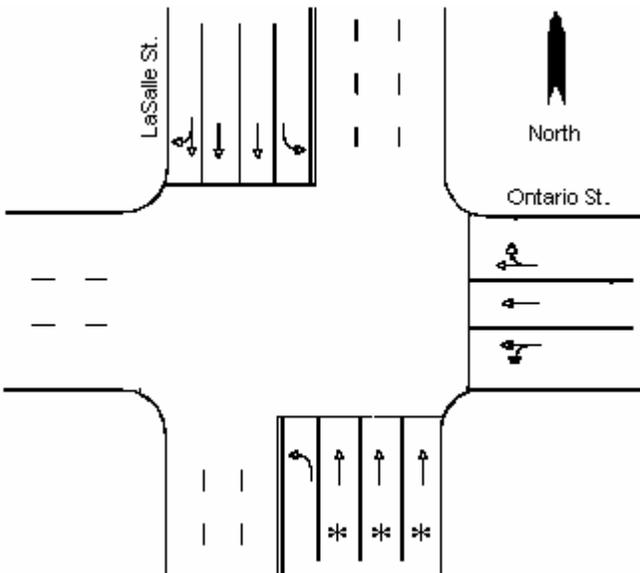
Figure 8. Bias-corrected predictions of reality (Model M2; NB LaSalle-Ontario). The solid curves are predictions as a function of volume for four different combinations of P, Q ; the dotted curves form the 90% uncertainty bounds; the dashed lines are the pure-model predictions..



**Signal Plan
(SB Wells St.)**

Green = 29 sec
 Yellow = 3 sec
 Red = 1 sec
 Cycle Length = 75 sec

a) Wells-Grand Intersection (M1)



**Signal Plan
(NB LaSalle St.)**

Green = 34 sec
 Yellow = 3 sec
 Red = 1 sec
 Cycle Length = 75 sec

b) LaSalle-Ontario Intersection (M2)

* Subject lane group

Figure 1 Intersection Configurations

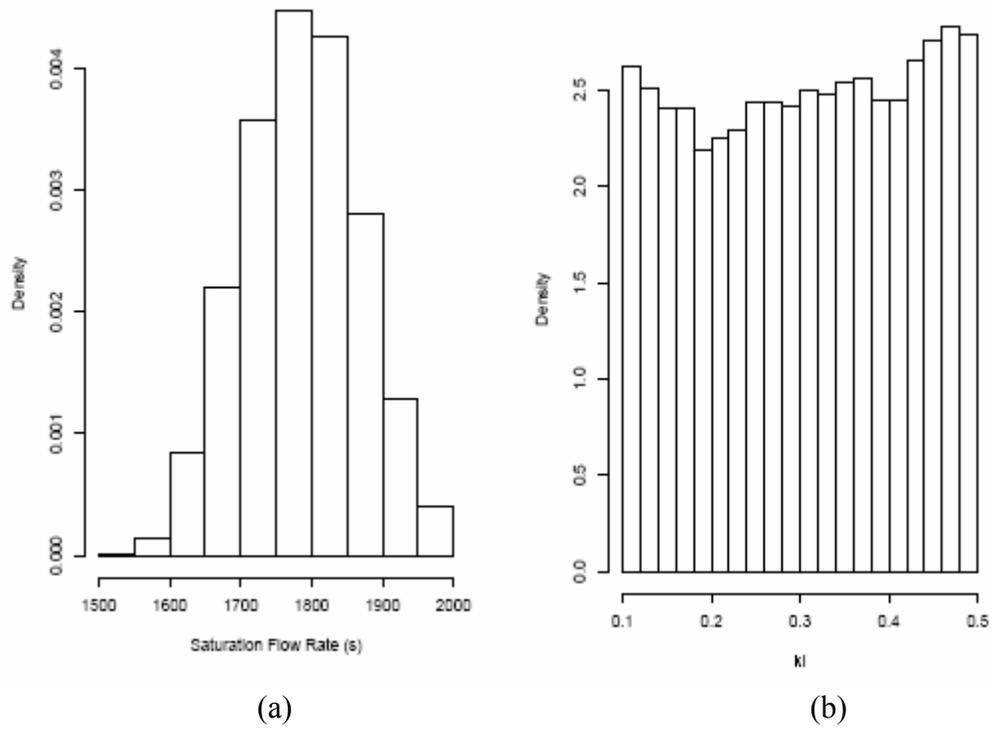


Figure 2. Histograms of the posterior densities of the tuning parameters s and kI . (Model M1; SB Wells-Grand)

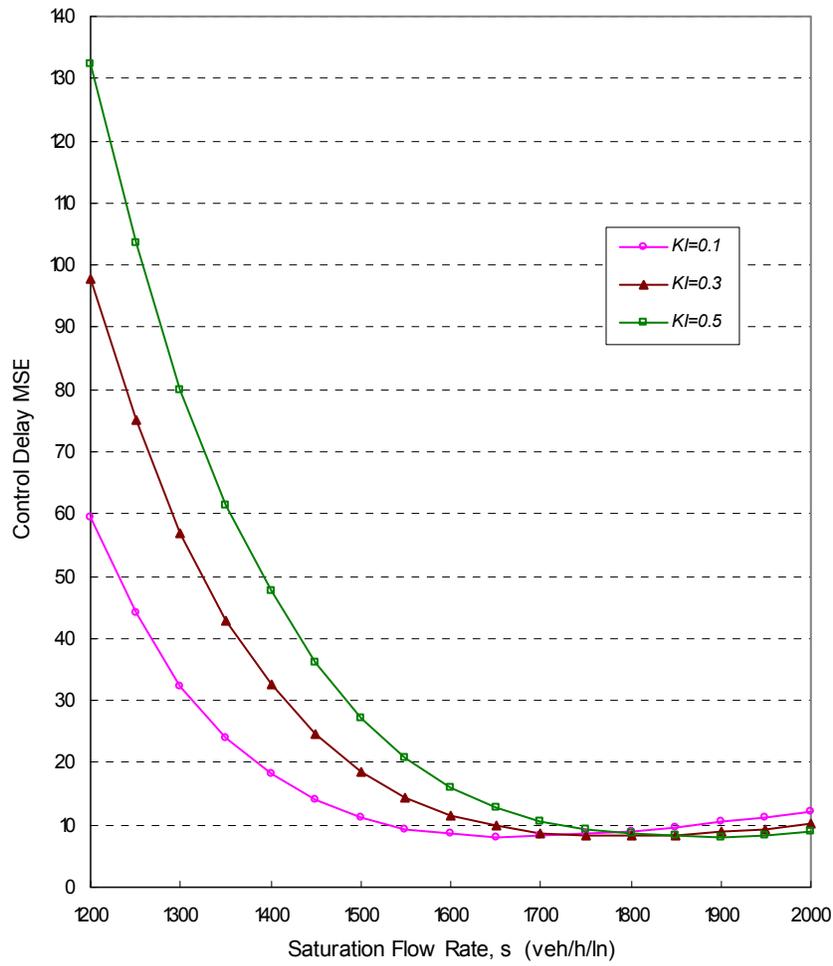
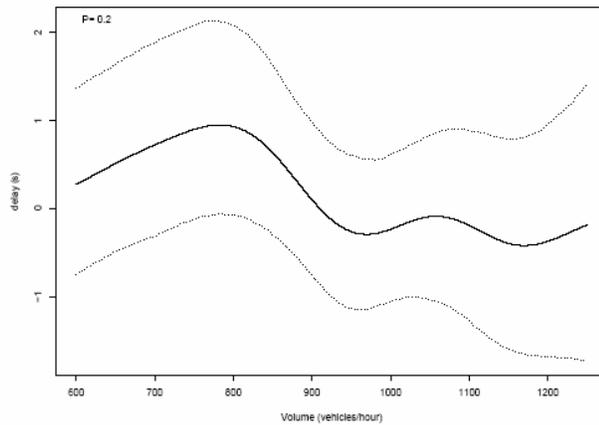
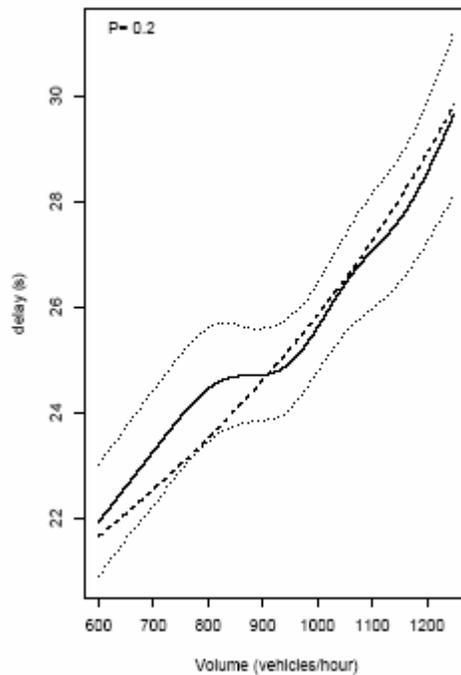


Figure 3 Control Delay MSE* vs. s and kI for Wells-Grand SB Lane Group

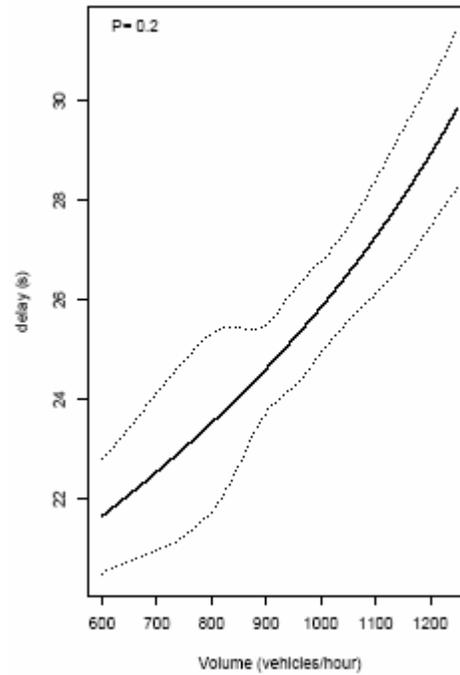
$$* MSE = \frac{1}{48} \sum_{i=1}^{48} (y_i^F - y_i^M)^2$$



(a)

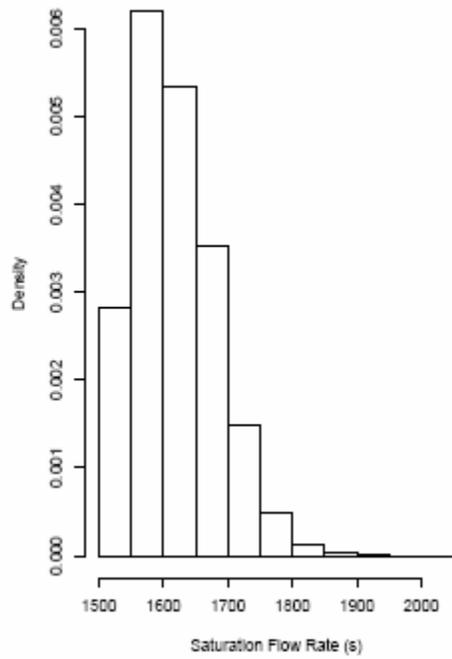


(b)

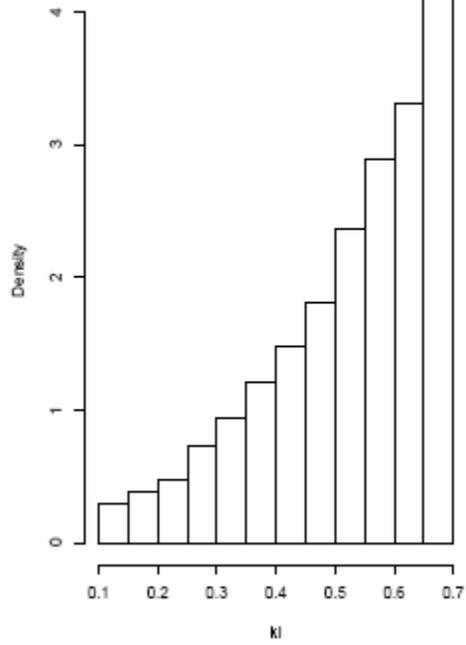


(c)

Figure 4. Estimates of bias, bias-corrected predictions and pure-model for Model M1 and SB Wells-Grand. For (a) the solid curve is the estimate of b and the dotted curves form 90% uncertainty bounds. For (b) the solid curve is the bias-corrected prediction of reality; the dotted curves form its 90% uncertainty bounds and the dashed curve is the pure-model prediction of reality. For (c) the solid curve is the pure-model prediction of reality and the dotted curves form its 90% uncertainty bounds.



(a)



(b)

Figure 5. Histograms for posterior densities of the tuning parameters s and kI . (Model M2; NB LaSalle-Ontario)

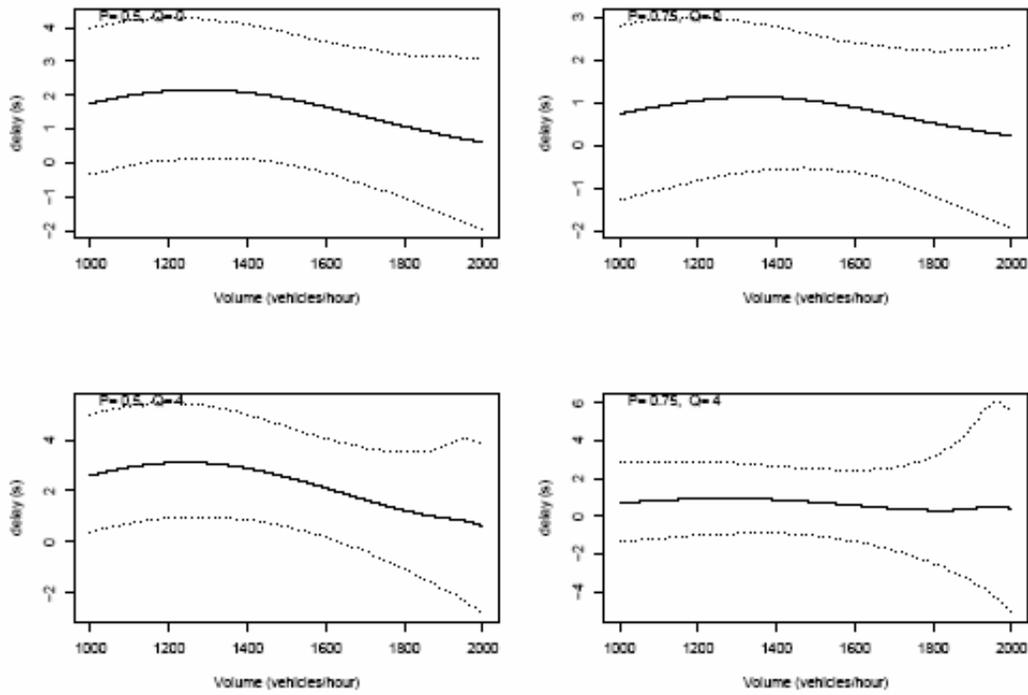


Figure 6. Estimates of bias (Model M2; NB LaSalle-Ontario). The solid curves are estimates of b as a function of volume for four different combinations of P, Q ; the dotted curves form the 90% uncertainty bounds.

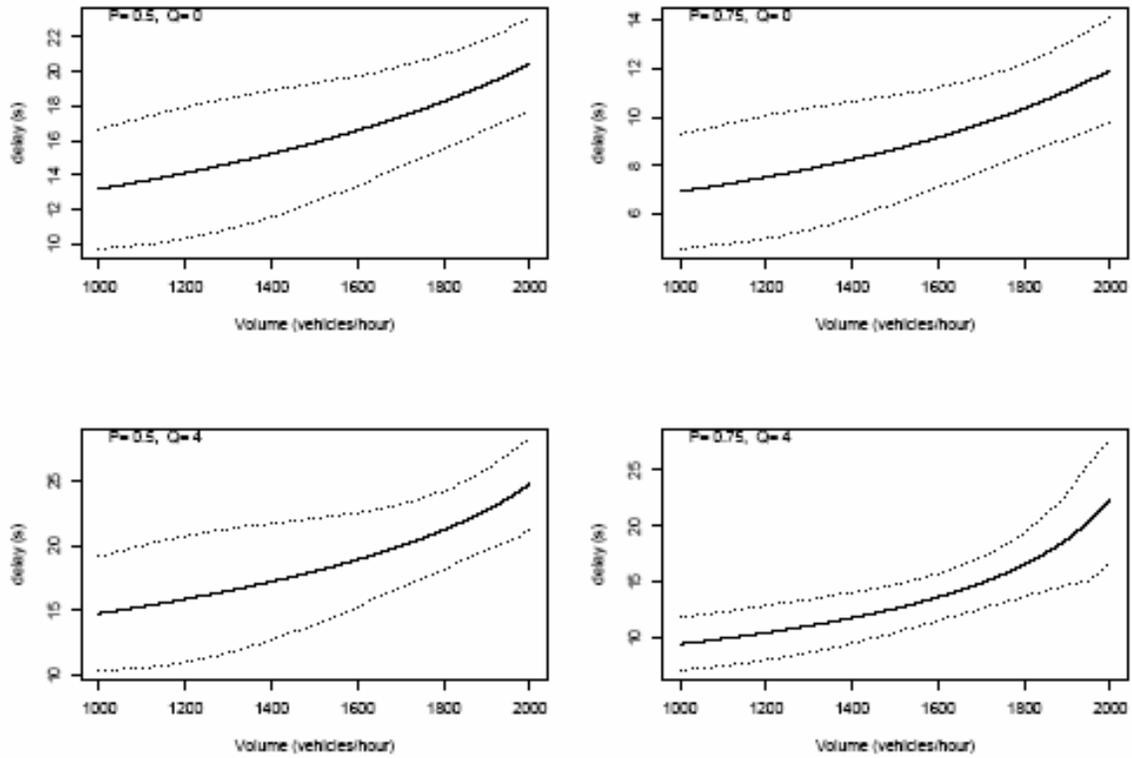


Figure 7. Pure-model predictions of reality (Model M2; NB LaSalle-Ontario). The solid curves are predictions as a function of volume for four different combinations of P, Q ; the dotted curves form the 90% uncertainty bounds.

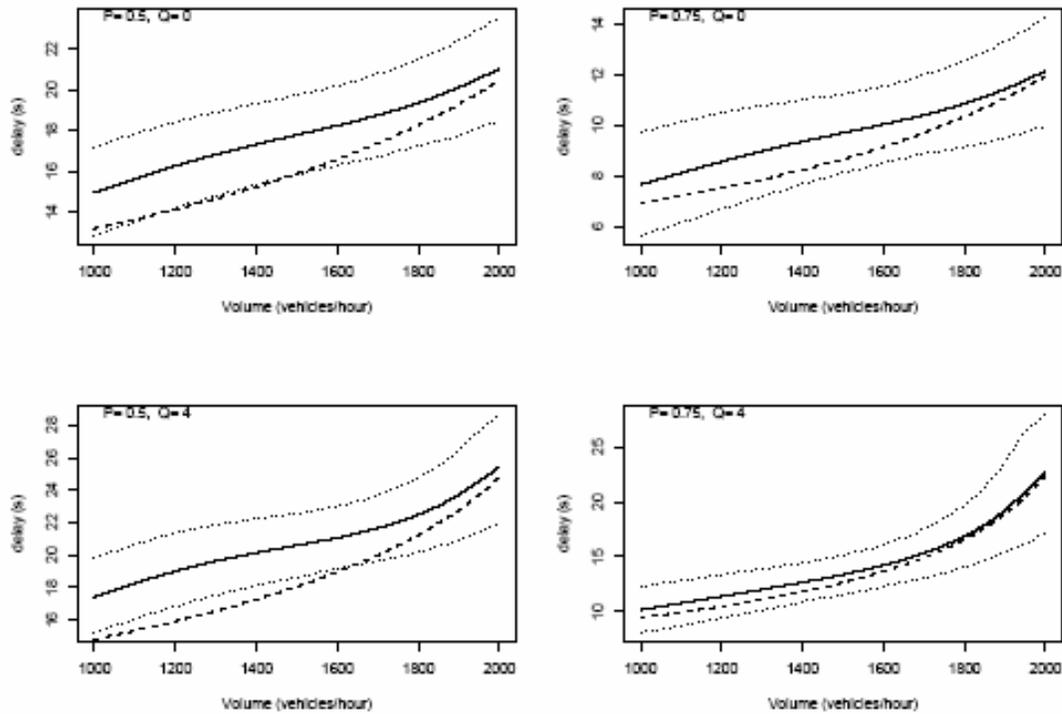


Figure 8. Bias-corrected predictions of reality (Model M2; NB LaSalle-Ontario). The solid curves are predictions as a function of volume for four different combinations of P, Q ; the dotted curves form the 90% uncertainty bounds; the dashed lines are the pure-model predictions..