Statistical Analysis and Validation of Multipopulation Traffic Simulation Experiments

SHIRISH S. JOSHI AND AJAY K. RATHI

Computer simulation has become a very powerful decision aid for varied facets of traffic engineering. Simulation experiments are often used to fit a metamodel of interest between the mean response and a selected set of input factors. This is done by carefully designing statistical experiments under alternative system designs, which are referred to as multipopulation simulation experiments. Validation and statistical analysis procedures are presented on linear metamodels from multipopulation traffic simulation networks under the common random number (CRN) strategy on three sample networks using the TRAF-NETSIM model. Under the CRN strategy, positive correlations are induced among the observations, and hence the usual statistical analysis cannot be applied to obtain point estimates and confidence intervals; therefore it must be modified. Before the statistical analysis is conducted, certain assumptions of the CRN strategy should be validated—those that, if violated, render the modified statistical analysis invalid.

Variance reduction techniques (VRTs) reduce the variance of the estimates of interest by replacing the original sampling procedure with a new procedure that yields the same expected value but with a smaller variance. Among the various correlation-induction techniques used as VRT, such as the common random numbers, antithetic variates, and Schruben-Margolin strategy, the common random number (CRN) strategy is perhaps one of the easiest to employ. Rathi (*1*) illustrated the effectiveness of the CRN strategy for the TRAF-NETSIM simulation model developed by FHWA.

TRAF-NETSIM is a microscopic, stochastic simulation model of traffic operations on urban street networks. This program has been applied extensively to a wide variety of problem areas by both practitioners and researchers and is the most widely used traffic simulation model (*2*). The availability of this model has enabled the development and testing of innovative traffic management concepts and designs (*3*). An important feature of this model is its amenability to control randomness from one simulation run to the next. This control can be used to induce desired correlations among the outputs and reduce the variance of estimates on the statistics of interest.

This paper presents validation and statistical analysis procedures on linear metamodels for multipopulation traffic simulation networks under the CRN strategy on three sample networks using the TRAF-NETSIM model. Under the CRN strategy, positive correlations are induced among the observations, and hence the usual statistical analysis cannot be applied to obtain point estimates and confidence intervals; therefore, it must be modified. Before the statistical analysis is conducted certain assumptions of the CRN strategy should be validated—assumptions that, if violated, render the modified statistical analysis invalid.

MULTIPOPULATION SIMULATION EXPERIMENTS

Often the purpose of a simulation experiment is to estimate a *metamodel* of a selected response, that is, a linear or nonlinear model of the mean response in terms of relevant decision variables for the simulated system. This fitted metamodel can be used in several ways. For example, it can be used to perform a sequential search in order to obtain better response values or make inferences on the behavior of the system. Consider a situation in which each simulation run yields a univariate response y. A particular run, j, called a *design point,* and denoted by x_{jl} $(l = 1,2,\ldots,k)$, is identified by a setting of k factors or decision variables that are used as inputs to the simulation model. Suppose there are r replications of the simulation experiment across the m design points composing the experiment; then the relation of the response y_{ij} for the ith replication and the jth design point to the level of the k factors can be represented as a linear-metamodel having the form

$$y_{ij} = \beta_0 + \sum_{l=1}^{k} \beta_l x_{jl} + \epsilon_{ij} \qquad \text{for } i = 1, 2,\ldots, r \text{ and } j = 1,2,\ldots,m \tag{1}$$

where β_l $(l = 1,2,\ldots,k)$ are the metamodel parameters and ϵ_{ij} is the experimental error at the ith replication at the jth design point. Across all m design points in the experiment, the metamodel in Equation 1 for the ith replicate can be written in matrix notation as

$$\mathbf{y}_i = X\beta + \epsilon_i \qquad \text{for } i = 1, 2, \ldots, r \tag{2}$$

where

$\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{im})'$,
$\beta = (\beta_0, \beta_1')' = (\beta_0, \beta_1, \ldots, \beta_k)'$,
$\epsilon_i = (\epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{im})'$, and
$\mathbf{X} = (\mathbf{1_m T})$ is the $m \times (k + 1)$ design matrix with ones in the first column and x_{jl} in the jth row and $(l + 1)$st column.

In classical statistics, a design point is also referred to as a population. Since this experimental setup has more than one design point, it is called a *multipopulation simulation experiment.* Often, the variances of the estimates of the metamodel coefficients can be reduced by inducing correlations of a desired sign between observations obtained from different runs. The induced correlations are obtained

Center for Transportation Analysis, Oak Ridge National Laboratory, P.O. Box 2008, MS 6206, Oak Ridge, Tenn. 37831-6206.

by controlling the random number streams that drive the simulation model. Unfortunately there is no general guarantee that the correlation-induction strategies produce the desired variance reduction. Therefore, careful implementation of these techniques is needed. CRN is one such useful correlation-induction technique.

CRN STRATEGY

The idea of the CRN strategy is to compare alternative simulation models under similar experimental conditions in order to improve confidence that observed differences in performance are due to the differences in the model structure rather than to differences in the experiment itself (*4*, p. 61). Under the CRN strategy, the same set of random number streams, $\mathbf{R}_i = (r_{i1}, r_{i2} \ldots, r_{ig})$ is applied to all m design points in the ith replicate where g is the number of streams used to drive the simulation model. Also, independent random number streams are used across replicates of the experimental design. Replications reduce the variance of the outputs and present means of computing pure error. For the CRN strategy applied to simulation experiments, the following assumptions are made:

1. The response variance is constant across all design points, so that

$$\sigma_j^2 = \text{var}[y_{ij}(R_i)] = \sigma^2 \qquad \text{for } j = 1,2,\ldots,m \text{ and } i = 1,2,\ldots,r \tag{3}$$

2. There is a constant nonnegative correlation between all pairs of responses within a given replicate, y_{ij} and y_{ik} $(j \neq k)$. That is,

$$\text{corr}(y_{ij} y_{ik}) = \rho_+ \text{ for } j \neq k \quad 1 < j,k < m \tag{4}$$

where $0 < \rho_+ < 1$.

3. The vector of responses composing the ith replicate has a multivariate normal distribution. Under the first two assumptions, the covariance matrix between observations within a replicate is given by

$$\sum{}^{(CRN)} = \sigma^2 \begin{bmatrix} 1 & \rho_+ & \cdot & \cdot & \cdot & \rho_+ \\ \rho_+ & 1 & \cdot & \cdot & \cdot & \rho_+ \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_+ & \rho_+ & \cdot & \cdot & \cdot & 1 \end{bmatrix} \tag{5}$$

VALIDATION AND STATISTICAL ANALYSIS PROCEDURES

The results for validation and statistical analysis procedures developed by Tew and Wilson (*5*) are presented. For a detailed theoretical framework, the reader is referred to Tew and Wilson (*5*). To perform statistical analysis and validation of the fitted metamodel under the CRN strategy, it is useful to transform the model to one with independent observations within each replicate. This is done by applying the $m \times m$ orthogonal transformation $\Gamma^{(\text{CRN})}$ given by

$$\Gamma^{(CRN)} = \begin{bmatrix} m^{-1/2} & 1'_m \\ & C' \end{bmatrix} \tag{6}$$

where $\mathbf{C}$ is an $m \times (m - 1)$ matrix such that $m^{-1/2}\, \mathbf{1}_m\, \mathbf{C}$ is orthogonal. (Note that in this case, the term $\mathbf{1}_m\, \mathbf{C}$ does not indicate matrix multiplication; instead, it indicates the $m \times m$ matrix whose first column is given by $\mathbf{1}_m$, and whose remaining $m - 1$ columns comprise the matrix $\mathbf{C}$.)

Validation

The validation consists of a three-step procedure in which each step checks a key assumption across all design points. The test in each step depends on validation of hypothesized properties of the previous steps; hence, these diagnostic checks on the experimental design and analysis must be performed in order. At each step a highly significant test statistic generally will indicate the need for some corrective action by the analyst. The following three diagnostic tests must be performed.

1. Test for multivariate normality.

H_0: $\mathbf{y}_i \sim N_m(\boldsymbol{\mu}, \Sigma)$ where Σ is positive definite but otherwise μ and Σ are unspecified

versus

H_1: $\mathbf{y}_i$ has any nonnormal, nonsingular m-dimensional distribution (7)

2. Test for induced covariance structure.

H_0: $\text{cov}(\mathbf{y}_i) = \Sigma^{(\text{CRN})}$ with σ^2 and ρ_+ as in Equation 5 so that $\Sigma^{(\text{CRN})}$ is positive definite and $0 < \rho_+ < 1$; otherwise σ^2 and ρ_+ are unspecified

versus

H_1: $\text{cov}(\mathbf{y})_i$ is positive definite but different from $\Sigma^{(\text{CRN})}$ (8)

3. Test for lack of fit in the linear model.

H_0: $E(\mathbf{y}_i) = \mathbf{X}\,\beta$

versus

H_i: $E(\mathbf{y}_i) \neq \mathbf{X}\,\beta$ (9)

The Shapiro-Wilk test (*5*, Section 2.1) will be used to test the normality of responses $\mathbf{y}_i$.

To test the covariance structure, the conventional likelihood ratio test statistic for H_0 has the form

$$L = \left[\frac{\det(r^{-1}A)}{\hat{\lambda}_1^2(\hat{\lambda}_2^2)^{(m-1)}}\right]^{r/2} \tag{10}$$

where

$$A = \sum_{i=1}^{r} (y_i - \bar{y})(y_i - \bar{y})' \tag{11}$$

and $\hat{\lambda}_1^2$ and $\hat{\lambda}_2^2$ are explicitly defined elsewhere (*25, 27* respectively). Also,

$$\bar{y} = \frac{1}{r}\sum_{i=1}^{r} y_i \tag{12}$$

is the sample mean of the original m-dimensional response vectors. If the responses are multinormal with the prescribed covariance structure given by Tew and Wilson (*5*), then the test statistic $N \equiv -2\ \ln(L)$ asymptotically has a chi-squared distribution with $1/2\ m(m+1) - 2$ degrees of freedom as $r \rightarrow \infty$ (*6*). However, the rate of convergence to this limiting distribution can be slow. To achieve adequate convergence to this limiting distribution of N with moderate values of r, Joshi and Tew (*6*) developed a modified likelihood ratio statistic, M, whose definition and use are described in the following.

Note that all the tests and analyses presented hereafter were derived using the transformed responses. They are, however, presented in terms of original responses to illustrate their application and ease of use to the simulation practitioner.

Reject the null hypothesis in Equation 8 if

$$M > \chi^2_{1-\alpha}\left[\frac{1}{2}m(m+1) - 2\right] \tag{13}$$

$$M = -2\psi_0 \ln(L) \tag{14}$$

with

$$\psi_0 = \frac{\dfrac{m(m+1)}{2} - 2}{\dfrac{m^2 - 3m + 2}{2\psi_1} + \dfrac{m-1}{\psi_2} + \dfrac{m-2}{\psi_3}} \tag{15}$$

and

$$\psi_1 = 1 - \frac{2m+3}{6r} \tag{16}$$

$$\psi_2 = 1 - \frac{3m^2 - 1}{6r(m-1)} \tag{17}$$

and

$$\psi_3 = 1 + \frac{m}{3r(m-1)} \tag{18}$$

The last stage of the validation procedure is to test for the lack of fit in the model. It uses a standard lack-of-fit test only applied to transformed responses. Define the error sum of squares, S_E, as

$$S_E = \sum_{i-1}^{r} \left\| y_i - X\hat{\beta}_1 \right\|^2 \tag{19}$$

where

$$\hat{\beta} = (X'X)^{-1} X'\bar{y} \tag{20}$$

is the ordinary least squares estimate of β. Also define S_E^*, the error sum of squares for the transformed responses, and S_{PE}^*, the pure error of the transformed responses respectively, as

$$S_E^* = S_E - m\sum^{r}(\bar{y}_{i.} - \bar{y}_{..})^2 \qquad \text{with } \nu_E^* = mr - k \tag{21}$$

$$S_{PE}^* = r(m-1)\lambda_2^2 \qquad \text{with } \nu_{PE}^* = m(r-1) \tag{22}$$

where $\bar{y}_{i.}$ is the average response vector at the ith replication taken over across all design points, and $\bar{y}_{..}$ is the average response taken over all design points and over all replications.

Reject H_0 in Equation 9 if

$$\frac{(S_E^* - S_{PE}^*)/(\nu_E^* - \nu_{PE}^*)}{S_{PE}^*/\nu_{PE}^*} > F^{\alpha}_{(\nu_E^* - \nu_{PE}^*, \nu_{PE}^*)} \tag{23}$$

where F is the quantile of order $1-\delta$ for the F-distribution with $\nu_E^* - \nu_{PE}^*$ and ν_{PE}^* degrees of freedom. Results on statistical analysis are presented next.

Statistical Analysis

The statistical analysis involves estimation of β and construction of simultaneous confidence intervals for the elements of β. The uniformly minimum variance unbiased (optimal) estimator of β is $\hat{\beta}$ and is as given by Equation 20. The model independent estimator of σ^2 is given by

$$\hat{\sigma}^2 = [m(r-1)]^{-1} \sum_{i=1}^{r}\sum_{j=1}^{m} (y_{ij} - \bar{y}_{.j})^2 \tag{24}$$

where $\mathbf{y}_{.j}$ is the average response at the jth design point taken over all replicates. The variance on the estimate of β_0 is given by $\hat{\lambda}_1^2$ and can be viewed as the *between replicate variation;* it is

$$\hat{\lambda}_1^2 = m \sum_{i=1}^{r} \frac{(\bar{y}_{i.} - \bar{y}_{..})^2}{r} \tag{25}$$

We have

$$\frac{(mr)^{1/2}(\beta_0 - \hat{\beta}_0)}{\hat{\lambda}_1} \sim t_{r-1} \tag{26}$$

which can be used to construct $100(1-\alpha)$ percent confidence interval for β_0.

Next, define $\hat{\lambda}_2^2$, the estimate of the pure error variance σ^2, as

$$\hat{\lambda}_2^2 = \frac{(mr - m)\hat{\sigma}^2 - m\sum_{i=1}^{r}(\bar{y}_{i.} - \bar{y}_{..})^2}{r(m-1)} \tag{27}$$

The joint $100(1-\alpha)$ percent simultaneous confidence interval for $\mathbf{l'H}\beta_1$ for all $l \in \mathbf{R}^h$ under the prescribed covariance structure where H is a known $h \times k$ matrix of constants with rank $h \leq (k+1)$ and is given by

$$\mathbf{l'H}\beta_1 \in \mathbf{l'H}\hat{\beta}_1 \pm \hat{\lambda}_2^2 \left[\frac{hF^{\alpha}_{h,(m-1)r-k-1} l'H(\mathbf{T'T})^{-1}\mathbf{H'l}}{r}\right]^{1/2} \tag{28}$$

ILLUSTRATIVE EXAMPLES

For the purpose of illustration, three sample TRAF-NETSIM networks were selected. The geometric conditions for Networks 1, 2, and 3 are depicted in Figure 1, 2, and 3, respectively. These data sets

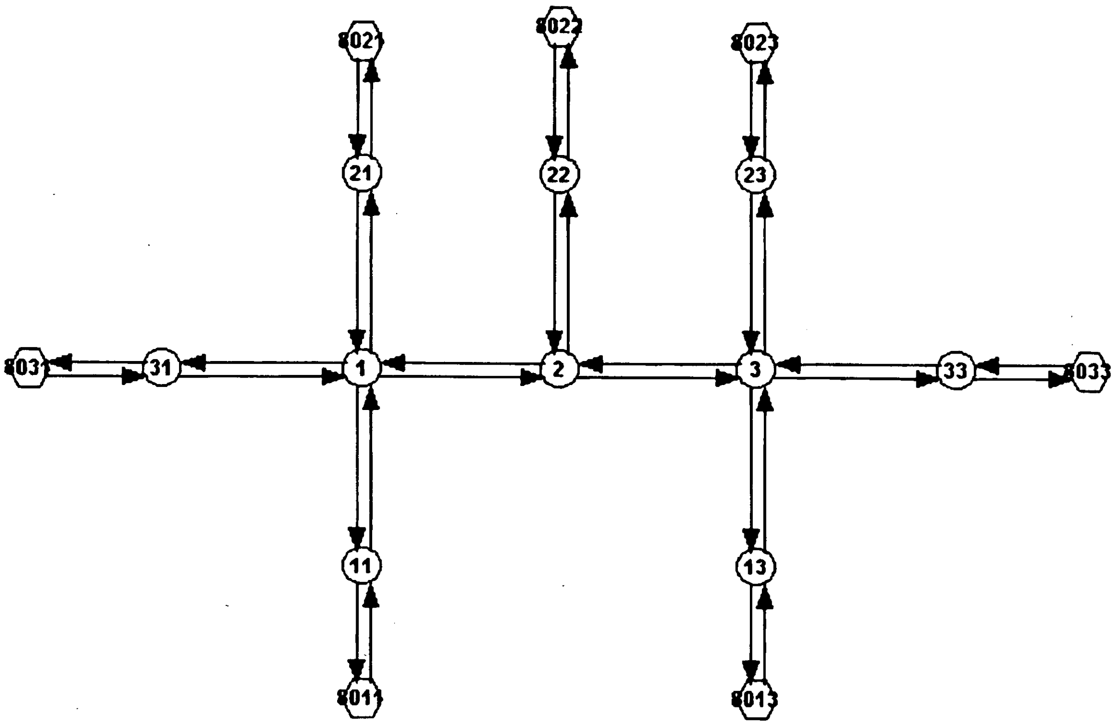

FIGURE 1 Graphical representation of Network 1.

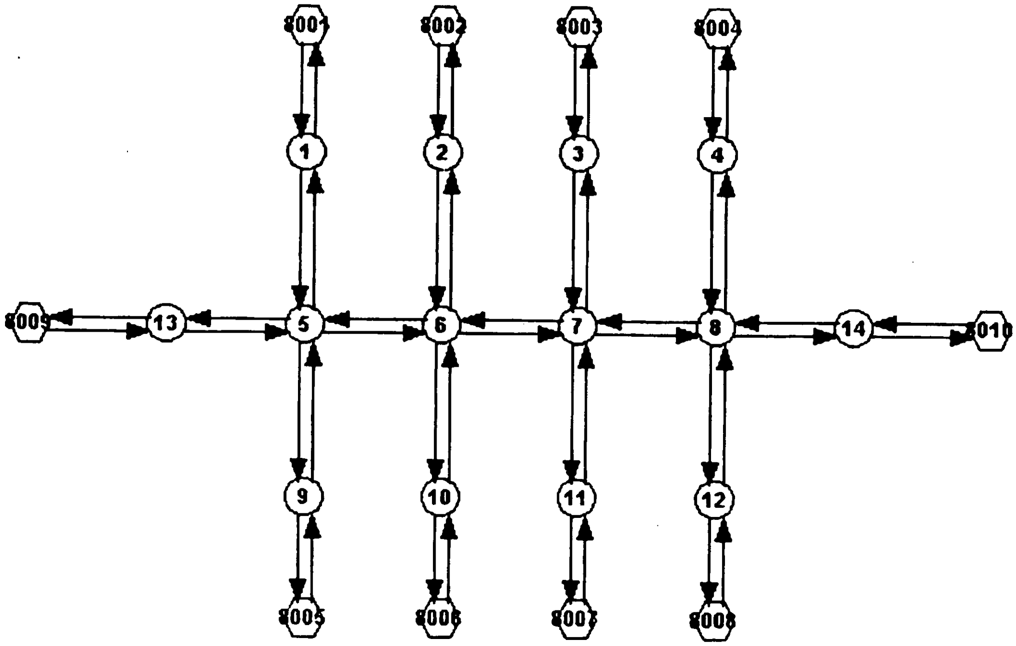

FIGURE 2 Graphical representation of Network 2.

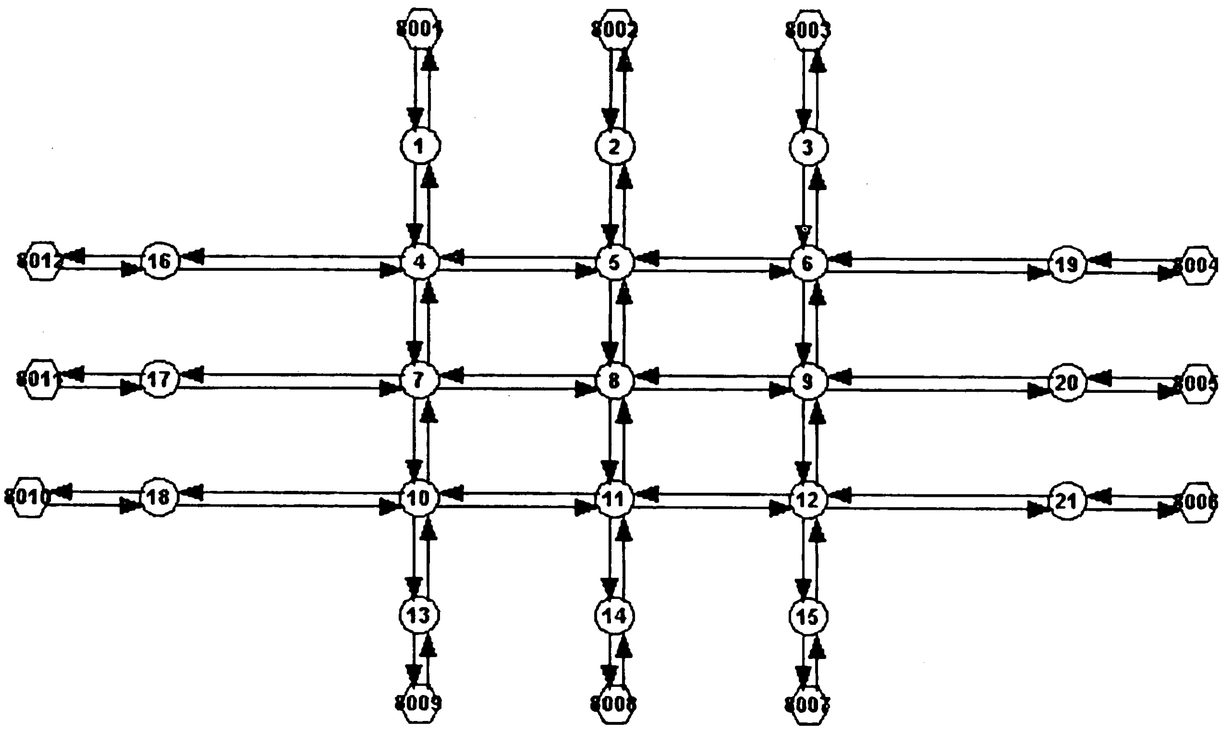

FIGURE 3 Graphical representation of Network 3.

represent traffic networks consisting of an isolated intersection (Figure 1), an arterial (Figure 2), and a grid (Figure 3). The input data for these networks regarding geometric length, signal control, number of lanes on each link, turning movements, and volume information are presented in Tables 1, 2, and 3 respectively. Because of the difference in characteristics of these three networks, different measures of effectiveness (MOEs) were chosen. The basic information pertaining to simulation experiments for the three networks is summarized in Table 4.

For all these networks, nodes numbered 8XXX, where X is an integer between 0 and 9, are called *entry/exit* nodes. That is, traffic enters and exits through these nodes *only.* The rest of the nodes are all *internal* nodes. The combination of selected traffic volume, network geometry, and control represents congested network. The purpose of the study was to estimate and validate the relationship between the MOE (y), and the decision variables (x's), for each network. Simulation experiments specified in Tables 1 through 4 were conducted, and MOEs, recorded.

For all these networks, the following first-order metamodel was used to describe the relationship between the response, y (MOE), and the decision variables, x_1 and x_2 (i = 1, 2, ..., 10 and j = 1, 2, 3, 4):

$$y_{ij} = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \beta_3 x_{j1} x_{j2} + \epsilon_{ij} \qquad (29)$$

where β_0 is the average delay of interest across all design points and replicates.

EXPERIMENTAL SETUP

A sample experimental setup for Network 2 under the CRN strategy is described to help the practitioner appreciate the ease of its use. Since there are two decision variables, or factors (green splits at Nodes 5 and 8), denoted by x_1 and x_2, a 2^2 factorial experiment is conducted—that is, the two factors can be set at two levels, low and a high. The respective low and high levels for this example are 27 and 37 sec as described in Table 4. There can, then, be four combinations to set these factors: x_1 at its low level and x_2 at its low level, x_1 at its high level and x_2 at its low level, and so on. Each of these four combinations is called a design point, and there are four design points for this simulation experiment. Under the CRN strategy, each design point is driven by the same random number stream. Replications of such simulations at each design point, however, use independent random number streams. If five replications are performed, five random number streams are used. However, the same five random number streams are used across the four design points. In practice, applying the CRN strategy is easier than performing "normal simulation" (independent streams), since fewer random number streams are required under the CRN strategy. Note that in this case, the "normal simulation" would need 20 random number streams.

SIMULATION ANALYSIS AND RESULTS

The results of the validation for Networks 1, 2, and 3 are presented in Tables 5, 6, and 7 respectively, and those for statistical analysis procedures (if applicable) are presented in Tables 8, 9, and 10, respectively.

A visual inspection of the correlation matrix in Table 5 indicates consistently induced positive correlations among responses between all pairs of design points in Network 1. It therefore guarantees that variance reduction is achieved, and using the CRN strategy is better for this network than using independent streams. The test for multivariate normality of responses fails to be rejected, and

TABLE 1 Input data for Network 1

Link length:	All links	500 ft.
Signal Control		
Nodes 1 and 3		Signal control
Node 2		Two-phase actuated control
All other nodes		Perpetual Green
Number of Lanes		
Links with 3 lanes		8031-31, 3-33, 31-1, 1-2, 2-3
Links with 1 lane		8022-22, 2-22, 22-2
Links with 2 lanes		All other links.
Entry Volume (Follows a Uniform Distribution)		
node 8011		700 vph
node 8013		750 vph
node 8021		500 vph
node 8022		300 vph
node 8023		650 vph
node 8031		1350 vph
node 8033		650 vph
Turning Movement		
Link 31-1		30% left, 60% through, 10 % right
Link 2-1		30% left, 54% through, 16% right
Link 11-1		14% left, 36% through, 50 % right
Link 21-1		40% left, 40% through, 20% right
Link 3-2		71% through, 29% right
Link 2-3		82% through, 18% right
Link 13-3		19% left, 48% through, 33% right
Link 23-3		27% left, 50% through, 23% right
All other links		100% through

so does the test for the covariance structure from equation 8. The linear model representation in Equation 29 is found to be an adequate representation for Network 1. Since all three stages of the validation procedure failed to reject the null hypotheses in Equations 7, 8, and 9, the analyst proceeds with the statistical analysis. This analysis yields point estimates and confidence intervals for the unknown parameters of the model in Equation 29. From the preceding confidence intervals, it is observed that the main effect for x_1, the green split at Node 1, and the interaction effect do not appear to influence the delay in any significant manner. However, increasing x_2, the green split at Node 3 will decrease the vehicle delay, at least in the vicinity of the current setting of the decision variable.

Table 6 illustrates consistently induced positive correlations among responses between all pairs of design points in Network 2. The use of CRN strategy is therefore justified for this network. The test for multivariate normality of responses fails to be rejected, and so does the test for the covariance structure from Equation 8. The linear model representation in Equation 29 is found to be an adequate representation for Network 2. As for the previous network, all three stages of the validation procedure failed to reject their respective null hypotheses. Statistical analysis can therefore be conducted, as prescribed. From the previous confidence intervals, it is observed that the main effects, x_1, the green split at Node 5, and x_2, the green split at Node 8, are both significant. The interaction term does not appear to contribute significantly to the fitted metamodel. Therefore, decreasing x_1, and increasing x_2, will decrease the vehicle delay, at least in the vicinity of the current setting of the decision variables.

Finally for Network 3, Table 7 indicates consistently induced positive correlations among responses between all pairs of design points in Network 3. Employing the CRN strategy will therefore improve metamodel estimation, provided that the assumptions are validated. The test for multivariate normality of responses fails to be rejected, and so does the test for the covariance structure from Equation 8. The linear model representation in Equation 29 is found to be an adequate representation for Network 3. Since all three stages of the validation procedure failed to reject their null hypotheses, the statistical analysis could be conducted. It yields point estimates and confidence intervals for the unknown parameters of the model in Equation 29. From the confidence intervals, it is observed

TABLE 2 Input data for Network 2

Link length:	All links	500 ft.
Signal Control		
Nodes 5,6, 7, and 8		Signal control
All other nodes		Perpetual Green
Number of Lanes	All links	Two lanes
Entry Volume (Follows a Uniform Distribution)		
All nodes		1600 vph
Turning Movement		
All links at four-way intersections		25% left, 50% through, 25 % right
All other links		100% through

TABLE 3 Input data for Network 3

Link length:	All links	500 ft.
Signal Control		
Nodes 4 through 12		Signal control
All other nodes		Perpetual Green
Number of Lanes		
All links		Two lanes
Entry Volume (Follows a Uniform Distribution)		
nodes 8001, 8005, 8009, and 8011		1000 vph
All other nodes		1600 vph
Turning Movement		
All links at four-way intersections		25% left, 50% through, 25 % right
All other links		100% through

TABLE 4 Information for TRAF-NETSIM Experiments, Networks 1–3, 2^2 Factorial

	Network		
	1	2	3
No. of replications	10	5	5
Duration (sec)	1,800	1,800	1,800
Decision variables[a]			
X_1	Nodes 1 and 31	Nodes 5 and 13	Nodes 4 and 16
X_2	Nodes 3 and 33	Nodes 8 and 14	Nodes 10 and 18
Levels (sec)			
X_1 (low/high)	8/12	27/37	27/47
X_2 (low/high)	13/17	27/37	27/47
MOE: y (sec)	Avg. delay for vehicles entering at Node 31 and exiting at Node 33	Avg. delay in network	Avg. delay in network

[a]Decision variables denote green split and approach nodes.

TABLE 5 Validation results for Network 1

Correlation matrix of responses

$$corr(y) = \begin{bmatrix} 1.0000 & 0.5924 & 0.7259 & 0.7354 \\ 0.5924 & 1.0000 & 0.7339 & 0.7630 \\ 0.7295 & 0.7339 & 1.0000 & 0.7211 \\ 0.7354 & 0.763\ 0 & 0.7211 & 1.0000 \end{bmatrix}$$

Test for multivariate normality	
W^*	0.67
$w^*_{0.05}(4,10)$	0.598
Test for the correlation matrix	
M	7.86
$\chi^2(8)$	20.09
Test for lack-of-fit of postulated model	
Test statistic	0.54
$F^{1-0.05}_{(1,26)}$	4.23
Validation complete. Fail to reject all null hypotheses. Therefore proceed with statistical analysis.	

that the main effect for x_1, the green split at Node 16 does not appear to influence the delay in any significant manner. However, increasing x_2, the green split at Node 10 will decrease the vehicle delay, at least near the current setting of the decision variable. The interaction term, however, can play a role in this situation, and hence the analyst should proceed with caution. Conducting a pilot study to explore the effects of increasing the green split at Node 10 may be a suitable alternative before reaching to any meaningful conclusions about this network.

The goal of this paper is to demonstrate the application of the validation and statistical analysis procedures of linear metamodels from multipopulation simulation experiments under the CRN strategy, so the authors do not conduct further analysis on the effect of the decision variables on the response but instead point out the ease of application of such procedures to the practitioner. The three networks selected for this study exhibit three different characteristics in their statistical analyses. Network 1 has only one factor significant, which is the main effect, x_2. The other main effect and the interaction terms

TABLE 6 Validation results for Network 2

Correlation matrix of responses

$$corr(y) = \begin{bmatrix} 1.0000 & 0.4728 & 0.5992 & 0.9048 \\ 0.4782 & 1.0000 & 0.0785 & 0.3184 \\ 0.5992 & 0.0785 & 1.0000 & 0.6362 \\ 0.9048 & 0.3184 & 0.6362 & 1.0000 \end{bmatrix}.$$

Test for multivariate normality	
W^*	0.72
$w^*_{0.05}(4,10)$	0.5
Test for the correlation matrix	
M	1.6873
$\chi^2(8)$	20.09
Test for lack-of-fit of postulated model	
Test statistic	0.16
$F^{1-0.05}_{(1,26)}$	4.23
Validation complete. Fail to reject all null hypotheses. Therefore proceed with statistical analysis.	

TABLE 7 Validation results for Network 3

Correlation matrix of responses

$$corr(y) = \begin{bmatrix} 1.0000 & 0.8373 & 0.7235 & 0.3418 \\ 0.8373 & 1.0000 & 0.9531 & 0.7992 \\ 0.7235 & 0.9531 & 1.0000 & 0.8538 \\ 0.3418 & 0.7992 & 0.8538 & 1.0000 \end{bmatrix}$$

Test for multivariate normality	
W^*	0.64
$w^*_{0.05}(4,10)$	0.5
Test for the correlation matrix	
M	7.9986
$\chi^2(8)$	20.09
Test for lack-of-fit of postulated model	
Test statistic	0.16
$F^{1-0.05}_{(1,11)}$	4.84

Validation complete. Fail to reject all null hypotheses. Therefore proceed with statistical analysis.

are not significant. Network 2 exhibits the significance of both major factors on the metamodel, but the interaction term is not significant. Network 3 has one main effect, x_2, and the interaction term to be significant in the fitted metamodel, but the other main effect appears insignificant. These networks are therefore interesting for further exploration in their own way. For example, exploring Network 3 along increasing values of the variable x_2 can be self-defeating if the interaction term increases the delay with an increase in x_2.

In applying the CRN strategy for metamodel estimation and analysis to any simulation experiment, a note of caution is warranted. There is no guarantee that the CRN strategy will produce the desired variance reduction. For this reason, the analyst should conduct a pilot study of the system before conducting an exhaustive simulation analysis. This can be done by computing the correlations obtained across design points for a smaller study and validating the assumptions. If negative correlations are observed across design points, then it is an indication that the CRN strategy may not be amenable for this particular problem.

CONCLUSIONS AND FUTURE RESEARCH

This paper demonstrates the statistical analysis and validation techniques for linear metamodels in multipopulation traffic simulation networks under the CRN correlation-induction strategy. This illustration comprises a three-stage validation procedure and comprehensive postvalidation statistical analysis. The examples show the ease of applying the CRN strategy for traffic simulation experiments and of performing estimation and analysis on a fitted metamodel. The TRAF-NETSIM model can be used to perform efficient simulation experiments that could help a traffic simulation analyst gain more confidence in the results.

TABLE 8 Statistical analysis results for Network 1

Optimal estimator of β

$$\hat{\beta} = \begin{bmatrix} 27.1339 \\ -0.0933 \\ -3.6866 \\ -0.9689 \end{bmatrix}$$

95% Confidence Interval for β_0: $25.82 \leq \beta_0 \leq 28.44$

Simultaneous 95% confidence interval for elements of β

$-1.94 \leq \beta_1 \leq 1.76$, $-5.54 \leq \beta_2 \leq -1.84$, $-2.82 \leq \beta_3 \leq 0.88$

TABLE 9 Statistical analysis results for Network 2

Optimal estimator of β

$$\hat{\beta} = \begin{bmatrix} 533.1096 \\ 18.3969 \\ -4.9419 \\ 3.7534 \end{bmatrix}$$

95% Confidence Interval for β_0: $524.47 \le \beta_0 \le 541.74$

Simultaneous 95% confidence interval for elements of β

$$14.53 \le \beta_1 \le 22.24, \quad -8.79 \le \beta_2 \le -1.09, \quad -0.1 \le \beta_3 \le 7.6$$

TABLE 10 Statistical analysis results for Network 3

Optimal estimator of β

$$\hat{\beta} = \begin{bmatrix} 526.6619 \\ 1.0044 \\ 14.3064 \\ -11.0461 \end{bmatrix}$$

95% Confidence Interval for β_0: $517.8 \le \beta_0 \le 535.51$

Simultaneous 95% confidence interval for elements of β

$$-2.46 \le \beta_1 \le 4.46, \quad 10.84 \le \beta_2 \le 17.76, \quad -14.5 \le \beta_3 \le -7.58$$

Future development in the TRAF-NETSIM model could include the statistical analysis procedures incorporated within the model for different types of simulation experimentation. Another avenue for research would be to develop statistical analysis and validation techniques for multiple responses instead of just a single response.

REFERENCES

1. Rathi, A. K. The Use of Common Random Numbers to Reduce the Variance in Network Simulation of Traffic. *Transportation Research,* Vol. 26B, 1992, pp. 357–363.
2. Traffic Network Analysis with NETSIM—A User Guide. Report FHWA-IP-80-3. FHWA, U.S. Department of Transportation, 1980.
3. Rathi, A. K., and E. B. Lieberman. Effectiveness of Traffic Restraint for a Congested Urban Network: A Simulation Study. In *Transportation Research Record 1232,* TRB, National Research Council, Washington, D.C., 1989, pp. 95–102.
4. Law, A. M., and W. D. Kelton. *Simulation Modeling and Analysis,* 2nd ed. McGraw Hill, New York, 1991.
5. Tew, J. D., and J. R. Wilson. Validation of Simulation Analysis Methods for the Schruben-Margolin Correlation-Induction Strategy. *Operations Research,* Vol. 40, 1992, pp. 87–103.
6. Joshi, S. S., and J. D. Tew. Validation and Statistical Analysis Procedures Under the Common Random Number Correlation-Induction Strategy for Multipopulation Simulation Experiments. *European Journal of Operational Research* (in preparation).