SEQUENTIAL ESTIMATION OF THE STEADY-STATE VARIANCE IN DISCRETE EVENT SIMULATION

Adriaan Schmidt
Institute for Theoretical Information Technology
RWTH Aachen University
D-52056 Aachen, Germany
Email: Adriaan.Schmidt@rwth-aachen.de

Krzysztof Pawlikowski
Computer Science and Software Engineering
University of Canterbury
Christchurch, New Zealand
Email: Krys.Pawlikowski@canterbury.ac.nz

Don McNickle
Department of Management
University of Canterbury
Christchurch, New Zealand
Email: Don.McNickle@canterbury.ac.nz

KEYWORDS

Steady-State Simulation, Variance Estimation, Sequential Output Analysis, Batch Means

ABSTRACT

For sequential output data analysis in non-terminating discrete-event simulation, we consider three methods of point and interval estimation of the steady-state variance. We assess their performance in the analysis of the output of queueing simulations by means of experimental coverage analysis. Over a range of models, estimating variances turns out to involve considerably more observations than estimating means. Thus, selecting estimators with good performance characteristics is even more important.

INTRODUCTION

The output sequence $\{x\} = x_1, x_2, \ldots$ of a simulation program is usually regarded as the realisation of a stochastic process $\{X\} = X_1, X_2, \ldots$. In the case of steady-state simulation, we assume this process to be stationary and ergodic.

Current analysis of output data from discrete event simulation focuses almost exclusively on the estimation of mean values. Thus, the literature on "variance estimation" mostly deals with the estimation of the *variance of the mean*, which is needed to construct a confidence interval of the estimated mean values.

In this paper, we are interested in finding point and interval estimates of the steady-state variance $\sigma^2 = \mathrm{Var}[X_i]$ and the *variance of the variance*, from which we can construct confidence intervals for the variance estimates. Similar to the estimation of mean values, one problem in variance estimation is caused by the fact that output data from steady-state simulation are usually correlated.

The variance we estimate is not to be confused with the quantity $\sigma_0^2 = \lim_{n\to\infty} n \,\mathrm{Var}[\overline{X}(n)]$, sometimes referred to as *variance parameter* (Chen and Sargent, 1990) or *steady-state variance constant* (Steiger and Wilson, 2001), and which is important in the methods of standardized time series (Schruben, 1983) and various methods using this concept.

Applications for the estimators we propose can be found in the performance analysis of communication networks. In audio or video streaming applications, for example, the actual packet delay is less important than the packet delay variation or jitter (see e.g. Tanenbaum, 2003). Other applications include estimation of safety stock or buffer sizes, and statistical process control.

Our estimation procedures are designed for sequential estimation. As more observations are generated, the estimates are continually updated, and simulation is stopped upon reaching the required precision.

In simulation practice, one observes an *initial transient phase* of the simulation output due to the initial conditions of the simulation program, which are usually not representative of its long-run behaviour. It is common practice to let the simulation "warm up" before collecting observations for analysis. For many processes, σ^2 converges to its steady-state value slower than the process mean; therefore, existing methods of detection of the initial transient period with regard to the mean value may sometimes not be applicable for variance estimation. A method based on distributions, which includes variance, is described in (Eickhoff et al., 2007). This is, however, not the focus of this paper, so we use a method described in (Pawlikowski, 1990).

In the next section we present three different methods of estimating the steady-state variance. We assessed these estimators experimentally in terms of the coverage of confidence intervals. The results of the experiments are presented in Section 3. The final section of the paper summarises our findings and gives an outlook on future research.

ESTIMATING THE STEADY-STATE VARIANCE

In the case of *independent and identically distributed* random variables, the well-known consistent estimate of the variance is

$$\hat{\sigma}^2(n) = s^2(n) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}(n))^2, \qquad (1)$$

and its variance is known to be (see e.g. Wilks, 1962, p. 200)

$$\mathrm{Var}[s^2(n)] = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\,\sigma^4\right), \qquad (2)$$

where $\bar{x}(n) = \sum_{j=1}^{i} x_j/i$, is the sample mean of the first n observations, and μ_4 is the fourth central moment of the steady-state distribution.

In the case of *correlated observations* usually encountered in simulation output data, s^2 is no longer unbiased, but in fact (see Anderson, 1971, p. 448)

$$\mathrm{E}[s^2(n)] = \sigma^2\left(1 - 2\,\frac{\sum_{j=1}^{n-1}(1-j/n)\rho_j}{n-1}\right), \qquad (3)$$

where ρ_j is the lag j autocorrelation coefficient of the sequence $\{x\}$.

We propose three estimators of σ^2: Method 1 treats it as a mean value and thus avoids dealing with the statistical problems associated with the estimate s^2, Method 2 uses uncorrelated observations to overcome the bias problem of s^2, and Method 3 compensates for this bias.

Method 1: Variance as a Mean Value

Variance is defined as the mean squared deviation of a random variable from its mean value: $\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]$. This suggests that it should be possible to estimate variance as a mean value. We estimate the variance of a sequence $\{x\}$ as the mean of the new sequence $\{y\}$, defined by $y_i = (x_i - \bar{x}(i))^2$. This makes the point estimate

$$\hat{\sigma}_1^2(n) = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

To obtain a confidence interval, any existing procedure of mean value estimation can be used. We use the method of spectral analysis, as proposed by (Heidelberger and Welch, 1981).

Method 2: Uncorrelated Observations

Processes typically encountered as simulation output data have a monotonically decreasing autocorrelation function; so, observations that are spaced "far apart" are less correlated, and if we choose a sufficiently large spacing, we can assume that the observations are (almost) uncorrelated. This enables us to use (1) as a point estimate of the variance.

We define a secondary sequence $\{y\}$ as $y_i = x_{k_0 i}$, where k_0 is the spacing distance needed to make the observations approximately uncorrelated. The ergodic property of the process $\{X\}$ ensures that $\mathrm{Var}[X_i] = \mathrm{Var}[Y_i] = \sigma^2$.

As point estimate we apply (1) to the sequence $\{y\}$:

$$\hat{\sigma}_2^2(n) = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \bar{y}(i))^2$$

where $m = \lfloor n/k_0 \rfloor$ is the length of the sequence $\{y\}$. To calculate the variance of the estimate, we use (2), replacing the actual values of σ^2 and μ_4 with their estimates:

$$\mathrm{Var}[\hat{\sigma}_2^2(n)] = \frac{1}{m}\left(\sum_{i=1}^{m}(y_i - \bar{y}(m))^4 - \frac{m-3}{m-1}\,\hat{\sigma}_2^4(n)\right)$$

The half-width of the confidence interval is then

$$\Delta_2 = z_{1-\alpha/2}\sqrt{\mathrm{Var}[\hat{\sigma}_2^2]}.$$

We do not know of any general results on the distribution of $S^2(n)$; however, we can use a normal distribution because we assume the y_i to be uncorrelated, and their sample size to be large.

To find an appropriate value for the spacing k_0, we successively test values by extracting the respective subsequences, and analysing their autocorrelation.

Method 3: Batch Means

We have seen in (3) that the sample variance of a finite, correlated sample is a biased estimate of σ^2. When considering batches of observations of size m, we know that their means have a variance of (see Law and Kelton, 1991, p. 285)

$$\mathrm{Var}[\bar{x}(m)] = \frac{\sigma^2}{m}\left(1 + 2\sum_{j=1}^{m-1}(1-j/m)\rho_j\right).$$

Based on an unpublished paper by (Feldman et al., 1996), we show that adding $s^2(m)$ and $\mathrm{Var}[\bar{x}(m)]$ yields a consistent estimate of σ^2. One can think of this as splitting the variance into two components, a *local variance* describing the short term variations of the process, and a *global variance* representing the long term variations. The local variance is calculated as the mean variance inside equal-sized batches, and the global variance is the variance of the means of the same batches.

To this end, we define a number of statistics: Given b equal-sized batches of size m, we calculate the sample mean and sample variance of each batch j, containing the observations $x_{j,1}, x_{j,2}, \ldots, x_{j,m}$ as

$$\bar{x}_j = \frac{1}{n}\sum_{k=1}^{n} x_{j,k}$$

$$s_j^2 = \frac{1}{n}\sum_{k=1}^{n}(x_{j,k} - \bar{x}_j)^2.$$

Furthermore, we calculate

$$\bar{\bar{x}} = \frac{1}{b}\sum_{j=1}^{b}\bar{x}_j$$

$$s_X^2 = \frac{1}{b-1}\sum_{j=1}^{b}(\bar{x}_j - \bar{\bar{x}})^2$$

$$\bar{v} = \frac{1}{b}\sum_{j=1}^{b}s_j^2.$$

$\bar{\bar{x}}$ and s_X^2 are the sample mean and sample variance of the batch means, and $\bar{v}$ is the sample mean of the batch variances.

We now define the point estimator as

$$\hat{\sigma}_3^2 = \bar{v} + s_X^2.$$

We know that $\bar{v}$ is an unbiased estimator of $\mathrm{E}[S_j^2] = \mathrm{E}[\frac{m-1}{m}S^2(m)]$, and that s_X^2 is an unbiased estimator of $\mathrm{Var}[\overline{X}]$. So we can see that

$$\begin{aligned}\mathrm{E}[\hat{\sigma}_3^2] &= \mathrm{E}[\bar{v}] + \mathrm{E}[s_X^2] \\ &= \mathrm{E}[\frac{m-1}{m}\,S^2(m)] + \mathrm{Var}[\overline{X}] \\ &= \sigma^2,\end{aligned}$$

To obtain a confidence interval, we consider the statistic

$$y_j = s_j^2 + \frac{b}{b-1}\left(\bar{x}_j - \bar{\bar{x}}\right)^2,$$

whose sample mean is equal to the point estimate $\hat{\sigma}_3^2$. Because we can express the variance estimate as the sample mean of a random sample, we can justify calculating the confidence interval using the sample variance of this sample:

$$s_Y^2 = \frac{1}{b-1}\sum_{j=1}^{b}\left(y_j - \hat{\sigma}^2\right)^2.$$

The half-width of the confidence interval is then

$$\Delta_3 = t_{b-1,1-\alpha/2}\sqrt{s_Y^2/b}.$$

Finding an appropriate batch size m is crucial in this estimator. If the batches are too small, the y_j will be correlated, and their sample variance s_Y^2 does not accurately describe the variance of the estimate $\hat{\sigma}_3^2$.

To determine the batch size we use a simple procedure which tests the first autocorrelation coefficients of the sequence of batch means, continually increasing the batch size until the correlation becomes negligible.

EXPERIMENTAL COVERAGE ANALYSIS

To assess the performance of the proposed estimators, we use the method of sequential coverage analysis, as described in (Pawlikowski et al., 1998). The estimators were implemented using the Akaroa2 framework (Ewing et al., 1999).

We analyse the steady-state variance of the waiting times of customers in the single server queueing models M/M/1, M/E_2/1, and M/H_2/1. The model parameters are selected such that the coefficient of variation of the service time is 1 for the M/M/1 model, $\sqrt{0.5}$ for M/E_2/1, and $\sqrt{5}$ for M/H_2/1.

Simulations are run sequentially and stopped upon reaching a relative precision of 0.05 at the 0.95 confidence level. The coverage of confidence intervals is then calculated from the frequency with which the generated confidence interval covers the actual (theoretical) variance, and a confidence interval for the coverage is calculated at the 0.95 confidence level. Our procedure to analyse coverage follows the rules proposed in (Pawlikowski et al., 1998):

1. Coverage analysis is done sequentially.
2. Our results include a minumum number of 200 “bad” confidence intervals.
3. Simulation runs shorter than the mean run length minus one standard deviation are discarded.

Each of the estimators is used on the three models at different system loads, ranging from 0.1 to 0.9. To deal with bias due to the initial transient period of the simulation we use the method described in (Pawlikowski, 1990), which is a combination of a heuristic and a statistical test for stationarity of the output sequence.

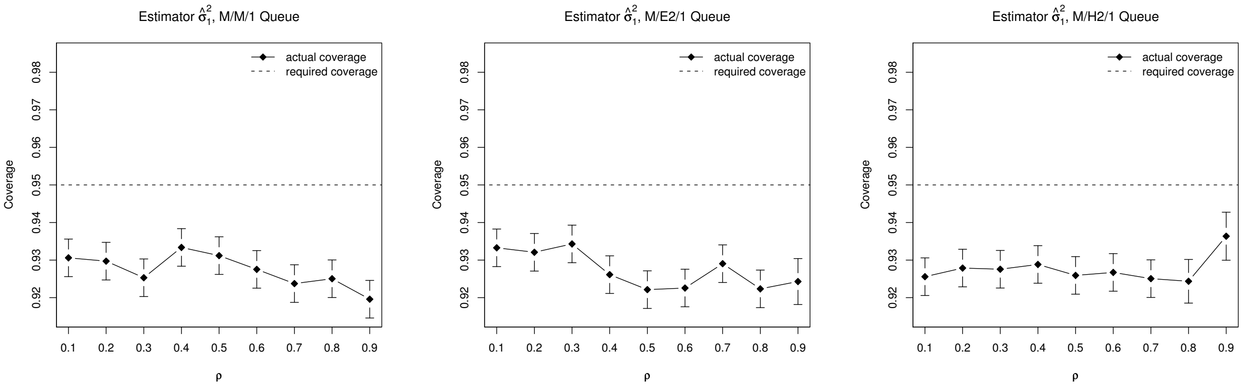
The Estimator $\hat{\sigma}_1^2$ (Figure 1) shows a coverage of around 0.93, which is slightly lower than the required 0.95, but still acceptable for most practical purposes. Because of its easy implementation (using an existing method of mean value estimation), this estimator can be interesting for practical use.

The estimator $\hat{\sigma}_2^2$ has generally good coverage (Figure 2), which shows that the method of “far apart” uncorrelated observations works.

Of the three estimators proposed, $\hat{\sigma}_3^2$ has the best coverage of confidence intervals (Figure 3).

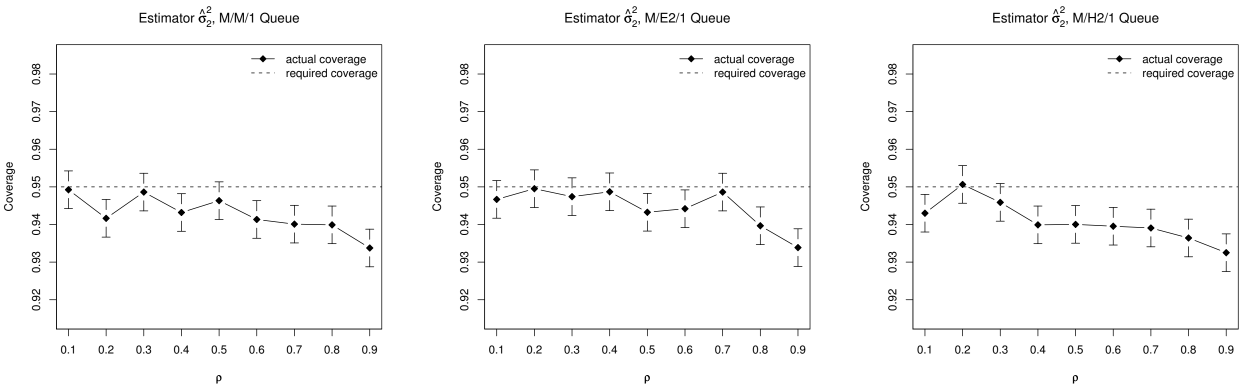
In addition to the coverage of confidence intervals, we examine the convergence of the estimators on the basis of the number of observations needed to reach the required precision. Table 1 shows the mean number of observations collected before reaching that precision. Overall, Method 3 needs the fewest observation before the stopping criterion is satisfied, Method 2 needs the most. With increasing system load the simulation run lengths of all three estimators show a similar behaviour. It is worth noting that the run lengths in Table 1 are roughly an order of magnitude greater than those that would be required for estimating the mean waiting times to the same precision.

For the estimator $\hat{\sigma}_2^2$, we analyse the value of the parameter k_0, which is automatically determined during the estimation process. Table 2 shows the average spacing used to produce almost independent observations. We see that the models with a higher coefficient of variation of the service



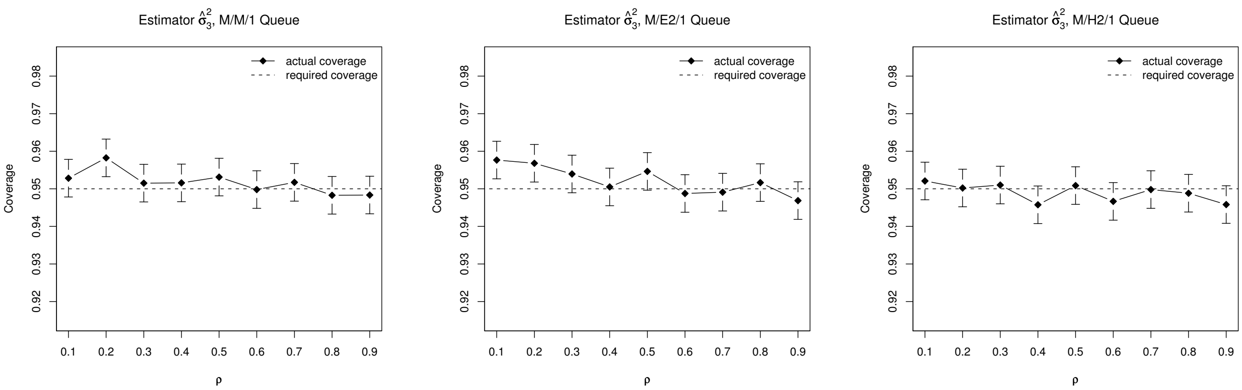
(a) M/M/1 Queue (b) M/E_2/1 Queue (c) M/H_2/1 Queue

Figure 1: Coverage of Estimator $\hat{\sigma}_1^2$



(a) M/M/1 Queue (b) M/E_2/1 Queue (c) M/H_2/1 Queue

Figure 2: Coverage of Estimator $\hat{\sigma}_2^2$



(a) M/M/1 Queue (b) M/E_2/1 Queue (c) M/H_2/1 Queue

Figure 3: Coverage of Estimator $\hat{\sigma}_3^2$

ρ	M/M/1 Queue			M/E_2/1 Queue			M/H_2/1 Queue		
	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_3^2$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_3^2$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_3^2$
0.1	160	570	155	117	407	112	577	1736	552
0.2	131	365	123	94	265	89	556	1302	517
0.3	140	339	131	101	248	95	624	1291	567
0.4	172	372	158	125	275	118	752	1437	672
0.5	234	461	214	172	346	160	973	1744	862
0.6	359	652	325	266	494	245	1383	2342	1222
0.7	636	1078	571	481	826	435	2317	3646	2010
0.8	1491	2331	1322	1134	1776	996	5019	7282	4317
0.9	6237	9096	5509	5079	6942	4146	22044	26930	16750

Table 1: Mean Number of Observations Needed to Reach Required Precision (in 1000 Observations)

ρ	M/M/1	M/E_2/1	M/H_2/1
0.1	6.6	6.2	10.7
0.2	8.6	7.8	17.4
0.3	11.9	10.4	27.4
0.4	17.4	14.6	43.3
0.5	26.4	21.7	70.2
0.6	43.5	35.1	119.8
0.7	80.5	64.0	228.3
0.8	188.2	146.5	535.8
0.9	769.4	589.5	2213.5

Table 2: Mean Value of k_0 in Estimator $\hat{\sigma}_2^2$

time need larger values of k_0 due to the higher correlation of the output sequence.

CONCLUSIONS AND FUTURE WORK

In this paper, we introduced three different approaches for estimating the steady-state variance in discrete event simulation. We applied the estimators to single-server queueing models and compared them in terms of coverage and sample size needed to obtain confidence intervals of a certain precision.

So far, the proposed estimators have only been tested on single-server queues. Although we expect them to be applicable to a much broader range of models, further study is required to confirm this.

Methods 1 and 3 estimate the variance as the mean of a secondary sequence of observations. This suggests that the method of Multiple Replications in Parallel, as proposed in (Pawlikowski et al., 1994), can be used to speed up variance estimation by utilising multiple computers in parallel. While this is indeed confirmed by preliminary experiments reported in (Schmidt, 2008), further study is needed.

Method 2 has the obvious weakness of only using a fraction of the generated observations. Investigation is needed to determine if the method can be improved so that it does not discard as many observations.

Method 3 required the smallest number of observations and produced estimates with good coverage properties.

The sample sizes required for a certain precision turned out to be substantially larger than those needed to estimate comparable means, so selecting the best estimator is an important problem.

The problem of the initial transient period has not been addressed with regard to the steady-state variance. The method used in this paper (Pawlikowski, 1990) was designed for the estimation of mean values. Analysis of the warm-up period in estimation of steady-state variance is a subject of our future studies.

ACKNOWLEDGEMENTS

The resarch leading to this paper was in part supported by the Summer Scholarship Programme of the University of Canterbury in Christchurch, New Zealand. In addition, the first author received financial support from the German Academic Exchange Service (DAAD).

References

Anderson, T. (1971). *The Statistical Analysis of Time Series*. John Wiley & Sons.

Chen, B. and Sargent, R. (1990). Confidence interval estimation for the variance parameter of stationary processes. *Management Science*, 36(2):200–211.

Eickhoff, M., Pawlikowski, K., and McNickle, D. (2007). Detecting the duration of initial transient in steady state simulation of arbitrary performance measures. In *Proceedings of ACM ValueTools07 (Nantes)*.

Ewing, G., Pawlikowski, K., and McNickle, D. (1999). Akaroa2: Exploiting network computing by distributing stochastic simulation. In *Proceedings of the European Simulation Multiconference (ESM 99, Warsaw)*, pages 175–181. International Society for Computer Simulation.

Feldman, R., Deuermeyer, B., and Yang, Y. (1996). Estimation of process variance by simulation. unpublished paper.

Heidelberger, P. and Welch, P. (1981). A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM*, 25:233–245.

Law, A. and Kelton, W. (1991). *Simulation Modeling and Analysis*. McGraw-Hill, 2nd edition.

Pawlikowski, K. (1990). Steady-state simulation of queueing processes: A survey of problems and solutions. *ACM Computing Surveys*, 22(2):123–170.

Pawlikowski, K., Ewing, G., and McNickle, D. (1998). Coverage of confidence intervals in sequential steady-state simulation. *Journal of Simulation Practice and Theory*, 6(3):255–267.

Pawlikowski, K., Yau, V., and McNickle, D. (1994). Distributed stochastic discrete-event simulation in parallel time streams. In *Proceedings of the 1994 Winter Simulation Conference*, pages 723–730.

Schmidt, A. (2008). Automated analysis of variance in quantitative simulation by means of multiple replications in parallel. Master's thesis, RWTH Aachen University.

Schruben, L. (1983). Confidence interval estimation using standardized time series. *Operations Research*, 31(6):1090–1108.

Steiger, N. and Wilson, J. (2001). Convergende properties of the batch means method for simulation output analysis. *Informs Journal on Computing*, 13(4):277–293.

Tanenbaum, A. (2003). *Computer Networks*. Prentice Hall, 4th edition.

Wilks, S. (1962). *Mathematical Statistics*. John Wiley & Sons, 2nd edition.

AUTHOR BIOGRAPHIES

ADRIAAN SCHMIDT is a 2008 graduate of Electrical Engineering, with a specialisation in Technical Computer Science, from RWTH Aachen University. His research interests include software for embedded and real-time systems, stochastic simulation, and sub-symbolic AI. His email address is <adriaan.schmidt@rwth-aachen.de>.

KRZYSZTOF PAWLIKOWSKI is a Professor of Computer Science and Software Engineering at the University of Canterbury. His research interests include quantitative stochastic simulation; and performance modelling of telecommunication networks. He received a PhD in Computer Engineering from the Technical University of Gdansk; Poland. He is a Senior Member of IEEE and a member of SCS and ACM. His email address is <krys.pawlikowski@canterbury.ac.nz> and his webpage is <http://www.cosc.canterbury.ac.nz/krys/>.

DON MCNICKLE is an Associate Professor of Management Science in the Department of Management at the University of Canterbury. His research interests include queueing theory; networks of queues and statistical aspects of stochastic simulation. He is a member of INFORMS. His email address is <don.mcnickle@canterbury.ac.nz>.