# Generating confidence intervals for spatial simulations – determining the number of replications for spatial terminating simulations

[1]**Robert M. Itami,** [2]**Darrel Zell,** [3]**Frank Grigel,** [4]**Randy Gimblett**

[1]**GeoDimensions Pty Ltd, E-Mail: bob.itami@geodimensions.com.au,** [2]**Parks Canada, E-Mail: Darrell.Zell@pc.gc.ca,** [3]**Parks Canada, E-Mail: Frank.Grigel@pc.gc.ca,** [4]**University of Arizona, E-Mail: gimblett@ag.arizona.edu**

*Keywords: Spatial Simulation, Simulation Statistics, Simulation Output Analysis*

## EXTENDED ABSTRACT

A standard problem in the analysis of outputs from terminating simulations is the need to determine the number of replications needed to construct confidence intervals for performance indicators from the simulation (Law and Kelton, 2000). In traditional industrial applications of simulation such as manufacturing and queuing simulations a single mean for each performance indicator is all that is needed. In spatial simulations however, the problem is more complex as performance indicators can vary spatially as in the case of travel simulations where performance indicators for each destination must be analysed. This paper presents three alternative methods recommended in the simulation literature for determining the number of replications required to obtain confidence intervals based for a given alpha level and user defined confidence interval half width or relative preceision. The problem of measuring multiple performance indicators is addressed with a short discussion of the Bonferroni Correction. These methods are then adapted to spatial simulations using a travel simulation for Banff, Yoho, Kootenay and Jasper National Parks as an example. Outputs for daily link Use and daily link encounters are examined applying different values for absolute accuracy and relative precision. Conclusions are then drawn on the relationship between the sensitivity of performance indicators to random variables in the simulation model and the specification of absolute accuracy and relative precision for spatial dynamic simulation models.

## 1. INTRODUCTION

Spatial simulation models use random numbers to generate input variables such as arrival times, durations at destinations, and trip selection. Because of this, it is not recommended to draw any conclusions from the output of a single replication of a simulation model since it represents only one realization of a stochastic process. This is a common characteristic of all simulation models, however spatial simulations have added complexity because output measures such as number of visits or average duration at nodes are measured across many different links and nodes. To solve the problem of variability in simulation outputs from random inputs, the method of independent replications (Goldman 1992) is used. In this method one observation per replication is generated (usually the mean value of the performance indicator). In the case of spatial simulations you would have one observation per link or node of interest. The simulation is replicated a number of times to generate confidence intervals that meet a certain reliability (usually 0.90 or 0.95) and a given accuracy (measured in the same units as the performance indicator). The problem is how to determine how many replications are required to meet the user specified reliability and accuracy.

## 2. STATISTICAL METHODS FOR DETERMINING THE NUMBER OF REPLICATIONS FOR TERMINATING SIMULATIONS

The method of independent replications requires the model is run for a "small" number of replications. In the case of probabilistic simulations, "small" may mean 10 to 15 replications.

The next step is to calculate the (1-α) confidence interval using equation 1. (See Centeno and Reyes, 1998 and Law and Kelton, 2000 pp 253-259)

**Equation 1.**

$$\overline{X}\left(n\right) \pm t_{1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}$$

WHERE $\overline{X}$ is the mean of the performance indicator for the current replication

$t_{1-\alpha/2}$ is the is the (1-α) percentile of the t-student distribution with n-1 degrees of freedom

$S^2(n)$ is the sample variance

In equation 1 the expression: $t_{1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}$ is referred to as the confidence interval half width. If this value is less than the user specified accuracy after the initial n replications for the "short run" then there is no need for further replications. However if this value is larger then the user specified accuracy, then n can be estimated using Equation 2.

**Equation 2.**

$$n^* = Round\left[ n \times \left(\frac{h}{h^*}\right)^2 \right]$$

WHERE $n^*$ is the estimated number of replications needed

h is the half width from the sample run

h* is the desired half width or absolute accuracy specified by the user.

Law and Kelton (2000, p512) suggest a modification of the above estimate in a method they call the iterative method. In this case the number of replications is increased by 1 each time and the confidence interval is recomputed after each iteration until the desired accuracy is achieved. This method assumes that the population variance will not change (appreciably) as the number of replications increase.

**Equation 3.**

$$n_a^*\left(\beta\right) = \min\left\{ i \geq n : t_{i-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{i}} \leq \beta \right\}$$

WHERE $n_a^*\left(\beta\right)$ is the estimated number of replications needed with absolute accuracy $\beta$

n is the number of replications from the "short run"

$t_{i-1,1-\alpha/2}$ is the is the (1-α) percentile of the t-student distribution with i-1 degrees of freedom

$S^2(n)$ is the sample variance from n replications

i is the iteration (greater than n)

A third method described by Law and Kelton (2000, p. 513) uses a measure called "relative accuracy". Relative accuracy is the Confidence

Interval Half Width from Equation 1 divided by $\overline{X}$. is an estimate of the actual relative error. In this method the user specifies a desired relative error $\lambda'$.

**Equation 4.**

$$n_r^*(\lambda) = \min\left\{ i \geq n : \frac{t_{i-1,1-\alpha/2}\sqrt{S^2(n)/i}}{|\overline{X}(n)|} \leq \lambda' \right\}$$

WHERE $n_r^*(\lambda)$ is the estimated number of replications needed with relative accuracy $\lambda$

N is the number of replications from the "short run"

$\lambda'$ is the user-specified relative accuracy

$t_{1-\alpha/2}$ is the is the $(1-\alpha)$ percentile of the t-student distribution with n-1 degrees of freedom

$S^2(n)$ is the sample variance from n replications

i is the iteration (greater than n)

## 3. MULTIPLE PERFORMANCE INDICATORS AND THE BONFERRONI CORRECTION

In a typical spatial simulation, there is normally more than one performance indicator being measured. If simulations are viewed as experiments where we are testing hypotheses about the system under study, then the alpha levels of the statistical tests applied in the previous section must be adjusted using the Bonferroni Correction (MathWorld, 2005). The Bonferroni Correction is used when several tests are being performed simultaneously. Where a given alpha level for a single performance indictor may be appropriate, it is not for the set of all comparisons. The simplest form of the Bonferroni Correction is to take the desired alpha level and divide by the number of performance indicators being tested. Thus, if the desired alpha level is 0.10 and there are five performance indicators, the adjusted alpha level would be 0.10/5 or 0.02 for each test.

Law and Kelton (2000) suggest another approach in which the sum of the alpha levels for each test equals the desired alpha level, suggesting that each performance indicator can have a unique alpha level in the statistical test. They also suggest that more than 10 performance indicators is impractical and that, given the stochastic nature of simulation it may be impractical to meet statistical requirements for all performance indicators simultaneously and that one must just have to accept that some indicators may not be used reliably in drawing conclusions from the simulation.

## 4. DETERMINING THE NUMBER OF REPLICATIONS FOR SPATIAL SIMULATIONS

In traditional industrial applications of simulation such as manufacturing and queuing simulations a single mean for each performance indicator is all that is needed. In spatial simulations however, the problem is more complex as performance indicators can vary spatially as in the case of travel simulations where performance indicators for each destination must be analysed simultaneously. Essentially, the approach required is to apply the same statistical methods described in section 2 to each and every location in the spatial simulation where performance indicators are to be measured. For instance, in travel simulations we may be interested in:

- the total visits per destination,
- average visit duration per destination, and
- average queuing times at parking facilities.

In this case, we have three performance indicators. Suppose our network has 10 destinations we wish to evaluate. We must first determine the alpha level we wish to test for. If the overall alpha is 0.10 then, according to the Bonferroni Correction we must use an alpha level of 0.10/3 or 0.03.

Next we decide which method will be used to determine the number of replications either by specifying the confidence interval half width for equations 2 and 3 or by specifying the relative accuracy for equation 4.

The simulation is then replicated for a "short run" of say, 10 replications and the outputs from each of the 10 destinations for the 3 performance indicators are gathered. Using this output we then apply the corresponding method (equations 2, 3 or 4) for each of the 10 destinations for the 3 performance indicators using an alpha level of 0.03.

## 5. EXAMPLE: PATTERN OF USE SIMULATION FOR CANADAS 4 MOUNTAIN PARKS

A probabilistic spatial simulation of tourist flows in Banff, Jasper, Kootenay and Yoho National Parks in the Canadian Rocky Mountains was developed using RBSim software. Data from independent travelers was collected in 2003 in a survey conducted by Parks Canada over the entire year. The 2003 Mountain Parks Study provided

the trip itinerary data for a trace simulation. Out of the 13,373 first time arrivals to the Mountain parks, 9348 respondents agreed to complete the questionnaire. Out of the 9348 respondents, 2383 questionnaires were returned, resulting in a 25.5 percent return rate. Out of those 2383 questionnaires, 1982 respondents actually completed the trip diary component (necessary to develop the trip itineraries for the trace simulation), resulting in a 21.2 percent return rate.

The survey trips represent a "single realization" of the full variation of travel patterns in the four parks. To use a pattern of use simulation for management, a probabilistic model of travel patterns must be developed from analysis of survey trips. Trip itineraries vary within the constraints of seasonal patterns, which will constrain certain activities and access to destinations because of weather. Trips must be grouped or clustered in order to classify trip itineraries according to seasonal variations in pattern of use. In order to do this, a number of different cluster analysis techniques were applied to the 1620 survey trips using a number of different clusters and against different definitions of a "season". The strongest clustering results came from a technique called "kmeans". In this technique, the analyst must nominate the number of clusters. 3 and 4 clusters were tested with 3 clusters differentiating the winter season trips most distinctly. The other trips were clustered on the basis of the type of activity: either active (hiking, biking etc) or inactive.

Based on analysis of the survey trips looking at sample size during each week of the year and the results of the cluster analysis, two periods of the year were selected to develop a probabilistic simulation from survey trip itineraries and traffic counts. The winter period selected is the month of January 2003. The summer period selected is 4 weeks beginning June 23, 2003 and ending July 24, 2003. We report on the January results in this paper.

The next step in the development of the probabilistic simulation is to randomly assign, for each day of the simulation, the correct number of trip arrivals to each day of the simulation and then scheduling the exact minute of arrival based on the hourly arrival distribution. A trip itinerary must be randomly selected from a pool of trips specified from the cluster analysis for the entry gate and time of year. This is done for each gate by proceeding from the first day of the simulation to the last day of the simulation according to the following procedure:

For each day,

1. The week of arrival is determined (1 through 53)

2. The week day is determined (Monday through Sunday).

3. The total number of arrivals for the week for the current gate is determined from the weighted traffic count data for the week.

4. The weekly arrival distribution for the current week is selected

5. The total number of arrivals for the week is multiplied by each day in the weekly arrival distribution. This determines the number of arrivals for each day of the week.

6. The hourly distribution is then selected

7. The exact time of arrival for the current trip is selected from the hourly arrival distribution by generating a random number and calculating the exact minute of arrival by piecewise linear interpolation from the hourly arrival distribution.

8. Finally, the trip clusters that occur in this week are looked up and then all trips that fall in these clusters for the current entry gate are selected to create the pool of trips. Each trip in the pool has an equal chance of being selected. A random number is generated and a single trip is selected from the pool

9. The trip selection is now complete and the process is repeated for the next trip until all trips for the current day are selected. If all trips are scheduled, the process progresses to the next day of the simulation and the process is repeated until the last day of the simulation.

The first two weeks of the simulation outputs must be discarded because the simulation starts with no visitors so any outputs during an initial "warm up" period are not representative of system behavior. To correct this problem, the simulation is allowed to run until the system comes to full capacity. In this case the period that was selected was two weeks since the longest visits are 11 days. The warm up period gives the simulation time to populate the system to capacity before collecting statistics from the simulation. For the winter simulation 1 week of simulation was obtained for analysis.

Two performance indicator were measured for links: Link Use and Link Encounters. Link use is a frequency count for the number of parties visiting each link for each day of the simulation. Link Encounters is the number of direct contacts between parties along a link.

There were 658 links visited and 493 links with encounters recorded in the simulation. Since daily link use and link encounters were generated, there were 3799 link use days and 2083 link encounter days recorded with confidence intervals calculated for each.

For purposes of understanding the implications of the three methods for estimating the number of replications needed to obtain user specified measures of confidence and reliability from the simulation, the results of two links are shown in the tables below. Confidence intervals for the 7 replications for daily link use and daily link encounters were calculated using different alpha values and different user-specified confidence half-width values to investigate the impact these values have on the three methods for estimating the number of replications.

|  | Date | Link Use | | Link Encounters | |
|---|---|---|---|---|---|
| Link | Jan 03 | Mean | StDev | Mean | StDev |
| 116 | 15 | 380.71 | 20.18 | 2.86 | 3.23 |
| 116 | 16 | 387.00 | 15.00 | 3.86 | 5.64 |
| 116 | 17 | 401.57 | 25.55 | 3.71 | 2.86 |
| 116 | 18 | 432.86 | 21.79 | 3.71 | 5.31 |
| 116 | 19 | 448.71 | 20.88 | 6.29 | 6.82 |
| 126 | 15 | 559.14 | 18.41 | 32.86 | 16.23 |
| 126 | 16 | 580.14 | 19.73 | 48.71 | 31.44 |
| 126 | 17 | 595.43 | 28.58 | 32.29 | 35.96 |
| 126 | 18 | 679.00 | 16.41 | 86.71 | 74.31 |
| 126 | 19 | 670.71 | 17.73 | 25.14 | 28.74 |

Table 1. Means and Standard Deviations for Link Use and Link Encounters

Table 1 shows the means and standard deviations for link use and link encounters for two links. Link 116 and Link 26. Link 116 is typified by low number of encounters as compared to Link 126. Note that standard deviations for link encounters are relatively high compared to link use. This is because link encounters are much more sensitive to random variations than link use since small changes in arrival times and volumes can change the number of encounters on a link.

Table 2 shows the confidence interval half width (CI Half) and relative precision (Rel prec) from confidence intervals calculated with an alpha of 0.10 (90% confidence). Law and Kelton (2000) suggest that an alpha of 0.10 is a reasonable level of confidence given that random number are used to generate the performance indicators. Note the large values for relative precision for link encounters as compared to link use. Remember that the relative precision is the ratio of the

confidence interval half width with the mean. This shows the value of this measure to indicate the amount of variation between simulation runs.

|  | Date | Link Use | | Link Encounters | |
|---|---|---|---|---|---|
| Link | Jan-03 | CI Half | Rel Prec | CI Half | Rel Prec |
| 116 | 15 | 14.82 | 0.04 | 2.37 | 0.83 |
| 116 | 16 | 11.02 | 0.03 | 4.14 | 1.07 |
| 116 | 17 | 18.76 | 0.05 | 2.10 | 0.57 |
| 116 | 18 | 16.00 | 0.04 | 3.90 | 1.05 |
| 116 | 19 | 15.33 | 0.03 | 5.01 | 0.80 |
| 126 | 15 | 13.52 | 0.02 | 11.92 | 0.36 |
| 126 | 16 | 14.49 | 0.02 | 23.09 | 0.47 |
| 126 | 17 | 20.99 | 0.04 | 26.41 | 0.82 |
| 126 | 18 | 12.05 | 0.02 | 54.57 | 0.63 |
| 126 | 19 | 13.02 | 0.02 | 21.10 | 0.84 |

Table 2. Confidence Half Intervals and Relative Precision for Link Use and Link Encounters, Alpha = 0.10

|  | Date | Link Use | | | Link Encounters | | |
|---|---|---|---|---|---|---|---|
| Link | Jan 03 | Eq. 2 | Eq. 3 | Eq. 4 | Eq. 2 | Eq. 3 | Eq. 4 |
| 116 | 15 | 1583 | 1113 | 7 | 39 | 31 | 156 |
| 116 | 16 | 850 | 616 | 7 | 120 | 88 | 260 |
| 116 | 17 | 2465 | 1784 | 7 | 31 | 25 | 74 |
| 116 | 18 | 1792 | 1297 | 7 | 106 | 79 | 249 |
| 116 | 19 | 1646 | 1192 | 7 | 176 | 128 | 144 |
| 126 | 15 | 1280 | 927 | 7 | 994 | 720 | 32 |
| 126 | 16 | 1470 | 1064 | 7 | 3731 | 2701 | 152 |
| 126 | 17 | 3084 | 2232 | 7 | 4883 | 3534 | 152 |
| 126 | 18 | 1016 | 736 | 7 | 20848 | 15089 | 91 |
| 126 | 19 | 1187 | 860 | 7 | 3118 | 2257 | 159 |

Table 3. Estimated replications using Equations 2, 3 and 4 for Link Use and Link Encounters, Alpha = 0.10, User CI Half Width = 1, User Relative Precision = 0.15

Table 3 shows the estimated number of replications for link use using the three different methods described in this paper. For equations 2 and 3 an absolute accuracy (the user specified confidence interval half width) of 1 is used. For Equation 4 a user specified relative accuracy of 0.15 is used. For link use equation 4 shows that there is no need for further replications whereas equations 2 and 3 show that many more replications are needed to reduce the confidence interval to a half width of 1. This is instructive,

because it may indicate that the criteria for absolute accuracy may be unrealistic given the high standard deviations for link use. Note that equation 3 (the iterative method) shows a consistently lower estimate for the number of replications required to generate confidence intervals with a half width of 1.

In Table 3 the number of replications required for link encounters is greater for Equation 4 than equations 2 and 3. This is because there are relatively few encounters per link (see Table 1) for link 116. However for relative accuracy, we see that we would require anywhere from 74 to 260 replications in order to reach a relative precision of 0.15 for link 116. This reflects the high standard deviations as compared to the means for link encounters for link 116 and the resulting high relative precision. For link 126 there were many more encounters per day than link 116. The result is that we have a larger confidence interval half width, which requires many more replications to achieve our user specified absolute accuracy of 1. This shows how important it is to carefully select appropriate standards of accuracy for each performance indicator.

To examine the impact of the Bonferroni Correction, we now assume that in order to achieve an overall alpha of 0.10 for the simulation, we need to estimate replications using an alpha level of 0.05 for each of our two performance indicators.

|      | Date   | Link Use |             | Link Encounters |             |
|------|--------|----------|-------------|-----------------|-------------|
| Link | Jan-03 | CI Half  | Rel Prec    | CI Half         | Rel Prec    |
| 116  | 15     | 18.67    | 0.05        | 2.98            | 1.04        |
| 116  | 16     | 13.88    | 0.04        | 5.22            | 1.35        |
| 116  | 17     | 23.63    | 0.06        | 2.65            | 0.71        |
| 116  | 18     | 20.15    | 0.05        | 4.91            | 1.32        |
| 116  | 19     | 19.31    | 0.04        | 6.31            | 1.00        |
| 126  | 15     | 17.03    | 0.03        | 15.01           | 0.46        |
| 126  | 16     | 18.25    | 0.03        | 29.07           | 0.60        |
| 126  | 17     | 26.43    | 0.04        | 33.26           | 1.03        |
| 126  | 18     | 15.17    | 0.02        | 68.73           | 0.79        |
| 126  | 19     | 16.40    | 0.02        | 26.58           | 1.06        |

Table 4.  Confidence Half Intervals and Relative Precision for Link Use and Link Encounters, Alpha = 0.05

Table 4 shows, as we would expect, that the confidence interval half widths and relative precisions have increased because of the lower alpha value.

|      | Date   | Link Use |          |          | Link Encounters |          |          |
|------|--------|----------|----------|----------|-----------------|----------|----------|
| Link | Jan-03 | Eq. 2    | Eq. 3    | Eq. 4    | Eq. 2           | Eq. 3    | Eq. 4    |
| 116  | 15     | 24       | 19       | 7        | 7               | 7        | 221      |
| 116  | 16     | 13       | 12       | 7        | 7               | 7        | 370      |
| 116  | 17     | 39       | 28       | 7        | 7               | 7        | 104      |
| 116  | 18     | 28       | 21       | 7        | 7               | 7        | 354      |
| 116  | 19     | 26       | 20       | 7        | 7               | 7        | 204      |
| 126  | 15     | 20       | 16       | 7        | 16              | 13       | 44       |
| 126  | 16     | 23       | 18       | 7        | 59              | 41       | 74       |
| 126  | 17     | 49       | 34       | 7        | 77              | 53       | 215      |
| 126  | 18     | 16       | 13       | 7        | 331             | 200      | 128      |
| 126  | 19     | 19       | 15       | 7        | 49              | 35       | 226      |

Table 5.  Estimated replications using Equations 2, 3 and 4 for Link Use and Link Encounters, Alpha = 0.05, User CI Half Width = 10, User Relative Precision = 0.15

In Table 5 we see a significant drop in the number of replications required from equations 2 and 3 for link use because we have now increased the user specified half width from 1 to 10. Equation 4 for link use shows no need for more replications because of the low relative precision values obtained as the result of the high use levels and small confidence interval half widths. For link encounters no further replications are required from equations 2 and 3 because of the low number of encounters and the high user specified confidence interval half width of 10. Equation 4 for node 116 still shows a large number of replications due to the high standard deviations as compared to the small confidence interval half widths (see Table 4). For Node 126 we see a large reduction in the number of replications as compared to table 3 primarily because of the increase in the user specified confidence interval half width of 10.

Now we look at the estimated replications for the entire network. Tables 6 and 7 show the summary of results for all links for link use and link encounters for an alpha of 0.05, and a confidence interval half width of 10. Note in table 6 it is possible to achieve the user desired accuracy for 90% of the links with less than 60 replications for link use whereas in table 7 it takes over 180 replications to achieve the same coverage for link encounters.

| Link Use Absolute Accuracy = 10 | | | |
|---|---|---|---|
| **Reps** | **Count** | **%Links** | **Accum%** |
| 1: 20 | 2926 | 77.02% | 77.02% |
| 21: 40 | 428 | 11.27% | 88.29% |
| 41: 60 | 111 | 2.92% | 91.21% |
| 61: 80 | 12 | 0.32% | 91.52% |
| 81:100 | 16 | 0.42% | 91.95% |
| 101:120 | 6 | 0.16% | 92.10% |
| 121:140 | 17 | 0.45% | 92.55% |
| 141:160 | 16 | 0.42% | 92.97% |
| 181:200 | 267 | 7.03% | 100.00% |
| **Total** | **3799** | **100.00%** | |

Number of replications for link use using equation 3 for alpha = 0.05, CI Half Width = 10 for entire network

| Link Encounters Absolute Accuracy = 10 | | | |
|---|---|---|---|
| **Reps** | **Count** | **%Links** | **Accum%** |
| 1: 20 | 1120 | 53.77% | 53.77% |
| 21: 40 | 169 | 8.11% | 61.88% |
| 41: 60 | 95 | 4.56% | 66.44% |
| 61: 80 | 66 | 3.17% | 69.61% |
| 81:100 | 61 | 2.93% | 72.54% |
| 101:120 | 45 | 2.16% | 74.70% |
| 121:140 | 36 | 1.73% | 76.43% |
| 141:160 | 33 | 1.58% | 78.01% |
| 161:180 | 26 | 1.25% | 79.26% |
| 181:200 | 432 | 20.74% | 100.00% |
| **Total** | **2083** | **100.00%** | |

*Table 6.   Number of replications for link encounters using equation 3 for alpha = 0.05, CI Half Width = 10 for entire network*

Tables 7 and 8 show the results of equation 4 with a relative accuracy of 0.15. A similar pattern is seen here with less than 40 replications needed for link use to achieve the relative precision for 0.15 for link use and over 500 replications required for link encounters. Since both link use and link encounters are generated in the same simulation the obvious method for selecting the required simulations is simply to take the maximum value from all estimates, which is this case is over 500 replications. However this ignores the costs of processing time. In this simulation, each replication takes around an hour to run. If we run the simulation for 500 replications this means 500 hours of computer time and with limitations on file sizes for output databases, it is likely the simulation will fail from reaching file size limits.

| Link Use - Relative Precision = 0.15 | | | |
|---|---|---|---|
| **Reps** | **Count** | **%** | **Accum%** |
| 1: 40 | 3493 | 91.95% | 91.95% |
| 41: 80 | 87 | 2.29% | 94.24% |
| 81:120 | 44 | 1.16% | 95.39% |
| 121:160 | 23 | 0.61% | 96.00% |
| 201:240 | 14 | 0.37% | 96.37% |
| 281:320 | 10 | 0.26% | 96.63% |
| 401:440 | 30 | 0.79% | 97.42% |
| 481:500 | 23 | 0.61% | 98.03% |
| 501: | 75 | 1.97% | 100.00% |
| **Total** | **3799** | **100.00%** | |

*Table 7.   Number of replications for link use using equation 4 for alpha = 0.05 relative accuracy = 0.15  for entire network*

| Link Encounters - Relative Precision = 0.15 | | | |
|---|---|---|---|
| **Reps** | **Count** | **%** | **Accum%** |
| 1: 40 | 57 | 2.74% | 2.74% |
| 41: 80 | 92 | 4.42% | 7.15% |
| 81:120 | 127 | 6.10% | 13.25% |
| 121:160 | 139 | 6.67% | 19.92% |
| 161:200 | 117 | 5.62% | 25.54% |
| 201:240 | 124 | 5.95% | 31.49% |
| 241:280 | 117 | 5.62% | 37.11% |
| 281:320 | 97 | 4.66% | 41.77% |
| 321:360 | 54 | 2.59% | 44.36% |
| 361:400 | 28 | 1.34% | 45.70% |
| 401:440 | 147 | 7.06% | 52.76% |
| 441:480 | 131 | 6.29% | 59.05% |
| 481:500 | 48 | 2.30% | 61.35% |
| 501: | 805 | 38.65% | 100.00% |
| **Total** | **2083** | **100.00%** | |

*Table 8.   Number of replications for link encounters using equation 4 for alpha = 0.05 relative accuracy = 0.15  for entire network*

## 6.   CONCLUSIONS

These results show a number of important characteristics of performance indicators and the impact on the number of replications required to meet user-specified reliability measures. First the more sensitive a performance measure is to random variation, the higher the variances and the

wider the confidence interval and therefore the greater number of replications required. Second, if using equations 2 and 3 for estimating the number of replications, it is important to carefully select the absolute accuracy (desired CI half width) for each performance indicator – it may be helpful to use the output analysis for the "short run" to help determine reasonable values for each performance indicator. Third, absolute accuracy and relative precision are two very different measures yielding very different results in terms of calculating the number of replications. It may be useful to evaluate both measures when estimating the number of replications for terminating simulations using the methods described in this paper.

When estimating the number of replications to obtain reliable results for management there is a tradeoff between the different performance measures described here and the processing time and limitations of output file sizes. The reliability of results will therefore vary with each performance indicator and with each location. If a trade-offs need to be made between accuracy and computing resources it is advisable to map out the reliability of the results for each performance indicator. The analyst and manager may simply have to live with a degree of uncertainty or accept that confidence interval widths will vary across space. At least, the methods presented in this paper give the analyst the tools needed to accurately measure this variation spatially.

## 7. REFERENCES

Centeno, Martha. and Reyes, M. Florencia (1998), "So you have your model: What to do next. A tutorial on simulation output analysis" in Proceedings of the 1998 winter Simulation conference, D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, editors. http://www.informs-cs.org

Goldsman, D. (1992), Simulation Output Analysis. In Proceedings of the 1992 Winter Simulation Conference, ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson, 97-103.

Law, Averill M. and Kelton, W. David (2000) Simulation Modeling and Analysis, Third Edition, McGraw Hill, Boston.