

The Impact of Using Excel Macros for Teaching Simulation Input and Output Analysis*

M. C. COURT

School of Industrial Engineering, University of Oklahoma, 202 West Boyd, Suite 124, Norman, OK 73019. E-mail: mcourt@ou.edu

Excel macros are used to teach simulation topics for an undergraduate operations research and simulation course, and a graduate simulation course emphasizing input and output analysis. The macros have the ability to perform simulation analysis when simulation packages are not available (i.e. they are independent of a specific simulation language) and do not require a large amount of CPU memory. These features have lent themselves well to supporting laptop and network independent learning environments. Statistical analysis shows that the macros have also improved overall grade point averages and have reduced the amount of instructional time required to teach these topics.

INTRODUCTION AND PROBLEM STATEMENT

IN 1995 the College of Engineering (COE) at the University of Oklahoma, began implementing a wireless, online-laptop program throughout all undergraduate engineering courses. The intrinsic consequence of switching from the traditional, lecture-based delivery system to the online-laptop classrooms was that changes had to be made in how courses were designed and taught. A common approach was to implement off-the-shelf software for use inside and outside the classroom, since access was no longer an obstacle for the instructors or the students.

However, access to simulation software for industrial engineering students was of great concern. In general, to support state-of-the-art learning in simulation modeling and analysis, professional versions of the software must be utilized. While industrial engineering students at COE had access to academic versions of simulation languages (e.g. Arena) and are allowed to load and maintain the software on their laptops, professional versions are maintained by the COE's Engineering Computer Services (ECS) via ECS network servers. The access to the professional versions of the software is limited to hardwired PCs in ECS computing labs where only a specific number (i.e. 25) of online users can simultaneously utilize the software.

These factors necessitated the development of laptop and online teaching tools aimed at equipping industrial engineering students with fundamental skills for performing simulation analysis. The tools needed to be independent of a specific

simulation language and could not require a large amount of CPU memory for execution. Excel macros were chosen as the best vehicle for developing the laptop simulation analysis and teaching tools since:

1. Access to Microsoft software during and after a student's academic career tends not be a concern.
2. In general, students are well adapted to the Excel environment and commands by their senior year of their undergraduate program.

Since 1997, Excel macros have been used by the instructor to:

- convey the analysis issues (and consequently ad-hoc decisions that must be made) when performing steady-state parameter estimation;
- compare different simulation output analysis techniques;
- demonstrate and explain coverage;
- demonstrate discrete-event simulation logic, etc.

Other macros address the analysis and decisions that must be made when fitting data to a distribution (input analysis). The macros are also used by students for supporting independent learning, to assist them in solving homework assignments and to answer exam questions.

A case study on the impact the Excel macros have had on the simulation courses at the School of Industrial Engineering is presented. An overview of the macros is presented along with the learning objective(s) that each macro or group of macros was designed to support. An analysis of student grades over the past six years on assignments and tests support the conclusions that the macros have improved overall grade point averages and have reduced the amount of instructional time required to teach these topics.

* Accepted 14 July 2004.

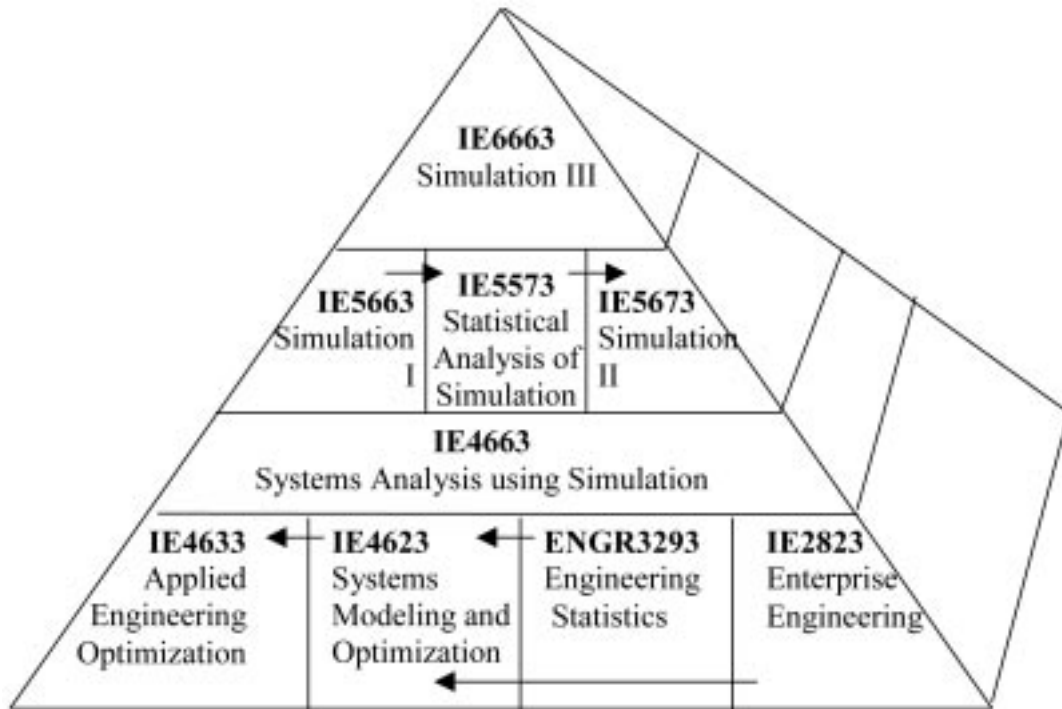


Fig. 1. Pyramid approach to simulation courses (modified from [1]).

SIMULATION COURSES, TOPICS, LEARNING OBJECTIVES AND MACROS

Figure 1 displays the current simulation course sequence as revised in the fall semester of 1995 and implemented in the fall semester of 1996 to support

the laptop program [1]. The simulation sequence can be viewed as a pyramid of knowledge, where the lowest level of the pyramid depicts the pre-requisite courses for the simulation course sequence. Notice that the pyramid represents the simulation sequence for both the undergraduate

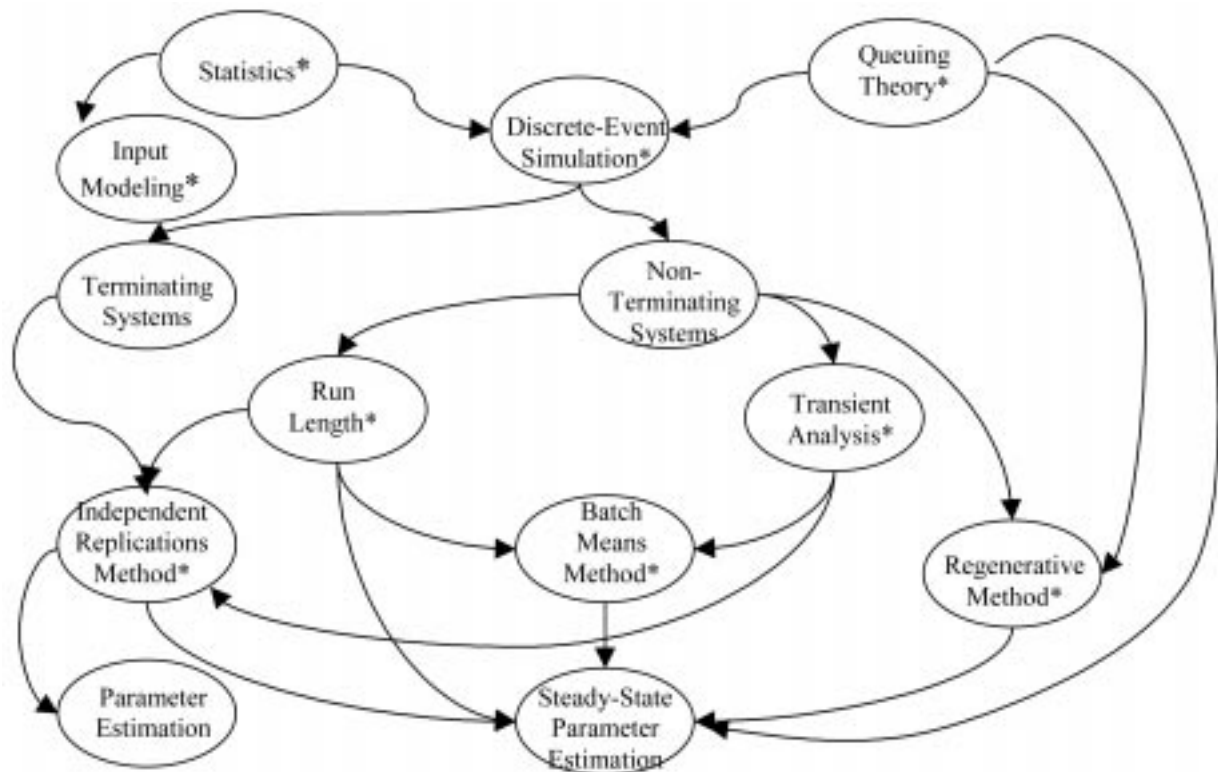


Fig. 2. Major topic links for statistics, simulation and operations research courses.

Table 1. Simulation learning objectives (* indicates support by Excel macros)

Course number and title	Major topic	Major simulation learning objective: At the end of this course, you should be able to . . .
IE4633, Engineering Optimization	Queuing theory	1. Generate a rate diagram to document the arrival process, queuing discipline and service mechanism of a system.
	Statistics	2. Use queuing theory to analyze for system parameters if a queuing model accurately represents the arrival process, queuing discipline and service mechanism of the system.
	Discrete-event simulation	3. Use the inverse transformation method to generate random variates.* 4. Understand that transient analysis must be performed before steady-state parameter estimation can be obtained via simulation analysis.* 5. Understand that transient analysis is ad-hoc and the transient itself is stochastic.* 6. Perform the method of independent replications to obtain confidence intervals on the mean of system parameters (e.g. Average queue time).* 7. Implement discrete-event logic to calculate unknown system parameters, e.g. Queue time, time in system and busy functions.* 8. Understand that simulation is a statistical experiment run on a computer and is thus, subjected to certain flaws (e.g. GIGO, one run of the model is only one observation of the system, not an optimization tool, etc.).
IE4663, Systems Analysis Using Simulation	Parameter estimation statistics	1. Analyze a system and the objectives of the simulation study to categorize the system as terminating or non-terminating.
		2. Understand that terminating systems have natural beginning and ending events that support the capturing of independent and identically distributed data from each run of the simulation model.
		3. Understand that the batch means method is a single-replication method for performing steady-state parameter estimation that requires transient and batch size analysis.*
		4. Use an autocorrelogram to identify 'lag0' for establishing the batch size when using the batch means method for steady-state parameter estimation.*
IE5573, Statistical Analysis of Simulation	Queuing theory statistics	1. Use the regenerative method to obtain steady-state parameter estimation for queues displaying regenerative behavior and understand that this method does not require transient analysis.*
		2. Understand that simulation run length and transient analysis impacts coverage, as well as the choice of the output analysis technique.*
		3. Understand that non-parametric tests must be performed on data before input modeling is performed.*
		4. Generate theoretical data sets to test the hypothesized data against the collected data.
		5. Perform of goodness-of-fit tests and interpret the results.*

and graduate programs. The lowest level consists of the operations research IE4623 (Systems Modeling and Optimization) and IE4633 (Applied Engineering Optimization) courses and a statistics course ENGR3293 (Engineering Statistics) taken in the sophomore and junior years. The second level is the simulation course IE4663 (Systems Analysis Using Simulation) taken in the senior year of the undergraduate program. The third level represents the first and second year of the graduate program in simulation. The first year's courses, IE5663 (Simulation I) and IE5573 (Statistical Analysis of Simulation), are also technical electives for the undergraduate program. The final level is the doctoral-level modeling course IE6663 (Simulation III) for the simulation sequence (only graduate students are allowed to enroll in the course).

Excel macros have been developed for simulation topics taught in IE4633, IE4663, and IE5573 (partial funding was provided by the National Science Foundation for the IE4663 macros). The Excel macros for IE4633 concentrate on supporting M/M/1 queuing analysis, random variate generation and steady-state parameter estimation via the independent replications method. The Excel macros for IE4663 are aimed at executing the steps required to perform the batch means method for steady-state parameter estimation.

The macros for IE5573 are to assist input analysis (statistical modeling and analysis tasks), to perform the regenerative method and are used to perform and compare simulation output analysis techniques in terms of coverage. Macros were first introduced in the IE4663 course in 1997, in the IE5573 course in 1998 and in the IE4633 course in 2002.

Figure 2 displays the major topic links that support learning objectives on simulation output analysis and input modeling. An asterisk in Fig. 2 indicates that a specific Excel macro has been developed to support a major simulation-learning objective (see Table 1). Learning objectives are active statements (e.g. at the end of this course, the student will be able to perform transient analysis) of what a student is supposed to accomplish.

Establishing clearly defined learning objectives assist the instructor in developing content and assessment tools for identifying knowledge gains and misconceptions (concepts learned incorrectly e.g., a student is able to calculate the effect but identifies the wrong cause for that effect). Learning objectives were first documented by Bloom in his handbook, *Taxonomy of Educational Objectives in the Cognitive Domain* [2]. Bloom's taxonomy states that there are hierarchical classifications of learning objectives where the student is expected to

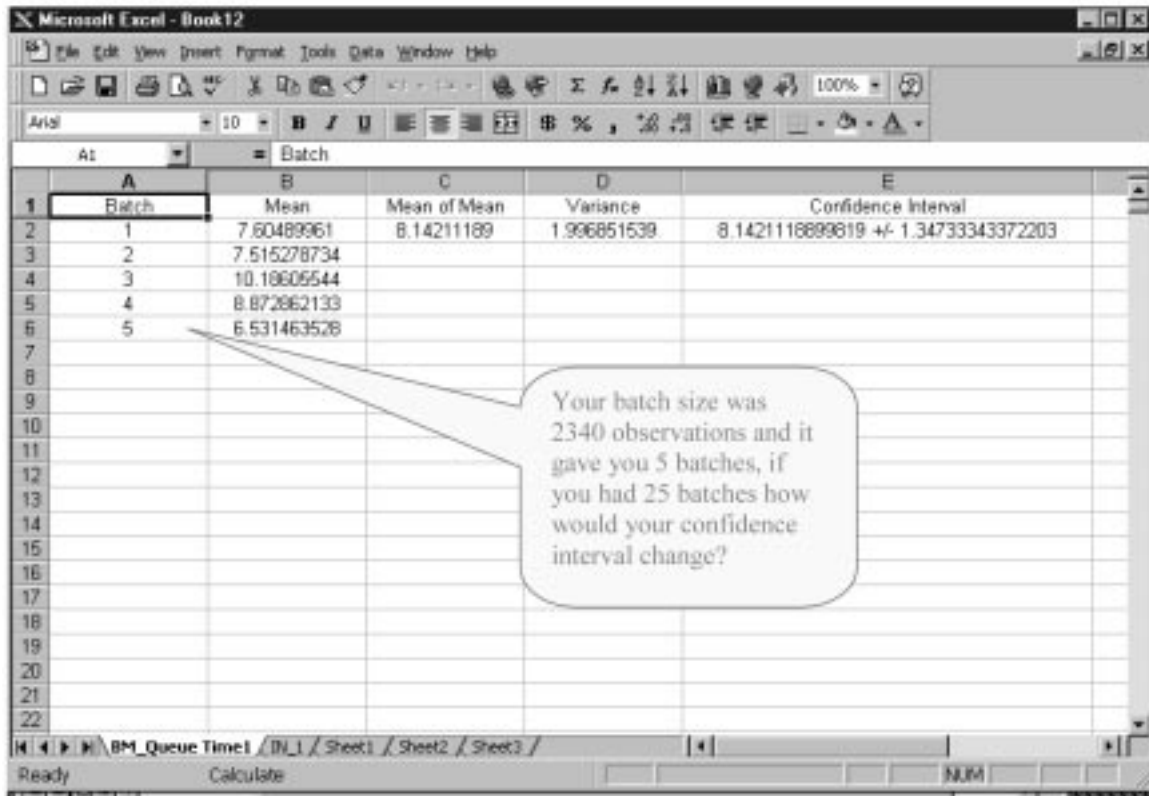


Fig. 3. Excel macro for performing batch means method.

complete the lower level learning before moving on to the next learning objective. The learning objectives from the lowest to the highest level of classification are:

1. *Knowledge*: the student can recall the information presented.
2. *Comprehension*: the student can restate the idea in different words.
3. *Application*: the student can apply the knowledge appropriately to solve a problem.
4. *Analysis*: the student can break a problem into its components and note the relationships of the components.
5. *Synthesis*: the student can rearrange component ideas into a new whole.
6. *Evaluation*: the student can make decisions based on the whole situation.

Even at the university level, most learning is found to be focused on the three lowest levels of learning [3, 4].

Consequently, another goal the instructor had for the development of the Excel macros was to bring the students to the three higher levels of learning. The macros are designed to assist the students with the application and prompt them with higher-level learning questions. For example, when a student utilizes the macro for performing the batch means method (a well-known output analysis technique in simulation analysis for generating point estimates and confidence intervals) in IE4663, a 'what-if' dialogue box opens, prompting

the student to ask him/herself about the impact the number of batches has on the confidence interval size (see Fig. 3). The 'what-if' impact analysis on the number of batches is necessary since the batch means method is ad-hoc. That is, determining the number of batches (and batch size) is a decision left to the simulation analyst (in this case, the IE4663 student). As the number of batches increases, the confidence interval length will decrease. However, in simulation analysis there is a danger that when the number of batches is increased, dependence will be introduced between the data points (batch means) and thus, the confidence interval generated will be invalid (the assumption of independent data points is not true).

Another example of higher-level learning support is shown in Fig. 4. A student in IE4633 is asked to choose different steady-state points (again, ad-hoc decisions for the analyst) from the cumulative average and investigate the impact on the estimate for the average time in queue. Prior to the implementation of the macros, the instructor would have to guide the student through this type of impact analysis. Typical software packages in simulation either do not support output analysis tasks (e.g. the generation of confidence intervals) or do not prompt the analyst to investigate their 'ad-hoc' decisions. In another example of how the macros support higher-level learning, a student in IE5573 is asked (see Fig. 5) to compare the regenerative method confidence interval to the batch means method confidence interval. Currently no

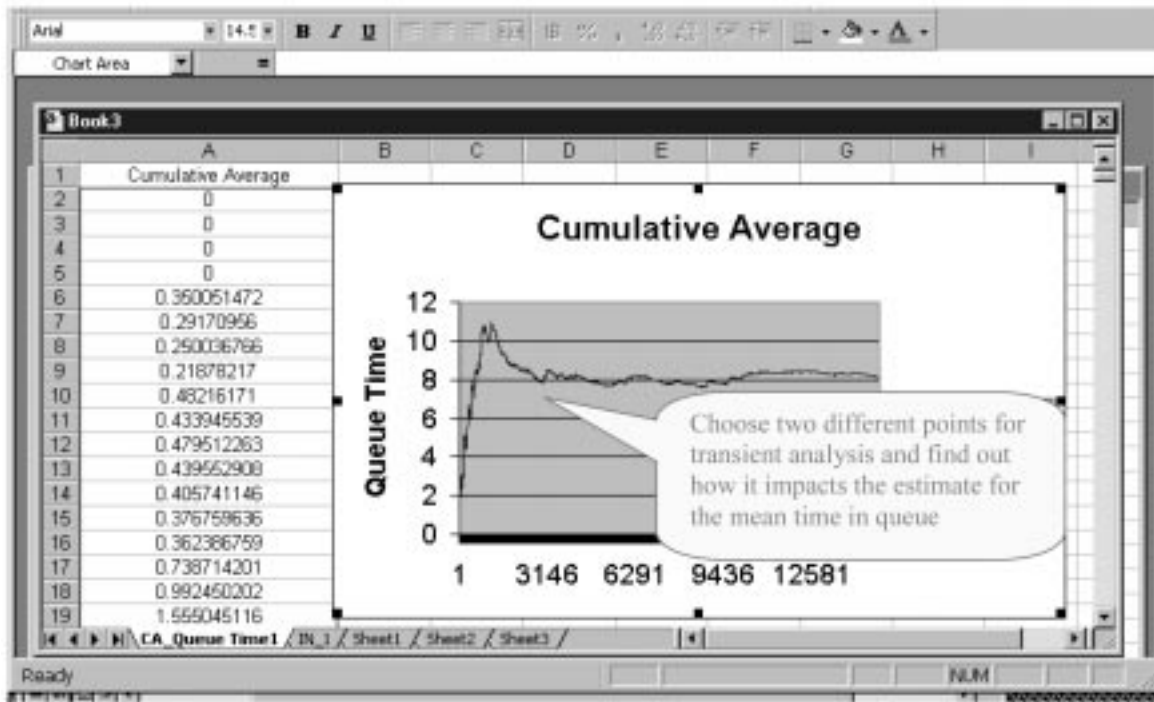


Fig. 4. Excel macro for performing transient analysis.

simulation package (including professional versions) supports confidence interval generation via the regenerative method (since it is a mathematically advanced methodology with limited practical application), or provides a coverage comparison tool for investigating confidence

intervals generated by different output analysis techniques (e.g. batch means versus independent replications versus the regenerative method).

All of the macros use 'pop-up' windows for the user to enter data. For example, Fig. 6 shows the input window used to generate the queue and

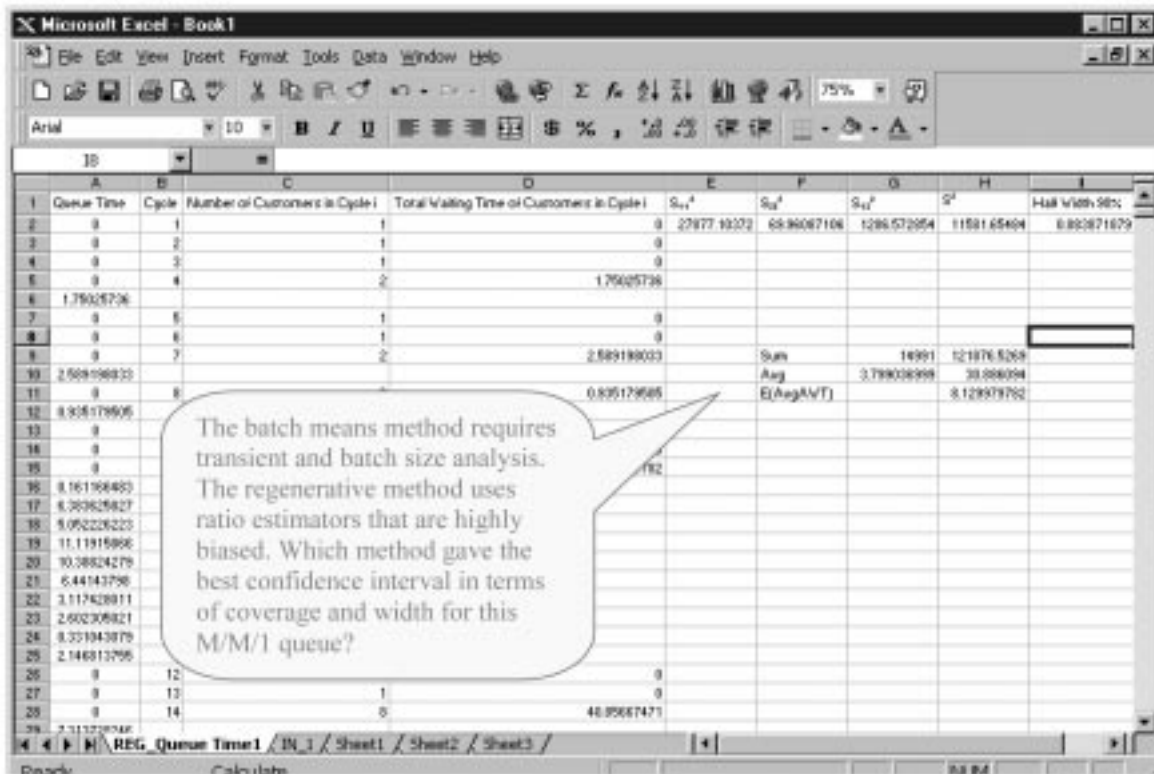


Fig. 5. An Excel macro prompting the student to compare output analysis techniques.

The image shows a dialog box titled "M/M/1 Input Form". It contains four input fields: "1/Lambda" with the value "4", "1/Mu" with the value "3", "Sample Size" with the value "15000", and "Replications" with the value "1". Below the fields are two buttons: "Cancel" and "OK". A message box at the bottom of the dialog says "Enter the Number of Observations" with a small icon.

Fig. 6. M/M/1 Input Form, data entry screen for generating 15,000 observations of an M/M/1 queuing system with $\rho=0.75$.

system times of an M/M/1 queue. The M/M/1 Input Form allows the user to enter the values into four fields for data entry: *1/Lambda*, *1/Mu*, *Sample Size*, and *Replications*. *1/Lambda* corresponds to the expected time between arrivals and is measured in time units. *1/Mu* corresponds to the expected time between service completions and is measured in time units. *Sample Size* determines how many data points (observations) the program will generate and *Replications* are the number of sets (spreadsheets) of independent data points the program will generate. Each data field is defined by moving the cursor into the corresponding data entry field where a definition screen is displayed within the 'pop-up' window. As shown in Fig. 6, the user is told that data entered into the *Sample Size* data entry field will result in the

generation of the number (in this case 15,000) of observations. When the user clicks on the OK button, Fig. 7 reveals the spreadsheet of observations generated by the M/M/1 Input Form corresponding to the data entered.

The spreadsheet contains seven columns of 15,000 observations each, as defined below:

1. Column A, *Interarrival Time*: The time between each arrival.
2. Column B, *Service Time*: The service time for each arrival.
3. Column C, *Queue Arrival Time*: The time each arrival enters the queue.
4. Column D, *Start Service Time*: The time each arrival begins its service.

The image shows a screenshot of Microsoft Excel with a spreadsheet containing 15,000 rows of data. The columns are labeled as follows: A (Interarrival Time), B (Service Time), C (Queue Arrival Time), D (Start Service Time), E (End Service Time), F (System Time), and G (Queue Time). The data consists of numerical values for each of these metrics across 15,000 observations.

Fig. 7. 15,000 observations of an M/M/1 Queuing System with $\rho=0.75$.

5. Column E, *End Service Time*: The time each arrival completes its service.
6. Column F, *System Time*: The total time each arrival spends in the system (queue time plus service time).
7. Column G, *Queue Time*: The time each arrival spends in queue.

Column F and G represent two system parameters (*Queue Time* and *System Time*) that may be further analyzed by other simulation output analysis macros (e.g., moving average plot, cumulative average plot, transient removal, regenerative method, correlation plot, and batch means method macros).

The instructor can also employ the spreadsheet to show how discrete-simulation logic is invoked. That is, each instance of the arrival and service time data can be tied to the events that trigger changes to the system state. For an M/M/1 queue there are two events: arrivals and departures. An arrival to the system increases the number in the system (system state) by 1, while a departure (service completion) decreases the number in the system by 1. Each event also triggers statistical data collection for unknown system parameters (e.g. average time in queue). For example, all arrivals will experience a queue time (including a time in queue of 0), and at the end of the simulation an average time in queue can be calculated based on the number of arrivals that have been generated and the output analysis technique chosen by the analyst (e.g. regenerative method). In simulation software packages the events are maintained on a future-event calendar (lists of events with the simulated time each event will occur) that is typically hidden from the user. Thus, another major advantage to the macro is that it reveals to the user lists that must be maintained for discrete-event simulation logic to

be implemented and how the events (lists) trigger statistical data collection to estimate unknown parameters of interest.

IMPACT OF EXCEL MACROS ON GRADES AND IN-CLASS TIME

Figures 3 through 7 are a sample of the types of macros that have been developed for the instructor and students use over the past six years. Statistical analysis on assignments and tests indicate that in general, the introduction of the macros has improved overall grades while reducing the amount of time spent on each topic. Table 2 indicates the post-behavior impact that the Excel macros have had on assignments and test questions used to assess the simulation learning objectives (see Table 1). Also shown is the impact on the in-class time spent on each topic. Please note, that prior to the introduction of the macros into the IE4633 course, the simulation topics were not taught and thus, IE4633 grades cannot be analyzed for pre versus post behavior. However, also note that the macros used in IE4633 assisted the instructor in maintaining the original IE4633 course content when adding the simulation topics into that course.

All statistical tests use the grade data from students prior to the introduction of the macro(s) as the basis (pre-behavior) for comparing post-behavior impact. The sample size for the base (pre-behavior) can be found in the first column under the course title. For IE4663 the post-behavior data collection is from 1997 to 2003, and for IE5573 it is from 1998–2003. The post-behavior sample size for assignment impact and test question impact is found in column 4 and 6 under ‘statistically significant differences’, respectively. Variance in the post-behavior sample sizes for assignment impact is due

Table 2. Impact of Excel macros on post-behavior average grades and in-class time

Course number and title (n-pre)	Simulation learning objective number (See Table 1)	Assignment Impact		Test Question Impact			
		Percentage increase in average grade	Statistically significant difference (n-post)	Percentage increase in average grade	Statistically significant difference (n-post)	Percentage reduction in in-class time	
IE4663, Systems Analysis using Simulation (32)	3	10%	Yes	7%	Yes	50%	
	4	6%	(53)	10%	(102)		
IE5573, Statistical Analysis of Simulation (10)	1	5%	No	10%	Yes	33%	
	2	8%	(31)	10%	(31)		
			Yes	(30)	Yes		(31)
	3	5%	Yes	(30)	8%		Yes
	4	7%	Yes	(31)	7%		(31)
5	3%	Yes	(31)	2%	Yes		
			No		No		
			(31)		(31)		

to students failing to turn in assignments designed to assess that specific learning objective. The larger sample size for the undergraduate (IE4663) test question impact is attributed to the fact that students are tested at least twice on the learning objective with a semester exam and a final exam. Graduate students in IE5573 only take a midterm exam while their final exam consists of a project report and presentation. This is why their post-behavior sample size for assignment and test question impact shows small deviations.

All of the statistics in Table 2 indicate that a statistically significant difference was realized in the percentage increase to average assignment grades and test questions except for learning objectives 1 and 5 of the graduate course (IE5573). Learning objective 1 (regenerative method) indicates that while there is not a statistically significant difference in the percentage increase to the average assignment grade, there is a statistically significant difference in the percentage increase to the average test grade. Thus the graduate students tended to do better on the exam than on the homework with the introduction of the macro. For learning objective 5 (goodness-of-fit tests) no statistically significant results were found in either the assignment or test question. This may be due to the fact that graduate students tend to take higher-level statistics courses before taking IE5573 and

consequently, the macros had little impact on improving their knowledge of the material.

Table 2 also reveals the percentage reduction in the amount of in-class time that the topics had in each course. IE4663 went from 4 class periods on the batch means method to 2, IE5573 reduced the number of class periods on the regenerative method from 2 to 1, and the input modeling topics in IE5573 were cut from 6 class periods to 4.

Acknowledgment—Excel macro development is partially funded by the National Science Foundation.

CONCLUSIONS AND FUTURE RESEARCH

While the macros have assisted students in attaining higher average homework and test grades, the reduction in class time has allowed the instructor to add depth-and-breadth to course content without impacting existing course content. In addition, the macros have given the students the ability to perform simulation analysis without access to costly simulation packages.

Future research is aimed at adding student feedback and query screens. The idea is to provide the student with the ability to ask his/her own 'what-if' questions of the macros and have the macro respond with either a new spreadsheet or a 'pop-up' window with the answer.

REFERENCES

1. M. C. Court, A case study on the impact of web-based technology in a simulation analysis course. *Simulation in Education and Education in Simulation, Simulation Special Issue*, **76**(4), 2001, pp. 207–213.
2. B. Bloom, *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*, Longman, New York, NY (1956).
3. J. E. Stice, A first step toward improved teaching. *Engineering Education*, **66**(5), 1976, pp. 394–398.
4. T. Rhoads, N. Hubele, T. Fernandez-Parker, and D. Rollier, Combining collaborative, traditional and computer-aided learning in an introductory engineering probability and statistics course, *Proc. American Society for Engineering Education-Pacific Southwest Section Annual Meeting*, April 5–6, 1995, San Jose, CA, pp. 171–182.

Mary Court is an Associate Professor in the School of Industrial Engineering at the University of Oklahoma. Her areas of research lie in simulation modeling and analysis and in electronic learning. She has taught advance statistics courses and simulation for over 9 years and has received several grants on simulation, statistics and electronic learning.