Proceedings of the 1994 Winter Simulation Conference
ed. J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila

OUTPUT ANALYSIS RESEARCH: WHY BOTHER?
A PANEL DISCUSSION

John M. Charnes
School of Business
University of Kansas USA

John S. Carson
AutoSimulations, Inc.
Atlanta, Georgia USA

Merriel Dewsnup
PROMODEL Corporation
Orem, Utah USA

Andrew F. Seila
Terry College of Business
University of Georgia USA

Jeffrey D. Tew
Consolidated Freightways, Inc.
Portland, Oregon USA

Randall P. Sadowski
Systems Modeling Corporation
Sewickley, Pennsylvania USA

1 BACKGROUND

In the last two decades, many different techniques have been developed and investigated by researchers for summarizing the output and drawing conclusions from simulation experiments. These techniques include (but are not limited to): classical statistical analysis of independent observations produced by replication of terminating systems; steady-state analytic methods such as nonoverlapping batch means, overlapping batch means, regenerative analysis, spectral analysis and other frequency-domain methods, autoregressive-moving-average time-series modeling, and standardized time-series; variance-reduction methods such as common random numbers, antithetic variates, control variates, indirect estimation, and conditioning; and methods for comparing system configurations such as pairwise comparisons, and ranking and selection methods. However, it is widely perceived by researchers that many of these techniques are not currently used in practice.

The lack of practitioner interest in using these techniques has been ignored by some researchers but has become a source of great existential angst among others. Some have suggested that the root cause of this disinterest is a lack of communication between researchers and practitioners. Because a large part of the success of previous Winter Simulation Conferences was due to attendance by both practitioners and researchers, it seems appropriate to hold at this year's conference a session devoted to airing the different views of output analysis research.

This session is an attempt to (*i*) determine the level of practitioner usage in the methods developed by output analysis researchers, (*ii*) find out how the output analytic research is perceived by practitioners of simulation, and (*iii*) find out what topics simulation practitioners think are important for further investigation or development by researchers. The discussion panel consists of researchers, educators, and practicing simulation professionals. The intent is to discuss these three main issues as well as any relevant issues raised by members of the audience, who are encouraged to participate.

2 DISCUSSION QUESTIONS

- Level of usage
  - To what extent are output analytic methods used in practice?
  - Are any statistical methods used?
  - If so, which methods? If not, is it because practitioners lack the knowledge to use statistical techniques, or because they do not find the techniques helpful?
- Perception of research
  - How do practitioners view the published research on output analytic methods?
  - Is the research too theoretical?
  - Is it too applied?
- Important research topics
  - Should we continue research into the development and investigation of output analytic methods?
  - If so, which topics should be investigated?
  - Which topics, if any, should be dropped?

3 POSITION STATEMENTS

3.1 John Charnes

It is disappointing to think that statistical methods are not widely used in practice for analyzing simulation output. When teaching university courses in simulation, I emphasize at the outset that because computer simulation models are mechanisms for generating observations on random variables, their output data require statistical techniques for proper analysis. I usually repeat this message so often throughout the course that I get tired of hearing myself say it, although I haven't yet had a student express the same sentiment (to me, anyway). Further, in the applied academic research that I have done with collaborators using simulation as a tool for investigating inventory policies or comparing alternate manufacturing layouts, statistical methods have been essential for drawing conclusions from the simulation models we have used.

In the more theoretical output analytic research I have done in using time-series methods to model multivariate output processes, I have found in many cases that these methods perform generally as well as conceptually simpler methods. However, at this point any potential improvement in performance may not be worth the additional intellectual burden required to use the more complicated techniques. Does this mean that they will never be useful? I think not. Even if they are never used in practice for analyzing simulation output data, the investigation of new techniques contributes to the evolution of knowledge.

It is probably inevitable that new output analytic techniques will seem arcane to the uninitiated, and some contributions might never amount to more than footnotes in the history of science, but that does not make them useless. The research performed today is part of the foundation upon which tomorrow's knowledge will be built. We cannot know at present which output analytic research stream will lead to changing practice in the future, but we can and do know that by discontinuing all output analytic research we will stagnate.

I see this discussion as a microcosm of the current debate in the U. S. on the social value of academic research. The tenure system might provide incentive for publishing results that otherwise would not see the light of day, but almost all academic research has some redeeming value. No one can know for sure at present where the work might lead in the future, and the past has shown that serendipitous results can develop from seemingly unrelated work done by researchers. Moreover, educators who also perform research can become better teachers if they are striving simultaneously to add to the body of knowledge they profess, and students employed as research assistants can benefit by developing investigative skills that will be useful in whatever career path they choose to follow.

Forums such as this one serve a useful purpose by keeping open the lines of communication between developers and prospective users of new simulation output analytic techniques. I hope that the discussion gives simulation practitioners some insight into the world of academic research, and that it gives researchers some insight into the difficulties faced by practitioners. If so, both groups will benefit.

3.2 John Carson

The purpose of statistical output analysis in stochastic simulation studies is to control the risk of drawing an incorrect conclusion from insufficient data. Professor Charnes asks, why is it so little used in practice? First, the knowledge and usage of statistical methods among practitioners ranges from zero to quite extensive. So we must careful over drawing broad conclusions. Second, the availability of good software to facilitate statistical calculation, data management, and graphical presentation of the results is as important a factor as training and education.

Let me address tho questions asked by Professor Charnes:

"To what extent are output analytic methods used in practice?" My impression, but I have no poll or scientific study to back it up, is that more and more people are using some of the basic methods provided the software languages and tools they use provide the necessary support. However, it may also be true that the population of practitioners (model users and model developers) is growing faster than usage of statistical methods among existing users and developers.

"Are any statistical methods used? If so, which methods?" From personal observation, I have seen use of Welch's moving average plots for the initial transient, replication and replication/deletion, design of experiments, and occasionally one of the ranking and selection methods. The most common are replication and Welch's moving average plots. The methods which are not used, or used rarely, ar those which require extensive academic background in esoteric subjects merely to understand them, and those that stand upon assumptions that never apply in practice or for which it could not be determined whether they apply or not.

If not being used, "is it because practitioners lack the knowledge to use statistical techniques, or because they do not find the techniques helpful?" Certainly, training and education play a role, first to convince users of the necessity of sound statistical analyses and then to train them in those broadly applicable methods that have proven to be sound in a wide range of circumstances.

"How do practitioners view the published research on output analytic methods? Is the research too theoretical? Is it too applied?" For the most part, I believe the overwhelming majority of practitioners arc not aware of the academic research. Far too much of it will never be applicable or even provide insight into the nature of the problem. Reasons include assumptions that are too narrow, too technical or too esoteric and methods that have only been tested on academic "toy" problems. Furthermore, academics should never expect a method to be adopted until it has been around for a while and been scrutinized and tested by other researchers and used by a few intrepid pioneers. Think of a newly proposed method as a newly discovered drug. Do you want to be the guinea pig? Is it safe? Is it effective?

"Should we continue research into the development and investigation of output analytic methods? If go, which topics?" Yes, research should be continued, but it should be re-focused on methods that can be broadly and safely applied to a wide range of models without having highly specialized knowledge. Methods that appear narrowly applicable if applicable at all to real-world problems, at least to most practitioners, include regenerative and spectral analyses.

One area that may need more attention is that of diagnostics similar to "goodness of fit" tests in regression. When using a given method, what assurances do I have that it is appropriate for the simulation at hand? That it isn't adding a false sense of security. For I know as a matter of fact that all the statistical methods are based on a number of assumptions, and that the outputs from the real-world models I develop do not meet those assumptions. A second related area for further research is the robustness of proposed methods under a range of conditions likely to be encountered in practice, and the rapid dissemination of the "bad news" when a proposed method is found to be non- robust or extremely sensitive to some "tuning" parameter or other user specified parameter. Such methods should never be used, nor should academics expect them to be used, in practice.

3.3 Merriel Dewsnup

The following general trends are continuing:

- Simulation models continue to be easier to build and at lower cost both for language processors and hardware. This is a result of the general use of C as the base programming language. The result is development of simulation products with higher functionality. The continuing trend toward lower per unit micro-computer cost has placed computing on almost every desk. This results in more people doing simulation with less formal training. The result, as I see it, is less output analysis.
- Formal engineering education programs are being asked to include more subject matter. Much of this is natural because of the ever expanding amount of knowledge. My experience has been that to require more than a one semester simulation course is infrequent. Only a few professors go the extra mile to offer extra classes that students may take.
- Large modeling teams with specialist are being replaced with much smaller and often one person team. I believe this is a result of the general downsizing trend continuing today.

The Past: Attempting to teach simulation through a language required a large startup time for both the students and faculty. By the time the quarter ended most students were just mastering the mechanics and syntax of the modeling language chosen. Little time was available to completely cover all material and early textbooks and manuals contained little information about output analysis.

The Present: The newer text books do attempt to address output analysis earlier rather than the last chapter or so and are doing a better job of helping students understand principles. Most languages now provide some statements and procedures to automate running multiple replications and scenarios.

The Future: Research and development needs to continue in order to make output analysis an easily applied and understood step in the simulation process. With the high cost of simulation analysts we will see more models completed by single person teams. Often they will not have or be allowed to take extra time to do all of the needed replications.

College level courses need to include output analysis as an essential topic. The continuing trend toward easier and faster modeling building should allow this topic introduction much earlier then in the past. Additional software and language improvements should continue that assist with output analysis.

Software developers should continue in the direction of ease of experimental design, analysis of output results and improved documentation and methods available.

Professional individuals and their organizations should accelerate the efforts to maintaining a sharp set of computer modeling and simulation tools and techniques.

3.4 Randy Sadowski

Although there are numerous published papers on output analysis, there are relatively few practitioners utilizing these techniques. In the past, most simulation practitioners were full time and could afford the time to understand and utilize such techniques fully. Today's practitioners use simulation only part time and, in many cases, only once for a specific project. Methods that are used tend to be limited to classical statistical techniques taught in vendor-offered courses. However, it is not uncommon to find isolated practitioners who employ more sophisticated techniques.

There are several reasons why only classical techniques are used by the majority of practitioners—when output analysis is used at all. The two major reasons are computer speed and a lack of statistical background. Today's computers and software provide runtimes that are relatively short; thus, users tend to overestimate the required runtime, as the penalty is negligible. In many cases, the required analysis is relatively simple and only requires a few runs and thus not justifying the more sophisticated techniques. The more common reason for only using simple techniques, or none at all, is an ignorance of the statistical implications by the user. Most simulation practitioners have only had minimal statistical training and that was acquired many years ago. Therefore, they tend to utilize only techniques that are presented in a simplified cookbook-approach method. Fear of statistics is probably the single largest reason for not using analytical techniques.

The perception of the published research by the majority of the practitioners is that it is simply unreadable. There are clearly reasons why these papers are written in this fashion. However, if the intent is to provide analytical techniques for practitioners, the research must be presented in a manner that allows the non-statistical reader to understand and implement the techniques fully. This is not to imply that the research should not be published in its current form, but to suggest that alternate write-ups be made available. These could be presented at conferences such as WSC, published in non-refereed practitioner journals, or made available to simulation software vendors for distribution.

It would be foolish to suggest that research in analytical methods should not be continued. The simple act of research often uncovers methods or techniques that become very useful, even though they may not have been the primary focus of the original research. Again, if researchers are interested in creating useful techniques for the average practitioner, they should first make available, in the proper form, the results of many years of research already completed. One might even consider this a classic problem of transfer of technology.

If researchers are truly concerned about usage, they should focus their research on the issues most often faced by the average practitioner:

1. When is the simulation in steady-state?
2. How long should the simulation be run?
3. How to compare easily many different scenarios with a minimum number of runs?

The common theme in all three of these items is the desire of the practitioner to analyze their systems accurately in a minimum amount of time.

3.5 Andy Seila

A simulation is a sampling experiment. Therefore, the output produced is a sample, and any quantities computed from the output to measure system performance will be subject to variation from sampling. While it may be true for many systems in practice that the sampling variation is small, this cannot be known for sure until it has been measured. The objective of most output analysis techniques is the measurement and control of sampling variation. In our simulation classes we teach that some type of statistical output analysis is appropriate for every simulation study. However, it seems that many simulation studies include little or no statistical analysis of output data. If this is the case, and I believe it is, I would like to pose two questions:

A. Why are the techniques that have been developed not used routinely on almost every simulation project?

B. What can we, as output analysis researchers, do to encourage our methods to be used routinely?

The statistical techniques for simulation output analysis (as well as other statistical techniques) will be widely used when two things happen: (1) users understand the value and use of these techniques, and (2) user-friendly software is available. By implication, these techniques are not used because one or both of these conditions is not met. The primary reason why output analysis techniques are not more widely used is that many decisionmakers, i.e., users of the simulation study, do not perceive the value of correct output analysis and therefore do not demand that it be a part of the simulation study. It is possible to spend tens of thousands or hundreds of thousands of dollars on a study and then make incorrect decisions because the data was not properly analyzed. Simulation is still a rather specialized activity that is performed by analysts rather than managers. If the results of simulation runs are incorrect because output analysis was done incorrectly, there is no indication of a problem and the red flags do not go up to the decision maker. To the analyst, output analysis requires additional time, but there is no additional reward for doing it correctly and there is no penalty for doing it incorrectly. One would have to ask — What therefore is the incentive to apply statistical techniques other than personal pride of doing a job right?

Many commercial software packages now include procedures for implementing the independent replications and batch means methods. These have proved to be the most useful methods for estimating the mean performance measure in terminating and nonterminating simulations. I suspect, however, that the level of use is not what it should be because the implementations are not as automated and user-friendly as they could be. If we want these methods to be used routinely, we must, in my opinion, make them routine to use. I think that if statistical techniques for output analysis are to fulfill their promise and be widely utilized, several things must be done. First, we must demonstrate to potential users that good output analysis methods will improve decision making based upon the simulation results. After all, if statistical output analysis techniques will not improve decision making, we must wonder what is their value. Second, we must seek to educate managers and potential managers (students) in basic statistical concepts and their value to decision making. In many business and engineering schools it seems that courses in statistics and applications of statistics in other courses are being de-emphasized. As a result, students may get the impression that the decision making environment is deterministic and statistical analysis is not informative.

Third, we as researchers must maintain a close working relationship with practitioners. We must find out exactly what are the problems they need to solve, then we must develop methods that quickly and reliably solve these problems. Our methods must work with real data in real simulation experiments, and they must provide consistent and reliable results. On the other hand, practitioners must be willing to work with us to test these methods in the laboratory that counts: the marketplace, and give us realistic feedback and suggestions for improvement. Fourth, we must seek to make our methods automated and easy to use. Even the batch means method, which is perhaps the simplest statistical method for estimating the mean, requires the user to specify two parameters: the truncation point and the batch size. Default values of these parameters should be offered to the user based upon the characteristics of the output data. We must remember that, unlike us, the users of these methods are not primarily interested in the theoretical underpinnings and normally will have only a vague understanding of the methods. Finally, since no statistical methodology will work all the time, we must provide users with some indication when the method is not working, so they will know that additional analysis is indicated. All confidence interval procedures will produce values for the upper and lower confidence limits whether they are valid or not. What is needed is an easily computed measure of the validity of the confidence interval.

In summary, I believe that we have a good product in the form of output analysis methodology. But, we must package it well in the form of user-friendly software and advertize it by educating potential users about its value if we want to see it be successful in the marketplace.

3.6 Jeff Tew

The *output analysis crisis* that has plagued the computer simulation community for at least the last 15 years is due to a lack of communication between "practitioners" and "researchers". As evidence, consider the the composition of the tracks in all of the recent WSC's. Further, this crisis is the direct result of a lack of identity for the field of computer simulation in both the industrial and academic communities.

This lack of identity is due to the diversity of fields which created, developed, and sustained computer simulation from its earliest years to the present. Although diversity is especially important in the formatitve years of a field of intellectual activity, it can become a roadblock to the maturation of that same field as it seeks, and requires, organizational "slots" (i.e., academic departments, corporate groups, etc.) for continued development and sustenance. It is during the maturation stage that the *identity* of a field should come into focus. This has not happened with computer simulation because the simulation community has failed to find the right "slots" for full maturation to take place. (Contrast what has happened to computer simulation to what happened earlier in this century to the field of statistics. How many academic and industrial statistics departments are there in the United States? How many academic and industrial computer simulation departments are there?)

The argument can be made that the field of statistics has a much stronger and clearer identity than does the field of computer simulation. The simulation community has failed to give computer simulation the identity it requires. Consequently, there has been little attention given to addressing the critical issues of developing standards for computer simulation curricula which, in turn, has left the field without a formal, broadly recognized approach to the education of the computer simulation student. Thus, within computer simulation, we are left with tremendous cultural boundaries that divide the community into factions that are all too often ignorant of the contributions each one is making to the field. This is at the heart of the output analysis crisis.

3.6.1 Level of Usage

In practice, output analytic methods and statistical methods are used very little, if at all. Unfortunately, when they are used they are usually used incorrectly. Very little attention is given, in practice, to the proper estimation of means and variances of interest. Also, point estimates are usually given in place of more meaningful interval estimates. There are two reasons for this. First, the vast majority of practitioners are not well educated in statistical methods and output analysis techniques. Many practitioners and users are unaware of the need for statistical methodologies in analyzing simulation results. Second, simulation researchers have not done a good job of communicating the need for these analytical techniques to the practitioners. Most analytical techniques are not readily available to the practitioner irrespective of his choice of software. Simply put, if it's not available, people won't use it.

3.6.2 Perception of Research

Practitoners view the published research on statistically oriented analytic methods as too theoretical, inaccessible, and not relevant to their simulation problems. Practitioners have this point of view because, in general, they lack sufficient knowledge regarding the statistical issues related to performing *any* simulation experiment, and the researchers have not done a good job of "taking the message to the people." Often, researchers are unaware of who the actual end-user is for their work. There is poor communication between practitioners and researchers.

3.6.3 Important Research Topics

Without question, research should continue on new and better output analytic methods. However, the lines of communication between *all* users of simulation (practitioners and researchers) must be strengthened so that the work of the researcher becomes more useful to the practitioner and, in turn, the practitioner becomes more aware of the need for these methods.

Computer simulation has been very fortunate to draw upon the work of many truly exceptional people who have come from very diverse educational backgrounds. This is one of computer simulation's great strengths. However, only through strong communication between these two parties can a clear identity for the field of computer simulation be forged.

AUTHOR BIOGRAPHIES

JOHN M. CHARNES is Associate Professor of Statistics and Quality Management in the School of Business at The University of Kansas. His research interests are in the area of statistical analysis of multivariate output. Prof. Charnes has taught at the University of Washington in Seattle, the University of Minnesota, Hamline University, and the University of Miami, Florida. He is a member of **ASA**, **ASQC**, **TIMS**, and is Editor of the *TIMS College on Simulation Newsletter*.

JOHN S. CARSON is with AutoSimulations, Inc. He has over 15 years experience in simulation in a wide range of application areas, including manufacturing, distribution, warehousing and material handling, transportation and rapid transit systems, port operations and shipping, and medical/health care systems. His current interests center on the simulation of transportation systems of various kinds (traditional material handling to rapid transit) and software for simulation output analysis. Previously, he was a co-founder and President of Carson/Banks & Associates, a simulation services firm. He has taught at Georgia Tech, the University of Florida, and the University of Wisconsin-Madison, and is the co-author of *Discrete Event System Simulation* (1984) and *Getting Started with GPSS/H* (1989). He holds a Ph.D. in Industrial Engineering and Operations Research from the University of Wisconsin-Madison. He is a member of IIE, SCS, ORSA and TIMS.

MERRIEL DEWSNUP received his formal training at the University of Utah where he earned a B.S. degree in Mathematics and a Master of Industrial Engineering Administration degree. While at the University Merriel worked as a Statistical Analyst for the Office of Long Range Planning, and as a programmer and as the Assistant Director of the computing Center. He was responsible for moving computing out of the engineering building and onto the general campus with remote centers established in business, medicine, biology and the library. He also taught Computer Modeling and Simulation as an Adjunct Assistant Professor of Mechanical and Industrial Engineering for many years. Simulation languages taught include GASP, GPSS, SLAM, SIMAN, and SIMSCRIPT.

Merriel also worked as the Modeling Systems Coordinator for Mountain Fuel Supply Company where several large financial models were developed. These models helped executive management by analyzing the impact of financial decisions on earnings per share before carrying out the decisions.

Merriel then joined Eaton-Kenway as the manager of the Simulation Engineering Department. At Eaton-Kenway he developed or supervised development of more than 500 models of AGV and AS/RS systems proposed for Eaton-Kenway's customers. Special interfaces were developed that allowed CAD drawing software to supply data to the simulation models and for the optimized AGV control tables to be moved from the models directly to the final software. This shortened system installation time and produced much happier customers, because systems worked the first time.

Merriel is currently the Vice President of Customer Support Services at PROMODEL Corporation where he is responsible for Technical Support, Training, and Consulting Services. Being in close daily contact with a large user base and just down the hall from the software development team has allowed development of modeling features and capability that match user needs.

RANDALL P. SADOWSKI is currently vice president of Systems Modeling Corporation in charge of consulting and user-education services. He was previously on the faculty at Purdue University and at the University of Massachusetts. He received his bachelors and masters degrees from Ohio University and his Ph.D. from Purdue. Dr. Sadowski's research interests are in manufacturing and production systems with emphasis on modeling, control, and applied scheduling.

ANDREW F. SEILA is Associate Professor in the Terry College of Business at the University of Georgia. He received the B.S. degree in Physics and the Ph.D. degree in Operations Research and Systems Analysis from the University of North Carolina at Chapel Hill, and was employed as by Bell Laboratories prior to joining the faculty of the University of Georgia. He has published in numerous journals including *Operations Research*, *Management Science*, *Operations Research Letters*, *Simulation* and others. His professional interests lie in all areas of simulation research and practice, especially statistical methodology for simulation. He is a member of ORSA, TIMS and ASA.

JEFFREY D. TEW is a senior systems engineer with the Intelligent Systems Engineering Group at Consolidated Freightways, Inc. and an adjunct associate professor of computer simulation at the Oregon Graduate Institute. Prior to these positions, he was on the faculty at Virginia Tech in the Department of Industrial and Systems Engineering. He is the 1994 WSC Proceedings Co-Editor and the Vice President/President Elect of the TIMS College on Simulation. He has published several articles in many journals such as *EJOR*, *IIE Transactions*, *JSCS*, *Management Science*, *Operations Research*, and *SCS Transactions*. He was the Publicity Chairman for the 1992 WSC and was the TIMS College on Simulation Newsletter editor from September 1992 to June 1994. His research interests are in the areas of applied statistics, design of computer simulation experiments, and output analysis and how these techniques can best be applied to industry. He is a member of ACM, ASA, IIE, IMS, Omega Rho, ORSA, SCS, Sigma Xi, Tau Beta Pi, and TIMS.