Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

HistoRIA: A NEW TOOL FOR SIMULATION INPUT ANALYSIS

Mohammadnaser Ansari
Ashkan Negahban
Fadel M. Megahed
Jeffrey S. Smith

Department of Industrial and Systems Engineering
Auburn University
3301 Shelby Center
Auburn, AL 36849, USA

ABSTRACT

An important step in input modeling is the assessment of data being independently and identically distributed (IID). While this is straightforward when modeling stationary stochastic processes, it becomes more challenging when the stochastic process follows a non-stationary pattern where the probability distribution or its parameters depend on time. In this paper, we first discuss the challenges faced by using traditional approaches. We then introduce the *Histograms and Rates for Input Analysis* (HistoRIA) as a tool to facilitate input modeling. The tool automates the analysis process and significantly reduces the amount of time and effort required to test the IID assumptions. The generated HistoRIA plot is capable of effectively illustrating changes in the rate and distribution over time. Although originally designed and developed for simulation input analysis, the paper demonstrates how the tool can potentially be applicable in other areas where non-stationarity in the data is also common.

1 INTRODUCTION

Simulation involves modeling the stochastic processes that exist in the real system under study. Examples of these stochastic processes include arrivals of customers to a restaurant, time to failure and time to repair for machines in a manufacturing system, processing time of a part on a machine, number of orders per day or number of items in an order received by a distribution center, etc. During the simulation run, random variables are generally used to mimic these processes through sampling from their probability distribution. Finding an appropriate distribution for these random variables is the main goal of input analysis. Input modeling is arguably one of the most expensive steps in most simulation studies and is essential to a successful simulation (Law 2009). The basic steps in input modeling are (Law and Kelton 2000): (1) data collection; (2) IID assessment; (3) hypothesizing potential distribution families; (4) estimating distribution parameters; and, (5) fitting the data to the hypothesized distributions and fit assessment.

There are various statistical packages that can be used for estimating the probability distribution of input data (i.e. steps 4 and 5 of input modeling) such as Minitab® (Ryan, Joiner, and Cryer 2012), Stat::Fit® (Benneyan 1998), EasyFit® (Schittkowski 2008) and even Microsoft Excel® (Zaixiang, Qingsong, and Chenwu 2005). Some simulation packages such as Arena® (Hammann and Markovitch 1995) also come with an input analyzer tool. Simulation-oriented issues of *ORMS Today* magazine provide biennial lists of these commercial packages and their capabilities related to input modeling as well as other topics in simulation modeling and analysis (Swain 2011). All of these packages base their analysis on the fundamental assumption that the data are IID without which the methods they use would be invalid. However, as discussed by Vincent (1998), these statistical packages lack the capabilities to do an analysis of the IID requirement.

Therefore, the analyst is responsible to use heuristic methods to test the IID assumptions before using these packages for distribution fitting. While in most cases where data is not independently distributed, correlation in the data can be detected using scatter diagrams or tables/plots of estimated lag (linear) correlations, detection of deviation from the stability of distribution is typically more involved to the point that even gross changes in the probability distribution can sometimes be difficult to detect (Vincent 1998).

Non-stationary stochastic processes are perhaps the most common form of deviation from the *identically distributed* assumption (Vincent 1998). A non-stationary process can be defined as a random variable whose distribution varies as a function of time in such a way that either the basic form/family of the probability distribution or at least one of its parameters depends on time. In this paper, we demonstrate the shortcomings of traditional heuristic methods for detecting and characterizing non-stationary processes. We then present a novel graphical tool to display the data in a way that facilitates IID assessment. More specifically, we propose a tool called Histograms and Rates for Input Analysis (HistoRIA) which enables us to perform data collection, IID assessment, and hypothesizing the underlying distribution of stochastic processes. Perhaps the most important advantage of the tool is that the generated plots allow the user to visually analyze the data and identify IID intervals. Moreover, by automating the IID assessment process, the proposed tool significantly reduces the burden of performing extensive trial and error attempts and thus enables the analyst to find a fairly accurate model for any non-stationary process in a short amount of time. Therefore, the main contribution of the this work is based on practical rather than theoretical grounds.

The remainder of the paper is organized as follows. Traditional approaches to detect data non-stationarity are described in Section 2. Section 3 describes the HistoRIA plot and its application in input analysis. We present the details of the proposed tool in Section 4. Section 5 provides a brief discussion on other potential application areas where the proposed tool can be useful. Finally, conclusions are provided in Section 6.

2 TRADITIONAL APPROACHES FOR MODELING NON-STATIONARY PROCESSES

In this section, we first discuss modeling of a simple stationary arrival process with a fixed arrival rate where the IID assumptions are known to be true. Next, through examples of non-stationary arrival processes, we discuss the traditional approaches and highlight the need for a new tool to facilitate the analysis.

The simplest case of input analysis for a stationary arrival process can be summarized in Figure 1. In the first step, 500 observations of inter-arrival times (IAT) are collected. The histogram of the data is then used to hypothesize the probability distribution of IAT. In this case, we hypothesize that an exponential distribution can be used based on the generated histogram. The only parameter to be estimated is the mean IAT which can be obtained by simply finding the average of the 500 data points. Finally, chi-squared goodness-of-fit test is performed to assess the fit of the hypothesized distribution by comparing the expected number of observations in each interval (np_j) with the the actual number of data points in that interval (N_j), where each interval is defined between a lower bound (LB) and an upper bound (UB). In this example, we use 8 equiprobable intervals to perform the test. Since the p-value is greater than 0.05, we fail to reject the null hypothesis at the 95% confidence level and thus we can claim that an exponential distribution with a mean of 44.8 minutes provides a good fit for the data and can be used in the simulation model.

Hypothesizing about the distribution of a random variable using a single histogram that combines all of the data that have been collected over time can be misleading if the stochastic process is non-stationary. Consider the operation of a hypothetical restaurant between 10 AM to 2 PM. After collecting data on customer arrivals, the IAT histogram is presented in Figure 2(a) where an exponential distribution with mean 2 minutes seems to provide a good fit for the data. However, the arrival rate into such system is most likely not constant during the 4 operating hours as we expect to see higher arrival rates during the lunch rush and lower arrival rates early in the day and after the lunch rush is over (for the sake of this example, we will assume that the actual arrival rate into the system is known to vary as shown in Figure 2(b)). Clearly, if we are interested in short-term operational behavior of such system, it would not be accurate to assume a stationary arrival process and use a single fixed distribution to generate inter-arrival times in the simulation model.

Observation No.	Inter-Arrival Time
1	11.4808
2	21.5138
3	101.01
4	79.239
5	155.745
6	18.3749
7	17.8944
8	48.2243
9	77.2024
10	3.21627
11	35.7252
12	10.3699
13	1.38124
⋮	⋮
493	12.4775
494	63.4998
495	7.92532
496	19.6303
497	3.76539
498	25.7873
499	18.1108
500	0.510225
Average	44.8265

H_0 : Inter-Arrival times are exponentialy distributed with mean 44.826

Chi-Squared Goodness of Fit Test Using 8 Eaqui-Probable Intervals					
Interval	LB	UB	N_j	np_j	statistic
1	-inf	5.986	59	62.5	0.196
2	5.986	12.896	74	62.5	2.116
3	12.896	21.069	67	62.5	0.324
4	21.069	31.071	69	62.5	0.676
5	31.071	43.967	58	62.5	0.324
6	43.967	62.143	46	62.5	4.356
7	62.143	93.214	54	62.5	1.156
8	93.214	inf	73	62.5	1.764

χ^2	10.912
$\chi^2_{k-1,\,1-\alpha}$	14.0671404
P-Value	0.85749613

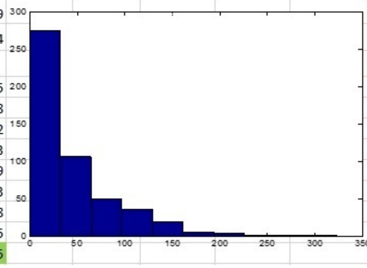
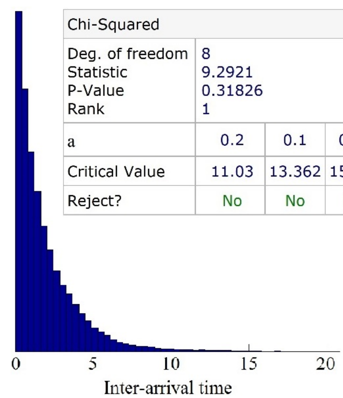
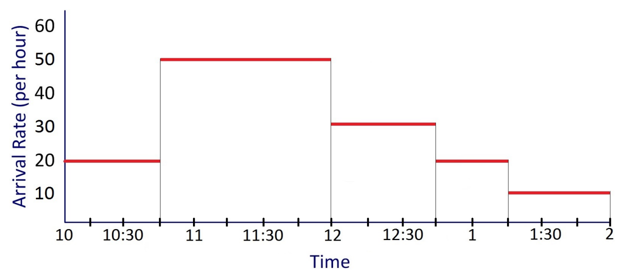


Figure 1: Input analysis of a stationary arrival process



Chi-Squared					
Deg. of freedom	8				
Statistic	9.2921				
P-Value	0.31826				
Rank	1				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	11.03	13.362	15.507	18.168	20.09
Reject?	No	No	No	No	No

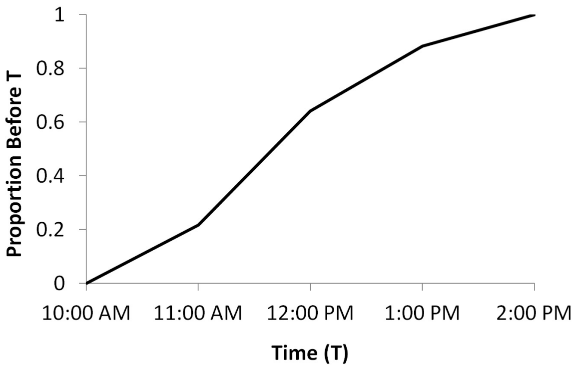
(a) Histogram of all inter-arrival times with goodness-of-fit test results from EasyFit®



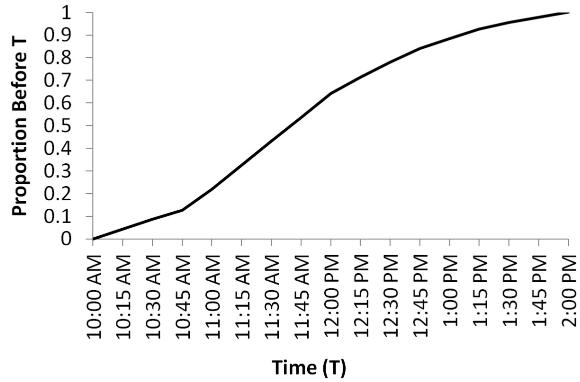
(b) The actual non-stationary arrival process

Figure 2: Data for the sample restaurant

In this case, our goal is to first identify IID time blocks and then hypothesize IAT distribution in each time block. However, the histogram provided in Figure 2(a) does not provide such information. Vincent (1998) describes the use of three types of plot to detect deviations from stability of distribution. The first approach involves plotting the proportion of arrivals up until a time T as a function of T. In the case of a constant arrival rate the plot is expected to look linear while a drastic deviation from linearity indicates a change in the rate. Clearly, the resulting plot is dependent on the choice of the interval width. Figure 3 illustrates the plot for 1-hour and 15-minute interval widths. There are two major problems associated with this approach that could potentially lead to an incorrect decision about the actual IID intervals. First, even with the 15-minute widths (which should reveal the actual intervals with constant arrival rate) it is still difficult to identify the exact transition points since it is difficult to decide whether the slight differences in the slopes are due to randomness or because of differences in the arrival rates. For example, one could mistakenly conclude from Figure 3(b) that the rate is constant from 12:00 to 2:00 p.m. (which we know is not true). Secondly, the plot does not provide any information about the form of the IAT distribution. This can cause problems if the rate remains constant over a period of time while the form of the distribution changes with time (as we will see later).



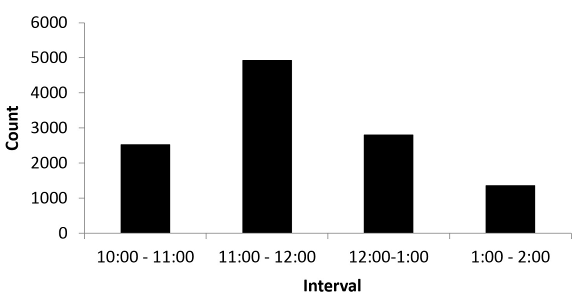
(a) 1-hour time widths



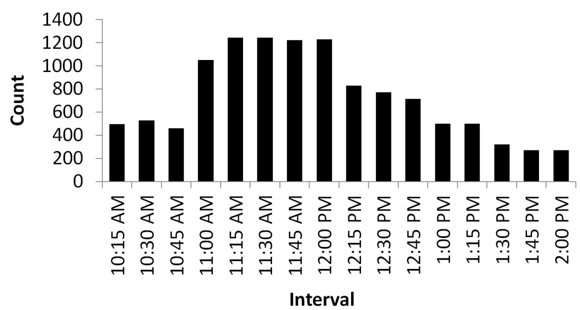
(b) 15-minute time widths

Figure 3: Proportional number of arrivals

The other traditional approach involves plotting the histogram of the number of observations (arrival rate in this case) per time interval. Again, the plot is dependent on the choice of interval width. Figure 4 illustrates the plot for the non-stationary arrival data for 1-hour and 15-minute time intervals. An advantage of this approach over the previous plot is that it shows the non-stationarity of the process more clearly. In fact, we will be using this positive aspect of the plot in our proposed tool. However, similar to the second problem discussed above, this plot is only capable of visualizing the change in the rate but not the form of the probability distribution. As a result, it is still possible to make incorrect assumptions about the IID intervals if the distribution changes over time while the rate remains the same.



(a) 1-hour time widths

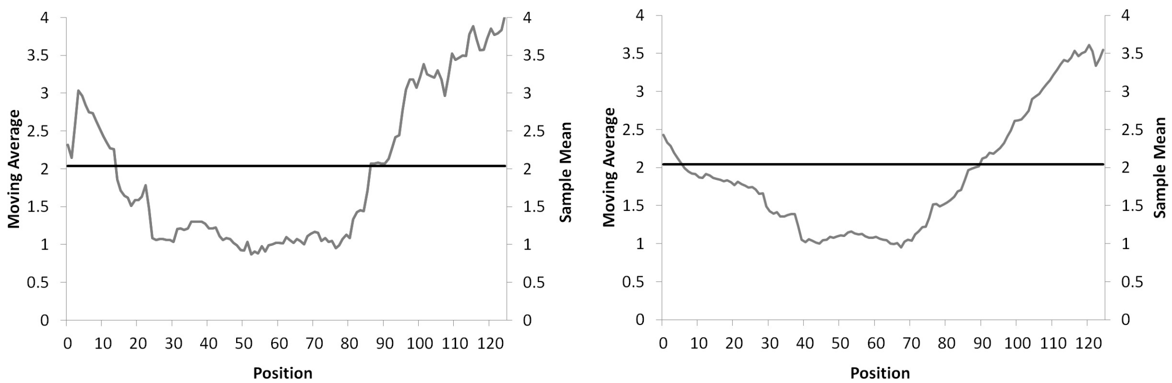


(b) 15-minute time widths

Figure 4: Histogram of arrival counts per time interval

The third traditional method involves plotting the moving average of the observations to identify deviations from the overall average. If the process is stationary, the plot of the moving average should be roughly constant (or oscillating around the overall average). The shape of this plot can also vary substantially with the choice of the moving average window (w). Figure 5 illustrates two plots of centered moving average of inter-arrival times with different windows where the sample average is shown as a solid horizontal line to help detect changes in the arrival rate. Although the plot clearly shows that the process is non-stationary, it still suffers from some important deficiencies. First, it is not possible to to visually identify the exact time of transition to a different rate. Secondly, similar to the other two approaches discussed above, the plot does not provide any information about the shape of the distribution. Moreover, it is not possible to make any judgments about the time intervals near the beginning and ending of the plot since fewer data points are available to calculate the moving average in these intervals.

Since the number and lengths of IID intervals are not known and the shape of the plots described above depends on the choice of their corresponding parameters, simulation analysts typically perform multiple



(a) $W = 10$ (b) $W = 25$

Figure 5: Moving average plot for the arrival data

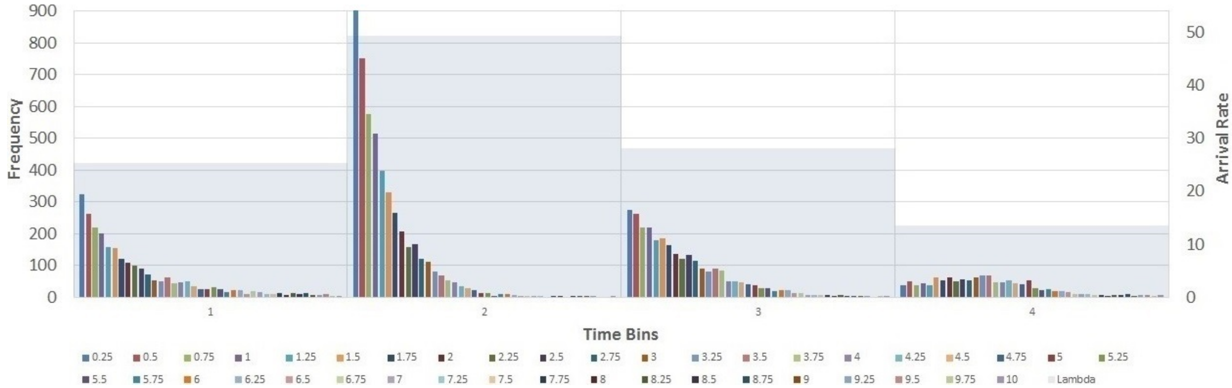
trial and error steps before making an assumption about the potential IID intervals. Also, since the general-purpose statistical packages are not specifically designed to do such experimentation, the burden is all on the simulation analyst to process and divide the data into groups based on the chosen intervals, perform the necessary calculations, and generate the resulting plot in each trial and error attempt. Moreover, once the analyst is confident about the chosen intervals, the histogram of the data for each identified stationary time block needs to be developed separately to hypothesize the underlying distribution in that interval since none of the above approaches provide visual information about the probability distribution. Therefore, this step of input analysis can be extremely time consuming and burdensome if traditional approaches are used.

3 HISTOGRAMS AND RATES FOR INPUT ANALYSIS: HistoRIA

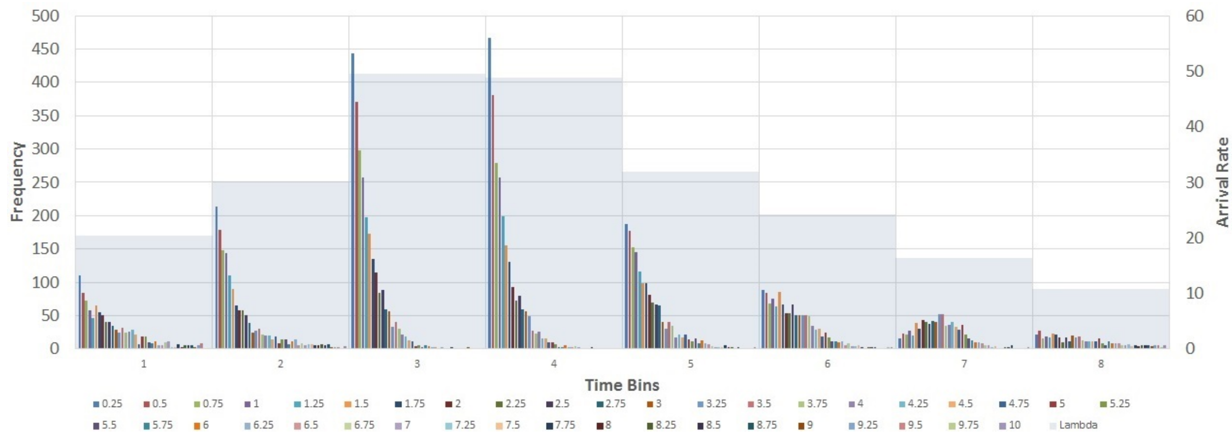
The proposed tool automates the above procedure and provides a plot that allows for simultaneous comparison of the rate and distribution among different intervals which can be extremely useful by reducing the burden and time to perform traditional IID assessment. Here, we introduce the HistoRIA plot as a possible display of the data without reference to the tool used to construct it which is described in detail in Section 4.

Three HistoRIA plots generated for the dataset is shown in Figure 6. In the first HistoRIA plot, each time block represents an hour (e.g. the first time block represents 10 a.m. to 11 a.m. and so forth). In the second and third plots the length of each time block is chosen to be half an hour and 15 minutes, respectively. The HistoRIA plot illustrates the arrival rate as well as the histogram of inter-arrival times for each time window. By sequentially reducing the length of each time block, i.e. transition from Figure 6(a) to Figure 6(c), we are able to capture the behavior of the non-stationary process in more details. In fact, by using 15-minute time blocks we were able to identify the actual time periods during which the arrival rate remains unchanged. Using the HistoRIA plot presented in Figure 6(c), we can also make hypotheses about the IAT distribution in each of these 15-minute time blocks. Based on the plot, we would hypothesize that the IAT in all time blocks are exponentially distributed except for the 12^{th} and 13^{th} intervals (i.e. from 12:45 PM to 1:15 PM) where it seems to follow a normal distribution. Therefore, the HistoRIA plot resolves the problem that was common among all of the existing traditional methods and allows us to identify the actual IID intervals where both the rate and the distribution remain unchanged.

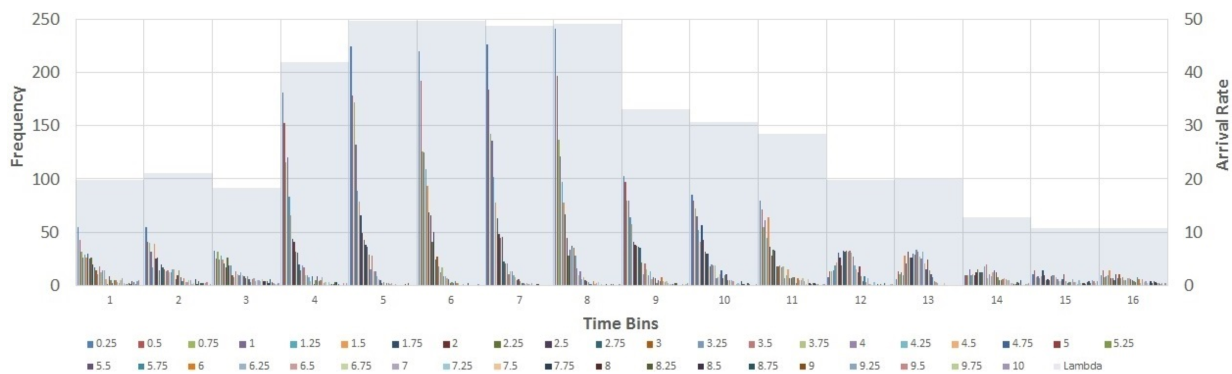
Clearly, the quality and representativeness of the resulting HistoRIA plot depends on the choice of its parameters (length of the time blocks and bin size in the histograms). Although the proposed tool (as will be discussed in the next section) fully automates the HistoRIA plot generation process, which enables the user to easily try different parameter settings in a short amount of time, a brief discussion on parameter selection could still be worthwhile especially for novice modelers. In general, the choice of an appropriate time block length and bin size depends on two main factors: the level of accuracy that is desired/necessary for identification of the IID time intervals and the amount of data available. The required



(a) 1-hour time blocks



(b) 30-minute time blocks



(c) 15-minute time blocks

Figure 6: The HistoRIA plots of the arrival process

level of accuracy essentially depends on the impact of the input on the validity of the simulation model as well as the likelihood and cost of making an incorrect decision due to an insufficiently accurate input model. The modeler needs to determine whether the difference among shorter time blocks has any practical significance. This is a matter of subjectivity and depends on the problem at hand. In our restaurant example, based on the type of decisions to be made using the simulation model, the modeler needs to decide whether identifying a difference between the arrival processes of 15-minute or 30-minute time intervals is going to have an impact on the final decision even if they are in fact found to be different.

It is also important to have enough observations/data points to construct histograms that can effectively represent the underlying distribution. The amount of data is not an issue in the case of a high-throughput process where many observations can be collected in a relatively short time or when sufficient pre-collected data are available through sensors and tracking devices such as RFIDs (Miller, Ferrin, Flynn, Ashby, White, and Mauer 2006). With the lack of historical data and for the case of a stochastic process with a slow rate, however, collecting the necessary amount of data to characterize the probability distribution for a short time block requires extensive data collection and thus may be practically infeasible in most cases. In general, the shorter the length of the time blocks, the more the data required to be able to make reasonable hypotheses about the characteristics of the stochastic process. To illustrate these issues, we have further decreased the length of each time block in the HistoRIA plot of our dataset to 5 minutes in Figure 7. It seems that for some of the time blocks, especially for those that are associated with lower arrival rates, we do not have enough observations of inter-arrival times to generate a meaningful histogram that clearly represents the shape of the underlying distribution. As a result, it is not really possible to hypothesize the IAT distribution for the last 10 or 15 time blocks by looking at their respective histograms.

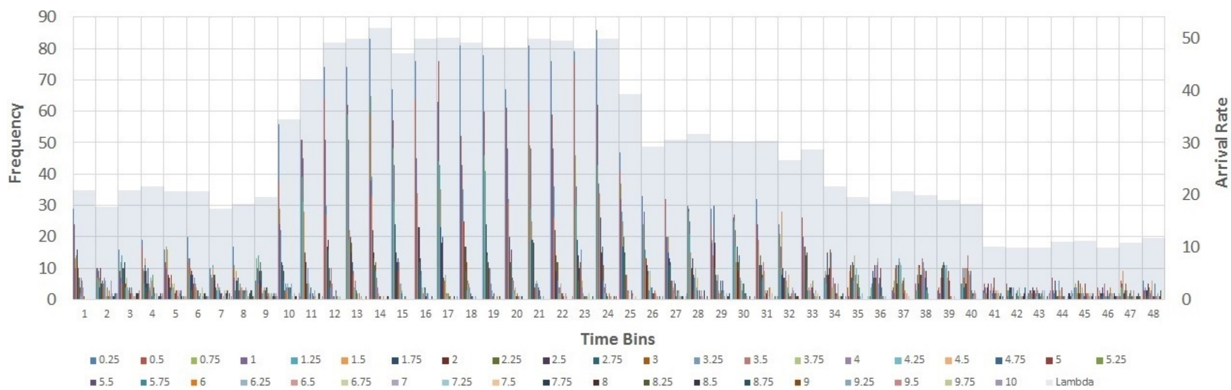


Figure 7: The HistoRIA plot of arrival data with 5-minute time blocks

There are no general guidelines or definitive rules for determining the *optimal* bin size for plotting a histogram (Vincent 1998). While too small a width will produce a ragged histogram, too wide a width will result in a block-like histogram that over-aggregates the data. As a rule of thumb, we suggest a time block length so that the number of observations in each time window (n) is at least 100 to 200 (Vincent 1998) while the Sturges' rule is used to select the number of bins (k) for a given time window as follows: $k = \lfloor 1+\log_2(n) \rfloor = \lfloor 1+3.322\log_{10}(n) \rfloor$ (Law and Kelton 2000). The bin size for each time block is then calculated by dividing the range of data in that time window (maximum - minimum) by k. Since we use a common bin size for all of the histograms in the HistoRIA plot, the largest bin size over all time blocks can then be used to generate the plot. However, the best advice is to experiment with different time blocks and bin sizes and choose a HistoRIA plot that well communicates the shape of the distributions in different time windows while the above rule can be used as a starting point for the experimentation.

Finally, the time blocks in the HistoRIA plot do not necessarily need to represent time of day. They can also represent different operating shifts, days of the week, weeks in a month, months in a year and so forth. For instance, Figure 8 illustrates the HistoRIA plot generated for a month of real data collected during lunch rush (11:00 a.m. to 1:00 p.m.) from a local restaurant in Auburn, AL. Here, the HistoRIA plot is used to investigate whether there is any difference in the arrival rate and the inter-arrival time distribution for different days of the week. In this plot, each bin of the histograms represents half a minute. Based on the generated HistoRIA plot, we could hypothesize that the arrival rate (customer per hour) varies from day to day with Friday having the maximum arrival rate during lunch rush. The plot also suggests that the exponential distribution seems to be a reasonable hypothesis for the distribution of inter-arrival time in each weekday.

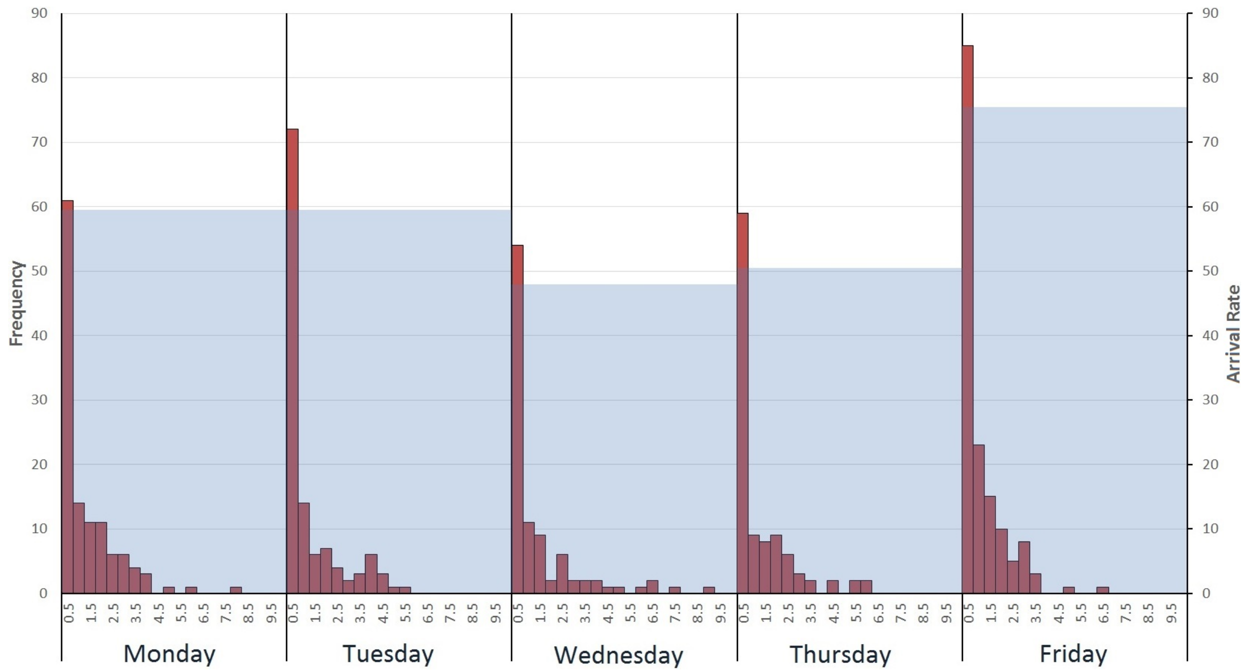


Figure 8: The HistoRIA plot of real-world data collected from a restaurant

4 DETAILS FOR HistoRIA CONSTRUCTION

We now turn to the construction details of the HistoRIA plot. The data is required to include time-date stamps so that each data point has the information on the actual date and time that it was collected. This will provide us with enough information to investigate any potential differences between time windows regardless of what time frame they represent. It is worth noting that the plot is not necessarily limited to studying arrival processes. The data may represent processing times of different resources (such as workers, machines, workstations, etc.) and the generated plot can be used to assess IID in different time blocks.

An Excel/Visual Basic for Applications (VBA) tool is developed to generate the HistoRIA plot which is available online at http://jsmith.co/node/128. Once the data are available, the user only needs to specify the length of each interval and the bin size in the histograms. The tool will then automatically assign each data point to its corresponding time block. For each interval, the rate is simply obtained by counting the number of observations which can be shown by a straight line or a bar similar to the second traditional approach discussed in Section 2. Next, the histogram of the data in each block is generated and the HistoRIA plot is then complete by placing the histograms side by side. Figure 9 summarizes the general steps of the code for generating the HistoRIA plot. Moreover, Figure 10 shows the steps that the user needs to follow to use the HistoRIA tool. To the best of our knowledge, this is the first tool that integrates the first three basic steps of input analysis described in Section 1.

Finally, the tool is designed such that it also provides the means for convenient data collection. In many simulation projects where pre-collected data is available, the data necessary to perform the simulation analysis still needs to be collected because the existing data can be: (1) recorded in an order other than that in which it was collected; (2) already grouped into intervals; (3) recorded with insufficient precision (for example, rounded to the closest integer); and, (4) stored without recording the time-date stamps which inhibits testing the IID assumptions (Vincent 1998). The developed tool can be used for data collection, allowing the user to record observations in the required time stamp format. The collected data can then be used by the tool to generate the HistoRIA plot.

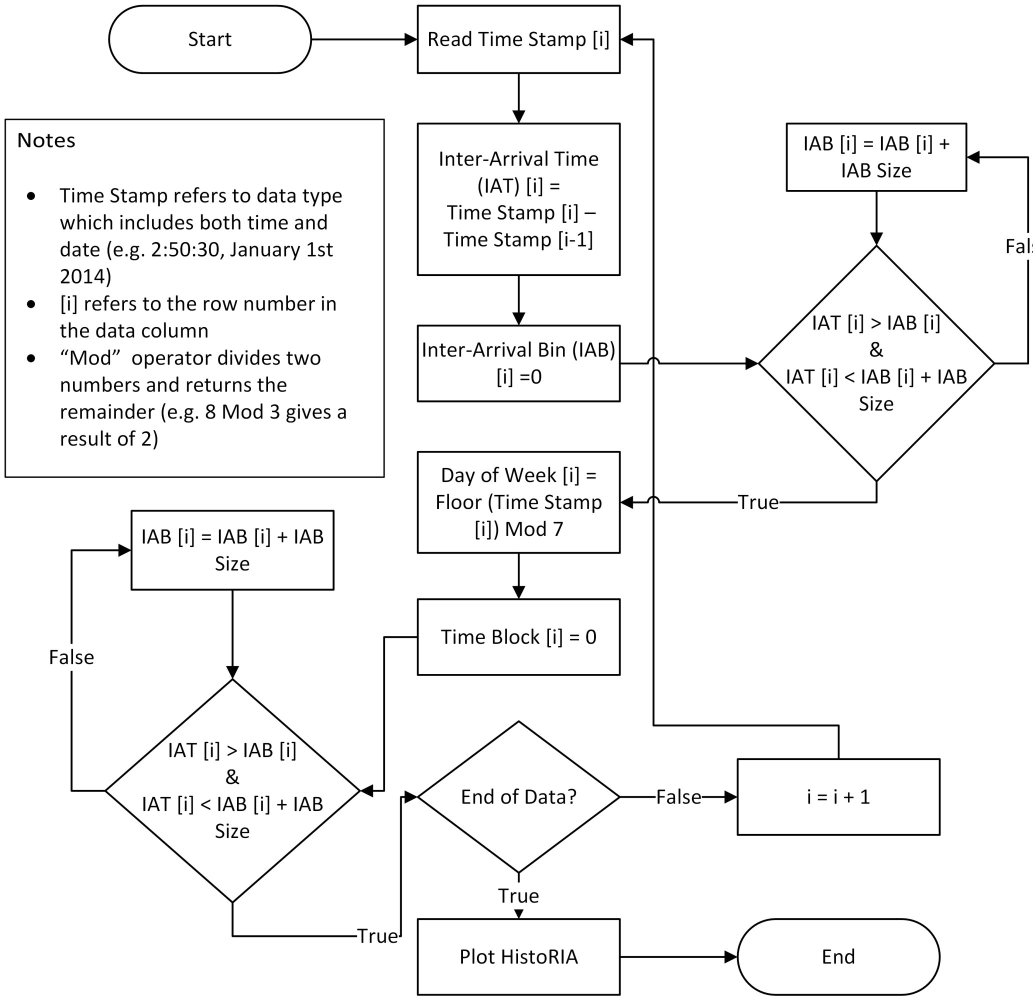


Figure 9: Flow Chart for constructing the HistoRIA plot

5 POTENTIAL NON-SIMULATION APPLICATION AREAS

To this point, we have illustrated the usefulness of the HistoRIA plot for simulation input analysis. However, there are other fields that can also benefit from this new graphical representation. Here, we illustrate how the HistoRIA plot can be useful in *quality control* and *transportation* as two potential non-simulation application areas. It is worth noting that since the spreadsheet-based tool was originally designed and developed for simulation input analysis, some modifications may be necessary to set up the tool to handle other types of data used in other fields.

In the field of quality control, histograms are typically used to illustrate the probability distribution of the number of defects per unit of product. The histogram shown in Figure 11 is obtained by inspecting 1000 products from each of the three operating shifts of a manufacturing line. Based on this histogram, one could assume that the number of defects per product follows a normal or triangular distribution truncated on both sides and base the rest of the analysis on this assumption. However, as shown before, what the histogram fails to demonstrate is how these defects occur over time (i.e., over the three shifts). The HistoRIA plot for the exact same data set, on the other hand, can reveal any pattern or change in the distribution of this metric over time. As shown in Figure 12, the HistoRIA plot suggests that in the first and third operating shifts, the number of defects per product seems to follow a uniform distribution while it seems to be triangularly distributed in the second shift. Moreover, the plot demonstrates how the defect

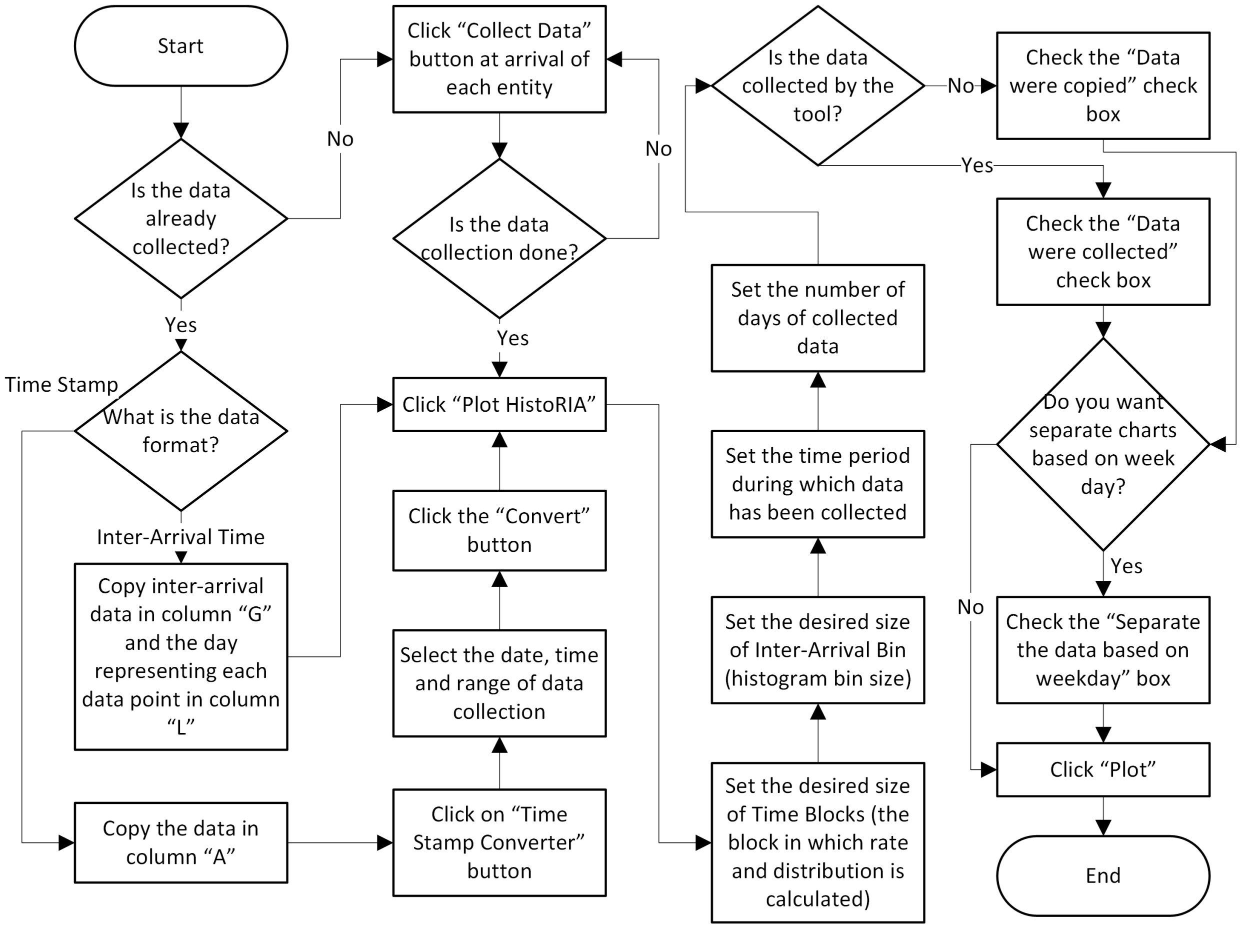


Figure 10: Developed tool user guide flow chart

rate (i.e., average defect per unit of product) increases with time (a behavior that is typically observed as the equipment and tools wear out or become out of calibration with time). As a result, the HistoRIA plot provides important information that allows for a more robust decision making about possible quality improvement strategies for for different shifts in this manufacturing system.

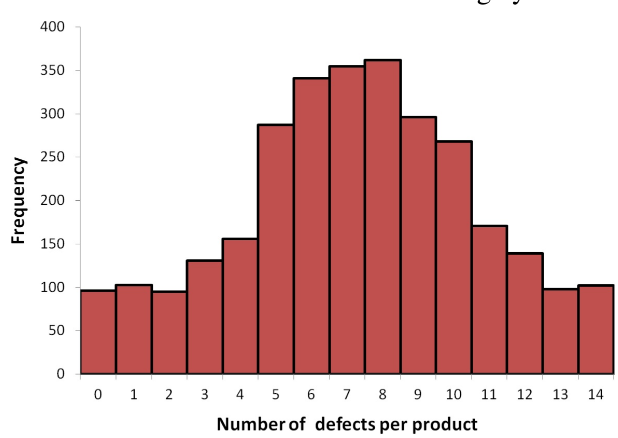


Figure 11: The histogram of the number of defects

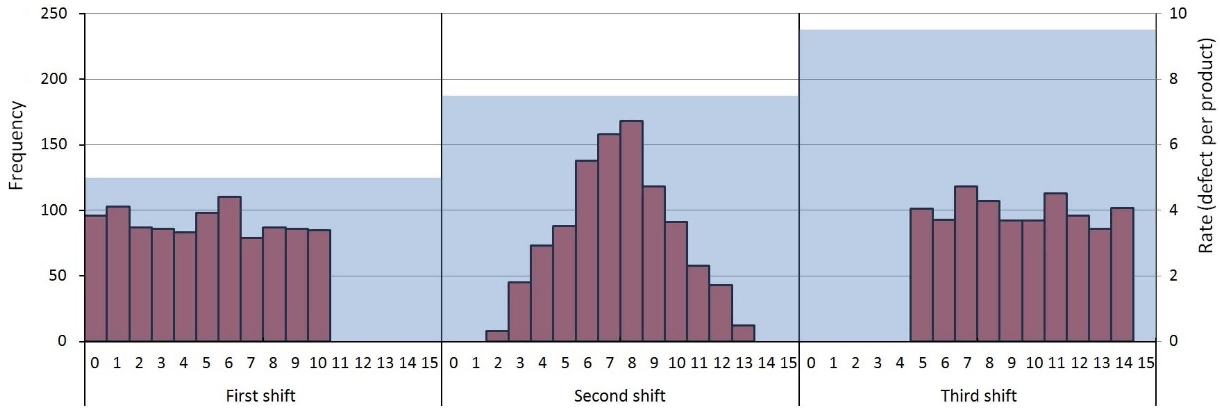


Figure 12: The HistoRIA plot of the rate and distribution of defects per unit

With relation to transportation management, due to the use of sensors and cameras, there is a massive amount of data available on the traffic on different roads which can be used for urban planning purposes. A HistoRIA plot of the number of cars traveling on a certain road or through an intersection at different times of the day can help identify the peak hours and support appropriate decision making. Moreover, by creating and comparing HistoRIA plots for different days of the week, it would be possible to capture any differences between time blocks of calendar days in terms of the traffic load.

6 CONCLUSIONS

A fundamental step in input analysis is concerned with identifying any inconsistencies with the IID assumptions. A general type of deviation from the IID requirement that is commonly found in simulation occurs when the stochastic process being modeled is non-stationary. The process of identifying non-stationarity of the data is generally known to be time-consuming requiring multiple trial and error steps which puts a significant burden on the analyst. In this paper, the HistoRIA plot is presented to facilitate identifying and modeling non-stationary stochastic processes.

By using the actual time that each data point is collected, the HistoRIA plot provides the necessary information to identify how the rate and distribution of the process being modeled changes over time and thus support sound input analysis and correct decision making. The usefulness of the HistoRIA plot for simulation input analysis as well as other application areas is illustrated through several hypothetical and real-world examples. A spreadsheet-based tool that facilitates data collection and automates the generation of the HistoRIA plot is also developed and made available online to simulation users. It is important to note that while the HistoRIA plot helps the analyst identify non-stationary stochastic processes and hypothesize the underlying distribution in each interval, it should not be regarded as a complete input analysis tool. Existing statistical techniques are still needed to estimate distribution parameters and evaluate the goodness-of-fit for the hypothesized distributions.

In general, the HistoRIA plot can be helpful whenever we are interested in the rate and distribution of a certain metric that can vary with time. Therefore, the authors believe that the list of non-simulation application areas will certainly grow as researchers and practitioners in different areas begin to adopt the plot to represent other types of data.

REFERENCES

Benneyan, J. C. 1998. “Software review: Stat:: Fit”. *OR/MS Today* 25 (1): 38–41.

Hammann, J. E., and N. A. Markovitch. 1995. “Introduction to ARENA”. In *Proceedings of the1995 Winter Simulation Conference*, edited by C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, 519–523. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Law, A. M. 2009. "How to build valid and credible simulation models". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 24–33. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*. New York, USA: McGrae-Hill Inc.

Miller, M. J., D. M. Ferrin, T. Flynn, M. Ashby, K. P. White, and M. G. Mauer. 2006. "Using RFID technologies to capture simulation data in a hospital emergency department". In *Proceedings of 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 1365–1371. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ryan, B. F., B. L. Joiner, and J. D. Cryer. 2012. *MINITAB Handbook: Update for Release 16*. Kentucky, USA: Brooks/Cole Publishing Co.

Schittkowski, K. 2008. "EASY-FIT User Guide".

Swain, J. J. 2011. "Discrete event simulation software tools: Back to the future". *OR/MS Today* 38 (5): 56–69.

Vincent, S. 1998. "Input Data Analysis". In *Handbook of Simulation*, edited by J. Banks, 55–90. Hoboken, NJ: John Wiley & Sons, Inc.

Zaixiang, T., G. Qingsong, and X. Chenwu. 2005. "Nonlinear Equation Fitting by Microsoft Office Excel and Its Application [J]". *Chinese Agricultural Science Bulletin* 3:087.

AUTHOR BIOGRAPHIES

MOHAMMADNASER ANSARI is a teaching assistant and PhD student in the Department of Industrial and Systems Engineering at Auburn University. He received his B.Sc. in Industrial Engineering from University of Tehran, Iran. His research interests are in simulation, visualization analytics, supply chain management, and distribution center design and operations. His email address is: ansarim@auburn.edu.

ASHKAN NEGAHBAN is a doctoral student at the Department of Industrial and Systems Engineering at Auburn University. He received his Master's and Bachelor of Science degrees both in Industrial Engineering from Auburn University and University of Tehran, respectively. He has taught courses in simulation modeling and analysis at the undergraduate level. His research interest is in the application of discrete event and agent-based simulation in manufacturing systems and marketing. He is a member of INFORMS and his email address is anegahban@auburn.edu.

FADEL M. MEGAHED is an Assistant Professor in the Department of Industrial and Systems Engineering at Auburn University. He received his Ph.D. and M.S. in Industrial and Systems Engineering from Virginia Tech, and his B.S. in Mechanical Engineering from the American University in Cairo. He is the recipient of the Mary G. and Joseph Natrella Scholarship (2012) from the American Statistical Association. His research interests are in the areas of data analytics, data visualization, statistical quality control, and reliability. His work in these areas has been funded by the *NIOSH Deep South Center for Occupational Safety and Ergonomics*, *Proctor and Gamble (P&G) Fund of the Greater Cincinnati Foundation*, *Amazon Web Services (AWS)*, *Windows Azure (Microsoft)*, and the *National Science Foundation*. His email address is fmegahed@auburn.edu.

JEFFREY S. SMITH is the Joe W. Forehand Professor of Industrial and Systems Engineering at Auburn University. He has served as the WSC Business Chair (2010) and General Chair (2004) and is currently on the WSC Board of Directors. He has a BIE from Auburn University and a MS and PhD (both in Industrial Engineering) from Penn State University. His email and web addresses are: jsmith@auburn.edu and http://jsmith.co.