Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

THE BIVARIATE MEASURE OF RISK AND ERROR (BMORE) PLOT

Mi Lim Lee

College of Business Administration
Hongik University
Sangsu-dong 72-1
Mapo-gu, Seoul, 121-791, SOUTH KOREA

Chuljin Park

Department of Industrial Engineering
Hanyang University
222 Wangshimniro
Seongdonggu, Seoul, 133-791, SOUTH KOREA

ABSTRACT

We develop a graphical method, namely the bivariate measure of risk and error (BMORE) plot, to visualize bivariate output data from the stochastic simulation. The BMORE plot consists of a sample mean, median, minimum/maximum values for each measure, an outlier, and the boundary of a certain percentile of the simulation data on a two-dimensional space. In addition, it depicts confidence regions of both the true mean and the percentile to show how accurate the two estimates are. From the BMORE plot, scholars, practitioners, and software engineers in simulation fields can understand the variability and potential risk of the simulation data intuitively, design simulation experiments effectively, and reduce a great deal of time and effort to analyze the simulation results.

1 INTRODUCTION

In the last decade, the simulation had more attentions to analyze complicated industrial and service systems under various uncertainties. Scholars, practitioners, and software engineers have developed numerous simulation software that can analyze the systems with cutting-edge techniques. Nevertheless, most results from the simulation software have focused on static numbers representing estimates of mean or long-run average values, and thus it is hard to intuitively understand risk measures from the results.

Since risk measures and their errors are important to understand uncertainties in the systems (Savage 2009), simulation researchers and practitioners have studied how to visualize them effectively. For univariate output data, Nelson (2008) introduces a very intuitive and easy-to-implement plot, namely the Measure of Risk and Error (MORE) plot. The MORE plot shows not only the sample mean and percentiles but also their confidence intervals based on the histogram of simulation data. From the plot, the users can catch the ideas of measures of risk and errors immediately and intuitively. A commercial simulation software, SIMIO provides the MORE plot as a graphical presentation of simulation results from multiple across-replications (Kelton et al. 2014).

Recently, many researchers have more interests in analyzing bivariate output data. Specifically, bi-objective optimization problems and stochastically constrained problems require estimation of multiple performance measures. For example, Bekker (2013) formulates a buffer allocation problem as a bi-objective optimization problem with two performance measures, the work-in-process (WIP) and the throughput. In addition, Park et al. (2014) considers a water quality monitoring network design problem as a stochastically constrained simulation optimization problem with two performance measures, the minimum elapsed time for contaminant detection and the probability of the detection in a real river system. When considering bivariate output data, one may want to know correlation and skewness, or decide whether we need more within and across replications or not. Therefore, it becomes necessary to develop a graphical method to show simulation output data for two performance measures simultaneously in the two-dimensional space.

In this paper, we newly develop a bivariate MORE (BMORE) plot that enables users to see bivariate results from a stochastic simulation immediately. The BMORE plot consists of a sample mean, median, minimum/maximum values, outlier, and boundary of a certain percentile of the result data that are close to the median in the Cartesian coordinate. In addition, it provides confidence regions of the true mean and the percentile to show how accurate the estimates are. After introducing the structure of the BMORE plot, we present how to construct the BMORE plot. Then, based on simulation results of a buffer allocation problem as an example, we discuss the experimental results, followed by concluding remarks.

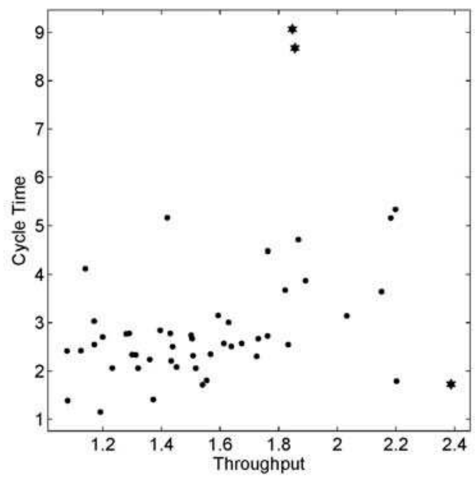
2 THE BMORE PLOT

The BMORE plot demonstrates simulation output data under two performance measures simultaneously. In this section, we introduce the main components of the BMORE plot.

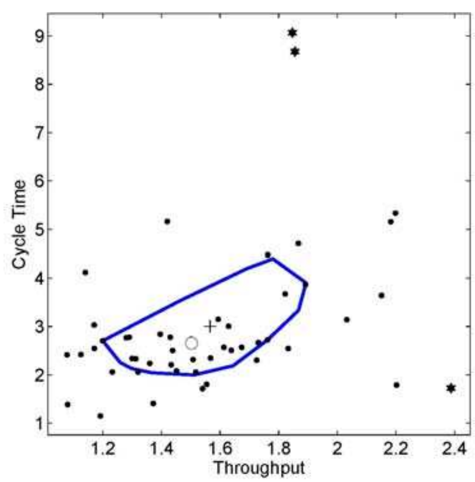
Since one of the most popular methods to visualize a set of bivariate data is a scatter plot, the BMORE plot employs the format of the scatter plot. The BMORE plot uses the Cartesian coordinate with a horizontal axis for one measure and a vertical axis for the other measure. Each bivariate observation is plotted as a point in the coordinate. Based on the plot, we may roughly capture the variability of simulation data and the correlation between two performance measures. For example, Figure 1 (a) shows a scatter plot of 50 bivariate observations from a three-buffer allocation problem (Pichitlamken and Nelson 2003). In Figure 1 (a), one might recognize that two measures, the throughput and the cycle time, are positively correlated.

In terms of basic statistical analysis, the next step that naturally comes after drawing a scatter plot is finding the sample mean and median of the bivariate data. In Figure 1 (b), a plus symbol marks the sample mean, and a white circle marks the median. The difference between locations of the sample mean and median indicates how much the distribution of the data is skewed. In addition to the sample mean and the median, we draw a boundary of $100\times\beta\%$ of data that are close to the median, where β is a user-specified value between 0 and 1. The boundary is called *β-bag*, and the 0.5-*bag* for the three buffer allocation problem is shown as a bold line in Figure 1 (b). The *β-bag* is a bivariate generalization of the percentile in the MORE plot (Nelson 2008) that considers two ranks of univariate data (i.e., observations with ranks $\lfloor n\cdot 0.25\rfloor$ and $\lceil n\cdot 0.75\rceil$ where n is the number of simulation runs). If the *β-bag* is large, the variability of the observation data is likely to be large, and thus users can perceive the future risk regarding two measures. In Figure 1 (b), star symbols represent outliers. The specific methods to draw the *β-bag* and select the outliers are discussed in the following section.

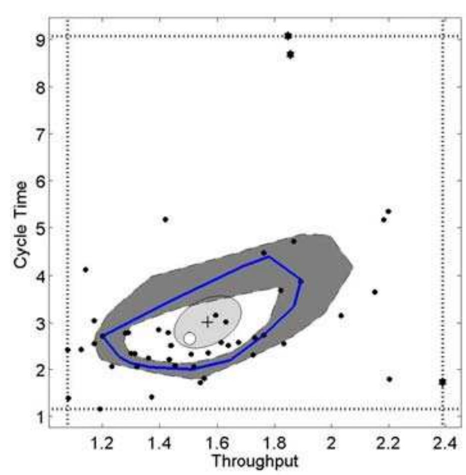
Even though the sample mean and the *β-bag* in the BMORE plot provide useful information about the central tendency, skewness and variability of the data, users may wonder how accurate the information is unless the number of across-replications is large enough. For the users, the BMORE plot shows confidence regions of the true mean and the *β-bag* as shown in Figure 1 (c). Let α be another user-specified value between 0 and 1, representing a significance level of a confidence region. In Figure 1 (c), the region in light grey around the sample mean represents an approximated $100\times(1-\alpha)\%$ confidence region of the mean, and the region in dark grey represents an approximated $100\times(1-\alpha)\%$ confidence region of the *0.5-bag*, where α equals 0.05. (The detailed methods to construct the confidence regions of the mean and the *β-bag* are discussed in the following section.) The BMORE plot also shows the minimum and maximum values of the observations for each measure as bold dotted lines as shown in Figure 1 (c).



(a) Scatter plot



(b) Sample mean, median, and the 0.5-*bag*



(c) Confidence regions of the mean and the 0.5-*bag*

Figure 1: Construction of a BMORE plot.

3 CONSTRUCTION OF THE BMORE PLOT

In this section, we describe specific methods to construct the BMORE plot. The BMORE plot consists of the sample mean, median, β-*bag*, minimum/maximum values for each measure, outliers, and confidence regions of the true mean and the β-*bag* on a scatter plot. Since the methods to obtain other values are trivial, we mainly focus on explaining how to construct the sample mean, the confidence region of the true mean, the β-*bag*, and the confidence region of the β-*bag*.

3.1. Sample Mean and the Confidence Region

Let X and Y be the two variables representing two performance measures that we are interested in. We define $\mathbf{Z} = (X, Y)$ and $\mathbf{Z}_i = (X_i, Y_i)$, representing the i^{th} simulation observation of Z. When n is the number of across-replications of the simulation, $\mathbf{Z}_i$ for $i=1,2,\ldots n$ are independent and identically distributed (i.i.d.). Then, we define a sample mean vector $\overline{\mathbf{Z}}_n = (\overline{X}_n, \overline{Y}_n)^T$, where $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

Regardless of the normality of $\mathbf{Z}_i$, $\overline{\mathbf{Z}}_n$ follows an approximately bivariate normal distribution by the central limit theorem if n is large enough. Therefore, we use the Hotelling's T^2statistics to construct a confidence region of the true mean of $\overline{\mathbf{Z}}_n$. Let $\boldsymbol{\mu} = (\mu_X, \mu_Y)^T$ define a true mean vector where μ_X and μ_Y are the unknown true mean values of X and Y respectively. Then, from Mason and Young (2002), the $100 \times (1 - \alpha)\%$ confidence region of the true mean is constructed by $\boldsymbol{\mu}$ satisfying the following inequality:

$$\left(\overline{\mathbf{Z}}_n - \boldsymbol{\mu}\right)^T \Sigma^{-1}\left(\overline{\mathbf{Z}}_n - \boldsymbol{\mu}\right) \leq \frac{2(n-1)}{n(n-2)} F_{(2,n-2)}(\alpha), \tag{1}$$

where Σ is the variance-covariance matrix of $\mathbf{Z}_i$ and $F_{(2,n-2)}(\alpha)$ is the $1-\alpha$ quantile of the F distribution value with 2 and $n-2$ degrees of freedom.

In the following subsection, we discuss the β-*bag* and its confidence region.

3.2. $\boldsymbol{\beta}$-*bag* and the Confidence Region

Since the distribution of the simulation observations is often unknown in practice, we first consider a nonparametric method to construct a β-*bag*. Rousseeuw et al. (2012) develop a bivariate generalization of the univariate box plot, namely bag plot. The bag of the bag plot is designed to contain 50% of the data points near the median. Based on the bag plot, the β-bag is designed to contain a certain percentage (i.e., $100 \times \beta\%$) of data points near the median. As in the original bag plot, we define a set of all simulation observations in two-dimensional real space as $\Omega = \{\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_n\}$ and use the same location depth function $ldepth(\boldsymbol{\theta}, \Omega)$ for some point $\boldsymbol{\theta} \in \mathrm{R}^2$ relative to Ω. The location depth function, $ldepth(\boldsymbol{\theta}, \Omega)$, represents the smallest number of simulation observations (i.e., $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_n$) in Ω that are included by any closed halfspace with a boundary line through $\boldsymbol{\theta}$. See Tukey (1975) for an introduction to the location depth function $ldepth(\boldsymbol{\theta}, \Omega)$ and also see Rousseeuw and Tukey (1996) for a detailed algorithm for calculating the $ldepth(\boldsymbol{\theta}, \Omega)$. For a non-negative integer k, let Λ_k be a set of $\boldsymbol{\theta}$ having $ldepth(\boldsymbol{\theta}, \Omega) \geq k$, and $\#\Lambda_k$ is the number of simulation observations $\boldsymbol{Z}_{\mathrm{i}}$ in Λ_k. Note that $\Lambda_k \subseteq \Lambda_{k-1}$ always hold for $k > 0$. Then, we can find a β-*bag* for bivariate observations by following the two steps:

Step 1: Find $k > 0$ such that $\#\Lambda_k \leq \lfloor n \cdot \beta \rfloor < \#\Lambda_{k-1}$.
Step 2: Linearly interpolate between $\Lambda_k$and Λ_{k-1}.

The median is defined as the $\boldsymbol{\theta}$ with the highest location depth function value if such $\boldsymbol{\theta}$ is unique. If there are multiple $\boldsymbol{\theta}s$ with the same highest location depth, the geometric center of the $\boldsymbol{\theta}$s is selected as the median. When we enlarge the *0.5-bag* by a scale factor 3 relative to the median and define its boundary as fence, then the points located outside of the fence are marked as outliers.

A confidence region of a β*-bag* consists of inner and outer boundaries. Let β_L and β_U be lower and upper bounds of β with a significance level α and n observations, then β_L and β_U can be approximated by the followings equations (Nelson 2008):

$$\beta_L = \beta - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\beta(1-\beta)}{n-1}} \text{ and } \beta_U = \beta + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\beta(1-\beta)}{n-1}}, \quad (2)$$

where $Z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution. By using β_L and β_U instead of β in the above steps, we can obtain inner and outer boundaries of the $100 \times (1-\alpha)$% confidence region of the β*-bag,* respectively.

As a special case, if the observation data follow a bivariate normal distribution or if the simulation observations are either within-replication averages with enough simulation or batch means, we can use the Hotelling's T^2 statistic to construct an approximated β-bag and its confidence region instead of the non-parametric methods. In order to assess if observations are normally distributed, numerous tests including the Shapiro-Wilk test, the Kolmogorov-Smirnov test, and Anderson-Darling test are available. In this case, similar to the construction method in Section 3.1, an approximated β-bag can be defined by the boundary of a set containing all Z values satisfying the following inequality (Mason and Young 2002):

$$\left(\boldsymbol{Z} - \overline{\mathbf{Z}}_n\right)^T \Sigma^{-1}\left(\boldsymbol{Z} - \overline{\mathbf{Z}}_n\right) \leq \frac{2(n+1)(n-1)}{n(n-2)} F_{(2,n-2)}(\beta). \quad (3)$$

In order to obtain the inner and outer boundaries of the $100 \times (1-\alpha)$% confidence region of the β-*bag*, one can insert β_L and β_U from (2) instead of β to the inequality (3). Note that constructing β-bag with the Hotelling's T^2 statistic under the normality assumption is much faster than constructing β-bag with the location depth function under the non-normality assumption.

4 EXPERIMENTAL RESULTS

As a test problem, we examine a three-buffer allocation problem (Buzacott and Shanthikumar 1993) to see how the BMORE plot works. Figure 2 shows the basic structure of the system having three buffers. In the problem, x_1, x_2, and x_3 represent service rates of the server 1, 2, and 3, and x_4 and x_5 represent the buffer sizes in front of the servers 2 and 3 respectively. Service rates and buffer sizes are assumed to be integers. We assume that all service times are exponentially distributed, and an infinite number of jobs exist in front of the server 1. If the buffer of station *i* is full, then the server *i*-1 is blocked, and the finished job on the server *i*-1 cannot move into the station *i*. Under the constraints on the total buffer size (i.e., $x_4 + x_5 \leq 20$) and total service rate (i.e., $x_1 + x_2 + x_3 \leq 20$) for the stations, Pichitlamken and Nelson (2003) finds an optimal design, $(x_1, x_2, x_3, x_4, x_5) = (7, 6, 6, 8, 12)$, for the maximum averaged throughput.

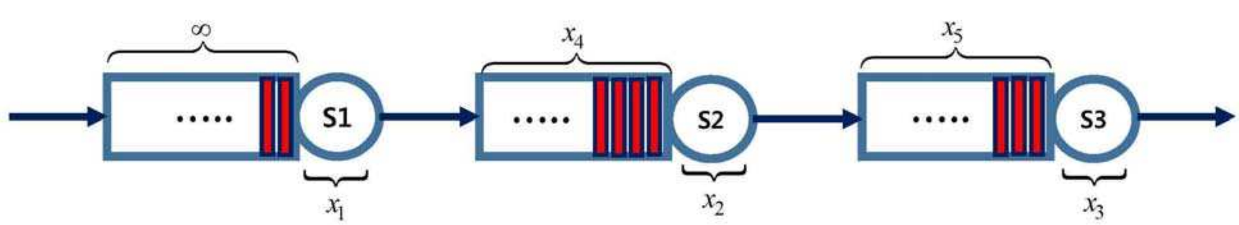


Figure 2: Three-buffer allocation problem.

In this section, we show the BMORE plots with the design, (x_1, x_2, x_3, x_4, x_5) = (7, 6, 6, 8, 12), when considering the averaged throughput and the averaged cycle time as performance measures. After discarding the first 2000 units simulated for a warm-up period, the throughput and cycle time are averaged over the subsequent m units released (i.e., the number of the within-replications is m). Let n represent the number of simulation runs (i.e., the number of the across-replications). We perform the simulation with four different choices of (m, n) and then depict the BMORE plots for α = 0.05 and β = 0.5.

When assuming the non-normality of the simulation observations, we draw the BMORE plots based on the nonparametric method as in Figure 3. As the number of within-replications, m, increases, the variance of observation data decreases under both performance measures, and the (negative) correlation between two measures becomes clearer. Outliers are marked by star symbols in the figure. As the number of across-replications, n, increases, the confidence regions become smaller, and thus we get more accurate estimates of the mean and the 0.5-*bag*. Note that the shapes of the β-*bag* and the boundaries of the confidence region of the 0.5-*bag* are different and the shapes change as the number of observations changes. The number of points in the 0.5-*bag* is almost exactly a half of the total number of observations. Similarly, inside areas of the inner and outer boundaries of the confidence region of the 0.5-*bag* also include almost exactly $\lfloor n \cdot \beta_L \rfloor$ and $\lfloor n \cdot \beta_U \rfloor$ number of observations respectively.

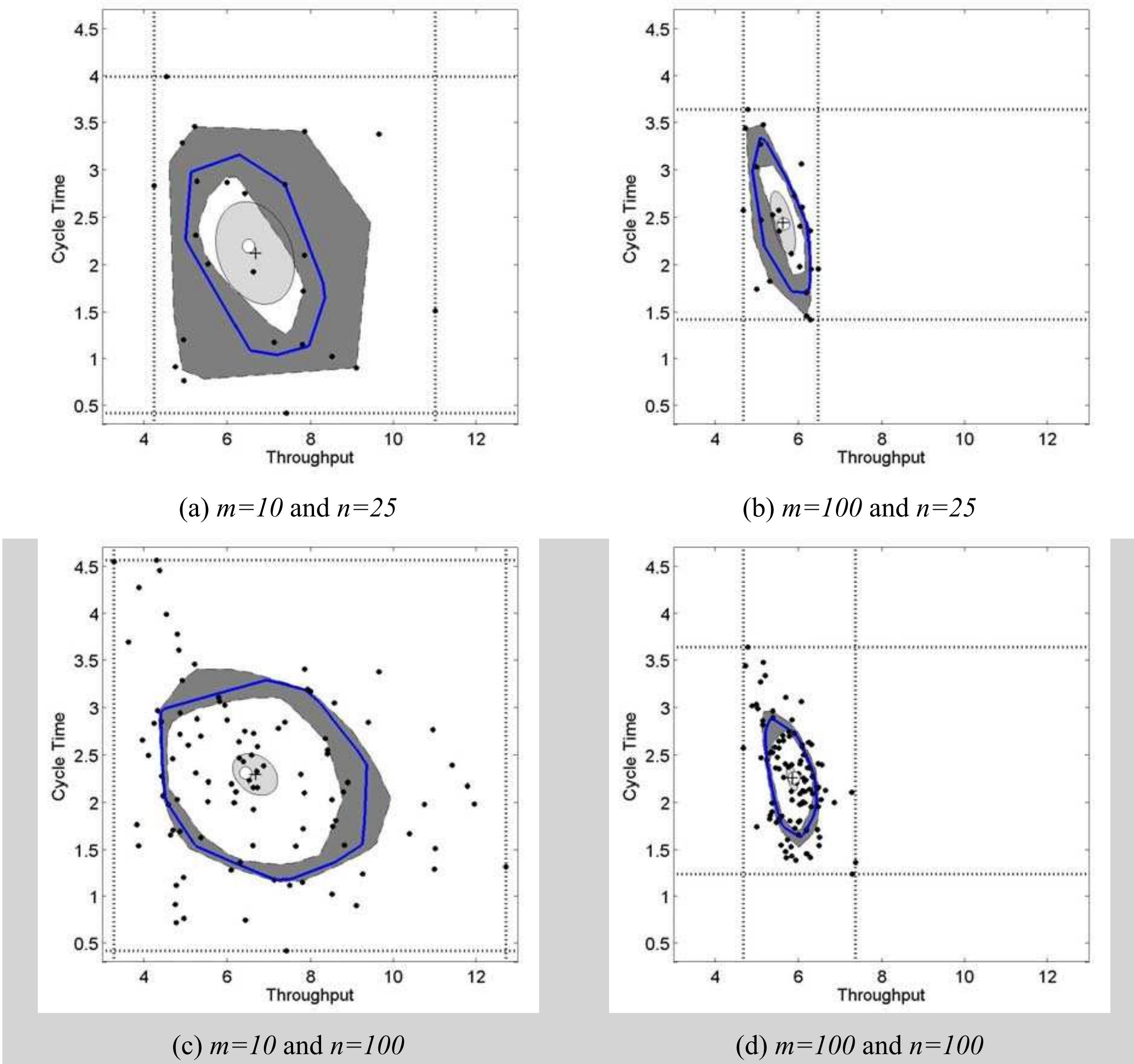
When assuming the normality of the simulation observations (although this assumption might not appropriate for the three-buffer allocation example with small m and n), the BMORE plot based on the Hotelling's T^2 statistic can be used to approximate the β-*bag* and its confidence region and the results are shown in Figure 4. In Figure 4, the features except the 0.5-*bag* and its confidence region remain same as in Figure 3. When compared to the piecewise-linear shapes in Figure 3, the shapes of the 0.5-*bag*s and the boundaries of their confidence regions change to ellipses in Figure 4. Note that the sizes of the ellipses change as m or n changes but the two boundaries of the confidence region of the β-*bag* are evenly apart from the β-*bag* under all pairs of m and n. For small m and n, the BMORE plots based on the location depth function and the Hotelling's T^2 statistic look different from each other as in Figure 3 (a) and Figure 4 (a), while they become more similar as m and n increases. Especially, for small n (i.e., n=25), the BMORE plot Figures 3 (a) and (b) show that the confidence region of the true mean is partly overlapped by the confidence region of the 0.5-*bag,* while Figures 4 (a) and (b) show that no overlapped region exists between two confidence regions. As a result, the BMORE plot based on the location depth function enforces users to increase the number of across replications to get a more accurate sample mean and 0.5-*bag*, while the BMORE plot based on the Hotelling's T^2 statistic can hardly provide such information.

5 SUMMARY AND CONCLUSION

We develop a new chart, namely the BMORE plot, to depict the variability and the measures of risk and error for bivariate simulation observation data. The BMORE plot consists of the sample mean, median, minimum/maximum values, outliers, β-*bag*, and confidence regions of the mean and the β-*bag* on the scatter plot.

By using the method in Section 3.2, the users can construct their β-*bag* and its confidence region no matter what distribution the observation data follow. However, if the simulation observation is either within-replication averages or batch means with enough simulation length and the number of simulation observations is large enough, then the β-*bag* and its confidence region can be constructed by the Hotelling's T^2 statistic.

Compared to many conventional simulation results providing a list of static numbers (such as sample means, long-run average values, standard deviations, and minimum/maximum values), the BMORE plot considers risk measures in addition to the static results and visualize them effectively for bivariate simulation output data. The BMORE plot does not make users spend a great deal of time to analyze the static numbers. Instead, by using the BMORE plot, the users may catch the idea of measures of risk and error at a glance and understand the potential risks as well as the setting of simulation experiments efficiently. If repeatedly applied to different sets of performance measures, the BMORE plot can be used as a tool to analyze the measure of risk and error for multivariate observation data as well. In this case, the setting of β for an overall error rate needs to be adjusted.



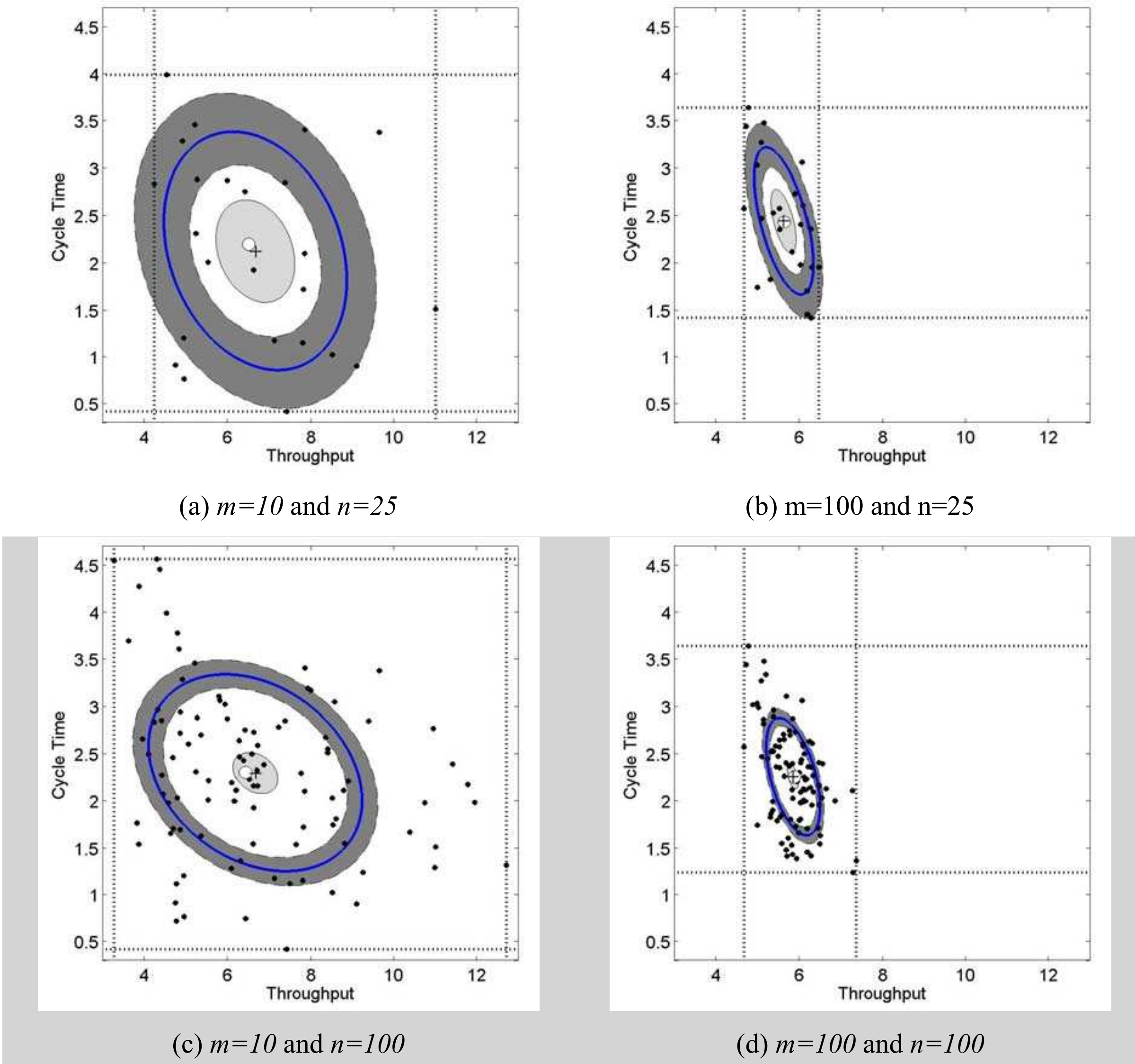
(a) *m=10* and *n=25*

(b) *m=100* and *n=25*

(c) *m=10* and *n=100*

(d) *m=100* and *n=100*

Figure 3: BMORE plots with the non-normality assumption**.**



(a) *m=10* and *n=25* (b) m=100 and n=25

(c) *m=10* and *n=100* (d) *m=100* and *n=100*

Figure 4: BMORE plots with the normality assumption**.**

REFERENCES

Bekker, J. 2013. "Multi-Objective Buffer Space Allocation with the Cross-Entropy Method." *International Journal of Simulation Modelling* 12:50 – 61.

Buzacott, J. A., and J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems (Vol. 4)*. Englewood Cliffs, NJ.

Kelton, W. D., J. S. Smith, and D. T. Sturrock. 2014. *Simio and Simulation: Modeling, Analysis, Applications*. SIMIO LLC.

Mason, R. L., and J. C. Young. 2002. *Multivariate Statistical Process Control with Industrial Applications*. ASA-SIAM.

Nelson, B. L. 2008. “The MORE plot: Displaying, Measures of Risk and Error from Simulation Output.” In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 413 – 416. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Pichitlamken, J., and B. L. Nelson. 2003. “A Combined Procedure for Optimization via Simulation.” *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 13:155-179.

Rousseeuw P. J., and I. Ruts. 1996. "Algorithm AS 307: Bivariate Location Depth.” *Applied Statistics (JRSS-C)*, 45:516-526.

Park, C., I. T. Telci, S.-H. Kim, and M. M. Aral. 2014. “Designing an Optimal Water Quality Monitoring Network for River Systems Using Constrained Discrete Optimization via Simulation.” *Engineering Optimization* 46:107 – 129.

Rousseeuw P. J., I. Ruts, and J. W. Tukey. 2012. “The Bag Plot: a Bivariate Boxplot.” *The American Statistician* 53:382-387.

Savage, S. L. 2009. *The Flaws of Averages: Why We Underestimate Risk in the Face of Uncertainty*. Wiley, Hoboken, New Jersey.

Tukey, J. W. 1975. “Mathematics and the picturing of data.” In *Proceedings of the international congress of mathematicians,* 523 – 531.

AUTHOR BIOGRAPHIES

MI LIM LEE is an Assistant Professor of the College of Business Administration at Hongik University. She received her B.S. and M.S. in Industrial Engineering from Korea Advanced Institute of Science and Technology in 2004 and 2006, respectively and received her Ph.D. in H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology in 2013. Her research interests include applied statistics, simulation modeling, output analysis, and applications of simulation techniques to healthcare service systems. Her e-mail is mllee@hongik.ac.kr.

CHULJIN PARK is an Assistant Professor of the Department of Industrial Engineering at Hanyang University. He received his B.S. in Mechanical Engineering from Korea Advanced Institute of Science and Technology in 2007 and received his M.S. and Ph.D. in H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology in 2010 and 2013, respectively. His research interests include stochastic simulation, optimization via simulation, statistical quality control, and applications of simulation techniques to environmental management. His e-mail is parkcj@hanyang.ac.kr and his webpage address is https://sites.google.com/site/stochsim.