Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

TWO NEW SUBJECTIVE VALIDATION METHODS USING DATA DISPLAYS

Husam Hamad

Electronic Engineering Department
Hijjawi College of Engineering Technology
Yarmouk University
Irbid, JORDAN

Sami Al-Hamdan

Computer Engineering Department
Hijjawi College of Engineering Technology
Yarmouk University
Irbid, JORDAN

ABSTRACT

Three graphical data displays of histograms, box plots, and behavior plots are used in existing literature for subjective model validation. In this paper, we present two additional plots that can be used for displaying graphs of data; these are the so-called circle plots and ordinal plots. These plots are easy to generate using model data and system data. Like the existing plot types, no statistical assumptions are made on the data that are represented. However, more expeditious subjective interpretations about model operational validity are made using the methods presented.

1 INTRODUCTION

Model operational validity is carried out to substantiate the model in terms of its accuracy in representing the system for the intended purposes, over the model's domain of applicability. This involves comparison of data for model and system responses using identical input conditions.

Model validation methods are generally classified into objective methods and subjective methods (Balci 1989, Sargent 1996, and Sargent 2004). Objective methods are based on statistical tests that make certain assumptions about the data's correlation, distribution, etc. A central assumption for the application of these statistics is related to the number of data points used. Even when other assumptions are satisfied, statistical tests are meaningful only if the data used is sufficient in number. It is often the case that, even if the system is completely observable, obtaining a sufficient number of observations is impractically expensive. For such cases, in particular, methods for subjective validation are used to assess model accuracy. One such technique is based on using graphical displays to plot and compare model data and system data.

Graphs displaying model data versus system data using histograms, box plots, and operational (scatter) plots are used in the existing literature (Sargent 1996, Sargent 2004, and Kleijnen and Deflandre 2005). Data are compared by simply 'eyeballing' the resulting plots to 'see' to what extent the model plots and the system plots are identical.

In this paper, we present two more plot types that can be used to display data for subjective comparison. These plots are referred to as *circles plot* and *ordinal plots*, in contrast with box plots and scatter plots, respectively. Like the three exiting graphical techniques, these plots are easy to generate from available data, and applicable to different types of models and systems. Furthermore, the presented approach enables the analyst to quickly develop a 'mental view' of how closely the model represents the system.

The remainder of this paper is organized as follows. In section 2, circles plots and ordinal plots are presented. The various plots discussed in this paper are generated and displayed for an illustrative example in section 3. The paper is then concluded by section 4.

2 GRAPHICAL DISPLAYS OF DATA

Graphical techniques that are used in the existing literature for model validation include histograms, box plots, and scatter plots, as mentioned above. (See Walpole and Myers (1993), Sargent (1996) and (2004), and Kleijnen and Deflandre (2005) for examples of these plots). In this section, we present graphical displays using circles plots and ordinal plots. These plots are contrasted to box plots and scatter plots for the example in the next section.

Suppose we have n pairs of data point $\hat{y}_i$ and y_i, for i = 1, 2, …, n, where $\hat{y}_i$ and y_i are the model and the system responses (outputs), respectively, for the same inputs x_i. Then, essentially, what matters for comparison purposes is the relative sizes $\hat{y}_i/y_i$. In particular, the following information are extraneous:

- The actual sizes of each of the n model responses and the corresponding n system responses.
- The location in space of the n inputs for which the model and the system responses are determined.
- The dimensionality of the inputs space.

Taking these preliminaries into account, circles plots and ordinal plots are constructed as described below.

2.1 Circles Plots

The circular (φ,r) coordinate system is used to generate circles plots; see figure 2d. Model and system data points are used to plot the n points ($2\pi i / n$, $\hat{y}_i/y_i$), for i = 1, 2, …, n. The ratios $\hat{y}_i/y_i$ are ordered before plotting. Points with radial distances $r = \hat{y}_i/y_i$ closer to unity indicate a better match between the model and the system. Ideally, a perfect model is represented by a unit circle centered at the origin. Note that points with system responses y_i equal to zero are simply removed from the data set; this should almost have no effect on the plots credentials, except, of course, if these points happen to be over a certain percentage threshold set by the analyst.

2.2 Ordinal Plots

The Cartesian (x,y) system is used for ordinal plots; see figure 2b. The n model-to-system ratios $\hat{y}_i/y_i$ are ordered, and then assigned the coordinates (x_i , $\hat{y}_i/y_i$), where x_i is the index of the i^{th} data point in the ordered ratios data. The more the resulting data lie closer to y = 1, the better the model data fit the system data. Again, points with system responses y_i equal to zero are simply removed.

Circles and ordinal plots are contrasted to box plots and scatter plots, in the following example.

3 EXAMPLE

The M/M/1 queuing example given in this section is taken from Shin et al. (2002). Figure 1 depicts the variation of the system's response W, the sojourn time, with the service rate η and the arrival rate λ . The relationship between the system response and its inputs is given by

$$W = 1/(\eta - \lambda). \qquad (1)$$

The arrival rate λ is set equal to 1.0, and the service rate η is varied between 1.1 and 10.0. Data for the system and the two models A and B were collected using a set of equally spaced values for η over its range.

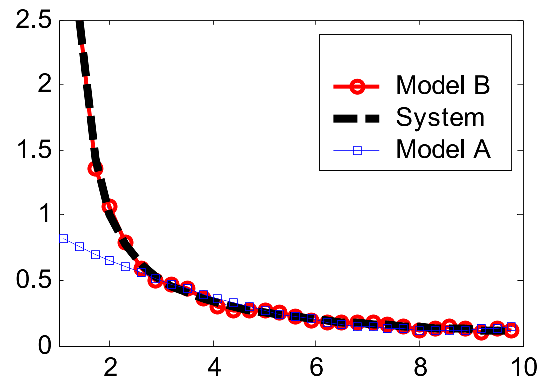

Figure 1: Plots of Model A, Model B, and System Data

Four plots are constructed for each of the two models. Plots for model A are shown in figures 2a-2d, while figures 3a-3d depict the corresponding plots for model B. Parts a, b, c, d of these two figures correspond to scatter plots, ordinal plots, circles plots, and box plots, respectively.

For scatter plots and box plots in both figures, the farthest outliers are not shown in order to clarify these plots. Also, note figure 3d has one open ° dot – this is the convention we use to indicate negative values of $\hat{y}_i/y_i$ when they occur.

When subjective validation is carried out using any of the four parts in figure 2 and the corresponding parts in figure 3, it is obvious that the data for model B more closely represents the system data. However, we claim that this conclusion is more readpparent from ordinal plots and circles plots, compared with scatter plots and box plots. Of course, these methods become even more appreciable when applied to systems with more than one input; e.g., when curves and surface plots cannot be developed for systems with more than two inputs.

4 CONCLUSION

This paper presented two new plot types for use in subjective model validation via graphical displays of data. We call these plots circles plots and ordinal plots, and contrast them to box plots and scatter plots – two of the types used in the literature. These four plots are used when model data and system data do not satisfy the statistical assumptions required for objective validation. Even when objective validation is possible, ordinal plots and circles plots are readily generated and interpreted.

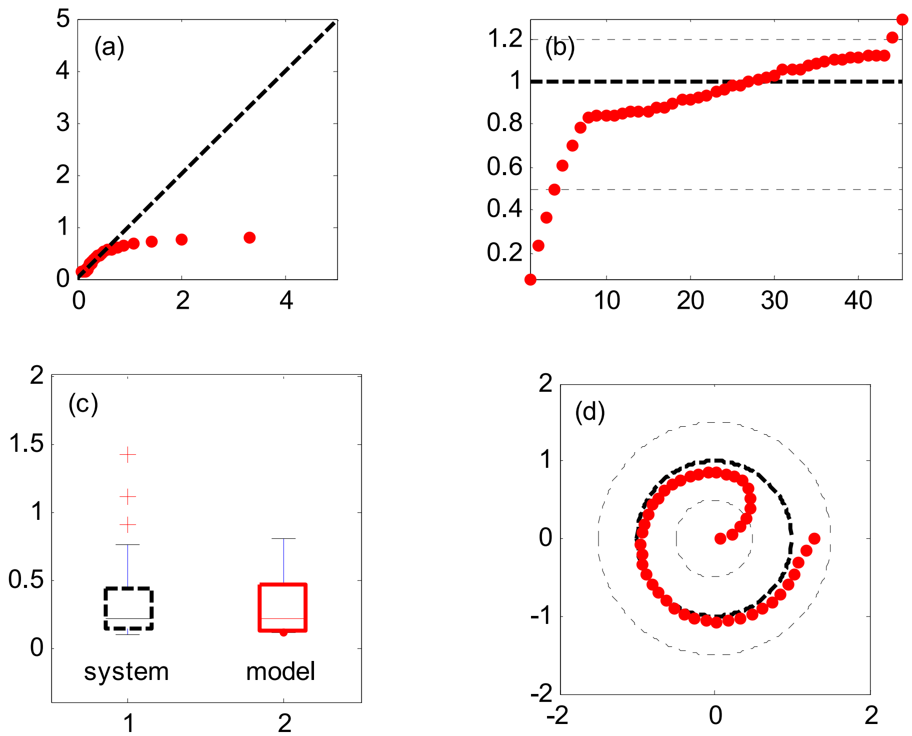

Figure 2: Graphical Displays for Model A (a) Scatter Plots (b) Ordinal Plots (c) Box Plots (d) Circles Plots

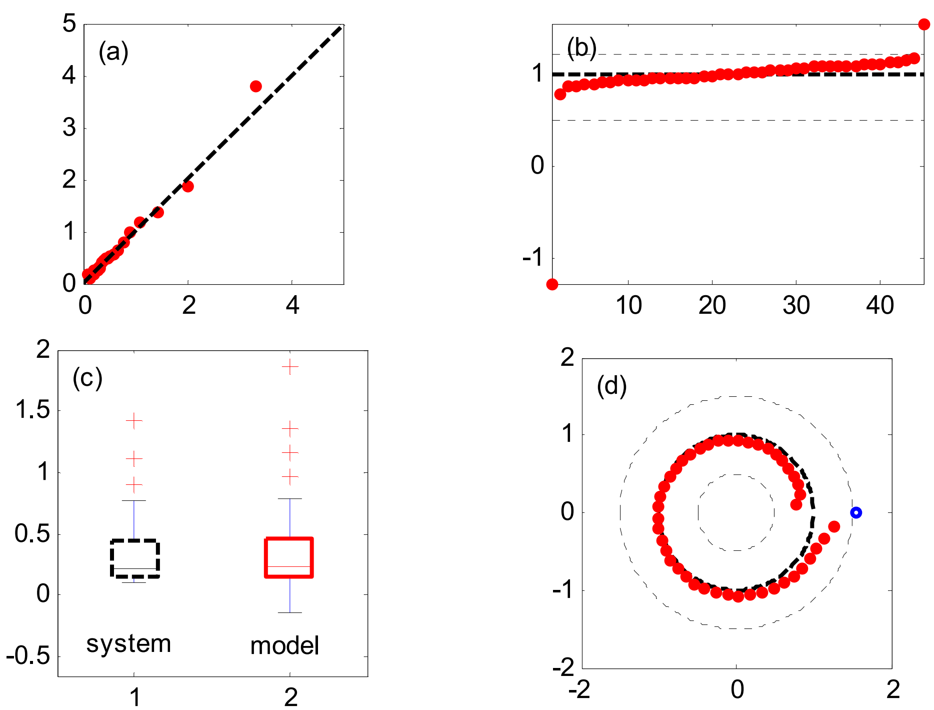

Figure 3: Graphical Displays for Model B (a) Scatter Plots (b) Ordinal Plots (c) Box PLots (d) Circles Plots

REFERENCES

Balci, Osman. 1989. How to assess the acceptability and credibility of simulation results. In *Proceedigns of the 1989 Winter Simulation Conf.,* ed. E. A. MacNair, K. J. Musselman, and P. Heidelberger, 62-71. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Kleijnen, J. P. C, and D. Deflandre. 2005 Validation of regression metamodels in simulation: bootstrap approach, E*uropean Journal of Operational Research* (accepted). [online]. Available via <http://center.kub.nl/staff/kleijnen> [accessed March, 2005].

Sargent, Robert G. 1996. Some subjective validation methods using graphical displays of data. In *Proceedings of the 1996 Winter Simulation Conf.,* ed. J. M. Charness, D. J. Morrice, D. T. Brunner, and J. J. Swain, 345-351. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Sargent, Robert G. 2004. Validation and verification of simulation models. In *Proceedings of the 2004 Winter Simulation Conf.,* ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 13-24. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Shin, Miyoung, R. G. Sargent, and A. L. Goel. 2002. Gaussian radial basis functions for simulation metamodeling. In *Proceedings of the 2002 Winter Simulation Conf.,* ed. E. Yucesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, 383-487. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers

Walpole, R. and R. Myers. 1993. *Probability and Statistics for Engineers and Scientists* (5^{th} ed.). Macmillan Publishing Company.

AUTHOR BIOGRAPHIES

HUSAM HAMAD is an assistant professor in the Electronic Engineering Department at Yarmouk University in Jordan. He received his B.S. in Electrical Engineering from Oklahoma State University in 1984, M.S. in Device Electronics from LSU in 1985, and PhD in Electronic Systems Engineering from the University of Essex, England, in 1995. He was a member of Phi Kappa Phi Honor Society during his study in the U.S. His research interests include analog integrated circuits analysis and design, electronic design automation, CAD, signal processing, and metamodel fitting and validation techniques. His e-mail address is <husam@yu.edu.jo>.

SAMI AL-HAMDAN is an assistant professor in the Computer Engineering Department at Yarmouk University in Jordan He received his B.S. in Electronic and Communication Engineering from Yarmouk University in 1984, and a PhD in Computer Engineering from Liverpool John Moores University, England, in 1996. His research interests include, electronic design automation, fringe analysis, signal processing, parallel processing, and metamodel fitting and validation techniques. His e-mail address is <shamdan@yu.edu.jo>.