

# Variance of System-Reliability Estimates With Arbitrarily Repeated Components

Tongdan Jin and David W. Coit, *Member IEEE*

**Abstract**—A model is developed to determine the variance of system reliability estimates and to estimate confidence intervals for series-parallel systems with arbitrarily repeated components. For these systems, different copies of the same component-type are used several or many times within the system, but only a single reliability estimate is available for each distinct component-type. The single estimate is used everywhere the component appears in the system design, and component estimation-error is then magnified at the system-level. The “system-reliability estimate” variance and confidence intervals are derived when the number of component failures follow the binomial distribution with an unknown, yet estimable, probability of failure. The “system-reliability estimate” variance and confidence intervals are obtained by expressing system reliability as a linear sum of products of higher order moments for component unreliability. The generating function is used to determine the moments of the component-unreliability estimates. This model is preferable for many system reliability estimation problems because it does not require independent component and subsystem reliability estimates; it is demonstrated with an example.

**Index Terms**—Binomial data, confidence interval, series-parallel system, system reliability.

## ACRONYMS<sup>1</sup>

CL	$s$ -confidence limit (bound)
gilb	greatest integer lower bound
liub	least integer upper bound
ML	maximum likelihood
MLM	ML model
r.v.	random variable
Var	variance

## NOTATION

$r_h$	reliability of component $h$
$q_h$	$1 - r_h$ : unreliability of component $h$
$\hat{r}_h$	estimate of $r_h$
$\hat{q}_h$	estimate of $q_h$
$s$	number of subsystems
$C_i$	set of components used in subsystem $i$ , $C_i \subset 1, 2, \dots, c$
$c$	number of component-types in the system

$k_{i,h}$	number of component $h$ used in subsystem $i$ $k_{i,h} = 0$ for $h \notin C_i$
$F_i$	$\prod_{h \in C_i} q_h^{k_{i,h}}$ : unreliability of subsystem $i$
$\hat{F}_i$	$\prod_{h \in C_i} \hat{q}_h^{k_{i,h}}$ : unreliability estimate of subsystem $i$
$R$	system reliability (unknown)
$\tilde{R}$	approximation of $R$
$\hat{R}$	estimate of $R$
$X$	$X \sim \text{binomial}(q, n)$ : number of component/system failures
$Y$	$Y \sim \text{binomial}(r, n)$ : number of component/system survivals
$n$	sample size
$\hat{n}$	pseudo sample size
$n_L$	$\lfloor \hat{n} \rfloor$ : gilb
$n_H$	$\lceil \hat{n} \rceil$ : liub

## I. INTRODUCTION

THE ESTIMATE OF  $R$  ( $\hat{R}$ ) are generally based on subsystem and component reliability information. For many system-design problems, there are limited failure and survival data at the component level; thus one must consider the uncertainty (measured here by variance—square of standard deviation) of the  $\hat{R}$ . It is also common to use the same component-type in different subsystems and in different system-locations, due to the need for similar performance requirements. For example, the same type of pump, operating with similar stress levels, can be used in many different locations within a nuclear power plant. For systems with repeated components, the variance of  $\hat{R}$  is relatively larger.

This paper presents a method to estimate variance of  $\hat{R}$  for series-parallel systems with arbitrarily repeated components. No parametric assumptions are required for component time-to-failure. The component reliability and variance are estimated based on the binomial distribution.

Early research focused on determining the CL with  $s$ -independent component-reliability estimates. Reference [1] initially determined exact CL estimation for products of 2  $s$ -independent binomial parameters. Reference [2] extended the model in [1] to calculate the upper CL on series and parallel systems with any number of subsystems. These CL are generally conservative because the number of failures is discrete. Reference [2] also presents a method for computing lower CL on series systems with  $s$ -independent components, based on determining equivalent or pseudo binomial failure data for the system. Then, the lower CL is based on the equivalent data. These CL are generally conservative.

Manuscript received August 20, 1999; revised May 31, 2000. The work of D. W. Coit was supported by the US-NSF CAREER Grant DMII-9874716.

The authors are with the Department of Industrial Engineering, Rutgers University, Piscataway, NJ 08854 USA (e-mail: JTD@rci.rutgers.edu; Coit@rci.rutgers.edu).

Publisher Item Identifier S 0018-9529(01)11345-X.

<sup>1</sup>The singular and plural of an acronym are always spelled the same.

Reference [3] derived the asymptotic  $\chi^2$  distribution of the logarithm of the likelihood ratio statistic, called the "LR model." Later [4] and [5] extended the LR model to complex systems. Reference [6] suggested another method, ML, for estimating the CL on  $\hat{R}$  based on the asymptotic  $s$ -normality of the ML estimate. Reference [7] suggested a modified MLM so that the upper CL never exceeds 1. These models generally do not allow component repetition or arbitrarily repeated components in the system.

Reference [8] proposed an approximate CL for complex-system reliability based on a log-normal approximation of system-reliability estimate. This approach applies to any system design that can be decomposed into series and/or parallel connections with  $s$ -independent components. Reference [9] extended this so that repeated components are allowed within a subsystem as long as subsystem reliability estimates remain  $s$ -independent.

Except for a few models, such as [9] and MLM, the available models require that the component and subsystem reliability estimates be  $s$ -independent. To apply one of the available models, it was often necessary to randomly-segment the available data to create  $s$ -independent subdata sets. This is a valid, but undesirable, approach because estimates from the segmented data-set can have high variance. More importantly, there often are small data sample-sizes initially, thus segmenting already limited data can be an unrealistic approach. A new model is proposed for estimating the  $\hat{R}$  variance and CL. It offers greater flexibility by allowing arbitrarily repeated components anywhere within a series-parallel system.

### Assumptions

- 1) Redundancy is active (failure behavior of redundant components is the same as active components).
- 2) There is no component (or system) repair or preventive maintenance.
- 3) Failures are recorded without error.
- 4) The number of component failures has a binomial distribution.
- 5) System failure is attributed to component failure.

## II. RELIABILITY MODEL FOR SYSTEMS WITH REPEATED COMPONENTS

Consider the 2 series-parallel systems in Figs. 1 and 2, and the empirical reliability data in Table I. The two systems have the same general configuration. For each component-type  $h$  used, there is a single  $\hat{r}_h$  that is used everywhere the component appears. For these 2 systems, the numerical  $\hat{r}_h$  are the same in each position. However, system A has  $s$ -independent component (and subsystem) reliability estimates, whereas, system B does not have  $s$ -independent estimates because of the repeated components.

For each sample system, the variance of  $\hat{R}$  is different, even though the  $\hat{r}_h$ , and the variance of those estimates, are the same. When repeated  $\hat{r}_h$  are used, the  $\hat{R}$  is biased; and the variance is larger when  $\hat{r}_h$  error is magnified at the system-level because of repeated components.

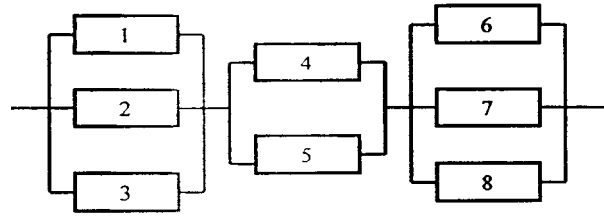


Fig. 1. System A.

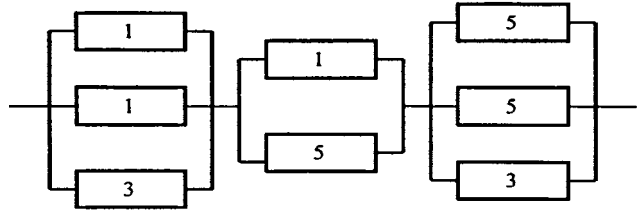


Fig. 2. System B (arbitrarily repeated components).

TABLE I  
COMPONENT-FAILURE DATA

$h$	$X$	$n$	$\hat{r}_h$	$\text{Var}[\hat{r}_h]$
1	3	25	0.880	0.004224
2	3	25	0.880	0.004224
3	4	32	0.875	0.003418
4	3	25	0.880	0.004224
5	1	10	0.900	0.009000
6	1	10	0.900	0.009000
7	1	10	0.900	0.009000
8	4	32	0.875	0.003418

Few previous articles have considered the variance and CL of reliability estimates for system B; yet this is a common system-design model in practice. In the following subsections, methods and mathematical models are established where arbitrarily repeated components exist among multiple subsystems.

### A. Model-Development Procedure

Fig. 3 shows the procedure for variance and CL estimation. An  $R$  expression is determined based on  $F_i$ , and  $q_h$  ( $h \in C_i$ ). To compute  $\text{Var}[\hat{R}]$ , the mean and the second moment of  $\hat{R}$  are calculated based on higher order moments of  $\hat{q}_h$ .

The system in Fig. 2 is a series-parallel with arbitrarily repeated components. For example, component #1 has 2 repetitions in subsystem #1 and 1 repetition in subsystem #2. Reliability estimates for subsystems #1 and #2 are  $s$ -dependent because the estimate for component #1 is from the same data set. Similarly, reliability estimates for subsystems #1 and #3 are  $s$ -dependent because of component #5.

For the example in Fig. 2

$$C_1 = \{1, 3\}, C_2 = \{1, 5\}, C_3 = \{3, 5\}.$$

Series-parallel system reliability can be written as

$$R = \prod_{i=1}^s \left( 1 - \prod_{h \in C_i} q_h^{k_{i,h}} \right) = \prod_{i=1}^s (1 - F_i). \quad (1)$$

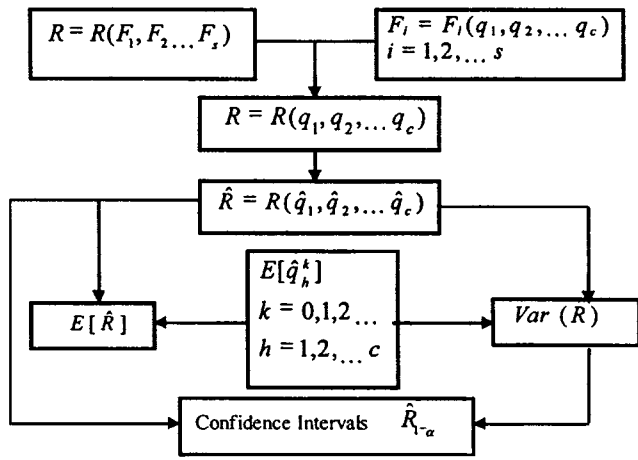


Fig. 3. Model-development procedure.

### B. System Reliability Approximation

System reliability can also be expressed as,

$$R = 1 - \sum_{i=1}^s F_i + \sum_{i<j} F_i \cdot F_j - \sum_{i<j<k} F_i \cdot F_j \cdot F_k + \dots + \prod_{i=1}^s (-F_i). \quad (2)$$

Many systems have relatively high reliability in practice. When  $R \geq 0.75$ , an accurate approximation for system reliability can be obtained using the first three terms of (2). Simulation shows that the approximation (3) is very accurate [9]

$$\tilde{R} = 1 - \sum_{i=1}^s F_i + \sum_{i<j} F_i \cdot F_j; \quad (3)$$

$$= 1 - \sum_{i=1}^s \left( \prod_{h \in C_i} q_h^{k_i, h} \right) + \sum_{i<j} \left( \prod_{h \in C_{i,j}} q_h^{k_i, h + k_j, h} \right); \quad (4)$$

$C_{i,j} \equiv C_i \cup C_j.$

In practice, the  $q_h$  are not known and must be estimated. Thus use  $\hat{q}_h$  in (4) instead of the  $q_h$ :

$$\hat{R} = 1 - \sum_{i=1}^s \left( \prod_{h \in C_i} \hat{q}_h^{k_i, h} \right) + \sum_{i<j} \left( \prod_{h \in C_{i,j}} \hat{q}_h^{k_i, h + k_j, h} \right). \quad (5)$$

### C. Mean and Variance for $\hat{R}$

The mean and variance of  $\hat{R}$  from (5) are

$$E[\hat{R}] = 1 - \sum_{i=1}^s \left( \prod_{h \in C_i} E[\hat{q}_h^{k_i, h}] \right) + \sum_{i<j} \left( \prod_{h \in C_{i,j}} E[\hat{q}_h^{k_i, h + k_j, h}] \right); \quad (6)$$

$$\text{Var}[\hat{R}] = \text{Var} \left[ 1 - \sum_{i=1}^s \hat{F}_i + \sum_{i<j} \hat{F}_i \cdot \hat{F}_j \right];$$

$$= \sum_{i=1}^s \text{Var}[\hat{F}_i] + 2 \sum_{i<j} \text{Cov}[\hat{F}_i, \hat{F}_j]$$

$$+ \sum_{i<j} \sum_{m<n} \text{Cov}[\hat{F}_i \cdot \hat{F}_j, \hat{F}_m \cdot \hat{F}_n]$$

$$+ 2 \sum_{i=1}^s \sum_{m<n} \text{Cov}[\hat{F}_i, \hat{F}_m \cdot \hat{F}_n]$$

$$= \sum_{i=1}^s \sum_{j=1}^s \left( E[\hat{F}_i \cdot \hat{F}_j] - E[\hat{F}_i] \cdot E[\hat{F}_j] \right)$$

$$+ \sum_{i<j} \sum_{m<n} \left( E[\hat{F}_i \cdot \hat{F}_j \cdot \hat{F}_m \cdot \hat{F}_n] - E[\hat{F}_i \cdot \hat{F}_j] \cdot E[\hat{F}_m \cdot \hat{F}_n] \right)$$

$$+ 2 \sum_{i=1}^s \sum_{m<n} \left( E[\hat{F}_i \cdot \hat{F}_m \cdot \hat{F}_n] - E[\hat{F}_i] \cdot E[\hat{F}_m \cdot \hat{F}_n] \right) \quad (7)$$

$$= \sum_{i=1}^s \sum_{j=1}^s \left( \prod_{h \in C_{i,j}} E[\hat{q}_h^{k_i, h + k_j, h}] - \prod_{h \in C_{i,j}} E[\hat{q}_h^{k_i, h}] \cdot E[\hat{q}_h^{k_j, h}] \right)$$

$$+ \sum_{i<j} \sum_{m<n} \left( \prod_{h \in C_{i,j,m,n}} E[\hat{q}_h^{k_i, h + k_j, h + k_m, h + k_n, h}] - \prod_{h \in C_{i,j,m,n}} E[\hat{q}_h^{k_i, h + k_j, h}] \cdot E[\hat{q}_h^{k_m, h + k_n, h}] \right)$$

$$+ 2 \sum_{i=1}^s \sum_{m<n} \left( \prod_{h \in C_{i,m,n}} E[\hat{q}_h^{k_i, h + k_m, h + k_n, h}] - \prod_{h \in C_{i,m,n}} E[\hat{q}_h^{k_i, h}] \cdot E[\hat{q}_h^{k_m, h + k_n, h}] \right); \quad (8)$$

$$C_{i,m,n} \equiv C_i \cup C_m \cup C_n,$$

$$C_{i,j,m,n} \equiv C_i \cup C_j \cup C_m \cup C_n.$$

The variance of  $\hat{R}$  is expressed as a linear combination of products of higher order moments of the  $\hat{q}_h$ . An iteration formula is developed in Section II-D to compute the higher-order moments of  $\hat{q}_h$ . Using these moments,  $E[\hat{R}]$  and  $\text{Var}[\hat{R}]$  can be estimated directly from (6) and (8).

### D. Component Unreliability Moments

The probability generating function for a discrete r.v.  $X$  is

$$G(z) = \sum_{i=0}^{\infty} P_i \cdot z^i;$$

$$P_i \equiv \Pr\{X = x_i\}, \quad \text{for } i = 1, 2, \dots$$

$$\frac{\partial^m G(z)}{\partial z^m} = \sum_{i=m}^{\infty} i \cdot (i-1) \cdot \dots \cdot (i-m+1) \cdot P_i$$

$$= \sum_{i=1}^m a_{m,i} \cdot E[X^i];$$

$$a_{m,i} \equiv \text{coefficient}, \quad \mathbf{a}_m = (a_{m,1}, a_{m,2}, \dots, a_{m,m}). \quad (9)$$

TABLE II  
 $a_{m,i}$  COEFFICIENTS ( $1 \leq m \leq 6$ )

$\mathbf{a}_m$	$a_{m,1}$	$a_{m,2}$	$a_{m,3}$	$a_{m,4}$	$a_{m,5}$	$a_{m,6}$
$\mathbf{a}_1$	1					
$\mathbf{a}_2$	-1	1				
$\mathbf{a}_3$	2	-3	1			
$\mathbf{a}_4$	-6	11	-6	1		
$\mathbf{a}_5$	24	-50	35	-10	1	
$\mathbf{a}_6$	-120	274	-225	85	-15	1

The  $a_{m,i}$  can be obtained by the discrete convolution operator ( $\otimes$ ). The  $\otimes$  increases the dimension of  $\mathbf{a}_m$  by 1 when  $m$  is increased by 1.

$$\begin{aligned} \mathbf{a}_m &= \mathbf{a}_{m-1} \otimes \mathbf{b}_{m-1}; \\ \mathbf{a}_{m-1} &\equiv (a_{m-1,1}, a_{m-1,2}, a_{m-1,3}, \dots, a_{m-1,m-1}), \\ \mathbf{b}_{m-1} &\equiv (1 - m, 1), \\ \mathbf{a}_m &\in I^m, \quad \mathbf{b}_m \in I^2. \end{aligned} \quad (10)$$

The  $\mathbf{a}_m$  can be determined iteratively for any  $m$  by using (10) and starting with  $\mathbf{a}_1 = (a_{1,1}) = (1)$ , and  $\mathbf{b}_1 = (-1, 1)$ . Table II presents example values of  $a_{m,i}$  coefficients for  $m \leq 6$ .

The number of component failures is a binomial r.v.

$$\begin{aligned} G(z) &= \sum_{i=0}^n z^i \cdot \binom{n}{i} \cdot q^i \cdot (1-q)^{n-i} \\ &= (q \cdot z + r)^n; \end{aligned} \quad (11)$$

$$\left. \frac{\partial^m G(z)}{\partial z^m} \right|_{z=1} = \frac{n!}{(n-m)!} \cdot q^m. \quad (12)$$

By equating (9) and (12), and noting that  $a_{m,m} = 1$  for all  $m$ , a recursive equation is derived to compute higher order moments based on lower-order moments

$$\sum_{i=1}^m a_{m,i} \cdot E[X^i] = \frac{n!}{(n-m)!} \cdot q^m; \quad (13)$$

$$E[X^m] = \frac{n!}{(n-m)!} \cdot q^m - \sum_{i=1}^{m-1} a_{m,i} \cdot E[X^i]. \quad (14)$$

$X$  has a binomial distribution; thus

$$\begin{aligned} \hat{q} &= \frac{X}{n}. \\ E[\hat{q}^m] &= \frac{E[X^m]}{n^m} \\ &= \frac{n!}{(n-m)!} \cdot \frac{q^m}{n^m} - \sum_{i=1}^{m-1} a_{m,i} \cdot \frac{E[X^i]}{n^m}. \end{aligned} \quad (15)$$

In practice,  $q$  is not known with certainty, but  $\hat{q}$  is available. Approximations of all moments can be determined by using  $\hat{q}$  instead of  $q$ . The variance of  $\hat{R}$  is computed iteratively using (10), (14), (15) to compute the moments for  $\hat{q}$ , and (8) is used to compute  $\text{Var}[\hat{R}]$ .

As the component binomial sample sizes increase, the approximate variance estimates and CL become more accurate. Alternatively, with very low binomial sample sizes ( $n_i < 10$  for all  $i$ ), the approximate variance estimates and CL can be appreciably in error. This method is very robust and offers many benefits, compared to alternative methods, but it can not always be applied. Reference [8] uses simulation to demonstrate the effect of component-data sample-sizes on an analogous problem. A similar effect is anticipated here.

TABLE III  
 CL WITH ( $\alpha = 0.05$ )

parameter	System A	System B
$\hat{R}$	.9880	.9880
$E[\hat{R}]$	.9880	.9863
$\text{Var}[\hat{R}]$	.000273	.000395
$\hat{n}$	43.35	29.97
$\hat{x}$	42.48	28.65
$R_L$	.9178	.8805
$R_U$	.9994	.9991

### III. CONFIDENCE INTERVALS

$s$ -Confidence intervals for  $\hat{R}$  can be estimated using an approach similar to that in [7]: regard  $\hat{R}$  as a binomial-distribution parameter associated with unknown sample size,  $n$ . Because the  $\text{Var}[\hat{R}]$  can be estimated (8), a useful pseudo sample-size is

$$\hat{n} \equiv \frac{\hat{R} \cdot (1 - \hat{R})}{\text{Var}[\hat{R}]}. \quad (16)$$

$Y$  has an approximate binomial distribution with parameters  $\hat{R}$  and  $n_L$ . Approximate  $(1 - \alpha)$  CL ( $R_L, R_U$ ) for  $R$  are:

$$R_L = \left\{ R; \sum_{i=0}^{x_L} \binom{n_L}{i} \cdot R^i \cdot (1-R)^{n_L-i} = 1 - \frac{\alpha}{2} \right\}, \quad (17)$$

$$R_U = \left\{ R; \sum_{i=x_H}^{n_L} \binom{n_L}{i} \cdot R^i \cdot (1-R)^{n_L-i} = 1 - \frac{\alpha}{2} \right\}; \quad (18)$$

$$\hat{x} \equiv \hat{R} \cdot [\hat{n}], \quad x_L \equiv [\hat{x}], \quad x_H \equiv [\hat{x}].$$

### IV. NUMERICAL EXAMPLES

Systems A and B (from Figs. 1 and 2, and Table I) are used to show how to compute the  $\text{Var}[\hat{R}]$  and its CL. System B has repeated components both within and among subsystems. Therefore, component and subsystem reliability estimates are  $s$ -dependent. There are no repeated components in system A, thus all components and subsystems are  $s$ -independent.

Table III presents the results. As anticipated, the variance for system B is larger than for system A. This is reasonable because all reliability estimates in system A are  $s$ -independent. The point estimates for  $R$  in both systems are the same, but their CL are quite different.

These results are very important because many system-designs use different copies of the same component-type within a system design, yet use the same reliability estimates each time. Previous methods either required  $s$ -independent component-reliability estimates, required assumed time-to-failure distributions, or required more rigid system-structures (*viz.* series, parallel).

### REFERENCES

- [1] R. Buehler, "Confidence intervals for product of two binomial parameters," *J. American Statistical Assoc.*, pp. 482-492, Dec. 1957.
- [2] D. K. Lloyd and M. Lipow, *Reliability: Management, Methods and Mathematics*. Englewood Cliffs, NJ: Prentice-Hall, 1964.
- [3] A. Madansky, "Approximate confidence limits for reliability of series and parallel systems," *Technometrics*, vol. 7, no. 4, pp. 495-503, 1965.

- [4] J. Myhre and M. Rennie, "Confidence bounds for reliability of coherent systems based on the binomial distributed component data," in *Reliability and Quality Control*, A. P. Basu, Ed. North Holland: Elsevier, 1984.
- [5] J. Myhre and C. Saunders, "On confidence limits for the reliability of systems," *Annals of Mathematical Statistics*, vol. 39, no. 5, pp. 1463–1472, 1968.
- [6] J. Rosenblatt, "Confidence limits for reliability of complex systems," in *Statistical Theory of Reliability*, M. Zelen, Ed: Univ. Wisconsin Press, 1963.
- [7] R. G. Easterling, "Approximate confidence limits for system reliability," *J. American Statistical Assoc.*, vol. 67, no. 337, pp. 220–222, 1972.
- [8] D. W. Coit, "System reliability confidence intervals for complex systems with estimated component reliability," *IEEE Trans. Reliability*, vol. 46, pp. 487–493, 1997.
- [9] D. W. Coit and T. Jin, "System reliability models with uncertain component reliability estimates," Rutgers Univ., Final Report, 1998.

**Tongdan Jin** received the M.S. degree (1994) in mechanical engineering from Beijing Institute of Technology, specializing in computer-aided design, and the B.S. degree (1991) in electrical engineering from Northwest Institute of Light Industry, PR China. He received the Ph.D. degree (2001) from the Industrial Engineering Department, Rutgers University, Piscataway, NJ.

His research interests are system and component reliability analysis and modeling.

**David W. Coit** received the B.S. degree in mechanical engineering from Cornell University, the M.B.A. from Rensselaer Polytechnic Institute, and the M.S. and Ph.D. degrees in industrial engineering from the University of Pittsburgh.

He is an Assistant Professor in the Industrial Engineering Department of Rutgers University, Piscataway, NJ. From 1980 to 1992, he was with IIT Research Institute (IITRI), Rome, NY; from 1980 to 1988, he was a reliability analyst and project manager; from 1988 to 1992, he was the Manager of Engineering at its Assurance Technology Center. His research involves reliability prediction and optimization.

Dr. Coit is a member of IIE, IEEE, and INFORMS.