Planning, Tracking, and Projecting Reliability Growth A Bayesian Approach

Richard Strunz, M.S., MBAe, Astrium GmbH

Jeffrey W. Herrmann, PhD, University of Maryland

Key Words: multilevel Bayesian data aggregation, reliability growth

SUMMARY & CONCLUSIONS

Liquid rocket engine reliability growth modeling is a blend of art and science because of data scarcity and heterogeneity, which result from the limited number of engine development programs as well as testing profiles that are much different from the actual mission profile. In particular, hot fire tests are shorter than full mission duration due to test facility limitations and some of them are performed at extreme load points to demonstrate robustness and design margin.

The well-known empirical Duane and analytical Crow/AMSAA models are therefore no longer best practice because the reliability growth rate is calculated using a MTBF estimate that is simply the total accumulated test time divided by the total number of failures. Therefore, we propose a new, fully Bayesian estimation based methodology that estimates the system reliability while taking into account the test profile characteristics and aggregating component, subsystem, and system level hot fire test data.

The methodology is applied to planning, tracking, and projecting reliability growth and illustrated using an example. In the example, a system reliability target must be demonstrated in a TAAF program. The system reliability target defines the scope of the hot fire test plan for the reliability growth planning using pseudo numbers for the planned hot fire tests. At each occurrence of a failure, the methodology is used in the context of reliability growth tracking, i.e. the attained system level reliability is estimated. The test plan is updated to reflect the need for additional tests to meet the system reliability target. Reliability growth projection is easily performed using either specific projection models or the prior distribution that features a knowledge factor to model the specified level of fix effectiveness.

1 INTRODUCTION

Reliability growth is typically attained through a formal TAAF program that discovers and corrects design deficits. Reliability growth models are used for test planning, tracking reliability throughout the program, and projecting the reliability when the tests are completed. The two most widely used reliability growth models are the empirical Duane and the analytical Crow/AMSAA, which both use the MTBF to estimate the reliability growth rate. The MTBF is calculated from the total accumulated test time divided by the total number of failures without considering the operational loads, durations, and sequences of the applied stresses, which highly affect the failure rate and as a consequence the MTBF metric [1]. Therefore, ignoring the applied stresses makes the Duane and Crow/AMSAA models questionable for cases in which the testing profiles differ, in terms of applied stresses, significantly from the stated mission profile [2, 3].

Modern liquid rocket engine hot fire testing profiles belong to such cases because the testing profile is a potpourri of tests that are shorter than full mission duration and tests performed at extreme load points to demonstrate robustness and design margins. Therefore, neither the Duane nor the Crow/AMSAA data analysis may be any longer best practice as the following brief discussion highlights.

Historically, liquid rocket engine hot fire testing profiles were used to comply with a formal reliability demonstration as it was the case for the F-1 and J-2 engines. These hot fire testing profiles followed adequately well the operational loads, and, as a consequence, the Golovin and empirical Duane models were successfully applied [4].

However, formal reliability demonstration hot fire testing profiles are lengthy and cost prohibitive, which led to the DVS approach that was applied to the SSME certification. The Crow/AMSAA model, one of the two reliability growth models used, initially estimated an increase of the MTBF (indicating reliability growth), but the system reliability declined towards the end of the testing profile although overall testing experience would have suggested an increase in the system reliability [5].

Most recently, an objective based variable test/time philosophy was used to qualify the RS-68 liquid rocket engine while lowering the development cost and reducing the development schedule. To achieve these objectives, the hot fire testing profile included extreme load points to demonstrate robustness and design margin [6]. Based on the SSME experience, the RS-68 engine testing profile should have been even more difficult to analyze with the Duane and Crow/AMSAA models and to estimate a system reliability that is based on the MTBF metric.

As response to the modern liquid rocket engine hot fire testing profiles, we propose a new, fully Bayesian estimation based methodology that estimates the system reliability without the MTBF metrics; instead, it takes into account all component, subsystem, and system level hot fire test data. The Bayesian estimation provides naturally the framework that is needed to apply the methodology in the three areas of reliability growth: planning, tracking, and projection because pseudo, actual, and the combination of both pseudo and actual hot firings test data can be used to estimate the system level reliability.

1.1 Acronyms

AMSAA	Army Materiel Systems Analysis Activity
CCCG	Common-cause Component Group
DVS	Design Verification Specification
MH	Metropolis-Hastings
MTBF	mean-time-between-failure
SSME	Space Shuttle Main Engine
TAAF	test, analyze, and fix

1.2 Notation

$AF_{i,j}$	acceleration factor at node level *i* of test group *j*
$c_{rel,i}$	reliable number of cycles at node level *i*
$EQL_{rel,i}$	reliable equivalent life at node level *i*
HW_i	number of hardware
MPC	mission profile cycles
MPD	mission profile duration
n_i	equivalent trials at node level *i*
$t_{rel,i}$	reliable hot fire time at node level *i*
$TD_{i,j}$	test duration at node level *i* of test group *j*
$TFT_{i,j}$	test failure time at node level *i* of test group *j*
TPC_i	testing profile cycles at node level *i*
$TPC_{i,j}$	testing profile cycles at node level *i* of test group *j*
TPD_i	testing profile duration at node level *i*
$TPD_{i,j}$	testing profile duration at node level *i* of test group *j*
$w_{i,j}$	weighting factor at node level *i* of test group *j*
x_i	equivalent successes at node level *i*
$\alpha_{1,i}$	weighting factor for cyclic loads at node level *i*
$\alpha_{2,i}$	weighting factor for temporal loads at node level *i*
α_i	shape parameter at functional node level *i*
β_i	shape parameter at functional node level *i*
ϕ_i	knowledge factor at functional node level *i*
$\pi_{0,i}$	prior distribution
$\underline{\pi}$	posterior parameter vector
π_i	reliability of functional node *i*

2 METHODOLOGY

The methodology is based on the Bayesian aggregation of multilevel binomial test data [7] but is extended with the notion of equivalent mission to account for the operational loads, durations, and sequences of the applied stresses that are present in the specific testing profiles but are unlike those in the mission profile [8].

The Bayesian aggregation of multilevel binomial test data uses a functional network that is based on the principles of the reliability block diagram technique [9]. The functional network serves two purposes: (1) It defines the fundamental test strategy that defines also the hot fire test configurations at the component, subsystem, and engine system levels and (2) it is used to derive the governing likelihood function that combines simultaneously all available multilevel hot fire test data. It should be noted that the functional component level nodes correspond to individual physical components or to a CCCG of the actual physical system architecture. Figure 1 depicts an example of such a functional network.

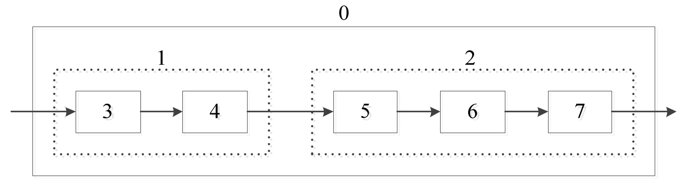

Figure 1 – Functional Network

The methodology begins with the set of prior distributions about the reliability of each functional node. Each prior is a modified Beta distribution with three parameters: α_i and β_i, which can be derived from previous engine reliability data as given in [10], and the knowledge factor (or relevance factor) ϕ_i, which measures the level of transformation of similar designs into new product designs and is derived from methods defined in [11, 12]. It can be determined qualitatively or quantitatively with methods described in [13]. Thus, the prior for node *i* is the following distribution:

$$\pi_{0,i} = f\left(\pi_i \mid \alpha_i, \beta_i, \phi_i\right) = \frac{\pi_i^{\alpha_i \phi_i - 1}(1-\pi_i)^{(\beta_i - 1)\phi_i}}{Beta\left(\alpha_i \phi_i, (\beta_i - 1)\phi_i + 1\right)} \quad (1)$$

In addition, the methodology requires for each functional node the number of equivalent trials, n_i, and the number of equivalent successes, x_i. The notion of equivalent mission is introduced because it captures the two fundamental failure mechanisms (characterized as stress-increased and strength-reduced) that are present in liquid rocket engine piece parts and subassemblies. The number of equivalent trials, n_i, is calculated as follows:

$$n_i = \alpha_{1,i} \frac{TPC_i}{MPC} + \alpha_{2,i} \frac{TPD_i}{MPD} \quad (2)$$

The first term relates the stress-increased (cyclic) and the second term the strength-reduced (time-dependent) failure mechanism, respectively. Both terms are weighted and relate the specific testing profiles to the mission profile.

These quantities are derived from the characteristics of the testing profiles as follows. For the number of equivalent trials, the testing profile duration TPD_i depends upon the test duration, an acceleration factor, which is introduced to model the extreme load points, and a weighting factor accounts for the hot fire tests that are shorter than full mission duration. Note that these different testing profiles at functional node *i* are accounted for by defining specific test groups *j*. The acceleration factor, $AF_{i,j}$, is based on the acceleration testing theory [14] and is not further discussed in the paper. The weighting factor, $w_{i,j}$, is based on a likelihood function that models the union of two mutually exclusive events: (1) a failure that takes place during the start-up and steady state operation (ordinary failure) and (2) a failure that takes place during the shutdown operation [15].

$$TPD_{i,j} = AF_{i,j}\ w_{i,j}\ TD_{i,j} \tag{3}$$

Because the individual hot fire test durations are usually different within each functional node i, which is reflected through subscript j, we use the following to calculate n_i:

$$n_i = \sum_{j=1}^{k}\left(\alpha_{1,i}\frac{TPC_{i,j}}{MPC} + \alpha_{2,i}\frac{TPC_{i,j}TPD_{i,j}}{MPD}\right) \tag{4}$$

For the number of equivalent successes at node i, the testing profile duration $TPD_{i,j}$ depends upon the actual failure time $TFT_{i,j}$, the acceleration factor, and the weighting factor:

$$TPD_{i,j} = AF_{i,j}\ w_{i,j}\ TFT_{i,j} \tag{5}$$

Then, the number of equivalent successes at node i is derived using an equation similar to equation (4):

$$x_i = n_i - \sum_{j=1}^{k}\left(\alpha_{1,i}\frac{TPC_{i,j}}{MPC} + \alpha_{2,i}\frac{TPC_{i,j}TPD_{i,j}}{MPD}\right) \tag{6}$$

After these quantities are derived, the Bayesian estimation uses Bayes' Theorem to define an unscaled posterior distribution for the parameters that must be estimated. The unscaled posterior distribution is defined through a likelihood function which models the data and a set of prior distributions for the parameters of the model (the likelihood function) that is given as

$$\pi\left(\underline{\pi} \mid Data\right) \propto \prod_i \pi_i^{x_i}\left(1-\pi_i\right)^{n_i - x_i} \prod_i \pi_{0,i} \tag{7}$$

The parameter vector, $\underline{\pi}$, of this unscaled posterior distribution is estimated with a one-variable-at-a-time MH algorithm. Important metrics of this solution strategy are the acceptance rate of the acceptance probability as well as the autocorrelation and convergence of the Markov chain of the proposed candidates. The candidates are drawn on a logit scale for which the proper acceptance rate is around 0.35. In order to obtain that rate, the burn-in period of the Markov chain is used to tune the standard deviation of the candidate density function. The autocorrelation function is used to obtain the lag at which the Markov chain is thinned. Finally, the convergence of the accepted Markov chain was visually inspected by means of trace plots.

The combined likelihood function of equation (7) is found as follows: The fundamental test strategy defines the test configurations that are expressed in terms of nodes. Using the example depicted in Figure 1, the system level is node 0, the two subsystem nodes would be 1 and 2, and the functional component level nodes are 3 to 7. The subsystem node 1 and 2 reliabilities are expressed as $\pi_1 = \pi_3\ \pi_4$ and $\pi_2 = \pi_5\pi_6\pi_7$. The system level node 0 reliability is given as $\pi_0 = \pi_1\pi_2$ or equivalently as $\pi_0 = \pi_3\pi_4\pi_5\pi_6\pi_7$. Finally, these functional component, subsystem, and system level reliabilities are inserted in equation (7) to combine simultaneously all level test data.

The probabilities of the mutually exclusive events that define the weighting factor, $w_{i,j}$, for the different testing profiles are also found by applying Bayes' Theorem to the likelihood function and a prior distribution for the model parameters. The likelihood function that describes the mutually exclusive events is based on a quasi multinomial distribution. The Dirichlet distribution is the conjugate prior for the multinomial distribution; therefore, it was also selected here.

Figure 2 depicts empirical evidences for the weighting factors for different liquid rocket engines using the data given in [16]. The figure includes also the weighting factors that are used in the illustrative example described in Section 3.

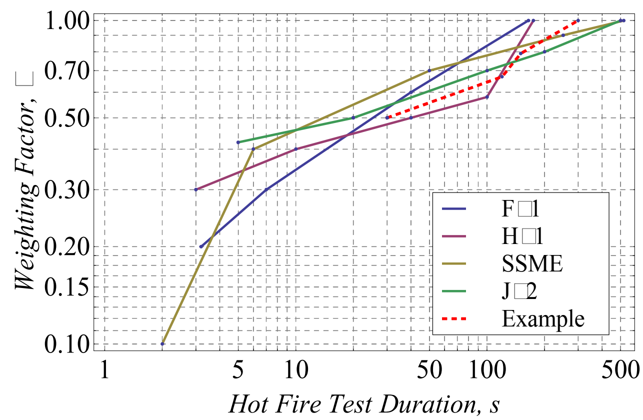

Figure 2 – Weighting Factor versus Hot Fire Test Duration

The equivalent trials, n_i, can be related to an equivalent life for the hardware components in order to estimate the number of hardware sets required to complete the specific testing profile. The reliable equivalent life is given as [8]

$$EQL_{rel,i} = \alpha_{1,i}\frac{c_{rel,i}}{MPC} + \alpha_{2,i}\frac{t_{rel,i}}{MPD} \tag{8}$$

The definitions of the reliable cycle, $c_{rel,i}$, and the reliable time, $t_{rel,i}$, may be either based on physics-of-failure models if available or on expert elicitation [8].

It should be noted that the structure of the reliable equivalent life is the same as for the equivalent mission. Therefore, the number of required hardware sets can be estimated with

$$HW_i = \frac{n_i}{EQL_{rel,i}} \tag{9}$$

Equation (9) can be applied using the overall number of equivalent trials or the equivalent trials that are associated with the relevant functional node i level.

The Bayesian estimation methodology is applied next to an illustrative example that describes the application in the context of reliability growth: planning and tracking. The area reliability growth projection is not explicitly demonstrated but once the system reliability is estimated various projection models can be applied [17].

3 ILLUSTRATIVE EXAMPLE

As an illustrative example, we consider a hypothetical liquid rocket engine TAAF program that includes a contractual reliability growth objective (system reliability target) for a cryogenic Gas Generator main stage engine.

The physical system architecture is similar to the RS-68 or Vulcain 2 liquid rocket engine. Therefore, the physical architecture can be described with nine functional component nodes in series [8].

The thrust class of the new engine is a significant increase compared to previous designs but our a priori knowledge is that the design authority has mastered a staged combustion engine at lower thrust scale. Based on this, we decided to use a knowledge factor ϕ_i of 0.80 for the functional component level node priors with distribution parameters $\alpha_i = 38$ and $\beta_i = 0.7$.

Furthermore, the stated engine mission profile consists of a 100 seconds acceptance test, a 10 seconds engine ground start hold-down with launch commit criteria abort, and a 300 seconds flight mission. The contractor and agency selected a specific testing profile (hot fire test plan) that includes component level, subsystem, and system level tests. Table 1 lists that specific testing profile in terms of number of tests, hot fire test duration, and acceleration factor to indicate the severity of the hot fire test conditions.

Table 1 – Testing Profile

Node	No. of Tests, -	Hot Fire Duration, s	Acceleration Factor, -
Gas Generator	60	50	1
Powerpack	10	100	1
Engine, Group 1	70	30	1
Engine, Group 2	50	120	1
Engine, Group 3a	35	150	1
Engine, Group 3b	35	150	5
Engine, Group 4a	20	300	1
Engine, Group 4b	20	300	5
Total / Accumulated	230	30,600	

Based on this data, setting the weights $\alpha_{1,i}$ and $\alpha_{2,i}$ to 0.5, and the application of equation (7), the average system level reliability estimate is 0.956.

We now consider the impact of failures. We will consider a scenario in which three failures occur (see Table 2). The failures are fully defined by means of the hot fire test order number, the failure time, and the affected physical component.

Table 2 – Assumed Failure Metrics

Node	No. of Tests, -	Failure Time, s	Component
Engine, Group 1	45	25	Turbopump, ox
Engine, Group 2	110	100	Gas Generator
Engine, Group 3a	150	30	Turbopump, fu

In this scenario, the TAAF program has started, the first couple of hot fire tests are successful, and then the failures occur. At each failure event, the following updating procedure is performed:

- the likelihood function for the weighting factor, $w_{i,j}$, is updated with the failure event and the Bayesian estimation calculates new weights that are used in equation (3) and (5),
- equation (4) and (6) are equated using the new weights and the actual failure event time,
- the a priori knowledge is considered as non existing for the failed component that modifies the prior distribution, and
- the recalculation of the functional component level reliabilities using equation (7) in order to update the system level reliability.

Table 3 lists the resulting system level reliability estimates at each failure occurrence and demonstrates the application of the methodology in the context of reliability growth tracking.

Table 3 – Reliability Growth Tracking

Tracking Steps	Test Number	Reliability Level
Failure 1	45	0.831
Failure 2	110	0.861
Failure 3	150	0.879

The next step in our TAAF program scenario is the definition of the remaining hot fire test effort given the failure occurrence in order to attain the contracted system reliability target (reliability growth planning). Either of two assumptions can be made: (1) no additional failures will occur during the remaining hot fire tests or (2) additional failures will occur and the number of the additional failures is estimated using reliability growth projection models. In this paper, we consider only the first case and update the reliability growth planning hot fire test scope at each time when an assumed failure occurred. Table 4 lists the consequences in terms of additional hot fire tests and as delta from the initial hot fire test plan to attain the contracted system reliability target, i.e. 0.956. Figure 3 depicts the described scenario graphically.

Table 4 – Test Scope Consequences

Events	Additional Hot Fire Tests, -	Delta from Initial Test Plan, -
Failure 1	20	20
Failure 2	30	50
Failure 3	25	75

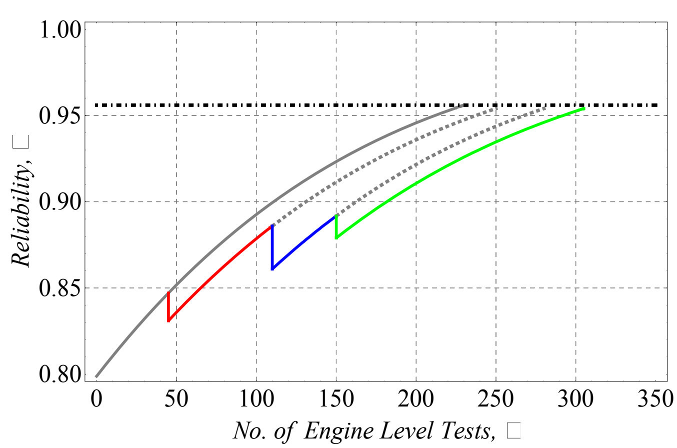

Figure 3 – Reliability Growth Planning and Tracking

The practical importance to both contractors and the space agency should be noted because the methodology not only estimates the attained or planned system reliability to assure the mission success but also provides the hot fire test scope during the requirements definition and after a failure occurrence. Thus, the presented Bayesian methodology in the context of reliability growth is also a valuable management tool for program managers.

4 CONCLUSION

In this paper, we presented a new, fully Bayesian estimation based methodology that provides a true alternative to the empirical Duane and analytical Crow/AMSAA models. The key features that distinguish the proposed methodology from the classical models are the aggregation of multilevel test data, the neutralization of the differences of the specific testing profile to the mission profile, the inclusion of a priori knowledge, and the capability to apply it to all three main areas of reliability growth: planning, tracking, and projection.

The illustrative example demonstrated the practical use of the proposed methodology by quantifying the impact of failures on the estimated system reliability in the context of reliability growth planning, tracking, and projection. Aside the reliability engineering aspects, the illustrative example highlighted the importance of the methodology as a risk management tool by providing quantitative figures for the hot fire test scope definition that drives both the development cost and development schedule.

REFERENCES

1. J. W. McPherson, "*Reliability Physics and Engineering*," Springer Science+Business Media, 2010.
2. M. Krasich, "Reliability Growth Test Design – Connecting Math to Physics," *Proc. Ann. Reliability & Maintainability Symp.*, (Jan.) 2011, pp 1–7.
3. L. H. Crow, "A Methodology for Managing Reliability Growth during Operational Mission Profile Testing," *Proc. Ann. Reliability & Maintainability Symp.*, (Jan.) 2008, pp 48–53.
4. E. O. Codier, "Reliability Growth in real Life," *Proc. 1968 Annual. Symp. Reliability*, pp 458-469
5. National Aeronautics and Space Administration (NASA), "*Report of the SSME Assessment Team*," 1993
6. B. K. Wood, "Propulsion for the 21st Century – RS-68," AIAA 2002-4324, *38th Joint Liquid Propulsion Conference*, Indianapolis, IN, 8 – 10 July 2002.
7. V. E. Johnson, T. L. Graves, M. S. Hamada, C. S. Reese, "A hierarchical Model for estimating the Reliability of complex Systems," in *Bayesian Statistics 7*, J. M. Bernardo, M. J. Bayarri, J. Berger, et al., Eds., pp. 199-213, Oxford University Press, London, UK, 2003
8. R. Strunz, J. W. Herrmann, "Reliability as an Independent Variable Applied to Liquid Rocket Engine Hot Fire Test Plans," *AIAA Journal of Propulsion and Power*, vol. 27 no. 5, 2011, pp 1032–1044.
9. S. C. Saunders, "*Reliability, Life Testing, and Prediction of Service Lives*," Springer Science+Business Media, 2007.
10. R. H. McFadden, Y. Shen, "An Analysis of the historical Reliability of US liquid-fuel Propulsion Systems," AIAA 90-2713, *26th AIAA/ASME/SAE/ASEE Joint Propulsion*, Orlando, FL, 16 – 18 July, 1990.
11. A. Krolo, "*Planung von Zuverlässigkeitstests mit weitreichender Berücksichtigung von Vorkenntnissen*," Dissertation, Universität Stuttgart, 2004.
12. A. V. Kleyner, "*Determining optimal Reliability Targets through Analysis of Product Validation Cost and Field Warranty Data*," PhD Dissertation, University of Maryland, MD, 2005.
13. T. Hitziger, "*Übertragbarkeit von Vorkenntnissen bei der Zuverlässigkeitstestplanung*," Dissertation, Universität Stuttgart, 2007.
14. W. B. Nelson, "*Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*," Wiley-Interscience, 1990.
15. D. K. Lloyd, M. Lipow, "*Reliability: Management, Methods, and Mathematics*," ASQ, 1984.
16. A. L. Worlund, J. C. Monk, F. D. Bachtel, "Next Launcher System Propulsion Design Considerations," AIAA 92-1181, *AIAA, Aerospace Design Conference*, Irvine, CA, Feb. 3 – 6, 1992
17. J. B. Hall, "*Methodology for evaluating Reliability Growth Programs of discrete Systems*," PhD Dissertation, University of Maryland, MD, 2008.

BIOGRAPHIES

Richard Strunz, M.S., MBAe
Astrium GmbH
Robert-Koch-Strasse 1
Taufkirchen, 82024, Germany

e-mail: richard.strunz@astrium.eads.net

R. Strunz is a Team Leader and Program Manager in the Advanced Programs & Systems Technology department at Astrium GmbH. In this position, Mr. Strunz plans and executes projects on liquid rocket engine system studies and combustion device technology maturations. He also lectures Project and Quality Management and Design and Analysis of Experiments at Munich University of Applied Sciences; Germany. Mr. Strunz earned a Diploma Engineering degree in Aircraft Design from Munich University of Applied Sciences. As a German Academic Exchange Service scholarship holder, he received his M.S. degree in Aerospace Engineering from San José State University; California. While full-time working, he earned a Master of Systems Engineering and an MBA and Engineering degree both awarded from Munich University of Applied Sciences. Mr. Strunz is also a certified Black Belt. Currently, he is pursuing a Ph.D. in Reliability Engineering at the University of Maryland; College Park, Maryland. His dissertation investigates risk-informed satisficed decision-making for liquid rocket engine trade-offs.

Jeffrey W. Herrmann, PhD

Department of Mechanical Engineering
Martin Hall Room 2181
University of Maryland
College Park, Maryland, 20742, USA

e-mail: jwh2@umd.edu

Jeffrey W. Herrmann is an associate professor at the University of Maryland, where he holds a joint appointment with the Department of Mechanical Engineering and the Institute for Systems Research. He is a member of INFORMS, ASME, IIE, SME, and ASEE. Dr. Herrmann earned his B.S. in applied mathematics from Georgia Institute of Technology. He received his Ph.D. in industrial and systems engineering from the University of Florida. His current research interests include emergency preparedness planning and response, health care operations, production scheduling, and engineering design decision-making.