

A Reliability Data Collection and Analysis System for Products under Development

Frank J. Groen

University of Maryland, College Park, USA
E-mail: fgroen@umd.edu

Enrique López Droguett

Federal University of Pernambuco – UFPE, Brazil
E-mail: ealopez@ufpe.br

Siyuan Jiang

Ford Motor Company, Dearborn, USA
E-mail: sjiang@ford.com

Ali Mosleh

University of Maryland, College Park, USA
E-mail: mosleh@umd.edu

Abstrac

This paper presents ReDCAS, the Reliability Data Collection and Analysis System. ReDCAS is a software tool that encapsulates a reliability data collection and analysis methodology for reliability assessment of products under development. The software, developed for Ford Motor Company, employs Bayesian data analysis techniques to estimate reliability measures based on various types of data such as warranty data, test data, and engineering judgments regarding the impact of design changes on the product reliability. It also provides the possibility of incorporating evidence concerning previous revisions of the same product or even information on products that are only similar to the one under development.

Key words: Design Reliability, Product Development, Automotive Reliability, Bayes Theorem

Introduction

The Reliability Data Collection and Analysis System (ReDCAS) software was developed for Ford Motor Company over the period 1995-1999, and has been used to perform reliability assessments for products under development. ReDCAS provides a working environment for engineers to incorporate reliability considerations into the design or redesign of products, even though data on the actual product under design is lacking or absent. This is achieved by basing reliability assessments on data available for different, yet similar products. By considering that these products will typically have similar reliability characteristics, this data can be considered (partially) relevant to the estimation of the new product's

reliability characteristics. Incorporation of the data into the assessment requires that corrections are made based on the anticipated reliability impact of design modifications. Similarly, prototype life test data is incorporated into reliability assessments, considering the effectiveness of the design corrective action to those failures observed during the tests.

Analyses are performed using Bayesian data analysis procedures. The reliability behavior is modeled via the Weibull and Increasing Decreasing Bathtub (IDB) models. The latter model is a three parameter model capable of representing bathtub-shaped failure rate functions. The procedures provide reliability estimates in the form of reliability and failure rate estimates, plus associated uncertainties.

The paper starts with a detailed discussion of the reliability assessment methodology, addressing the purpose and approach of the methodology as well as of the software tool, characterization of the various sources of evidence, an overview of the different steps that comprehend the proposed analysis procedure for assessing the product reliability, and the aggregation of individual components reliability estimates for the reliability assessment at system level. Next, the implementation of the Bayesian estimation procedure is developed, providing modeling details concerning the failure process based on both a Weibull and IDB distributions, and the treatment of warranty data, partially relevant evidence, design credit information, and engineering judgments concerning the impact of design changes. Concluding remarks are then provided.

The REDCAS Methodology

The ReDCAS methodology and software tool was designed to make assessments of the reliability behavior of components that are still in the design stage of their life cycle. Despite of the lack of data originating from the component itself, it is often possible to find alternative data sources that, even though only partially relevant, provide information on which an assessment of a component's reliability can be based.

For instance, in many practical situations, parts are not typically designed from scratch, but rather evolve as a series of designs that are put into operation. While the reliability of products at different stages in such design evolutions is not necessarily identical, earlier designs typically provide a meaningful indication of the reliability behavior of future products, on the basis that their design, manufacturing, and operation are largely the same. If a reliability assessment for future products is desired, the reliability behavior observed for existing products can therefore well serve as a useful source of evidence, as long as the perceived differences between component designs are accounted for.

Approach

The approach of the methodology is to examine and analyze the relevant evidence in a chronological order. To estimate the reliability of a product under development, a baseline reliability assessment is first established from real world usage data of the most representative product that is already in the market. Second, the methodology leads engineers to analyze any intermediate products, if any, which act as bridges to connect the baseline product and the product of interest. Lastly, ReDCAS methodology analyzes the evidence of the product of interest.

In terms of assessment, it is relative easy for an engineer to conduct a relative assessment rather than an absolute reliability assessment. A relative assessment measure, such as percent of improvement, quantifies the difference between two products. Therefore, the overall assessment is built on a solid baseline with subsequent adjustment from the analysis difference.

Sources of Evidence

The ReDCAS methodology allows the user to incorporate various types of evidence to generate reliability assessments. These types of evidence include both recorded test and warranty data as well engineering judgments.

In fact, a first source of data consists of the *warranty data* that is collected for components that are part of the product that have already been released to the market (in the following discussion, vehicles correspond to the product under development). In the Ford context, the data describing the failures of such components is collected in a rather detailed form. Repairs and replacements performed under warranty are documented in a database, which also lists the age and mileage of the vehicle at the time of the repair. The population size and age distribution of the entire population is also approximately known.

Rather than using the raw data records, ReDCAS uses failure rate curves, which describe the components' failure rate as a function of the vehicle's age. These curves take the form of tables, which describe the failure rates on a per month basis, and are compiled based on the recorded failures as well as the estimated population size and age distribution (see Figure 1). Naturally, the time period for which this type of data is available depends on the time that the particular product has been in the market.

ReDCAS allows an analyst to use warranty data collected for different product versions and possibly from multiple years. Usually, the most representative product is the newest in the market that may have very limited time in service. An earlier product therefore has a longer time in service. Further adjustments to the data, for instance due to the warranty policy, are also built into ReDCAS for realistic assessment, as shown in Figure 1.

In order to correct for differences between the tested prototypes and the final products, ReDCAS provides the option of assigning design credit numbers to test records corresponding to failures, representing the engineers degree of belief that the cause of the failure has been addressed in design modifications following the life test, such that the failure would not have occurred in the modified design.

A last main type of evidence consists of engineering assessment of the reliability impact resulted from the planned *design changes*. The measure of the reliability impact could be failure rate ratio or time-to-failure ratio. Take failure rate ratio, *Fr*, as an example. *Fr* is defined as:

$$Fr = \frac{\text{Failure Rate of New Product}}{\text{Failure Rate of Current Product}} \quad (1)$$

where $Fr = 1$ indicates no change, $Fr < 1$ indicates an improvement, and $Fr > 1$, a worsening (risk). *Fr* is usually derived from failure mode analysis or material testing. The uncertainty of *Fr* is described using a probability distribution and further simplified as a three point discrete distribution, known as, Optimistic, Best and Pessimistic Estimates. Figure 3 shows the screen used for entry of this type of information in the case of a Ford application.

The data types described above are supplemented with assessments regarding the relevance of specific data sets to the reliability assessment problem. These assessments will be described in a later section.

Purpose of Change <input type="checkbox"/> Veh. Packaging Req. <input type="checkbox"/> Design Requirement <input checked="" type="checkbox"/> Cost Reduction <input type="checkbox"/> Complexity Reduction <input type="checkbox"/> Manufacturing Request <input type="checkbox"/> Emission Control		<input type="checkbox"/> Fuel Economy <input checked="" type="checkbox"/> Warranty Reduction <input type="checkbox"/> Regulated (mandated item) <input type="checkbox"/> Other Product Change																	
Type of Change <input checked="" type="checkbox"/> Design Change <input type="checkbox"/> Supplier Change <input type="checkbox"/> Process Change		<input type="checkbox"/> Added Feature <input type="checkbox"/> Material Change <input type="checkbox"/> Deleted Feature																	
Type of Reliability Impact <input type="checkbox"/> Time to Failure Change <input checked="" type="checkbox"/> Failure Rate Change		Estimate of Reliability Impact <table border="1"> <tr> <td>Adjustment Factor</td> <td>0.0</td> <td>0.0</td> <td></td> </tr> <tr> <td>Optimistic</td> <td></td> <td></td> <td></td> </tr> <tr> <td>Best Estimate</td> <td></td> <td></td> <td></td> </tr> <tr> <td>Pessimistic</td> <td></td> <td></td> <td></td> </tr> </table>		Adjustment Factor	0.0	0.0		Optimistic				Best Estimate				Pessimistic			
Adjustment Factor	0.0	0.0																	
Optimistic																			
Best Estimate																			
Pessimistic																			
		Detailed Documentation																	
		Engineering Information																	

Figure 3 - Design change impact evidence for the Ford Motor Company case.

Analysis Overview

ReDCAS implements an analysis procedure which breaks down the problem of assessing the reliability of future products into a number of analysis steps that represent stages in the component's design evolution. The main elements of this analysis procedure have also been described in (Lin, 2002).

Each analysis step consists of a Bayesian analysis, and corresponds to a particular stage in the projected design evolution. Therefore, using the evidence from the sources described in the previous section, a different reliability function is estimated at each step in the analysis. The result of the estimation at each step consists of uncertainty distributions over the failure rate as a function of time.

Different percentiles of uncertainty distributions are computed, such as the 5th, 50th and 95th percentiles of the distributions $\pi(\lambda(t))$, see Figure 4. Therefore, the results provide a direct indication of the extent of the uncertainty surrounding the estimate. Note that the results are generated by using the reliability estimates obtained for one stage in the design evolution as a starting point for the next. Also, at each step, the transformations of the reliability functions corresponding to the expected impact of projected design changes are applied, and possibly the inclusion into these estimates of prototype data as it becomes available.

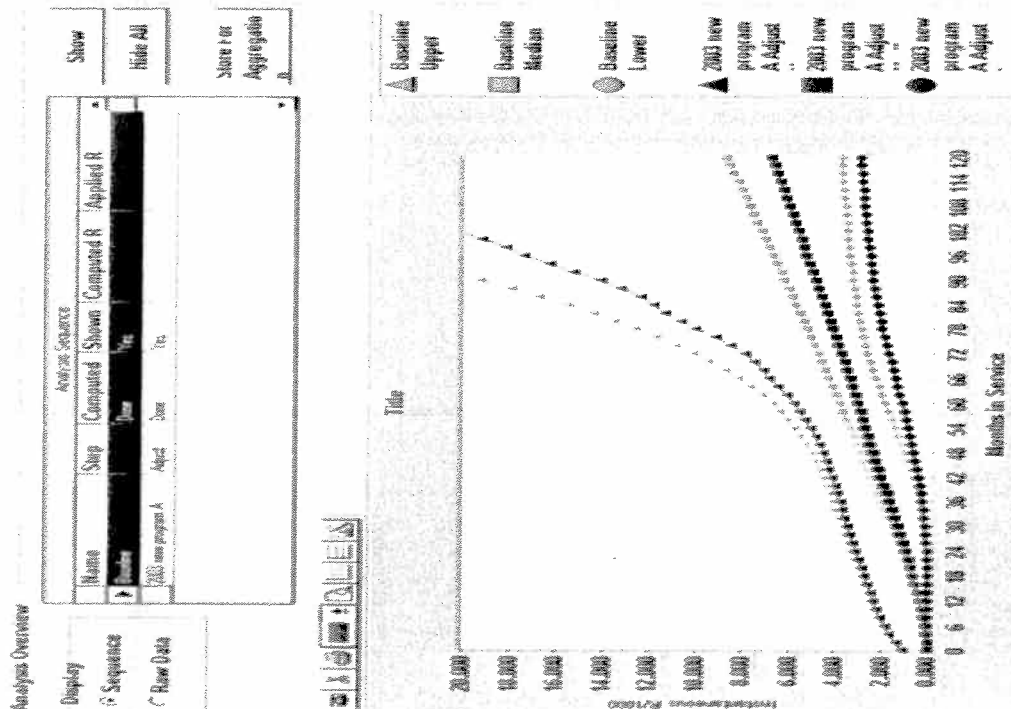


Figure 4 - An example of instantaneous failure intensity predictions.

The analysis steps representing the design evolution of the product under development are detailed in Figure 5. The first step in this analysis flow is to establish a reliability assessment of the baseline comparator. It is usually the newest, most relevant product in the market. To do so, multiple model years of comparator data can be used. These data are considered relevant to the Baseline Comparator. In order to be able to scale the impact of the data on the baseline estimate, a relevance factor, ranging between 0 and 1, is assigned to data originating from the comparators.

Following the baseline analysis, the Figure 5 shows two design programs, separated by the dotted lines. For each design program, three analysis steps are possible. The 'Design Changes' step modifies the result of the previous design step corresponding to the anticipated impact of the design changes. This step therefore does not consist of a Bayesian update in the conventional sense, where data is added to update the estimate of a given quantity, but rather transforms the results from earlier steps in order to estimate a new quantity, in this case the reliability behavior of a new design.

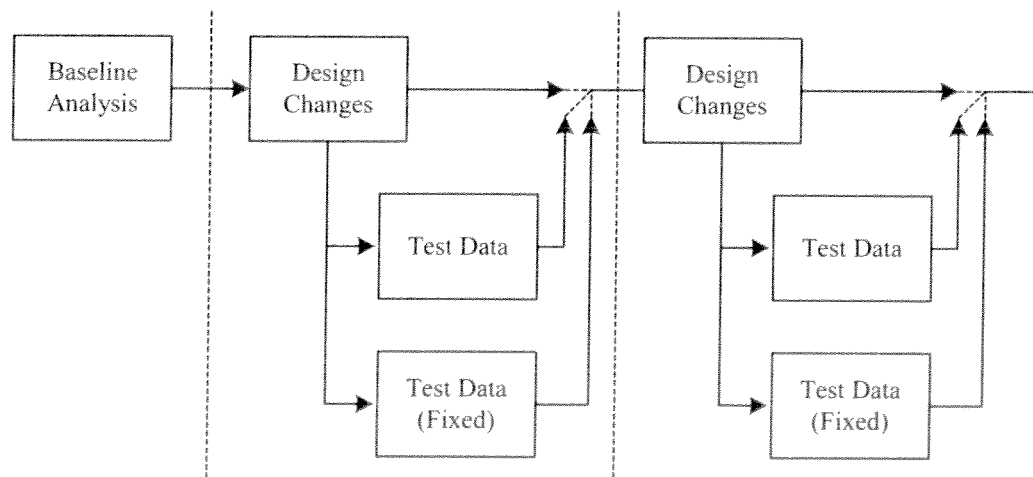


Figure 5 - Overview of the analysis steps.

The 'Test Data' and 'Test Data (Fixed)' steps are used to validate these results based on prototype data. These analysis steps include a check to see whether the test data indicates a reliability behavior significantly different from the behavior that was estimated based on the anticipated impact of design changes, and simultaneously add the test data as additional data to the reliability behavior assessment problem.

The difference between the 'Test Data' and 'Test Data (Fixed)' options is whether the design credits are taken into account or not. Together with the 'Design Changes' step, they form the three analysis steps that are carried out for each design round. Depending on

which of the steps have been performed, one of the three analysis steps is used as the base-line point for the analysis of the next design round.

As shown in Figure 4, for each of these steps, several reliability measures can be computed such as instantaneous failure intensity, failure fraction, and reliability. Note also that reliability estimates for different product revisions can be summarized in a product evolution plot, which clearly illustrate the trends in the estimated values as well as the associated uncertainty distributions.

Aggregation

The reliability estimates obtained for individual components can be used to estimate the reliability at the subsystem and system levels. For instance, the powertrain reliability may be obtained by aggregating reliability results found for the engine and transmission, which in turn may be found by aggregating the results for their respective components.

The aggregation is performed by propagating uncertainty distributions up through a system hierarchy, which corresponds to a structural breakdown of (part of) the vehicle. For each element in the hierarchy, an analysis can be designated as representative of the reliability behavior of that element. Alternatively, aggregation results from lower level analysis can be propagated upwards.

Aggregation takes place under the assumption that the system can be represented as a series system, in which the failure of any component counts as a failure of the system. This assumption is appropriate for warranty return estimation purposes.

Implementation of the Bayesian Estimation Procedure

Failure Process Modeling Assumptions

The failure behavior of vehicles under warranty is considered to be a time-distributed failure process, which includes repair or replacement of parts upon failure. For our estimation purposes, the repair time and availability measures are not of interest. The process is therefore represented using a point process, in which an as-good-as-old repair assumption is used. Further, it is assumed that for each time interval the available data corresponds to a homogeneous population, i.e., a constant failure rate can be considered for each time step. Note, however, that the product's failure can vary across different time intervals, and this variation is captured by making use of an appropriate time-to-failure distribution with varying failure rate (for a detailed discussion on the Bayesian assessment of failure rate from homogeneous and non-homogeneous populations, please refer to (Droguett, 2004)).

ReDCAS methodology is, in principle, capable of handling any parametric time-to-failure distribution. For example, in the context of Ford Motor Company, ReDCAS incorporates the two-parameter Weibull model

$$h(t) = \frac{\beta}{\alpha^\beta} \cdot t^{\beta-1} \tag{2}$$

where α and β are the scale and shape parameters, respectively, and the three parameter IDB model (Hjorth, 1980),

$$h(t) = \delta \cdot t + \frac{\theta}{1 + \beta \cdot t} \tag{3}$$

where θ is the scale parameter, δ and β are both shape parameters. Both models have the capability to represent aging effects. The latter model is also capable of representing bath-tub shaped distributions.

Basic Estimation Procedure

Given the available evidence sources, E , the estimation procedure first generates a sample distribution $\{\theta_1, \theta_2, \dots, \theta_v\}$ representing the posterior density $\pi(\underline{\theta} | E)$, where $\underline{\theta}$ is the set of parameters of the selected time-to-failure distribution. For instance, in the case of the Weibull distribution, we have that $\underline{\theta} = \{\alpha, \beta\}$. In terms of the Bayes theorem,

$$\pi(\underline{\theta} | E) = \frac{\Pr(E | \underline{\theta}) \cdot \pi_o(\underline{\theta})}{\int_{\theta} \Pr(E | \underline{\theta}) \cdot \pi_o(\underline{\theta}) \cdot d\underline{\theta}} \tag{4}$$

The sampling distribution accounts for correlation of the model parameters, which was typically found to be quite strong.

The form of the likelihood function $\Pr(E | \underline{\theta})$ depends on the type of evidence. For a warranty data record, describing the number of failure k recorded amongst n units over an time interval (t_{start}, t_{end}) , a constant failure rate assumption over the relatively short time interval leads to

$$\Pr(E | \underline{\theta}) = \frac{1}{k!} \cdot e^{-\lambda_{\underline{\theta}} \cdot n \cdot (t_{end} - t_{start})} \cdot (\lambda_{\underline{\theta}} \cdot n \cdot (t_{end} - t_{start}))^k \tag{5}$$

Here, values for $\lambda_{\underline{\theta}}$ are determined as failure rate at the interval's mid-point.

In the case of test data records, the likelihood function consists of the failure distribution model's density function or reliability function, depending on whether the test was terminated after failure or suspended without failure. The treatment of other sources of information will be discussed in later sections.

Distributions of such reliability measures as failure rate $h(t)$ and expected cumulative number of failures $c(t)$

$$c(t) = \int_{\tau=0}^t h(\tau) \cdot d\tau \tag{6}$$

are computed in the form of sample distributions $h_i(t)$ and $c_i(t)$, $i = 1, \dots, n$ by evaluating the respective functions for each value $\underline{\theta}_i$, e.g.,

$$h_i(t) = h(t | \underline{\theta}_i), \quad i = 1, \dots, n \tag{7}$$

Treatment of Baseline Comparator Relevance

The baseline comparator relevance factors are used to scale the impact of data originating from the baseline comparator products. Scaling is achieved using one of two methods: the weighted posterior method and the weighted likelihood method.

The weighted posterior is a method in which the distribution over $\underline{\theta}$ is computed as a weighted mix of posterior distributions,

$$\pi(\theta) = \sum_i \omega_i \cdot \pi(\theta | C_i) \tag{8}$$

where each posterior is computed based on the data taken from a unique combination C_i of comparator data sets, leading to a total of 2^m distributions. This weighting method interprets the weight factor w_i assigned to comparators as probabilities that the corresponding data set is permissible as evidence in the analysis. Table 1 illustrates the construction of the data set combinations and the computation of the corresponding weight factors for an example of two comparators.

Table 1 - Data set combinations generated by the Weighted Posterior method.

Evidence Set	Comparator A	Comparator B	Weight w
C_1	.	.	$w_A w_B$
C_2	.		$w_A(1-w_B)$
C_3		.	$(1-w_A)w_B$
C_4			$(1-w_A)(1-w_B)$

The weighted likelihood method consists of a geometric weighting of the likelihood function

$$L_{WL}(\underline{\theta}, w) = (\text{Pr}(E | \underline{\theta}))^w \tag{9}$$

This weighting method can be interpreted as a form of data averaging, in which the likelihood is computed for a hypothetical data set obtained by discounting the statistical strength of the original data set by a factor w . For more details, see (Groen, 1999).

Treatment of Design Credit Information

Design credit information is provided to discount failures observed during prototype testing. A design credit dc can be assigned to each record representing a test terminated upon failure. A design credit of 1 indicates complete confidence that the failure would have been prevented by subsequent design changes, where a design credit of 0 indicates the opposite case.

Two alternative interpretations are applied to the corresponding data set records. In one, the record is interpreted as if the failure did indeed occur at t . In the other, the record is interpreted as if the test was suspended without failure. Data averaging, by means of the weighted likelihood method, is then used to arrive at the likelihood function for test records to which a design credit is applied:

$$L(t) = (1 - R_{\theta}(t))^{dc} \cdot (p_{\theta}(t))^{1-dc} \quad (10)$$

where $R_{\theta}(t)$ and $P_{\theta}(t)$ are respectively the reliability and failure density corresponding to θ .

Treatment of Design Change Information

Design change information consists of engineering judgments regarding the impact of design changes on the reliability behavior. The impact is defined as a modification factor for either time to failure or failure rate. In both cases, the impact is specified in terms of three values, representing the optimistic, pessimistic, and best estimates.

Both types of design impact correction are incorporated by modifying the original data, such that it reflects the intended correction. For time-to-failure corrections, this means that warranty data time interval, as well as test data termination times, are scaled using the specified impact factors. For example, given time to failure correction c , all time entries in warranty data and test data records are replaced by

$$t' = c \cdot t \quad (11)$$

after which the likelihood functions that were discussed earlier are applied.

Failure rate corrections are applied to warranty data records by modifying the number of failures k observed during the time interval by the specified impact factor

$$k' = c \cdot k \quad (12)$$

In case of test data, a failure rate correction is not directly possible. An approximation is used by which the failure rate increase or decrease is translated into a time to failure correction under a constant failure rate assumption.

The procedure is repeated for the pessimistic, optimistic, and best estimate cases, leading to three separate posterior distributions. The weighted posterior approach is then applied to combine the three posterior distributions. Weight factors of 0.25 are applied to the posteriors corresponding to the pessimistic and optimistic corrections; a weight factor of 0.5 to the best estimate posterior.

Concluding Remarks

The ReDCAS methodology presented in this article allows for the reliability assessment of products under development. By making use of the methodology, the analyst is able to obtain product reliability estimates based on various sources of evidence and updated as new evidence becomes available at different stages of the development cycle. In particular, the methodology was discussed in the context of the Ford Motor Company applications, and put forward for the treatment of the following types of evidence: warranty data in terms instantaneous failure rate, test data, engineering judgments regarding impact of design changes and design credits.

An important aspect of the methodology refers to the flexibility of using (partially relevant) evidence gathered from previous revisions of the product under development or even from products that are only similar to the one being developed.

Although the assessment methodology was developed for the case of test data obtained under normal operational conditions, it is directly applicable to the case of accelerated lifetime testing data. For details, see (Droguett, 1999).

Acknowledgment

This paper describes the work that was funded by Ford Motor Company.

References

- Droguett, E.L., Groen, F.J. and Mosleh, A. (2004) The Combined Use of Data and Expert Estimates in Population Variability Analysis. Accepted for publication in the Reliability Engineering and System Safety Journal.
- Droguett, E.L. and Mosleh, A. (1999) Development of a Bayesian Procedure to the Power - Weibull Model with Nonconstant Shape Parameter, Proceedings of ESREL'99 - The Tenth European Conference on Safety and Reliability. Munich, Germany.
- Groen, F.J. and Mosleh, A. (1999) Behavior of Weighted Likelihood and Weighted Posterior Methods for Treatment of Uncertain Data, Proceedings of ESREL'99 - The Tenth European Conference on Safety and Reliability, Munich, Germany.
- Hjorth, U. (1980) A Reliability Distribution with Increasing, Decreasing, Constant and Bathtub-Shaped Failure Rates, Technometrics, Vol. 22, No. 1, pp. 99-107.

Lin, J.C. (2002) The Use of Indirect Evidence for Bayesian Reliability Analysis, Proceedings of the Annual Reliability and Maintainability Symposium - RAMS, Seattle, pp. 429-433.

Biography

Frank Groen is a Research Associate at the University of Maryland. His research interests include Bayesian data analysis and simulation methods for risk analysis. Frank is also president of Prediction Technologies, a company focused on the application of Bayesian methods in support of making under uncertainty. He received his PhD in reliability engineering from the University of Maryland in 2000.

Enrique López is an Associate Professor at the Federal University of Pernambuco (Brazil), Department of Production Engineering. His research interests include Bayesian data analysis, Probabilistic Risk Assessment, and simulation methods for reliability and risk analysis. He received his PhD from the University of Maryland in 1999.

Siyuan Jiang is Supervisor, Advanced Reliability Section, Global Core Quality. He has worked for Powertrain Operations, Global Core Engineering and Global Core Quality since he joined Ford Motor Company in 1990. He is responsible for reliability processes and implementation strategy development, reliability mathematical modeling, and reliability methods implementation. Dr. Jiang received a MSME from Beijing Institute of Iron and Steel Technology, China, in 1983; and received his MS in 1989 and Ph.D. in 1991 in Mechanical Engineering with specialty in Reliability Engineering from The University of Arizona.

Ali Mosleh is a Professor in the Reliability Engineering Program at the University of Maryland, College Park, MD. He received his Ph.D. in Nuclear Engineering from UCLA. His areas of research include Probabilistic Risk Assessment, Human Reliability, Data Analysis and Expert Opinion.

Call for Papers

ENEGEP 2004 and X Conference on Industrial Engineering and Operations Management

3-5 November, 2004
Florianópolis, SC, Brazil
Full paper submission: 12th May, 2004
www.abepro.org.br

2nd International Conference on Production Research – America's 2004

1-3 August, 2004
Santiago, Chile
Article submission: 16th April, 2004
Notification to the authors: 28th May, 2004
www.e-diind.usach.cl/icpr-america04/

18th International Conference on Production Research

31st July to 4th August, 2005
Fiscano - SA, Italy
Abstract submission: 5th November, 2004
Notification to the authors: 20th December, 2004
Paper submission: 18th March, 2005
Notification to the authors: 20th May, 2005
Final paper submission: 24th June, 2005
<http://icpr18.unisa.it>

ABEPRO Executive and Fiscal Board

IVE BOARD

President

Nivaldo L. Coppini, UNIMEP, SP, Brazil

First Vice-President

Paulo M. Selig, UFSC, SC, Brazil

Second Vice-President

Maria Rita P.A. Alves, UFSCar, SP, Brazil

Financial Director

Francisco C. Marcondes, Sandvik, SP, Brazil

Administrative Director

Osvaldo L.G. Quelhas, UFF, RJ, Brazil

Technical Director

Gilberto D. da Cunha, PUCRS, RS, Brazil

Student Director

Jair T. de Souza, UFJF, MG, Brazil

Executive Board Substitutes

Marly M. de Carvalho, USP, SP, Brazil

Luiz G.M. de Souza, UNIFEI, MG, Brazil

Sérgio E. Gouvêa da Costa, PUCPR, PR, Brazil

Fiscal Board

Adriana F. de Faria, UNIMINAS, MG, Brazil

Ricardo M. Naveiro, UFRJ, RJ, Brazil

Patrícia A. Cardoso, UVV, ES, Brazil

Fiscal Board Substitutes

Bernardete Trindade, UFSM, RS, Brazil

Marco Antonio Buseti de Paula, PUCPR, PR, Brazil

Francisco José Costa Araújo, UPE, PE, Brazil

Notes for Contributors

http://www.bjopm.abepro.org.br/bjopm_notes.aps