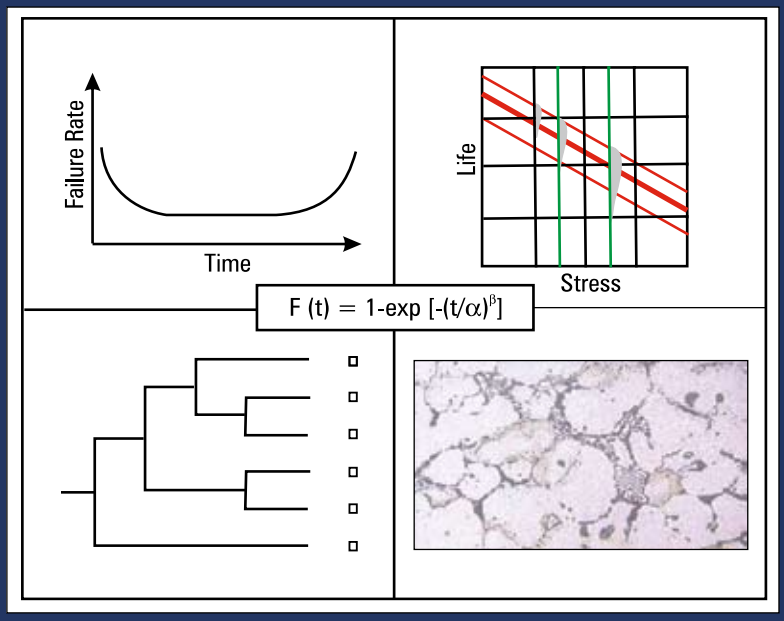
SRESA Journal of

LIFE CYCLE RELIABILITY AND SAFETY ENGINEERING

Vol. 2 Issue No. 2 April – June 2012 ISSN – 2250 0820



Chief-Editors

P.V. Varde
A.K. Verma
Michael G. Pecht

Society for Reliability and Safety

SRESA Journal of Life Cycle Reliability and Safety Engineering

Extensive work is being performed world over on assessment of Reliability and Safety for engineering systems in support of decisions. The increasing number of risk-based / risk-informed applications being developed world over is a testimony to the growth of this field. Here, along with probabilistic methods, deterministic methods including Physics-of-Failure based approach is playing an important role. The International Journal of Life Cycle Reliability and Safety Engineering provides a unique medium for researchers and academicians to contribute articles based on their R&D work, applied work and review work, in the area of Reliability, Safety and related fields. Articles based on technology development will also be published as Technical Notes. Review articles on Books published in the subject area of the journal will also form part of the publication.

Society for Reliability and Safety has been actively working for developing means and methods for improving system reliability. Publications of quarterly News Letters and this journal are some of the areas the society is vigorously pursuing for societal benefits. Manuscript in the subject areas can be communicated to the Chief Editors. Manuscript will be reviewed by the experts in the respective area of the work and comments will be communicated to the corresponding author. The reviewed final manuscript will be published and the author will be communicated the publication details. Instruction for preparing the manuscript has been given on inside page of the end cover page of each issue. The rights of publication rest with the Chief-Editors.

SCOPE OF JOURNAL

System Reliability analysis	**Structural Reliability**	**Risk-based applications**
Statistical tools and methods	Remaining life prediction	Technical specification optimization
Probabilistic Safety Assessment	Reliability based design	Risk-informed approach
Quantitative methods	Physics-of-Failure methods	Risk-based ISI
Human factor modeling	Probabilistic Fracture Mechanics	Risk-based maintenance
Common Cause Failure analysis	Passive system reliability	Risk-monitor
Life testing methods	Precursor event analysis	Prognostics & health management
Software reliability	Bayesian modeling	Severe accident management
Uncertainty modeling	Artificial intelligence in risk and reliability modeling	Risk-based Operator support systems
Dynamic reliability models	Design of Experiments	Role of risk-based approach in Regulatory reviews
Sensitivity analysis	Fuzzy approach in risk analysis	Advanced electronic systems reliability modeling
Decision support systems	Cognitive framework	Risk-informed asset management

SRESA AND ITS OBJECTIVES

a) To promote and develop the science of reliability and safety.

b) To encourage research in the area of reliability and safety engineering technology & allied fields.

c) To hold meetings for presentation and discussion of scientific and technical issues related to safety and reliability.

d) To evolve a unified standard code of practice in safety and reliability engineering for assurance of quality based professional engineering services.

e) To publish journals, books, reports and other information, alone or in collaboration with other organizations, and to disseminate information, knowledge and practice of ensuring quality services in the field of Reliability and Safety.

f) To organize reliability and safety engineering courses and / or services for any kind of energy systems like nuclear and thermal power plants, research reactors, other nuclear and radiation facilities, conventional process and chemical industries.

g) To co-operate with government agencies, educational institutions and research oganisations

SRESA Journal of

LIFE CYCLE RELIABILITY AND SAFETY ENGINEERING

Vol.2 Issue No.2 April – June 2012 ISSN – 2250 0820

Chief-Editors

P.V. Varde
A.K. Verma
Michael G. Pecht

SOCIETY FOR RELIABILITY AND SAFETY

Typeset & Printed

EBENEZER PRINTING HOUSE
Unit No. 5 & 11, 2nd Floor, Hind Services Industries,
Veer Savarkar Marg,
Dadar (west), Mumbai -28
Tel.: 2446 2632/ 3872
E-mail: outwork@gmail.com

Life Cycle Reliability and Safety Engineering
Vol.2 Issue 2 (2012) 01-15

Physical perspective towards stochastic optimal controls of engineering structures

Jie Li[1,2*], Yong-Bo Peng[1,3]

[1] *State Key Laboratory of Disaster Reduction in Civil Engineering, Tongji University, Shanghai 200092, China*
[2] *School of Civil Engineering, Tongji University, Shanghai 200092, China*
[3] *Shanghai Institute of Disaster Prevention and Relief, Tongji University, Shanghai 200092, China*

ABSTRACT

In the past few years, starting with the thought of physical stochastic systems and the principle of preservation of probability, a family of probability density evolution methods (PDEM) has been developed. It provides a new perspective towards the accurate design and optimization of structural performance under random engineering excitations such as earthquake ground motions and strong winds. On this basis, a physical approach to structural stochastic optimal control is proposed in the present paper. A family of probabilistic criteria, including the criterion based on mean and standard deviation of responses, the criterion based on exceeding probability, and the criterion based on global reliability of systems, is elaborated. The stochastic optimal control of a randomly base-excited single-degree-of-freedom system with active tendon is investigated for illustrative purposes. The results indicate that the control effect relies upon control criteria of which the control criterion in global reliability operates efficiently and gains the desirable structural performance. The results obtained by the proposed method are also compared against those by the LQG control, revealing that the PDEM-based stochastic optimal control exhibits significant benefits over the classical LQG control. Besides, the stochastic optimal control, using the global reliability criterion, of an eight-storey shear frame structure is carried out. The numerical example elucidates the validity and applicability of the developed physical stochastic optimal control methodology.

Keywords: *probability density evolution method; stochastic optimal control; control criteria; global reliability; LQG control*

1 Introduction

Stochastic dynamics has gained increasing interests and has been extensively studied. However, although the original thought may date back to Einstein (1905) and Langevin (1908) and then studied in rigorous formulations by mathematicians (Kolmogorov, 1931; Wiener, 1923; Itô, 1942), the random vibration theory, a component of stochastic dynamics, was only regarded as a branch of engineering science until the early of 1960's (Crandall, 1958; Lin, 1967). Till early 1990's, the theory and pragmatic approaches for random vibration of linear structures were well developed. Meanwhile, researchers were challenged by nonlinear random vibration, despite great efforts devoted coming up with a variety of methods, including the stochastic linearization, equivalent non-linearization, stochastic averaging, path-integration method, FPK equation and the Monte Carlo simulation and so on (see, e.g., Zhu, 1992; Lin & Cai, 1995; Lutes & Sarkarni, 2004). The challenge still existed. On the other hand, investigations on stochastic structural analysis (or referred to stochastic finite element method by some researchers), as a critical component of stochastic dynamics, in which the randomness of structural parameters is dealt with, started a little later from the late 1960's.

Till middle 1990's, a serious of approaches were presented, among which three were dominant, including the Monte Carlo simulation (Shinozuka & Jan, 1972; Shinozuka & Deodatis, 1991), the random perturbation technique (Kleiber & Hien, 1992, Haldar & Mahadevar, 2000) and the orthogonal polynomial expansion (Ghanem & Spanos, 1991; Li, 1996). Likewise with the random vibration, here the analysis of nonlinear stochastic structures encountered huge challenges as well (Schenk & Schuëller, 2005).

In the past ten years, starting with the thought of physical stochastic systems (Li, 2006) and the principle of preservation of probability (Li & Chen, 2008), a family of probability density evolution methods (PDEM) has been developed, in which a generalized density evolution equation was established. The generalized density evolution equation profoundly reveals the essential relationship between the stochastic and deterministic systems. It is successfully employed in stochastic dynamic response analysis of multi-degree-of-freedom systems (Li and Chen, 2009), and therefore provides a new perspective towards serious problems such as the dynamic reliability of structures, the stochastic stability of dynamical systems and the stochastic optimal control of engineering structures.

In this paper, the application of PDEM on the stochastic optimal control of structures will be summarised. Therefore, the fundamental theory of the generalized density evolution equation is firstly revisited. A physical approach to stochastic optimal control of structures is then presented. The optimal control criteria, including those based on mean and standard deviation of responses and those based on exceeding probability and global reliability of systems, are elaborated. The stochastic optimal control of a randomly base-excited single-degree-of-freedom system with active tendon is investigated for illustrative purposes. Comparative studies of these probabilistic criteria and the developed control methodology against the classical LQG control are carried out. The optimal control strategy is then further employed in the investigation of the stochastic optimal control of an eight-storey shear frame. Some concluding remarks are included.

2 Principle of Preservation of Probability And Generalized Density Evolution Equation

Principle of preservation of probability revisited

It is noted that the probability evolution in a stochastic dynamical system admits the principle of preservation of probability, which can be stated as: if the random factors involved in a stochastic system are retained, the probability will be preserved in the evolution process of the system. Although this principle may be faintly cognized quite long ago (see, e.g., Syski, 1967), the physical meaning has been only clarified in the past few years from the state description and random event description, respectively (Li & Chen, 2003, 2006, 2008). The fundamental logic position of the principle of preservation of probability was then solidly established with the development of a new family of generalized density evolution equations that integrates the ever-proposed probability density evolution equations, including the classic Liouville equation, Dostupov-Pugachev equation and the FPK equation (Li & Chen, 2009).

To re-visit the principle of preservation of probability, consider an n-dimensional stochastic dynamical system governed by the following state equation

$$\dot{\mathbf{Y}} = \mathbf{A}(\mathbf{Y}, t), \mathbf{Y}(t_0) = \mathbf{Y}_0 \tag{1}$$

where $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)^{\mathrm{T}}$ denotes the n-dimensional state vector; $\mathbf{Y}_0 = (Y_{0,1}, Y_{0,2}, \cdots, Y_{0,n})^{\mathrm{T}}$ denotes the corresponding initial vector; $\mathbf{A}(\cdot)$ is a deterministic operator vector. Evidently, in the case that $\mathbf{Y}_0$ is a random vector, $\mathbf{Y}(t)$ will be a stochastic process vector.

The state equation (1) essentially establishes a mapping from $\mathbf{Y}_0$ to $\mathbf{Y}(t)$, which can be expressed as

$$\mathbf{Y}(t) = g(\mathbf{Y}_0, t) = \mathscr{G}_t(\mathbf{Y}_0) \tag{2}$$

where $g(\cdot), \mathscr{G}_t(\cdot)$ are both mapping operators from $\mathbf{Y}_0$ to $\mathbf{Y}(t)$.

Since $\mathbf{Y}_0$ denotes a random vector, $\{\mathbf{Y}_0 \in \Omega_{t_0}\}$ is a random event. Here Ω_{t_0} is any arbitrary domain in the distribution range of $\mathbf{Y}_0$. According to the stochastic state equation (1), $\mathbf{Y}_0$ will be changed to $\mathbf{Y}(t)$ at time t. The domain Ω_{t_0} to which $\mathbf{Y}_0$ belongs at time t_0 is accordingly changed to Ω_t to which $\mathbf{Y}(t)$ belongs at time t; see Figure 1.

$$\Omega_t = g(\Omega_{t_0}, t) = \mathscr{G}_t(\Omega_{t_0}) \tag{3}$$

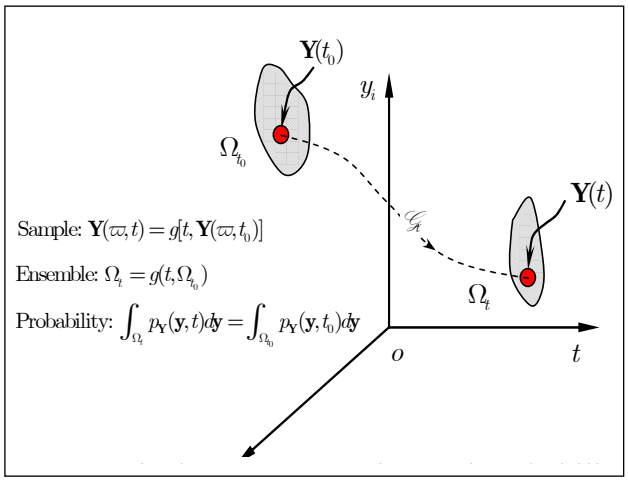


Figure 1. Dynamical system, mapping and probability evolution

Since the probability is preserved in the mapping of any arbitrary element events, we have

$$\int_{\Omega_{t_0}} p_{\mathbf{Y}_0}(\mathbf{y}, t_0)d\mathbf{y} = \int_{\Omega_t} p_{\mathbf{Y}}(\mathbf{y}, t)d\mathbf{y} \tag{4}$$

It is understood that Eq. (4) also holds at $t + \Delta t$, which will then result in

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t} p_{\mathbf{Y}}(\mathbf{y}, t)d\mathbf{y} = 0 \tag{5}$$

where operates its arguments with denotation of total derivative.

Eq. (5) is clearly the mathematical formulation of the principle of preservation of probability in a stochastic dynamical system. Since the fact of probability invariability of a random event is recognized here, we refer to Eq. (5) as the random event description of the principle of preservation of probability. The meaning of the principle of preservation of probability can also be clarified from the state space description. These two descriptions are somehow analogous to the Lagrangian and Eulerian descriptions in the continuum mechanics, although there also some distinctive properties particularly in whether overlapping is allowed. For details, refer to Li & Chen (2006a; 2008).

Generalized density evolution equation (GDEE)

Without loss of generality, consider the equation of motion of a multi-degree-of-freedom (MDOF) system as follows

$$\mathbf{M}(\boldsymbol{\eta})\ddot{\mathbf{X}} + \mathbf{C}(\boldsymbol{\eta})\dot{\mathbf{X}} + \mathbf{f}(\boldsymbol{\eta}, \mathbf{X}) = \boldsymbol{\Gamma}\xi(t) \tag{6}$$

where $\boldsymbol{\eta} = (\eta_1, \eta_2, \cdots, \eta_{s_1})$ are the random parameters involved in the physical properties of the system. If the excitation is a stochastic ground accelerogram $\xi(t) = \ddot{X}_g(t)$, , for example, then $\boldsymbol{\Gamma} = -\mathbf{M}\mathbf{1}$, $\mathbf{1} = (1,1,\cdots,1)^{\mathrm{T}}$. Here $\ddot{\mathbf{X}}, \dot{\mathbf{X}}, \mathbf{X}$ are the accelerations, velocities and displacements of the structure relative to ground. $\mathbf{M}(\cdot), \mathbf{C}(\cdot), \mathbf{f}(\cdot)$ denote the mass, damping and stiffness matrices of the structural system, respectively.

In the modelling of stochastic dynamic excitations such as earthquake ground motions, strong winds and sea waves, the thought of physical stochastic process can be employed (Li & Ai, 2006; Li, 2008; Li et al, 2011b). For general stochastic processes or random fields, the double stage orthogonal decomposition can be adopted such that the excitation could be represented by a random function (Li & Liu, 2006)

$$\ddot{X}_g(t) = \ddot{X}_g(\boldsymbol{\zeta}, t) \tag{7}$$

where . $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \cdots, \zeta_{s_2})$.

For notational consistency, denote

$$\boldsymbol{\Theta} = (\boldsymbol{\eta}, \boldsymbol{\zeta}) = (\eta_1, \eta_2, \cdots, \eta_{s_1}, \zeta_1, \zeta_2, \cdots, \zeta_{s_2}) = (\Theta_1, \Theta_2, \cdots, \Theta_s) \tag{8}$$

in which $s = s_1 + s_2$ is the total number of the basic random variables involved in the system. Eq. (6) can thus be rewritten into

$$\mathbf{M}(\boldsymbol{\Theta})\ddot{\mathbf{X}} + \mathbf{C}(\boldsymbol{\Theta})\dot{\mathbf{X}} + \mathbf{f}(\boldsymbol{\Theta}, \mathbf{X}) = \mathbf{F}(\boldsymbol{\Theta}, t) \tag{9}$$

where $\mathbf{F}(\boldsymbol{\Theta}, t) = \boldsymbol{\Gamma}\ddot{X}_{\mathrm{g}}(\zeta, t)$.

This is the equation to be resolved in which all the randomness from the initial conditions, excitations and system parameters is involved and exposed in a unified manner. Such a stochastic equation of motion can be further rewritten into a stochastic state equation which was firstly formulated by Dostupov & Pugachev (1957).

If, besides the displacements and velocities, we are also interested in other physical quantities $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_m)^{\mathrm{T}}$ in the system (e.g. the stress, internal forces, etc.), then the augmented system $(\mathbf{Z}, \boldsymbol{\Theta})$ is probability preserved because all the random factors are involved, thus according to Eq. (5) we have (Li & Chen, 2008)

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega_t \times \Omega_\theta} p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t) d\mathbf{z} d\boldsymbol{\theta} = 0 \tag{10}$$

where $\Omega_t \times \Omega_\theta$ is any arbitrary domain in the augmented state space $\Omega \times \Omega_{\boldsymbol{\Theta}}$, $\Omega_{\boldsymbol{\Theta}}$ is the distribution range of the random vector $\boldsymbol{\Theta}$, $p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)$ is the joint probability density function (PDF) of $(\mathbf{Z}(t), \boldsymbol{\Theta})$.

After a series of mathematical manipulations, including the use of Reynold's transfer theorem, we have

$$\int_{\Omega_{t_0} \times \Omega_\theta} \left(\frac{\partial p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)}{\partial t} + \sum_{j=1}^{m} \dot{Z}_j(\boldsymbol{\theta}, t) \frac{\partial p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)}{\partial z_j} \right) d\mathbf{z} d\boldsymbol{\theta} = 0 \tag{11}$$

which holds for any arbitrary $\Omega_{t_0} \times \Omega_\theta \in \Omega \times \Omega_{\boldsymbol{\Theta}}$. Thus we have for any arbitrary $\Omega_\theta \in \Omega_{\boldsymbol{\Theta}}$

$$\int_{\Omega_\theta} \left(\frac{\partial p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)}{\partial t} + \sum_{j=1}^{m} \dot{Z}_j(\boldsymbol{\theta}, t) \frac{\partial p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)}{\partial z_j} \right) d\boldsymbol{\theta} = 0 \tag{12}$$

and also the following partial differential equation

$$\frac{\partial p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)}{\partial t} + \sum_{j=1}^{m} \dot{Z}_j(\boldsymbol{\theta}, t) \frac{\partial p_{\mathbf{Z}\boldsymbol{\Theta}}(\mathbf{z}, \boldsymbol{\theta}, t)}{\partial z_j} = 0 \tag{13}$$

Specifically, as $m = 1$ Eqs. (12) and (13) become, respectively

$$\int_{\Omega_\theta} \left(\frac{\partial p_{Z\boldsymbol{\Theta}}(z, \boldsymbol{\theta}, t)}{\partial t} + \dot{Z}(\boldsymbol{\theta}, t) \frac{\partial p_{Z\boldsymbol{\Theta}}(z, \boldsymbol{\theta}, t)}{\partial z} \right) d\boldsymbol{\theta} = 0 \tag{14}$$

and

$$\frac{\partial p_{Z\boldsymbol{\Theta}}(z, \boldsymbol{\theta}, t)}{\partial t} + \dot{Z}(\boldsymbol{\theta}, t) \frac{\partial p_{Z\boldsymbol{\Theta}}(z, \boldsymbol{\theta}, t)}{\partial z} = 0 \tag{15}$$

which is a one-dimensional partial differential equation.

Eqs. (13) and (15) are referred to as generalized density evolution equations (GDEEs). They reveal the intrinsic connections between a stochastic dynamical system and its deterministic counterpart. It is remarkable that the dimension of a GDEE is not relevant to the dimension (or degree-of-freedom) of the original system; see Eq. (9). This distinguishes GDEEs from the traditional probability density evolution equations (e.g. Liouville, Dostupov-Pugachev and FPK equations), of which the dimension must be identical to the dimension of the original state equation (twice the degree-of-freedom).

Clearly, Eq. (14) is mathematically equivalent to Eq. (15). But it will be seen later that Eq. (14) itself may provide additional insight into the problem. Particularly, if the physical quantity of Z interest is the displacement X of the system, Eq. (15) becomes

$$\frac{\partial p_{X\boldsymbol{\Theta}}(x, \boldsymbol{\theta}, t)}{\partial t} = -\dot{X}(\boldsymbol{\theta}, t) \frac{\partial p_{X\boldsymbol{\Theta}}(x, \boldsymbol{\theta}, t)}{\partial x} \tag{16}$$

Here we can see the rule clearly revealed by the GDEE: in the evolution of a general dynamical system, the time variant rate of the joint PDF of displacement and source random parameters is proportional to the space variant rate with the coefficient being instantaneous velocity. In other words, the flow of probability is determined by the change of physical states. This demonstrates strongly that the evolution of probability density is not disordered, but admits a restrictive physical law. Clearly, this holds for the general physical system with underlying randomness. This rule could not be exposed in such an explicit way in the traditional probability density evolution equations.

Although in principle the GDEE holds for any arbitrary dimension, in most cases one or two-dimensional GDEEs are adequate. For simplicity and clarity, in the following sections we will be focused on the one-dimensional GDEE. Generally, the boundary condition for Eq. (15) is

$$p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)\big|_{z\to\pm\infty}=0\text{, or }\ p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)=0,\quad z\in\Omega_{\mathrm{f}} \tag{17}$$

the latter of which is usually adopted in first-passage reliability evaluation where Ω_{f} is the failure domain, while the initial condition is usually

$$p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)\big|_{t=t_0}=\delta(z-z_0)p_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \tag{18}$$

where z_0 is the deterministic initial value.

Solving Eq. (15), the instantaneous PDF of $Z(t)$ can be obtained by

$$p_Z(z,t)=\int_{\Omega_{\boldsymbol{\Theta}}}p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)d\boldsymbol{\theta} \tag{19}$$

The GDEE was firstly obtained as the uncoupled version of the parametric Liouville equation for linear systems (Li & Chen, 2003). Then for nonlinear systems, the GDEE was reached when the formal solution was employed (Li & Chen, 2006b). It is from the above derivation that the meanings of the GDEE were thoroughly clarified and a solid physical foundation was laid (Li & Chen, 2008).

Point evolution and ensemble evolution

Since Eq. (14) holds for any arbitrary $\Omega_\theta\in\Omega_{\boldsymbol{\Theta}}$, then for any arbitrary partition of probability-assigned space (Chen et al, 2009), of which the sub-domains are Ω_q's, $q=1,2,\cdots,n_{\mathrm{pt}}$ satisfying $\Omega_i\bigcap\Omega_j=\varnothing,\forall i\neq j$ and $\bigcup_{q=1}^{n_{\mathrm{pt}}}\Omega_q=\Omega_{\boldsymbol{\Theta}}$, Eq. (14) constructed in the sub-domain then becomes

$$\int_{\Omega_q}\left(\frac{\partial p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)}{\partial t}+\dot{Z}(\boldsymbol{\theta},t)\frac{\partial p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)}{\partial z}\right)d\boldsymbol{\theta}=0,\quad q=1,2,\cdots,n_{\mathrm{pt}} \tag{20}$$

It is noted that

$$P_q=\int_{\Omega_q}p_{\boldsymbol{\Theta}}(\boldsymbol{\theta})d\boldsymbol{\theta},\quad q=1,2,\cdots,n_{\mathrm{pt}} \tag{21}$$

is the assigned probability over Ω_q (Chen et al, 2009), and

$$p_q(z,t)=\int_{\Omega_q}p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)d\boldsymbol{\theta},\quad q=1,2,\cdots,n_{\mathrm{pt}} \tag{22}$$

then Eq. (20) becomes

$$\frac{\partial p_q(z,t)}{\partial t}+\int_{\Omega_q}\left[\dot{Z}(\boldsymbol{\theta},t)\frac{\partial p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)}{\partial z}\right]d\boldsymbol{\theta}=0,\quad q=1,2,\cdots,n_{\mathrm{pt}} \tag{23}$$

According to Eq. (19), it follows that

$$p_Z(z,t)=\sum_{q=1}^{n_{\mathrm{pt}}}p_q(z,t) \tag{24}$$

There are two important properties that can be observed here:

(i) Partition of probability-assigned space and the property of independent evolution

The functions $p_q(z,t)$ defined in Eq. (22) themselves are not probability density functions because $\int_{-\infty}^{\infty}p_q(z,t)dz=P_q\neq 1$, i.e. the consistency condition is not satisfied. However, except for this violation, they are very similar to probability density functions in many aspects. Actually, a normalized function $\tilde{p}_q(z,t)=p_q(z,t)/P_q$ meets all the conditions of a probability density function, which might be called the partial-probability density function over Ω_q. Eq. (24) can then be rewritten into

$$p_Z(z,t)=\sum_{q=1}^{n_{\mathrm{pt}}}P_q\cdot\tilde{p}_q(z,t) \tag{25}$$

It is noted that P_q's are specified by the partition and are time invariant. Thus, the probability density function of $Z(t)$ could be regarded as the weighted sum of a set of partial-probability density functions. What is interesting regarding the partial-probability density functions is that they are in a sense mutually independent, i.e. once a partition of probability-assigned space is determined (consequently Ω_q's are specified), then a partial probability density function $\tilde{p}_q(z,t)$ is completely governed by Eq. (23) (it is of course true if the function $p_q(z,t)$ is substituted by $\tilde{p}_q(z,t)$), the evolution of other partial probability density functions, $\tilde{p}_r(z,t),r\neq q$, has no effects on the evolution of $\tilde{p}_q(z,t)$. This property of independent evolution of partial probability density function means that the original problem can be partitioned into a series of independent sub-problems, which

are usually easier than the original problem. Thus, the possibility of new approaches is implied but still to be explored. It is also stressed that such a property of independent evolution is not conditioned on any assumption of mutual independence of basic random variables.

(ii) Relationship between point evolution and ensemble evolution

The second term in Eq. (23) usually cannot be integrated explicitly. It is seen from this term that to capture the partial-probability density function $\tilde{p}_q(z,t)$ over Ω_q, the exact information of the velocity dependency on $\boldsymbol{\theta} \in \Omega_q$ is required. This means that the evolution of $\tilde{p}_q(z,t)$ depends on all the exact information in Ω_q, in other words, the evolution of $\tilde{p}_q(z,t)$ is determined by the evolution of information of the ensemble over Ω_q. This manner could be called ensemble evolution.

To uncouple the second term in Eq. (23), we can assume

$$\dot{X}(\boldsymbol{\theta},t) \doteq \dot{X}(\boldsymbol{\theta}_q,t), \quad \text{for } \boldsymbol{\theta} \in \Omega_q \tag{26}$$

where $\boldsymbol{\theta}_q \in \Omega$ is a representative point of Ω_q. For instance, $\boldsymbol{\theta}_q$ could be determined by the Voronoi cell (Chen et al, 2009), by the average $\boldsymbol{\theta}_q = \frac{1}{P_q}\int_{\Omega_q} \boldsymbol{\theta} p_{\boldsymbol{\Theta}}(\boldsymbol{\theta})d\boldsymbol{\theta}$, or in some other appropriate manners. By doing this, Eq. (23) becomes

$$\frac{\partial p_q(z,t)}{\partial t} + \dot{Z}(\boldsymbol{\theta}_q,t)\frac{\partial p_q(z,t)}{\partial z} = 0, \quad q = 1,2,\cdots,n_{\text{pt}} \tag{27}$$

The meaning of Eq. (26) is clear that the ensemble evolution in Eq. (23) is represented by the information of a representative point in the sub-domain, i.e. the ensemble evolution in a sub-domain is represented by a point evolution.

Another possible manner of uncoupling the second term in Eq. (23) implies a small variation of $p_{Z\boldsymbol{\Theta}}(z,\boldsymbol{\theta},t)$ over the sub-domain Ω_q. In this case, it follows that

$$\frac{\partial p_q(z,t)}{\partial t} + E_q[\dot{Z}(\boldsymbol{\theta},t)]\frac{\partial p_q(z,t)}{\partial z} = 0, \quad q = 1,2,\cdots,n_{\text{pt}} \tag{28}$$

where $E_q[\dot{Z}(\boldsymbol{\theta},t)] = \frac{1}{P_q}\int_{\Omega_q} \dot{Z}(\boldsymbol{\theta},t)p_{\boldsymbol{\Theta}}(\boldsymbol{\theta})d\boldsymbol{\theta}$ is the average of $\dot{Z}(\boldsymbol{\theta},t)$ over Ω_q. In some cases, $E_q[\dot{Z}(\boldsymbol{\theta},t)]$ might be close to $\dot{Z}(\boldsymbol{\theta}_q,t)$ and thus Eqs. (27) and (28) coincide.

Numerical procedure for the GDEE

In the probability density evolution method, Eq. (9) is the physical equation while Eq. (15) is the GDEE with initial and boundary conditions specified by Eqs. (17) and (18). Hence, solving the problem needs to incorporate physical equations and the GDEE. For some very simple cases, a closed-form solution might be obtained, say, by the method of characteristics (Li & Chen, 2006b). While for most practical engineering problems, numerical method is needed. To this end, we start with Eq. (14) instead of Eq. (15), because from the standpoint of numerical solution usually an equation in the form of an integral may have some advantages over an equation in the form of differentiation.

According to the discussions in the preceding section, Eqs. (23), (27) or (28) could be adopted as the governing equation for numerical solution. Eq. (23) is an exact equation equivalent to the original equations (14) and (15). In the present stage, numerical algorithms for Eq. (27) were extensively studied and will be outlined here.

It is seen that Eq. (27) is a linear partial differential equation. To obtain the solution the coefficients should be determined first, while these coefficients are time rates of the physical quantity of interest as $\{\boldsymbol{\Theta} = \boldsymbol{\theta}\}$ and thus can be obtained through solving Eq. (9). Therefore, the GDEE can be solved in the following steps:

Step 1: Select representative points (RPs for short) in the probability-assigned space and determine their assigned probability. Select a set of representative points in the distribution domain $\Omega_{\boldsymbol{\Theta}}$. Denote them by $\boldsymbol{\theta}_q = (\theta_{q,1},\theta_{q,2},\cdots,\theta_{q,s})$; $q = 1,2,\cdots,n_{\text{pt}}$ where n_{pt} is the number of the selected points. Simultaneously, determine the assigned probability of each point according to Eq. (22) using the Voronoi cells (Chen et al, 2009).

Step 2: Solve deterministic dynamical systems. For the specified $\boldsymbol{\Theta} = \boldsymbol{\theta}_q$, $q = 1,2,\cdots,n_{\text{pt}}$, solve the physical equation (Eq. (9)) to obtain time rate (velocity) of the physical quantities $\dot{Z}(\boldsymbol{\theta}_q,t)$.

Through Steps 1 and 2, the ensemble evolution is replaced by point evolution as representatives.

Step 3: Solve the GDEE (Eq. (27)) under the initial condition, as a discretized version of Eq. (18),

$$p_q(z,t)\big|_{t=t_0} = \delta(z-z_0)P_q \tag{29}$$

by the finite difference method with TVD scheme to acquire the numerical solution of $p_q(z,t)$.

Step 4: Sum up all the results to obtain the probability density function of $Z(t)$ via the Eq. (24).

It is seen clearly that the solving process of the GDEE is to incorporate a series of deterministic analysis (point evolution) and numerical solving of partial differential equations, which is just the essential of the basic thought that the physical mechanism of probability density evolution is the evolution of the physical system.

3 Performance Evolution of Controlled Systems

Extensive studies have been done on the structural optimal control, which serves as one of the most effective measures to mitigate damage and loss of structures induced by disastrous actions such as earthquake ground motions and strong winds (Housner et al, 1997). However, the randomness inherent in the dynamics of the system or its operational environment and coupled with the nonlinearity of structural behaviors should be taken into account so as to gain a precise control of structures. The reliability of structures, otherwise, associated with structural performance still cannot be guaranteed even if the responses are greatly reduced compared to the uncontrolled counterparts. Thus, the methods of stochastic optimal control have usually been relied upon to provide a rational mathematical context for analyzing and describing the problem.

Actually, pioneering investigations of stochastic optimal control by mathematician were dated back to semi-century ago and resulted in fruitful theorems and approaches (Yong & Zhou, 1999). These advances mainly hinge on the models of Itô stochastic differential equations (e.g. LQG control). They limit themselves in application to white noise or filtered white noise that is quite different from practical engineering excitations. The seismic ground motion, for example, exhibits strongly non-stationary and non-Gaussian properties. In addition, stochastic optimal control of multi-dimensional nonlinear systems is still a challenging problem in open. It is clear that the above two challenges both stem from the classical framework of stochastic dynamics. Therefore, a revolutionary scheme through physical control methodology based on PDEM is developed in the last few years (Li & Peng, 2007; Li et al, 2008; 2010; 2011a).

Consider the multi-degree-of-freedom (MDOF) system represented by Eq. (9) is exerted a control action, of which the equation of motion is given by

$$\mathbf{M}(\boldsymbol{\Theta})\ddot{\mathbf{X}} + \mathbf{C}(\boldsymbol{\Theta})\dot{\mathbf{X}} + \mathbf{f}(\boldsymbol{\Theta},\mathbf{X}) = \mathbf{B}_s\mathbf{U}(\boldsymbol{\Theta},t) + \mathbf{D}_s\mathbf{F}(\boldsymbol{\Theta},t) \tag{30}$$

where $\mathbf{U}(\boldsymbol{\Theta},t)$ is the control gain vector provided by the control action; $\mathbf{B}_s$ is a matrix denoting the location of controllers; $\mathbf{D}_s$ is a matrix denoting the location of excitations.

In the state space, Eq. (30) becomes

$$\dot{\mathbf{Z}}(t) = \mathbf{A}\mathbf{Z}(t) + \mathbf{B}\mathbf{U}(t) + \mathbf{D}\mathbf{F}(\boldsymbol{\Theta},t) \tag{31}$$

where $\mathbf{A}$ is a system matrix; $\mathbf{B}$ is a controllers location matrix, and $\mathbf{D}$ is a excitation location vector.

In most cases, Eq. (30) is a well-posed equation, and relationship between the state vector $\mathbf{Z}(t)$ and control gain $\mathbf{U}(t)$ can be determined uniquely. Clearly, it is a function of $\boldsymbol{\Theta}$ and might be assumed to take the form

$$\mathbf{Z}(t) = \mathbf{H}_{\mathbf{Z}}(\boldsymbol{\Theta},t) \tag{32}$$

$$\mathbf{U}(t) = \mathbf{H}_{\mathbf{U}}(\boldsymbol{\Theta},t) \tag{33}$$

It is seen that all the randomness involved in this system comes from $\boldsymbol{\Theta}$, thus, the augmented systems of components of state and control force vectors $(Z(t),\boldsymbol{\Theta})$, $(U(t),\boldsymbol{\Theta})$ are both probability preserved, and satisfy the GDEEs, respectively, as follows (Li et al, 2010)

$$\frac{\partial p_{Z\Theta}(z,\boldsymbol{\theta},t)}{\partial t}+\dot{Z}(\boldsymbol{\theta},t)\frac{\partial p_{Z\Theta}(z,\boldsymbol{\theta},t)}{\partial z}=0 \tag{34}$$

$$\frac{\partial p_{U\Theta}(u,\boldsymbol{\theta},t)}{\partial t}+\dot{U}(\boldsymbol{\theta},t)\frac{\partial p_{U\Theta}(u,\boldsymbol{\theta},t)}{\partial u}=0 \tag{35}$$

The corresponding instantaneous PDFs of $Z(t)$ and $U(t)$ can be obtained by solving the above partial differential equations with given initial conditions

$$p_Z(z,t)=\int_{\Omega_\Theta} p_{Z\Theta}(z,\boldsymbol{\theta},t)\mathrm{d}\boldsymbol{\theta} \tag{36}$$

$$p_U(u,t)=\int_{\Omega_\Theta} p_{U\Theta}(u,\boldsymbol{\theta},t)\mathrm{d}\boldsymbol{\theta} \tag{37}$$

where Ω_Θ is the distribution domain of $\boldsymbol{\Theta}$; the joint PDFs $p_{Z\Theta}(z,\boldsymbol{\theta},t)$ and $p_{U\Theta}(u,\boldsymbol{\theta},t)$ are the solutions of Eqs. (34) and (35), respectively.

As mentioned in the previous sections, the GDEEs reveal the intrinsic relationship between stochastic systems and deterministic systems via the random event description of the principle of preservation of probability. It is thus indicated, according to the relationship between point evolution and ensemble evolution, that the structural stochastic optimal control can be implemented through a collection of representative deterministic optimal controls and their synthesis on evolution of probability densities. Distinguished from the classical stochastic optimal control scheme, the control methodology based on the PDEM is termed as the physical scheme of structural stochastic optimal control.

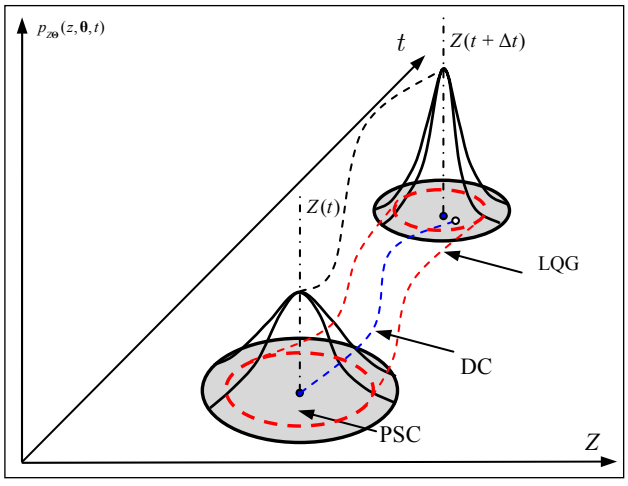


Figure 2. Performance evolution of optimal control systems: comparison of determinate control (DC), LQG control and physical stochastic optimal control (PSC).

Figure 2 shows the discrepancy among the deterministic control (DC), the LQG control and the physical stochastic optimal control (PSC) tracing the performance evolution of optimal control systems. One might realize that the performance trajectory of the deterministic control is point to point, and obviously it lacks of ability of governing the system performance due to the randomness of external excitations. The performance trajectory of the LQG control, meanwhile, is circle to circle. It is remarked here that the classical stochastic optimal control is essentially to govern the system statistics to the general stochastic dynamical systems since there still lacks of efficient methods to solve the response process of the stochastic systems with strong nonlinearities in the context of classical random mechanics. The LQG control, therefore, just holds the system performance in mean-square sense, and cannot reach its high-order statistics. The performance trajectory of the PSC control, however, is domain to domain, which can achieve the accurate control of the system performance since the system quantities of interest all admit the GDEEs, Eqs. (34) and (35).

4 Probabilistic Criteria of Structural Stochastic optimal control

The structural stochastic optimal control involves maximizing or minimizing the specified cost function, whose generalized form is typically the quadratic combination of displacement, velocity, acceleration and control force. A standard quadratic cost function is given by the following expression (Soong, 1990)

$$J_1(\mathbf{Z},\mathbf{U},\boldsymbol{\Theta},t)=\frac{1}{2}\mathbf{Z}^{\mathrm{T}}(t_f)\mathbf{P}(t_f)\mathbf{Z}(t_f)+\int_{t_0}^{t_f}[\mathbf{Z}^{\mathrm{T}}(t)\mathbf{Q}\mathbf{Z}(t)+\mathbf{U}^{\mathrm{T}}(t)\mathbf{R}\mathbf{U}(t)]dt \tag{38}$$

where $\mathbf{Q}$ is a positive semi-definite matrix, $\mathbf{R}$ is a positive definite matrix, and t_f is the terminal time, usually longer than that of the excitation. As should be noted, the cost function of the classical LQG control is defined as the ensemble-expected formula of Eq. (38) that is a deterministic function in dependence upon the time argument. Its minimization is to obtain the

minimum second-order statistics of the state as the given parameters of control policy and construct the corresponding control gain under Gaussian process assumptions. In many cases of practical interests, the probability distribution function of the state related to structural performance is unknown, and the control gain essentially relies on second-order statistics. While the cost function represented by Eq. (38) is a stochastic process, of which minimization is to make the representative solution of the system state globally optimized in case of the given parameters of control policy. This treatment would result in a minimum second-order statistics or the optimum shape of the PDF of system quantities of interests. It is thus practicable to construct a control gain relevant to a predetermined performance of engineering structures since the procedure developed in this paper adapts to the optimal control of general stochastic systems. In brief, the procedure involves two step optimizations; see Figure 3. In the first step, for each realization $\boldsymbol{\theta}_q$ of the stochastic parameter $\boldsymbol{\Theta}$, the minimization of the cost function Eq. (38) is carried out to build up a functional mapping from the set of parameters of control policy to the set of control gains. In the second step, the specified parameters of control policy to be used are obtained by optimizing the control gain according to the objective structural performance.

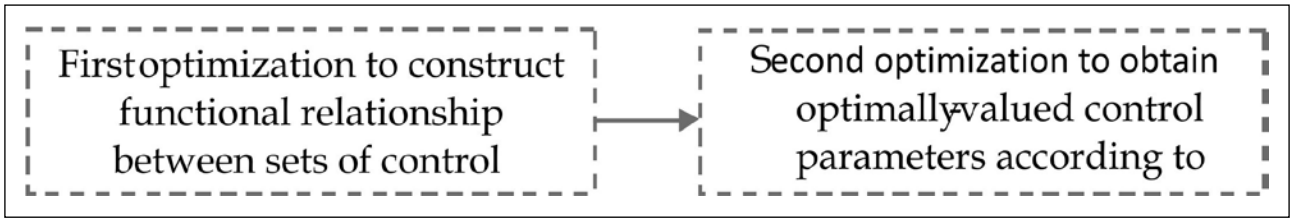


Figure 3. Two step optimizations included in the physical stochastic optimal control.

Therefore, viewed from representative realizations, the minimum of J_1 results in a solution of the conditional extreme value of cost function. The functional mapping, for a closed-loop control system, from the set of control parameters to the set of control gains is yield by (Li et al, 2010)

$$\mathbf{U}(\boldsymbol{\Theta},t) = -\mathbf{R}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{Z}(\boldsymbol{\Theta},t) \tag{39}$$

where is the Riccati matrix function.

As indicated previously, the control effectiveness of stochastic optimal control relies on the specified control policy related to the objective performance of the structure. The critical procedure of designing control system actually is the determination of parameters of control policy, i.e. weighting matrices $\mathbf{Q}$ and $\mathbf{R}$ in Eq. (38). There were a couple of strategies regarding to the weighting matrix choice in the context of classical LQG control such as, system statistics assessment based on the mathematical expectation of the quantity of interest (Zhang & Xu, 2001), system robustness analysis in probabilistic optimal sense (Stengel et al, 1992), and comparison of weighting matrices in the context of Hamilton theoretical framework (Zhu et al, 2001). We are attempting to, nevertheless, develop a family of probabilistic criteria of weight matrices optimization in the context of the physical stochastic optimal control of structures.

System Second-Order Statistics Assessment (SSSA)

A probabilistic criterion of weight matrices optimization based on the system second-order statistics assessment, including constraint quantities and assessment quantities, is proposed as follows

$$\min\left(J_2\right) = \arg\min_{\mathbf{Q},\mathbf{R}}\{E[\tilde{Y}] \ or \ \sigma[\tilde{Y}] \Big| F[\tilde{X}] \le \tilde{X}_{con}\} \tag{40}$$

whered J_2 enotes a performance function; $\tilde{Y} = \max_t[\max_i \left|Y_i(\boldsymbol{\Theta},t)\right|]$ is the equivalent extreme-value vector of the quantities to be assessed; $\tilde{X} = \max_t[\max_i \left|X_i(\boldsymbol{\Theta},t)\right|]$ is the equivalent extreme-value vector of the quantities to be used as the constraint; $\tilde{X}_{con}$ is the threshold of the constraint; The hat '~' on symbols indicates the equivalent

extreme-value vector or equivalent extreme-value process (Li et al, 2007); $F[\cdot]$ is the characteristic value function indicating confidence level. The employment of the control criterion of Eq. (40) is to seek the optimal weighting matrices such that the mean or standard deviation of the assessment quantity $\tilde{Y}$ is minimized when the characteristic value of constraint quantity $\tilde{X}$ less than its threshold $\tilde{X}_{con}$.

Minimum of Exceedance Probability of Single System Quantity (MESS)

An exceedance probability criterion in the context of first-passage failure of single system quantity can be specified as follows

$$\min(J_2) = \arg\min_{\mathbf{Q},\mathbf{R}} \left\{ \Pr\left(\tilde{Y} - \tilde{Y}_{thd} > 0\right) + \left(H(\tilde{X}_{\max} - \tilde{X}_{con})\right) \right\} \tag{41}$$

where $\Pr(\cdot)$ operates its arguments with denotation of exceedance probability; equivalent extreme-value vector $\tilde{Y}$ is the objective system quantity; $H(\cdot)$ is the Heaviside step function. The physical meaning of this criterion is that the exceedance probability of the system quantity is minimized (Li et al, 2011a).

Minimum of Exceedance Probability of Multiple System Quantities (MEMS)

An exceedance probability criterion in the context of global failure of multiple system quantities is defined as follows

$$\min(J_2) = \arg\min_{\mathbf{Q},\mathbf{R}} \left\{ \frac{1}{2}[\Pr_{\tilde{Z}}^{\mathrm{T}}(\tilde{Z} - \tilde{Z}_{thd} > \mathbf{0}) \Pr_{\tilde{Z}}(\tilde{Z} - \tilde{Z}_{thd} > \mathbf{0}) + \Pr_{\tilde{U}}^{\mathrm{T}}(\tilde{U} - \tilde{U}_{thd} > \mathbf{0}) \Pr_{\tilde{U}}(\tilde{U} - \tilde{U}_{thd} > \mathbf{0})] + \left(H(\tilde{X}_{\max} - \tilde{X}_{con})\right) \right\} \tag{42}$$

where equivalent extreme-value vectors of state and control force $\tilde{Z}, \tilde{U}$ are the objective system quantities. It is indicated that this control criterion characterizes system safety (indicated in the controlled inter-story drift), system serviceability (indicated in the controlled inter-story velocity), system comfortability (indicated in the constrained storey acceleration), controller workability (indicated in the limit control force) and their trade-off.

5 Comparative Studies

A base-excited single-storey structure with an active tendon control system (see Figure 4) is considered as a case for comparative studies of the control policies deduced from the above probabilistic criteria, and the developed control methodology against the classical LQG. The properties of the system are as follows: the mass of the storey is m=1×10^5 kg; The natural circular frequency of the uncontrolled structural system is ω_0=11.22 rad/sec; the control force of the actuator is denoted by $f(t)$, representing the inclination angle of the tendon with respect to the base, and the acting force $u(t)$ on the structure is simulated; The damping ratio is assumed to be 0.05. A stochastic earthquake ground motion model is used in this case (Li and Ai, 2006), and the mean-valued time history of ground acceleration with peak 0.11 g is shown in Figure 5.

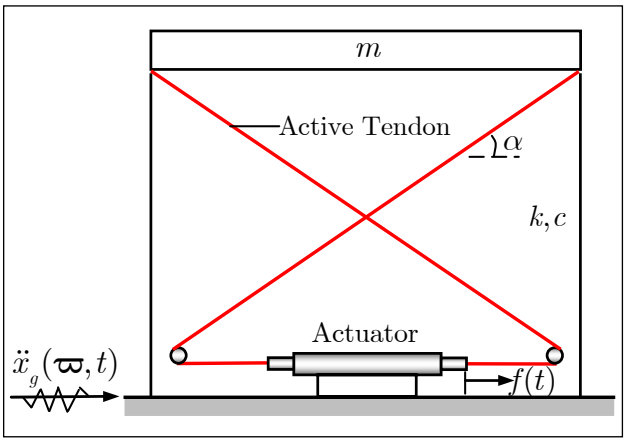


Figure 4. Base-excited single-story structure with active tendon control system.

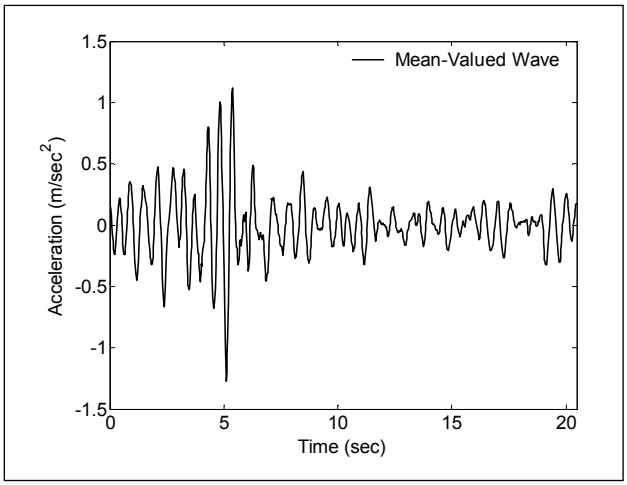


Figure 5. Mean-valued time history of ground motion.

The objective of stochastic optimal control is to limit the inter-story drift such that the system locates the reliability state, to limit the inter-story velocity such that the system provides the desired serviceability, to limit the storey acceleration such that the system provides the desired comfortability, and to limit the control force such that the controller sustains its workability. The thresholds/constraint values of the inter-story drift, of the inter-story velocity, of the storey acceleration, and of the control force are 10 mm, 100 mm/sec, 3000 mm/sec^2 and 200 kN, respectively.

Advantages in Global Reliability Based Probabilistic Criterion

For the control criterion of system second-order statistics assessment (SSSA), the inter-story drift is set as the constraint, and the assessment quantities include the inter-story drift, the storey acceleration, and the control force. The characteristic value function is defined as mean plus three times of standard deviation of equivalent extreme-value variables. For the control criterion of minimum of exceedance probability of single system quantity (MESS), the inter-story drift is set as the objective system quantity, and the constraint quantities include the storey acceleration and the control force. For the control criterion of minimum of exceedance probability of multiple system quantities (MEMS), the inter-story drift, inter-story velocity and control force are set as the objective system quantities, while the constraint quantity is the storey acceleration.

Table 1 Comparison of control policies

Ext. Values	SSSA: Q=diag {80,80}, R=10^{-12}	MESS: Q=diag{101.0,195.4}, R=10^{-10}	MEMS: Q=diag{1073.6,505.0}, R=10^{-10}
Dis-Mn (mm)	28.47[a]	28.47	28.47
	1.16[b]	6.23	4.15
	95.93%[c]	78.12	85.42%
Dis-Std (mm)	13.78	13.78	13.78
	0.20	1.41	0.81
	98.55%	89.77%	94.12%
Acc-Mn (mm/sec^2)	3602.66	3602.66	3602.66
	1069.79	1235.60	1141.00
	70.31%	65.70%	68.33%
Acc-Std (mm/sec^2)	1745.59	1745.59	1745.59
	360.81	348.92	331.78
	79.33%	80.01%	80.99%
CF-Mn (kN)	105.56	86.55	94.93
CF-Std (kN)	35.59	30.71	32.46

a indicates the uncontrolled system quantities.

b indicates the controlled system quantities.

c indicates the control efficiency defined as (a-b)/a.

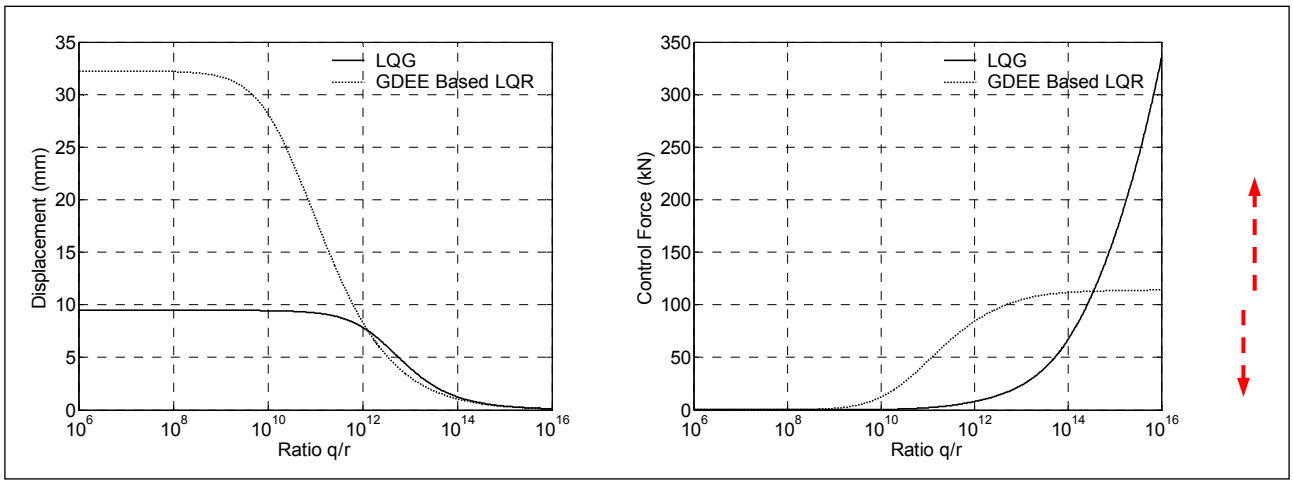
The comparison between the three control policies is investigated. The numerical results are listed in Table 1. It is seen that the effectiveness of response control hinges on the physical meanings of the optimal control criteria. As indicated in this case, the control criterion SSSA exhibits the larger control force due to the inter-storey drift being only considered as the constraint quantity, which thus has lower inter-story drift. The control criterion MESS, however, exhibits the smaller control force due to the storey acceleration and control force being simultaneously considered

as the constraint quantities that result in a less reduction on the inter-storey drift. The control criterion MEMS, as seen from Table 1, achieves the best trade-off between control effectiveness and economy in that the objective system quantities includes the inter-storey drift, together with inter-storey velocity and control force. It thus has reason to believe that the multi-objective criterion in the global reliability sense is the primary criterion of structural performance controls.

Control Gains against the classical LQG

It is noted that the classical stochastic optimal control strategies also could be applied to a class of stochastic dynamical systems, and synthesize the moments or the PDFs of the controlled quantities. The class of systems is typically driven by independent additive Gaussian white noise, and usually modelled as the Itô stochastic differential equations. The response processes, meanwhile, exhibit Markov property, of which the transition probabilities are governed by the Foker-Planck-Kolmogorov equation (FPK equation). It remains an open challenge in the civil engineering system driven by non-Gaussian noise. The proposed physical stochastic optimal control methodology, however, occupies the validity and applicability to the civil engineering system. As a comparative study, Figure 6 shows the discrepancy of root-mean-square quantity vs. weight ratio, using the control criterion of SSSA, between the advocated method and the LQG control.

One could see that the LQG control would underestimate the desired control force when the coefficient ratio of weighting matrices locates at the lower value, and it would overestimate the desired control force when the coefficient ratio of weighting matrices locates at the higher value. It is thus remarked that the LQG control using the nominal Gaussian white noise as the input cannot design the rational control system for civil engineering structures.



(a) Equivalent-extreme relative displacement. (b) Equivalent-extreme control force.
Figure 6. Comparison of root-mean-square equivalent-extreme quantity vs. weight ratio between PSC and LQG.

6 Numerical Example

An eight-storey single-span shear frame fully controlled by active tendons is taken as a numerical example, of which the properties of the uncontrolled structure are identified according to Yang et al (1987). The floor mass of each storey unit is m = 3.456×10^5 kg; the elastic stiffness of each storey is k = 3.404×10^2 kN/mm; the internal damping coefficient of each storey unit c = 2.937 kN×sec/mm, which corresponds to a 2% damping ratio for the first vibrational mode of the entire building. The external damping is assumed to be zero. The computed natural frequencies are 5.79, 17.18, 27.98, 37.82, 46.38, 53.36, 58.53 and 61.69 rad/sec, respectively. The earthquake ground motion model is the same as that of the preceding SDOF system, and the peak acceleration is 0.30 g. The control criterion MEMS is employed, and the thresholds/constraint values of the structural

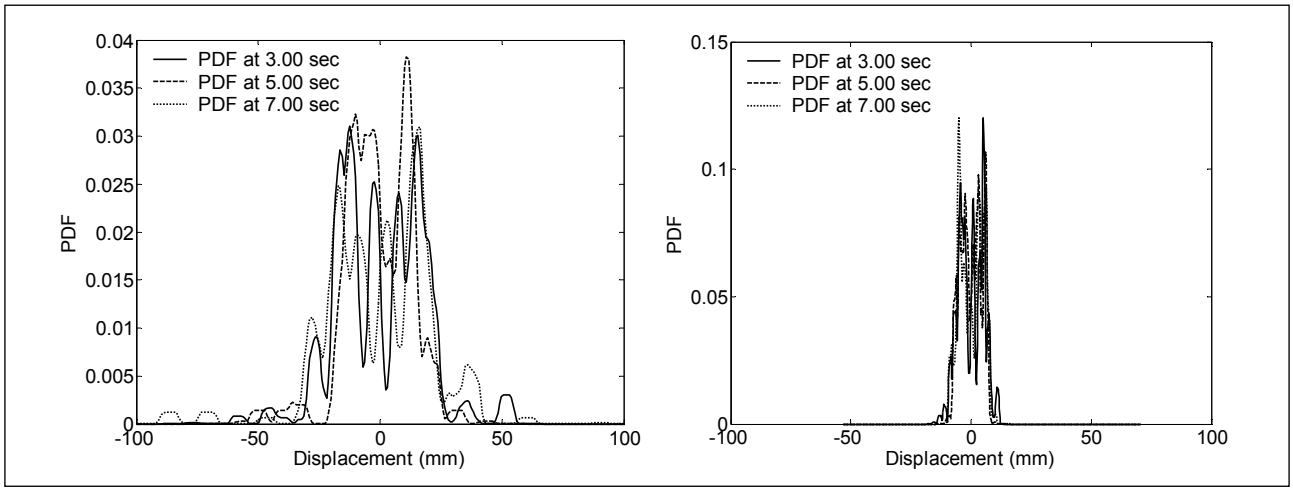
inter-story drifts, inter-story velocities, storey acceleration, and the control forces are 15 mm, 150 mm/sec and 2000 kN and 8000 mm/sec^2, respectively. For simplicity, the form of the weighting matrices in this case takes

$$\mathbf{Q} = diag\{Q_d,...,Q_d;Q_v,...,Q_v\},\ \ \mathbf{R} = diag\{R_u,...,R_u\} \tag{43}$$

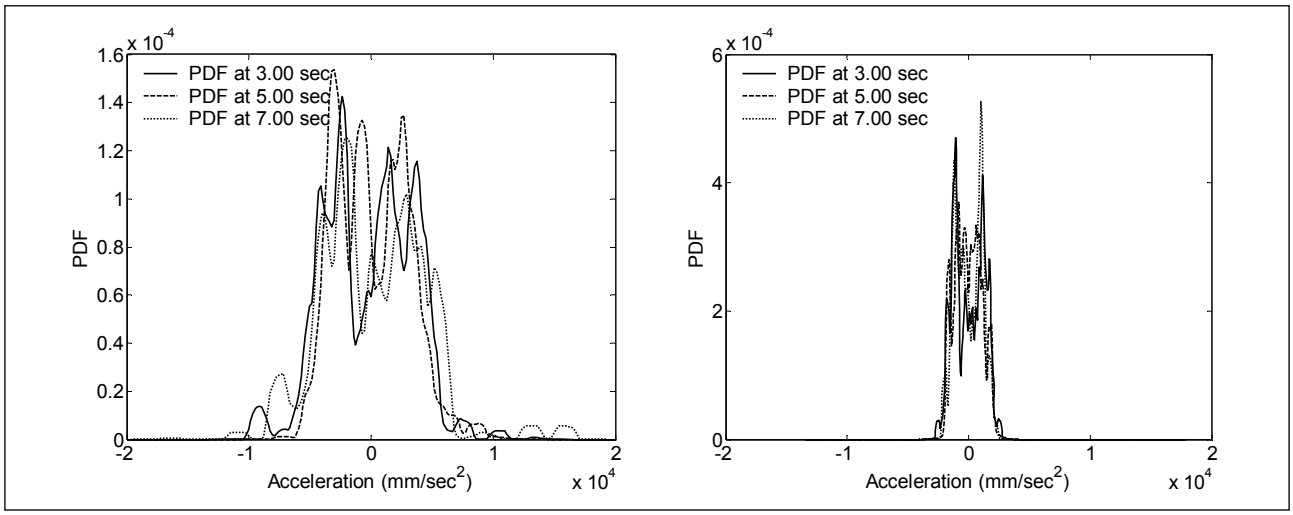
The optimization results of the numerical example are shown in Table 2. It is seen that the exceedance probability of system quantities, rather than the ratio of reduction of responses, is provided when the objective value of performance function reaches to the minimum, indicating an accurate control of structural performance implemented. The optimization results also show that the stochastic optimal control achieves a best trade-off between effectiveness and economy.

Table 2 Optimization results of example

Parameters	Q_d	Q_v	R_u
Initial value	100	100	10^{-12}
Optimal value	102.8	163.7	10^{-12}
Objective value	11.22×10^{-6} ($P_{f,d}$ =0.0023, $P_{f,v}$=0.0035, $P_{f,u}$=0.0022)		



(a) Without control. (b) With control.
Figure 7. Typical PDFs of inter-0-1 drift at typical instants of time.



(a) Without control. (b) With control.
Figure 8. Typical PDFs of the 8th storey acceleration at typical instants of time.

Figure 7 shows typical PDFs of the inter-0-1-story drift of the controlled/uncontrolled structures at typical instants of time. One can see that the variation of the inter-story drift is obviously reduced. Likewise, the PDFs of the 8th storey acceleration at typical instants show a reduction of system response since that distribution of the storey acceleration has been narrowed (see Figure 8). It is indicated that the seismic performance of the structure is improved significantly in case that the stochastic optimal control employing the exceedance probability criterion is applied.

7 Concluding Remarks

In this paper, the fundamental theory of the generalized density evolution equation is firstly revisited. Then a physical scheme of structural stochastic optimal control based on the probability density evolution method is presented for the stochastic optimal controls of engineering structures excited by general non-stationary and non-Gaussian processes. It extends the classical stochastic optimal control approaches, such as the LQG control, of which the random dynamic excitations are exclusively assumed as independent white noises or filter white noises. A family of optimal control criteria for designing the controller parameter, including the criterion based on mean and standard deviation of responses, the criterion based on exceeding probability, and the criterion based on global reliability of systems, is elaborated by investigating the stochastic optimal control of a base-excited single-storey structure with an active tendon control system. It is indicated that the control effect relies upon the probabilistic criteria of which the control criterion in global reliability operates efficiently and gains the desirable structural performance. The proposed stochastic optimal control scheme, meanwhile, of structures exhibits significant benefits over the classical LQG control. An eight-storey shear frame controlled by active tendons is further investigated employing the control criterion in global reliability of the system quantities. It is revealed in the numerical example that the seismic performance of the structure is improved significantly, indicating the validity and applicability of the developed PDEM-based stochastic optimal control methodology for the accurate control of structural performance.

Acknowledgements

The supports of the National Natural Science Foundation of China (Grant Nos. 50621062, 51108344) and the Exploratory Program of State Key Laboratory of Disaster Reduction in Civil Engineering at Tongji University (Grant No. SLDRCE11-B-04) are highly appreciated.

References

1. Chen JB, Ghanem R, Li J, 2009. Partition of the probability-assigned space in probability density evolution analysis of nonlinear stochastic structures. *Probabilistic Engineering Mechanics*, **24**(1): 27-42.
2. Dostupov BG, Pugachev VS, 1957. The equation for the integral of a system of ordinary differential equations containing random parameters. *Automatikai Telemekhanika*, **18**: 620-630.
3. Einstein A, 1905. Über Die Von Der Molecular-Kinetischen Theorie Der Wärme Geforderte Bewegung Von in Rhuenden Flüssigkeiten Sus-Pendierten Teilchen. *Ann. Phys*. (Leipzig), **17**: 549-560.
4. Ghanem RG, Spanos PD, 1991. *Stochastic Finite Elements: A Spectral Approach*. Berlin: Springer-Verlag.
5. Halder A, Mahadevar S, 2000. *Reliability Assessment using Stochastic Finite Element Analysis*. John Wiley & Sons.
6. Housner GW, Bergman LA, Caughey TK, et al, 1997. Structural control: past, present, and future. *Journal of Engineering Mechanics*, **123**(9): 897-971.
7. Itô K, 1942. Differential equations determining a Markoff process. *Zenkoku Sizyo Sugaku Danwakasi*, 1077.
8. Kleiber M, Hien TD, 1992. *The Stochastic Finite Element Method*. Wiley Chichester.
9. Kolmogorov A, 1931. über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Mathematische Annalen*, **104**(1): 415-458,
10. Langevin P, 1908. Sur La Theorie Du Mouvement Brownien. *C. R. Acad. Sci., Paris*, 530-532.
11. Li J, 1996. *Stochastic Structural Systems: Analysis and Modeling*. Beijing, Science Press (in Chinese).
12. Li J, 2006. A physical approach to stochastic dynamical systems. *Science Paper Online*, **1**(2): 93-104 (in Chinese).
13. Li J, 2008. Physical stochastic models for the dynamic excitations of engineering structures. *Advances in Theory and Applications of Random Vibration*, 119-132 (in Chinese).
14. Li J, Ai XQ, 2006. Study on random model of earthquake ground motion based on physical process. *Earthquake Engineering and Engineering Vibration*, **26**(5): 21-26. (in Chinese)
15. Li J, Chen JB, 2003. Probability density evolution method for dynamic response analysis of stochastic structures. *Proceeding of the Fifth International Conference on Stochastic Structural Dynamics*, 309-316, Hangzhou, China.
16. Li J, Chen JB, 2006a. Generalized density evolution equations for stochastic dynamical systems. *Progress in Natural Science*, **16**(6): 712-719.

17. Li J, Chen JB, 2006b. The probability density evolution method for dynamic response analysis of non-linear stochastic structures. *International Journal for Numerical Methods in Engineering*, **65**: 882-903.
18. Li J, Chen JB, 2008. The principle of preservation of probability and the generalized density evolution equation. *Structural Safety*, **30**: 65-77.
19. Li J, Chen JB, 2009. *Stochastic Dynamics of Structures*. John Wiley & Sons.
20. Li J, Chen JB, Fan WL, 2007. The equivalent extreme-value event and evaluation of the structural system reliability. *Structural Safety*, **29**(2): 112-131.
21. Li J, Liu ZJ, 2006. Expansion method of stochastic processes based on normalized orthogonal bases. *Journal of Tongji University (Natural Science)*, **34**(10): 1279-1283.
22. Li J, Peng YB, 2007. Stochastic optimal control of earthquake-excited linear systems. *Proceedings of 8th Pacific Conference on Earthquake Engineering*, Dec. 5-7, Singapore.
23. Li J, Peng YB, Chen JB, 2008. GDEE-based stochastic control strategy of MR damping systems. *Proceedings of 10th International Symposium on Structural Engineering for Young Experts*, Changsha, China, pp. 1207-1212.
24. Li J, Peng YB, Chen JB, 2010. A physical approach to structural stochastic optimal controls. *Probabilistic Engineering Mechanics*, **25**(1): 127-141.
25. Li J, Peng YB, Chen JB, 2011a. Probabilistic criteria of structural stochastic optimal controls. *Probabilistic Engineering Mechanics*, **26**(2): 240-253.
26. Li J, Yan Q, Chen JB, 2011b. Stochastic modeling of engineering dynamic excitations for stochastic dynamics of structures. *Probabilistic Engineering Mechanics*, in press (doi: 10.1016/j.probengmech.2011.05.004).
27. Lin YK, Cai GQ, 1995. *Probabilistic Structural Dynamics: Advanced Theory and Applications*. New York, McGraw-Hill.
28. Lutes LD, Sarkani S, 2004. *Random vibrations: Analysis of structural and mechanical systems*. Butterworth-Heinemann.
29. Schenk CA, Schuëller GI, 2005. *Uncertainty Assessment of Large Finite Element Systems*. Springer, Berlin.
30. Shinozuka M, Deodatis G, 1991. Simulation of stochastic processes by spectral representation. *Applied Mechanics Review*, **44**(4): 191-204.
31. Shinozuka M, Jan CM, 1972. Digital simulation of random processes and its applications. *Journal of Sound and Vibration*, **25**: 111-128.
32. Soong TT, 1990. *Active Structural Control: Theory and Practice*. Longman Scientific &Technical, New York.
33. Stengel RF, Ray LR, Marrison CI, 1992. Probabilistic evaluation of control system robustness. *IMA Workshop on Control Systems Design for Advanced Engineering Systems: Complexity, Uncertainty, Information and Organization*, Oct. 12-16, Minneapolis, MN.
34. Syski R, 1967. Stochastic differential equations. In: Saaty TL (ed), *Modern Nonlinear Equations*, Chapter 8. McGraw-Hill, New York.
35. Wiener N, 1923. Differential space. *Journal of Mathematical Physics*, **2**(13): 131-174.
36. Yang JN, Akbarpour A, Ghaemmaghami P, 1987. New optimal control algorithms for structural control. *Journal of Engineering Mechanics*, **113**(9): 1369-1386.
37. Yong JM, Zhou XY, 1999. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, New York.
38. Zhang WS, Xu YL, 2001. Closed form solution for along-wind response of actively controlled tall buildings with LQG controllers. *Journal of Wind Engineering and Industrial Aerodynamics*, **89**: 785-807.
39. Zhu WQ, 1992. *Random Vibration. Beijing*. Science Press (in Chinese).
40. Zhu WQ, Ying ZG, Soong TT, 2001. An optimal nonlinear feedback control strategy for randomly excited structural systems. *Nonlinear Dynamics*, **24**: 31-51.

Life Cycle Reliability and Safety Engineering
Vol. 2 Issue 2 (2012) 16-26

Susceptibility of Solid State Logic Circuit To High Power Microwave Shots

Authors: A. R. Ramakrishnan[1], R. Kumaran[2], Amitava Roy[3], Rakhee Menon[3], S. Mitra[3], Ankur Patel[3], Vishnu Sharma[3], Archana Sharma[3], D. P. Chakravarthy[3] and P.V. Varde[2]
varde@barc.gov.in,

Address: 1).
2). *RRSD, Bhabha Atomic Research Centre, Trombay, Mumbai 400 085, India.*
3). Accelerator *and Pulse Power Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400 085, India.*

Abstract

Intense electro-magnetic pulse when directed on electronic and electrical systems can cause malfunction and failure in its components. The feasibility and possibility of high power microwave electromagnetic strikes on vital installations has increased due to tremendous progress in microwave generation technology over the past few decades .The vulnerability of electronic and electrical equipment to HPM threats must therefore be taken into consideration in design of safety-critical systems. This becomes even more relevant with increasing dependence of systems on electronics and growing miniaturization in solid state technologies. The effects of HPM on a system depend on both HPM environment parameters and system parameters and topology. This paper presents a study of the effect of HPM on protection systems of nuclear power plants. Safety critical electronic circuits of the nuclear plant protection system were identified and their susceptibility to failure against radiated high power microwave environments was experimentally assessed using a Vircator and Marx Generator based microwave generation system.

Introduction

In the recent past various high-power electromagnetic (HPEM) environments that can adversely affect the operation of electrical systems have been developed [1]. Such a concept could be used against civil systems in what is known as an Intentional Electromagnetic Interference (IEMI) [2] & [3]. Typically, this HPEM energy arrives at the system in the form of an incident electromagnetic field. It can occur in the form of a pulsed waveform of microwave energy referred to as high power microwave (HPM) pulse or in the form of a broadband pulse of EM energy, referred to as an ultra wideband (UWB) pulse. The frequency of high power microwave (HPM) ranges from tens of MHz up to several GHz and the power of radiated HPM sources ranges from kilowatts up to several GHz [4].

Intentional electromagnetic environments can in general be radiated or conducted; a single pulse or a burst of many repetitions and it is now well established that radiated HPEM generators producing sufficiently intense electromagnetic signals in the frequency range of 200 MHz to 5 GHz (operating wavelength 1.5 m to 6 cm) can cause electronic damage in many systems because typical apertures, slots, holes, rivet spacings and hatch openings provide inadvertent coupling paths in this frequency range and because system antennas operating in this frequency range could provide coupling path into the system[1]. Conducted HPM effects may be limited by cable and wire transfer functions which limit the propagation of high frequency.

Even as technologies for suitable HPEM generation are evolving, a number of susceptibility tests have been conducted for systems and

components like semiconductor NAND & Inverter TTL and CMOS technologies [5] &[6] ,micro-processors[7]&[8], micro-controllers [9] &[10] , MOSFET devices [11], automobile circuits [12], computer networks [13], Analog systems like op amps[14], network control and measurement system[15], alarm systems [16] and PCs [17] over the last four five years. However system level response to HPEM among other factors is also based system layout, parameters, topology and components [18], [19] and [20].Hence the HPM radiation response of protection systems in general and nuclear plant protection system in specific would differ from other systems. In this study safety critical circuits of nuclear plant protection systems have been identified, modeled and subjected to HPM radiated environment for studying the effects. Analysis of the literature indicate that broadly there are two methods being used for susceptibility tests - HPEM injection methods as used in [14],[15] &[21] and HPEM irradiation method as used in [8] ,[13],[17] &[22]. Experimental measurements in the latter pose more difficulty but it is more realistic and suited for radiated threat tests and hence has been adopted in this project. The Marx generator and virtual cathode oscillator (Vircator) based 1kJ HPM system at APPD,BARC has been used for generating HPM for experiments.

II. Identification Of Safety Critical Circuits Of Nuclear Plant Protection Systems

A. Elements of nuclear plant protection system

The protection system of nuclear power plant [23] are designed to automatically shut down the reactor in an efficient manner and maintain it in safe shut down state on demand in case of any deviation from normal operation, anticipated abnormal situations and accident conditions. Reactor trip signals are generated in the event of any malfunction in the system. The main sub-systems of a reactor protection system are instrumentation comprising of a network of basic sensors to monitor various neutronic and process parameters; Protection Logic Circuit (PLC) which processes the information received from instrumentation channels and generates the signal to actuate shut down devices whenever any neutronic or process parameter crosses its stipulated limits; Shut Down Devices (SDD) comprising of elements like fast acting shut-off rods and slower but reliable back-up moderator dumping system consisting of fast acting valves and an Alarm Annunciation System (AAS) which displays the health status of all systems/ subsystems and components.

The protection logic circuit is implemented using various types of electronic circuits for various classes of reactor trips including Absolute trips (ATs), comprising of those parameters which are essential to initiate immediate protective action, irrespective of reactor power, Auxiliary absolute trips (XTs), for certain vital safety parameters to initiate back-up protection action, Conditional trips(CTs) which are effective above 1% of rated power and Emergency trips (ETs) generated on very high log rate or very high linear Power. In order to achieve a high degree of reliability both from the point of view of reactor safety and also to reduce the possibility of spurious reactor trips, a 2/3 majority voting logic is usually built in the system. The trip signals are usually processed in three independent but identical channels. Fail-safe philosophy is generally adopted to the maximum extent possible. Thus situations leading to failure of power supplies, breakage of electrical connections etc. result in a reactor trip. As large numbers of components are used in the trip logic unit, it is necessary to detect unsafe failure of components immediately hence a Testing Unit is usually incorporated such that any unsafe failure within the solid state logic is immediately annunciated.

B. Reliability Of Protection System in HPM Environment

As per definition in [24] Design Basis Threat comprises of the attributes and characteristics of potential insider and/or external adversaries, who might attempt unauthorized removal of nuclear material or sabotage, against which a physical protection system is designed and evaluated.

In case of an HPEM risk , the plant protection system which is entirely based on electrical and electronic components, is most crucial, since it

must ensure the safety of the plant during and after the risk. In the worst case, any damage to the plant protection system should not result in an unsafe failure state.

Structural and metallic enclosures reduce HPEM vulnerability of the components enclosed. Hence, electronic components of the protection system like instrumentation sensors and SSD within the protection and structural enclosure of reactor building are better protected against radiated HPEM risk than the PLC and the AAS are housed outside it. But from the point of view of criticality, PLC is most critical for the protection system.

C. Review of PSA Study of a Reference Plant in HPM Context

Though the PSA study[25] is limited to safety against internal initiating events (IEs) and the aspect of treatment to external initiating events like HPM has not been considered in it, the results of Failure Mode and Mechanisms Effect Analysis (FMMEA) performed in the PSA for electronic circuits can be used to identify critical electronic components of the Protection Logic Circuit (PLC). The aim of this analysis carried as a part of the PSA is to generate a list of components and their failure modes and its overall effect on the system in terms of safe and unsafe failure. From the review of FMMEA it is clear that failure of some components of Protection Logic circuit of nuclear plant protection system may result in an unsafe situation. For example failure of some ICs in SPC and some ICs and transistors in LPCPC. During normal operation, the failure frequency of these critical ICs and transistors is seen to be very low. Even if such a failure occurs, it is even more highly unlikely that it occurs simultaneously in more than one channel. Also, any such failures are deducted by the Testing Unit which constantly and sequentially monitors the health of each sub unit of the protection logic circuit once every few minutes and failure of a component if any is indicated, which can then be replaced. However, since HPM is a common cause risk, it can not only effect components in all the three channels but also potentially simultaneously effect the Testing Unit and such a situation is not only undesirable but could also be unsafe.

III. Modelling of Protection Logic Circuit

The objective of the modeling was to put all the essential and safety critical elements of the protection logic circuit including its protection devices on one board for the purpose of experiments.

A. Overview of the PLC

The protection logic system is built around various types of functional circuits built on small boards. Of these three types of functional circuits are most important and for ease these will be referred to as SCC- Signal Conditioning Circuit, the SPC- Signal Product Circuit and the LPCPC- Low Power Condition Processing Circuit. The simplified block diagram for a single channel of protection logic is as shown below in Figure 1.

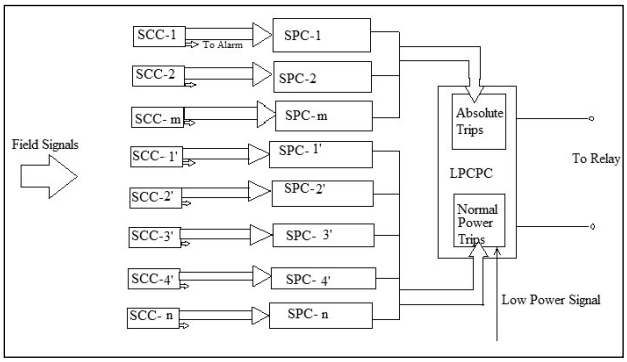


Figure 1 : Simplified schematic of Protection Logic Circuit

B. Working of the PLC

The PLC receives signals for reactor shut down in the form of potential free contacts when the output of the field instrumentation sensor exceeds the safe limit. The input contacts are fed to a set of SCCs. Each SCC can process a fixed number of signal contacts. For each input contact this circuit generates a current loop for the SPC and also sends a potential free contact to the AAS. The SPC processes all the loop signals from the corresponding SCC and generates a logic signal at its output.

(a) When all the input contacts to a SCC are closed representing a 'NO TRIP' condition, the output of the corresponding SPC is logic 'high' (logic '1').

(b) Whenever any of the input contacts opens representing a 'TRIP' condition, the output of

the corresponding SPC becomes logic 'low' (logic '0').

The output signals from one or more SPCs are fed to a Low Power Condition Processing Circuit (LPCPC) which provides a double ended drive to a 12V DC relay. The relay remains energised if all inputs to the LPCPC from the SPCs are logic '1' representing a 'NO TRIP' condition. However, if any of the SPC output becomes logic '0', the relay is de-energised and results in a trip leading to a reactor shutdown. In the LPCPC, the CT signals are made ineffective in low power condition. The output of the LPCPC for CTs will be '0' only when the trip occurs in condition of normal power, unlike the ATs which are not dependent on power condition.

C. Development of the IPLC

Since the protection logic circuit for all the channels have the same composition, the test circuit can be built around requirements of one channel alone. For each channel, the Protection Logic circuit processes a number of CT signals, AT signals, XT signals and ET signals making use of SCC, SPC and LPCPC. Hence the test circuit can be reduced to one that has one SCC, one SPC and one LPCPC. One set of SCC- SPC can process many field trip signals which is then fed to a LPCPC which drives the relay. The test circuit has been made without any loss of functionality, composition and logic, by putting all the circuit elements pertaining to one field trip signal along with their protection features from SCC, SPC and LPCPC including the 12 V relay on one board called the Integrated Protection Logic Circuit(IPLC),which represents the protection logic circuit of a nuclear plant protection system. For the field trip signal a 18 V battery operated trip signal has been used. The operating voltage of the circuit is 12 V (V_{DD}) supplied by another battery. The use of opto-isolator isolates the 18V supply fed to contacts and 12V logic circuit supply.

The IPLC was first modeled on a general purpose PCB using soldered wires, referred to as IPLC-1 and then modeled as a double sided industrial grade PCB with no jumpers or connecting wires, referred to as IPLC-2. The 18V and 12 V power supplies, switch connections for setting Normal Power/ Low Power and the Relay Reset switch connection were provided with a one metre long wire so that these switches and power supply batteries could be placed at any suitable location during the tests. Figures 2,3 and 4 show the IPLC-1,IPLC-2 and its setup respectively. The IPLC has been used as the circuit under test (CUT) in the experiments.

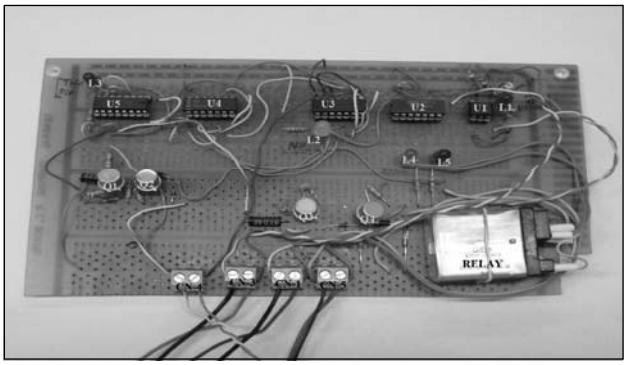


Figure-2 Integrated Protection Logic Circuit IPLC-1

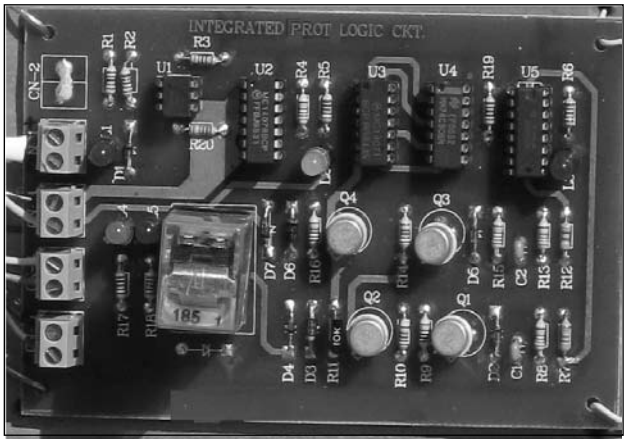


Figure 3 – Integrated Protection Logic Circuit IPLC-2

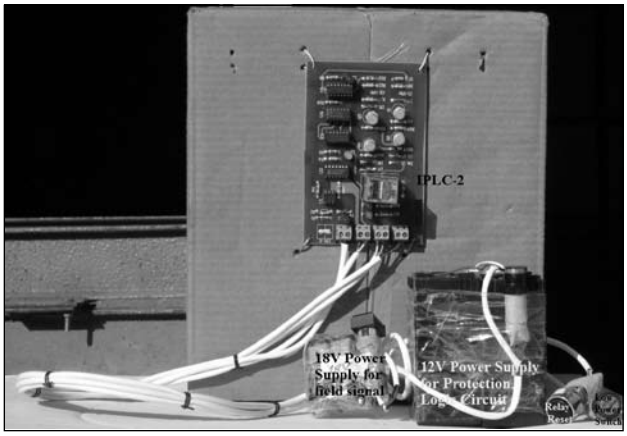


Figure 4- IPLC Set-up

IV. Design and Conduct of Experiments

A. Selection of testing Method

There are essentially therefore two broad testing strategies:-(a) HPEM injection into the CUT [6],[14] &[15] and (b) HPEM irradiation into the CUT [8],[10],[13],[22] &[26]. For this project HPM irradiation test method has been selected since radiated tests can physically and realistically demonstrate the effects of HPM on electronic components and circuits.

B. Identification of suitable HPM generation system

For HPM systems, two of the most important subsystems are the (a) Microwave generator/ radiation source and (b) High-voltage pulse generator. The high voltage generator commonly used is the HPM systems is based on a Marx generator. A conventional Marx generator comprises of a series-parallel connected capacitor-resistor bank. The capacitors are charged in parallel and discharged in series through spark gaps to boost the voltage. In order to reduce the size of the high-voltage pulse generators, explosive-driven systems using magnetic flux compression generators[27] are also being developed, they however have the limitation of being inherently single-shot. A number of microwave sources have been developed over the last few decades. The microwave sources can broadly be divided into the following two categories[28]:- (a) Impulsive sources like the various UWB sources and the LC Oscillator (b) Linear Beam sources like Magnetrons, Klystrons, Virtual Cathode Oscillators (Vircators) and Gyrotrons. Unlike HPM sources like Magnetrons and klystrons, Vircator source does not require any external magnetic fields (pulsed/permanent) for guiding electron beam propagation in vacuum. The efficiency of vircator source is typically limited to 5 to 15 % for the ratio of HPM power out to electrical power in, yet the vircator due to its relative simplicity overshadows alternative complex HPM sources, if all factors are considered[27]. Vircator has a combination of characteristics because of which it is recommended for high frequency use [29] and has been the most promising area of research among HPM sources[30].Feasibility of compact system using reflex triode vircator system and Marx generator is brought out in [31],[32],[33]. A vircator and Marx generator based HPM generator is therefore appropriate for the conduct of HPEM risk and susceptibility tests.

The 1kJ – Marx generator – reflex triode vircator based HPM generator system[34] of APPD,BARC was used for conducting the experiments. It is capable of producing a maximum output voltage of 300 kV into a matched load of 25 Ω with a pulse duration of 300 ns FWHM. A vacuum explosive electron emission (EEE) diode is used to generate an intense relativistic electron beam (IREB) and a vacuum level of the order of $< 1 \times 10^{-5}$ mbar is maintained in the vircator chamber by a diffusion pump backed by a rotary pump[35].The high-voltage pulse generated from the pulsed power system is applied to the anode. Electron emission occurs from the plasma that is formed on the cathode surface when a strong electric field E $\geq 10^7$ V/cm is applied to the AK gap.The diode consists of a planar cylindrical graphite cathode and a transparent SS wire-mesh anode. An open ended waveguide is used to radiate the output signal into the atmosphere. The 1kJ system is capable generating fields of 18kV/m field at 1.75m [35] . Higher field and power levels can be achieved at even closer distances. It is known that these field levels can temporarily or permanently damage electronic components like CMOS based ICs [5],[44] and [10].

C. Experimental Set Up

The schematic of the experimental set-up is drawn in Figure 5. The experimental set up broadly consists of four parts:-

(a) The 1 kJ HPM generator system facility at APPD, BARC.

(b) The Circuit Under Test ie.. the IPLC

(c) Measurement set up.

(d) Testing Set-up for CUT.

D. Measurement set-up

A resistive $CuSO_4$ voltage divider and a self-integrating Rogowski coil is used to measure

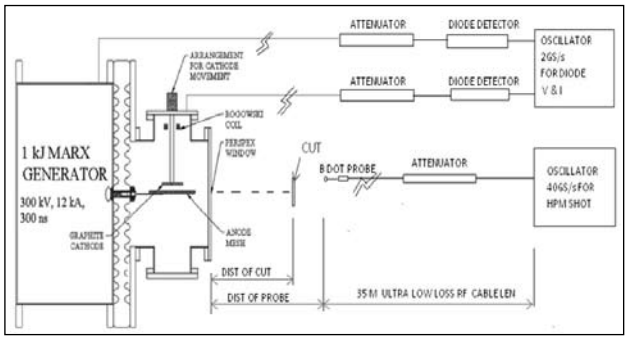


Figure 5 Schematic of the experimental set-up

the Vircator diode voltage and current pulses respectively [35]. HPM pulse for each shot is being detected, measured and recorded by using a B-Dot probe , a RF cable and a 6 GHz oscilloscope with sampling rate of 40 G samples per second. During experiments conducted in [36] the suitability of RF cable for measurements in the HPM environment has been established. Ultra Low loss flexible RF cable is being used in this measurement set-up. The B-Dot probe gives us the variation of the HPM signal in time domain. The peak amplitude of the signal in time domain $V(t)_{\max}$ is obtained from this. The Time domain to

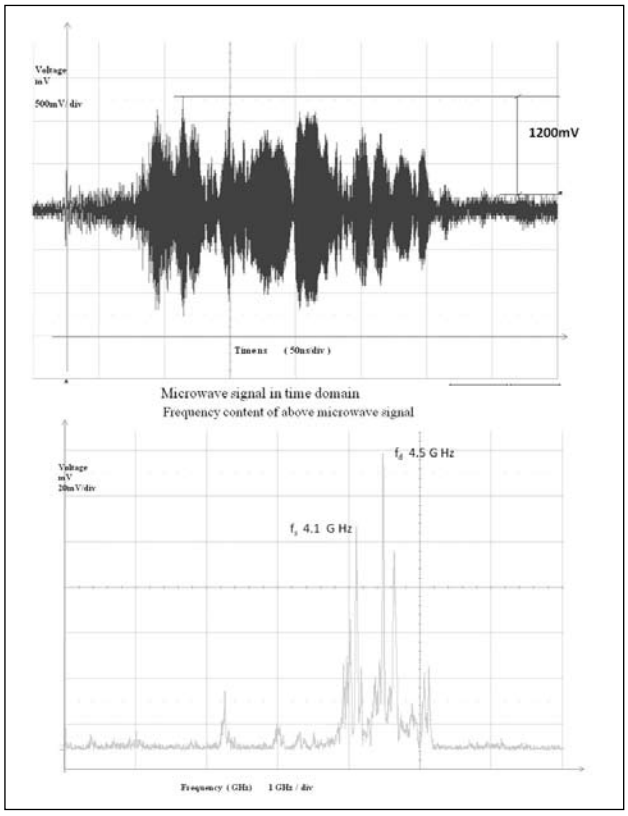


Figure 6 – Typical HPM shot : f_d 4.5GHz corresponds to 1200mV

Frequency domain transform function available in the oscilloscope is used to get the frequency composition of the HPM signal. From the values of $V(t)_{\max}$ and dominant frequency, f_d the magnetic field at the location of the probe is found [37]. The HPM pulse time and frequency characteristic of a typical shot is shown in Figure 6.

E. Development of Experiments

HEMP susceptibility level of a system is dependent on (a) Component susceptibility (b) Shielding and (c) Coupling to Cables [14]. In order to first analyse a worst case situation, the experiments have been conducted without shielding. However in order to emphasize the need for shielding, a few tests have been conducted with shielding.

The basic answer to the question "How does one know that the system will survive a given environment?" is to operate the system in the environment and see if it survives and that while full operation in the environment of concern is the ultimate test, one would like to have some test which if passed, would imply survival in real environment[38]. In our experiments the IPLC in regular operation has been subjected to HPM irradiation from the 1 kJ HPM generator system and the effect on the functionality of the IPLC after each shot has been checked through a testing procedure.

In order to determine the HPM effects on complex systems we need to answer the question "What does it mean for a system to survive a given environment?" [38]. In case of Nuclear Reactor Protection Systems, if HPM event results in an unscheduled trip then the failure would be a safe failure. However, if the failure doesn't result in a trip it would be an unsafe condition since in that case the reactor would continue to work and any trip generated from the field instrumentation will not result in a trip unless the trip circuit is diagnosed and repaired. Though this failure may eventually be identified by the testing unit of the system, yet if this type of failure occurs in more than one channel then it was designed for. Hence the occurrence of unsafe failure in the trip logic system would indicate its inability to survive the risk. Therefore the following is defined:-

(a) **Unsafe failure** Permanent failure of a component (say an IC) of the IPLC in HPM environment in such a way that the failure doesn't result in a trip.

(b) **Safe Failure** Permanent failure of a component (say an IC) of the IPLC in HPM environment in such a way that the failure result in a trip and reset of the trip relay doesn't restore the functionality of the circuit.

(c) **Malfunction** Temporary failure of a component (say an IC) of the IPLC in HPEM environment in such a way that the failure result in trip, but reset of the trip relay, restores the functionality of the IPLC.

E. Experiment Procedure

HPM pulses are characterized by the maximum electric field strength, their duration and their center frequency [13]. For risk level tests one need to measure failure levels as a function of frequency, amplitude, pulse width, direction of incidence, and polarization [38]. Single HPM shot of pulse duration of approx 300 ns have been used in the experiments. The dominant frequency in each shot has been assumed to characterize the HPM for purpose of experimental analysis. Since the direction of incident radiation will have an effect on the coupling, this has been fixed by aligning the axis of the source waveguide and the CUT. All the ICs on the IPLC were oriented with their longer side horizontal. All tests have been conducted for the operating condition of normal power. Since an open wave guide has been used, the polarization of the shots have not been fixed and hence there could be a shot to shot variation in the component of the electric and/ or the magnetic field along the path of the wave and that will not contribute to the coupled energy.

The HPEM susceptibility levels is usually specified in terms of electric field of the EM radiation at the location of the device under test that causes failure in it[5], [8]&[13]. For a given frequency, the HPM susceptibility level can be obtained as the electric field level at which failure/ malfunction of the circuit occurs. The electric field E at the CUT varies inversely with distance *r* of the CUT from the source, in the experiments we vary the field by bringing the CUT closer in steps to the source. While maintaining the same charging voltage and AK gap for a set of experiments we vary the electric field and power indirectly by varying distance of CUT from HPM source in steps of say 10 cm. Where ever effect was seen tests were repeated to check reproducibility of effects. Therefore, if the IPLC is kept at distances of less than 1 m from the 1kJ HPM source, effects are anticipated. It has been brought out in [38] that risk level tests with real environment can be done using lower power sources by bringing the test object closer to radiating source. The Magnetic Field value *B* obtained from the B-Dot probe has been used to find the electric field *E* at that point making use of the far field relationship $E = Bc$, where *c* is the speed of light. Power density at a distance *r* meters from the source, where the field is E_r is calculated using the relation $P_d = E_r^2/Z$ W/m^2. The far field area is related to the wavelength λ of the emitted radiation and is roughly taken as the distance beyond 2λ from the source.

G. Functionality Test

Testing of the functionality of the circuit is done during and after each shot and effects including failure type: - temporary/permanent, safe/unsafe if any is recorded. The operation worthiness of the IPLC was verified using a functionality test routine after each shot. In case of damage of ICs, the faulty ICs were identified and replaced, the circuit functionality was restored and rechecked before the circuit was subjected to next shot. The ICs were tested for functionality using the Linear cum Digital IC Testing system This equipment is already fed with programs for testing of all ICs used in the protection system.

The IPLC has been electrically configured in its operational mode for risk tests and the circuit with its LED indicators itself is then used to give information about its performance . The advantage of this self-diagnosing scheme[38] is that one does not need to measure signal at various points of the IPLC to determine the effect on it.

Two sets of experiments were conducted, set 1 using IPLC-1 and set 2 using IPLC-2 as the CUT. To obtain threshold levels for IPLC , a

frequency-electric field graph has been plotted for set-1 and set-2 (refer Figures 10 and 11). The Highest and smallest level susceptibilities (HL and SL) [8] have been determined. Malfunction, safe failure and unsafe failure was observed in CMOS based ICs of the IPLC. Shielding tests was also conducted with IPLC-2 first by using a mild steel enclosure 1.5mm thick without vents and then by using the same shields with vents size 3mm x 60mm facing the source. Experimental arrangements are shown in Figures 6,7 and 8. Results are summarized in Tables-1 to 3.

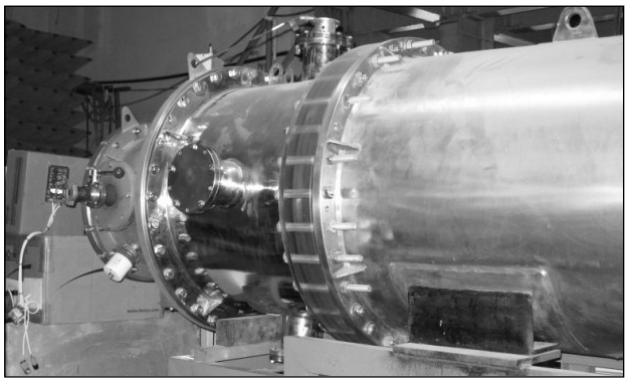


Figure 7 -Arrangement of IPLC in front of the HPM system

Figure 8 -Experimental Set Up for Shielding Test: mild steel enclosure 1.5mm thick

V. Analysis of Results

Based on analysis of 33 HPM shots, the susceptibility values so obtained for the soldered wired version of IPLC-1 are SL- 15.5kV/m and HL- 91.3kV/m, hence HPM Electric field threshold for failure of IPLC-1 in the experiments is 15.5kV/m. Based on analysis of 15 HPM shots on PCB based IPLC-2, the susceptibility values for IPLC-2 are SL- 31.8kV/m and HL- 62.4kV/m, hence HPM

Figure 9 Experimental Set-Up for tests with shielding enclosure having vents

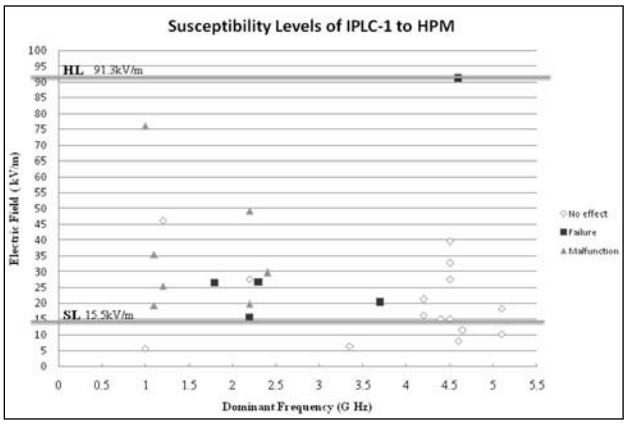


Figure 10 Plot for susceptibility levels of soldered wired IPLC -1 to HPM shots

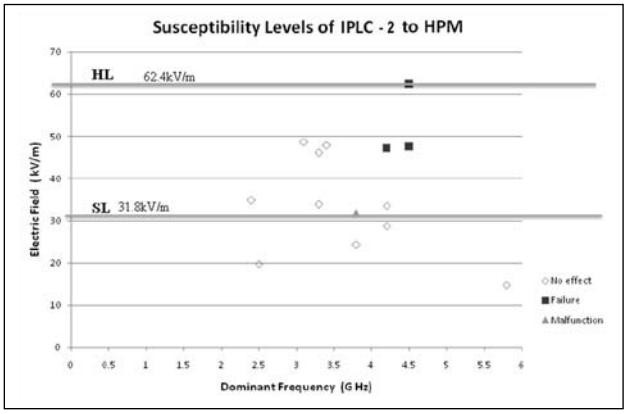


Figure 11 Plot for susceptibility levels of PCB IPLC -2 to HPM shots

Electric field threshold for failure of IPLC-2 in the experiments is 31.8 kV/m.

Both malfunction and failure observed with IPLC under shielding with slots.Four CMOS based ICs- 4071, 74C906, 4078 and 4050 used in the circuit were found to fail in the shots, however IC-4071 has been found to be most vulnerable. A

Table-1 Summary of Results of Threshold Tests

IPLC version	Total No of shots	Frequency Range (Dom) GHz	Electric Field Range kV/m	Mal-function or Failure Total cases	Malfunction => TRIP		Failure			
					Total cases	Thres-hold Electric Field (kV/m)	Total cases	Thres-hold Electric Field (kV/m)	Safe Cases=> TRIP	Unsafe cases => NO TRIP
1-Solder wired	33	1.0 to 5.1	5.7 to 91.3	12	7	19.5	5	15.5	4	1
2-PCB	15	2.5 to 5.8	14.7 to 62.4	5	1	31.8	4	47.1	3	1
Total	48	-	-	17	8	-	9	-	7	2

Table-2 Summary of Results of Shielding Tests

IPLC -2 PCB based	Total No of shots	Frequency Range (Dom) GHz	Electric Field Range kV/m	Mal-function or Failure Total cases	Malfunction cases	Failure	
						Safe cases	Unsafe cases
Shielding 1.5mm thk	6	3.3 to 4.6	5.7 to 91.3	0	0	0	0
Shielding with 3mmx6mm vents	5	3.3 to 4.2	14.7 to 62.4	4	3	0	1
Total	11	-	-	4	3	0	1

Table-3 Summary of Results – Distance and Power Density

IPLC	Maximum distance at which Malfunction occurred	Minimum power density at which Malfunction occurred	Maximum distance at which Failure occurred	Minimum power density at which failure occurred	Remarks
IPLC -1 Solder Wired	90cm	1 MW/m^2	70 cm *	0.6MW/ m^2 *	*Corresponds to Threshold Electric Field for malfunction/ failure
IPLC-2 PCB	60cm *	2.7 MW/ m^2 *	40cm	5.9MW/ m^2	

Table-4 Summary of Results : Failure of ICs

Failure type	No of cases	ICs Failure sets for Total no of HPM shots -59			
		4071 Alone	4071 & 74C906	4071, 74C906 & 4078	4071,74C906, 4078 & 4050
Safe Failures => Trip	7	3	2	1	1
Unsafe Failures => No trip	3	3	0	0	0
Total failures	10	6	2	1	1

total of 59 HPM shots taken in this project , out of which 10 shots resulted in failure of ICs. Details of ICs that failed during these shots is summarized below in Table -4.

Conclusion

In the context of mission critical military systems and safety critical nuclear power plant protection system, the need to incorporate HPEM hardness considerations in the design philosophy of their components , their critical logic circuits, their protection devices and their shielding requires to be emphasized since the failure of solid state components can lead to unsafe failure. For nuclear plants this becomes even more relevant since EMP is a common cause phenomenon and can potentially effect protection logic in all channels, testing unit and alarm annunciation systems simultaneously.

Circuits consisting of CMOS based solid state logic devices are vulnerable to radiated HPM environments. The vulnerability of soldered wired circuits is more than industrial grade PCB based circuits. The vulnerability can be eliminated by providing shielding enclosures around the circuits. To ensure effective shielding, the siting of vents and apertures in the shielding has to take into account HPEM risk concerns.

To begin with simple steps like avoiding of using jumper wires in sensitive equipment, provision of appropriate shielding enclosure at equipment and system levels, locating vents based on EM considerations and shielding of cables connected to these system could go a long way to reduce susceptibility to insignificant levels.

References

1. IEC Technical Report 61000-1-5 : *Electromagnetic Compatibility- HPEM Effects on Civil Systems,* First Edition, 2004-11.
2. D V Giri and F M Tesche : *Classification Of Intentional Electromagnetic interference* , IEEE EMC Transactions Special Issue Paper No B1.
3. *Special Issue on Intentional Electromagnetic Interference (IEMI),* IEEE Transactions on EMC, August 2004.
4. Leaflet 257 ,AECTP-250 : *Electrical and Electromagnetic Environmental conditions – High Power Microwave (HPM),* Edition 1 Allied Environmental Conditions and Tests Publications Feb 2009, NATO International Staff Defence Investment Division
5. Heyno Garbe and Michael Camp: *Susceptibility of different semiconductor technologies to EMP and UWB,* IEEE symposium on EMC, EMC conference 2002,19-23 Aug 2002, Vol 1, pg 87-92.
6. Joo-Hong, SunMook Hwang, Kwang Yong Kim, ChangSu Huh, UkYoulHuh and JinSooChol : *Susceptibility of TTL Logic Devices to Narrow-band High Power Electromagnetic Threats,* PIERS Online, Vol No 5.
7. Michael Camp, D Nitsch and Heyno Garbe : *UWB and EMP Susceptibility of Modern Electronics.*
8. Daniel Nitsch, Frank Sabath, Hans-Uirich Schmidt and Christian Braun: *Comparison Of HPM and UWB Susceptibility of Modern Microprocessor Boards* , Proceedings of EMC Symposium Zurich 2003 Pages 121-126.
9. F Sonnemann and J Bohl : *Susceptibility and Vulnerablility of Semiconductor Components and Circuits against HPM*, XXVIIth URSI General Assembly Mastrich 2002.
10. Sun-Mook Hwang, Joo-Il Hong, Seung-Moon Han,Chang-Su Huh, Uk-Youl Huh, and Jin-Soo Choi : *The Susceptibility of Microcontroller Device with Coupling Caused by UWB-HPEM,* Progress in Electromagnetics Research Symposium Proceedings, Moscow, Russia, August 18-21, 2009 , Pgs 756 to 760.
11. K Kim, A A Iliadis and V L Granatstein: *Effects Of Microwave Interference On The Operational Parameters Of N-Channel Enhancement Mode MOSFET devices in CMOS Integrated Circuits*, Solid-State Electronics Edition 48 (2004) 1795–1799.
12. Bernd Deutshmann,Etienne Sicard and Sonen Dhia : *On effects of transient Electro -magnetic Interference on Integrated Circuits* 2006.
13. D Nitsch, F Sabath , J L Haseborg, Michael Camp, and Heyno Garbe: *Susceptibility of some Electronic Equipment to HPEM threats,* IEEE Transactions on Electromagnetic Compatibility, Vol 46, No 3, Aug 2004, Pg 380-389.
14. Gunnar Garansson : *HPM Effects on Electronic Components and the Importance of This Knowledge in Evaluation of System Susceptibility,* Proceedings of the 1999 IEEE EMC Symposium, Seattle, Washington, IEEE Code No 0-7803-5057-W99, 1999.
15. Shi Li-hua, YuYang, Zhou Ying-hui and Zhou Bi-hua : *Statistical Results on the EMP Conductive Susceptibility of a Serial Network Measurement and Control Device* Proceedings of CEEN 2009, Xian.
16. Jostein Godø, Odd H. Arnesen, Mats Bäckström, Brian A. Kerr and Ernst Krogager : *High Power Microwave Effects On Alarm Systems And Components*.
17. C Mojert, D Nitsch, H Friedhoff, J Maack, and M Camp : *UWB And EMP Susceptibility Of Modern Computer Networks*, Proc. EMC Zürich 2001, Zurich, Switzerland, Feb. 2001.
18. IEC Technical Specification 61000-5-9: *Electromagnetic Compatibility -Installation And Mitigation Guidelines- System Level Susceptibility Assessments For HEMP and HPEM,* Edition-1.0 , 2009-07.
19. Brian G and Ruth: *A Nuclear Electromagnetic Pulse (EMP) Vulnerability/Lethality (V/L) Taxonomy With Focus on Component Assessment* , Army research Lab ,November 1994.
20. George Baker, J Philip Castillo and Edward F Vance, "Potential for a unified Topological Approach to Electromagnetic Effects Protection," *IEEE Transactions on Electromagnetic compatibility*, Vol. 50, No. 3, pp. 267-274, Aug 1992.
21. D Nitsch, F Sabath, R Sablehaus and T Wieting : *Determination*

of the WB and UWB Susceptibility of Microprocessors and Electronic Components New Measurement Setup and First Results.

22. Michael Camp and Heyno Garbe : *Reproducibility Of The Destruction Effects In Integrated Circuits.*
23. Design Manual : *Reactor Instrumentation System* , RRSD/ DESS/55000/M-1 September, 2008.
24. IAEA publication : *Recommendations for Physical Protection of Nuclear materials and Nuclear Facilities*, INFCIRC/225/ Rev 4.
25. PV Varde, NS Bhamra, V Gopika, PK Datta, V Late, Nilesh Goel and Arvind Singh : *Probabilistic Safety Assessment of Research Reactor,* Reactor Group, Apr 2002.
26. M Ratna Rajul , Rama Sarmal , Sandeep M Satavl , K Rajeshwar Raol ,K Surya Narayanal and ZH Sholapurwala : *Fast Transient High Voltage Pulse Radiating System for Vulnerability studies of NEMP on Electronic Systems*, Electromagnetic Interference and Compatibility- 2008, INCEMIC 2008, Issue date 26-27 Nov 2008, pp. 223-228.
27. A Neuber, A Young, M Elsayed, J Dickens, M Giesselmann and M Kristiansen : *Compact High Power Microwave Generation,* Center for Pulsed Power & Power Electronics, Department of Electrical & Computer Engineering, Texas Tech University, ADM002187, Proceedings of 26th Army Science Conference, Orlando, Florida Dec 2008.
28. Larry D Bacon and Larry F Rinehart: *High Power Microwave Sources,* Sandia National Laboratories.
29. Donald Sullivan, "High Power Microwave Generation From a virtual Cathode Oscillator (Vircator)," *IEEE Transactions On Nuclear Science,* Vol. NS-30, No. 4, pp. 3426 – 3428, August 1983.
30. A N Didenko and VI Rashchikov : *High Power Microwave Generation in Virtual Cathode Systems,* IEEE code 0-7803-0135-8/91.
31. P Appelgren, McAkyuz, Mattias Elfsberg,Tomas Hurtig, A Larsson, Sten Nyholm and Cecilia Moller, "Study of Compact HPM System with reflex Triode and a Marx Generator," *IEEE Transactions On Plasma Science,* Vol 34, No 5, pp. 1796-1805, Oct 2006.
32. Libor Drazan and Roman Vrana, "Axial Vircator for Electronic Warfare Application," *Radio Engineering,* Vol. 18, No. 4, pp. 618-628, December 2009.
33. L Austrin, T Torabzadeh and A Larsson: *Analysis Of A Feasible Pulse Power Supply System For An Unmanned Aerial Vehicle* ,25th International Congress of the Aeronautical sciences.
34. Archana Sharma, Senthil Kumar, Sabyasachi Mitra, Vishnu Sharma, Ankur Patel, Amitava Roy,Rakhee Menon, K V Nagesh and D P Chakravarthy, "Development and Characterization of Repetitive 1-kJ Marx-Generator-Driven Reflex Triode System for High-Power Microwave Generation," *IEEE Transactions On Plasma Science*, Vol. 39, No. 5, pp. 1262-1267, May. 2011.
35. Amitava Roy, Archana Sharma,Sabyasachi Mitra, Rakhee Menon, Vishnu Sharma, KV Nagesh and DP Chakravarthy. "Oscillation Frequency of a Reflex Triode Virtual Cathode Oscillator," *IEEE Transactions On Electron Devices*, Vol. 58, No. 2, pp. 553-561, Feb. 2011.
36. Amitava Roy, S K Singh, R Menon, D Senthil Kumar, R Venkateswaran, M r Kulkarni, P C Saroj, K V Nagesh, K Mittal and D P Chakravarthy, "Measurement Of High Power Microwave Pulse Under Intense Electromagnetic Noise," *Pramana, Journal of Physics,* Vol. 74, No. 1, pp. 123-133, Jan 2010.
37. P Applegren,Cecilia Nylander and Sten Nyholm *Magnetic Field Measurement Systems for Microwave Frequencies,* Swedish Defence Research Agency ,Sep 2004.
38. Carl E Baum, "From The Electromagnetic Pulse to High Power Electromagnetics," *Proceedings of IEEE,* Vol. 80, No. 6, pp. 789-817, Jun 1992.

Life Cycle Reliability and Safety Engineering
Vol. 2 Issue 2 (2012) 27-43

Geotechnics In The 21st Century, Uncertainties And Other Challenges - With Particular Reference To Landslide Hazard And Risk Assessment

Robin Chowdhury[1*], Phil Flentje[2] and Gautam Bhattacharya[3]

[1] *Emeritus Professor, University of Wollongong, Australia, E-mail: robin@uow.edu.au*
(Invited Keynote Speaker, ISEUSAM-2012 Conference, January 2012, BESU, Shibpur, India)*
[2]*Senior Research Fellow, University of Wollongong Australia, E-mail: pflentje@uow.edu.au*
[3]*Professor of Civil Engg., Bengal Engineering & Science University, Shibpur, Howrah 711103, E-mail: gautam@civil.becs.ac.in*

ABSTRACT

This paper addresses emerging challenges in geotechnics in the context of the significant challenges posed by hazards, both natural and human-induced. The tremendous importance of dealing with uncertainties in an organized and systematic way is highlighted. The paper includes reflections on responding to the need for multi-disciplinary approaches .While the concepts and ideas are pertinent to diverse applications of geotechnics or to the whole of geotechnical engineering , illustrative examples will be limited to research trends in slope stability and landslide management.

From time to time researchers, academics and practicing engineers refer to the need for interdisciplinary approaches in geotechnical engineering. However, surveys of the relevant literature reveal few examples of documented research studies based within an interdisciplinary framework. Meanwhile there is a broad acceptance of the significant role on uncertainties in geotechnics.

This paper includes reflections on what steps might be taken to develop better approaches for analysis and improved strategies for managing emerging challenges in geotechnical engineering. For example, one might start with the need to highlight different types of uncertainties such as geotechnical, geological and hydrological. Very often, geotechnical engineers focus on variability of soil properties such as shear strength parameters and on systematic uncertainties. Yet there may be more important factors in the state of nature which are ignored because of the lack of a multi-disciplinary focus. For example the understanding of the potential for progressive failure within a soil mass or a slope may requires careful consideration of the geological context and of the history of stress and strain. The latter may be a consequence of previous seismic activity and fluctuations in rainfall and groundwater flow.

The frequency and consequences of geotechnical failures involving soil and rock continue to increase globally. The most significant failures and disasters are often associated with major natural events but not exclusively so.

It is expected that climate change will lead to even more unfavorable conditions for geotechnical projects and thus to increasing susceptibility and hazard of landsliding. This is primarily because of the expected increase in the variability of rainfall and the expected increase in sea levels. Responding to the effects of climate change will thus require more flexible and robust strategies for assessment of landslide susceptibility and to innovative engineering solutions.

Introduction

There are significant challenges for the future development and application of geotechnical engineering. Developments in research, analysis and practice have taken place to advance knowledge and practice. While the scope of the profession and its discipline areas is already vast, significant extension is required in the areas of hazard and risk assessment and management. In particular, the field of natural disaster reduction requires the development of innovative approaches within a multi-

disciplinary framework. Very useful and up-to-date information on the occurrence frequency and impact of different natural disasters is being assessed and analyzed by a number of organizations around the world. However, geotechnical engineers have not played a prominent part in such activities so far. Reference may be made to the research and educational materials developed on a regular basis by the Global Alliance for Disaster Reduction (GADR) with the aim of information dissemination and training for disaster reduction. Some selected illustrations from GADR are presented in an Appendix to this paper .The role of geotechnical engineers in implementing such goals is obvious from these illustrations.

The variability of soil and rock masses and other uncertainties have always posed unique challenges to geotechnical engineers. In the last few decades, the need to identify and quantify uncertainties on a systematic basis has been widely accepted. Methods for inclusion of such data in formal ways include reliability analysis within a probabilistic framework. Considerable progress has been made in complementing traditional deterministic methods with probabilistic studies. Nevertheless, the rate of consequent change to geotechnical practice has been relatively slow and sometimes half-hearted. Reviewing all the developments in geotechnical engineering which have taken place over the last thirty years or more would require painstaking and critical reviews from a team of experts over a considerable period of time and the subsequent reporting of the findings in a series of books. In comparison, the scope of this keynote paper is humble. Experienced academics who have been engaged in serious scholarship, research and consulting over several decades should be able to reflect on recent and continuing trends as well as warning signs of complacency or lack of vision. In this spirit, an attempt is made to highlight some pertinent issues and challenges for the assessment and management of geotechnical risk with particular reference to slope stability and landslides.

The writers of the present paper present some highlights of their own research through case study examples. These relate to aspects of regional slope stability and hazard assessment such as a landslide inventory map, elements of a relational database, rainfall intensity duration for triggering landslides, continuous monitoring of landslide sites in near-real time, landslide susceptibility and hazard maps. The paper concludes with reflections on continuing and emerging challenges. For further details the reader may refer to Chowdhury & Flentje (2008), Flentje et al (2007, 2010) and a comprehensive book (Chowdhury et al, 2010).

In order to get a sense of global trends in geotechnical analysis and the assessment and management of risk, reference may be made to the work of experts and professionals in different countries as reported in recent publications. The applications include the safety of foundations, dams and slopes against triggering events such as rainstorms, floods earthquakes and tsunamis.

The following is a sample of 5 papers from a 2011 conference related to geotechnical risk assessment and management, GeoRisk- 2011. Despite covering a wide range of topics and techniques, it is interesting that GIS-based regional analysis for susceptibility and hazard zoning is not included amongst these publications. Such gaps are often noted and reveal that far greater effort is required to establish multi-disciplinary focus in geotechnical research. This is clearly a continuing challenge for geotechnics in the 21st century.

- A comprehensive paper on geo-hazard assessment and management involving the need for integration of hazard, vulnerability and consequences and the consideration of acceptable and tolerable risk levels (Lacasse and Nadim, 2011).
- Risk assessment of Success Dam, California is discussed by Bowles et al (2011) with particular reference to the evaluation of operating restrictions as an interim measure to mitigate earthquake risk. The potential modes of failure related to earthquake events and flood events are discussed in two companion papers.
- The practical application of risk assessment in dam safety (the practice in U.S.A) is discussed in a paper by Scott (2011).

- Unresolved issues in Geotechnical Risk and Reliability (Christian and Baecher, 2011)
- Development of a risk-based landslide warning system (Tang and Zhang, 2011)

The first paper (Lacasse and Nadim, 2011) has a wide scope of topics and discusses the following six case studies:

Hazard assessment and early warning for a rock slope over a fjord arm on the west coast of Norway – the slope is subject to frequent rockslides usually with volumes in the range 0.5-5 million cubic meters.

Vulnerability assessment – Norwegian clay slopes in an urban area on the South coast of Norway Risk assessment – 2004 Tsunami in the Indian Ocean Risk mitigation – quick clay in the city of Drammen along the Drammensfjord and the Drammen River.

Risk mitigation – Early Warning System for landslide dams, Lake Sarez in the Pamir Mountain Range in eastern Tajikistan Risk of tailings dam break – probability of non-performance of a tailings management facility at Rosia Montana in Romania

Uncertainties Affecting Geotechnics

The major challenges in geotechnical engineering arise from uncertainties and the need to incorporate them in analysis, design and practice. The geotechnical performance of a specific site, facility, system or regional geotechnical project may be affected by different types of uncertainty such as the following (with examples in brackets):

- Geological uncertainty (geological details)
- Geotechnical parameter uncertainty (variability of shear strength parameters and of pore water pressure)
- Hydrological uncertainty (aspects of groundwater flow)
- Uncertainty related to historical data (frequency of slides, falls or flows)
- Uncertainty related to natural or external events (magnitude, location and timing of rainstorm, flood, earthquake and tsunami)
- Project uncertainty (construction quality, construction delays)
- Uncertainty due to unknown factors (effects of climate change)

On some projects, depending on the aims, geotechnical engineers may be justified in restricting their attention to uncertainties arising from geological, geotechnical and hydrological factors. For example, the limited aim may be to complement deterministic methods of analysis with probabilistic studies to account for imperfect knowledge of geological details and limited data concerning measured soil properties and pore water pressures. It is necessary to recognize that often pore pressures change over time and, therefore, pore pressure uncertainty has both spatial and temporal aspects which can be critically important.

During the early development of probabilistic analysis methods researchers often focused on the variability of soil properties in order to develop the tools for probabilistic analysis. It was soon realized that natural variability of geotechnical parameters such as shear strength must be separated from systematic uncertainties such as measurement error and limited number of samples. Another advance in understanding has been that the variability of a parameter, measured by its standard deviation, is a function of the spatial dimension over which the variability is considered. In some problems, consideration of spatial variability on a formal basis is important and leads to significant insights.

An important issue relates to the choice of geotechnical parameters and their number for inclusion in an uncertainty analysis. The selection is often based on experience and can be justified by performing sensitivity studies. A more difficult issue is the consideration of 'new' geotechnical parameters not used in traditional deterministic or even in probabilistic studies. Thus one must think 'outside the box' for 'new' parameters which might have significant influence on geotechnical reliability. Otherwise the utility and benefits of reliability analyses may not be fully realized. As an example, the 'residual factor' (defined as proportion of a slip surface over which shear strength has

decreased to a residual value) is rarely used as a variable in geotechnical slope analysis. Recently, interesting results have been revealed from a consideration of 'residual factor' in slope stability as a random variable (Chowdhury and Bhattacharya 2011, Bhattacharya and Chowdhury 2011). Ignoring the residual factor can lead to overestimate of reliability and thus lead to unsafe or unconservative practice.

For regional studies such as zoning for landslide susceptibility and hazard assessment, historical data about previous events are very important. Therefore uncertainties with respect to historical data must be considered and analyzed carefully. Such regional studies are different in concept and implementation from traditional site-specific deterministic and probabilistic studies and often make use of different data-sets. A successful knowledge-based approach for assessment of landslide susceptibility and hazard has been described by Flentje (2009).

If the aim of a geotechnical project is to evaluate geotechnical risk, it is necessary to consider the uncertainty related to the occurrence of an external event or events that may affect the site or the project over an appropriate period of time such as the life of the project.

Consideration of project uncertainty would require consideration of economic, financial and administrative factors in addition to the relevant technical factors considered above. In this regard the reader may refer to a recent paper on georisks in the business environment by Brumund (2011); the paper also makes reference to unknown risk factors.

For projects which are very important because of their size, location, economic significance, or environmental impact, efforts must be made to consider uncertainty due to unknown factors. Suitable experts may be co-opted by the project team for such an exercise.

Slope Analysis Methods

Limit Equilibrium and Stress deformation Approaches

Deterministic methods can be categorized as limit equilibrium methods and stress-deformation methods. Starting from simple and approximate limit equilibrium methods based on simplifying assumptions, several advanced and relatively rigorous methods have been developed.

The use of advanced numerical methods for stress-deformation analysis is essential when the estimation of strains and deformations within a slope is required. In most cases, two-dimensional (2D) stress-deformation analyses would suffice. However, there are significant problems which need to be modeled and analyzed in three-dimensions. Methods appropriate for 3D stress-deformation analysis have been developed and used successfully. Advanced stress-deformation approaches include the finite-difference method, the finite-element method, the boundary element method, the distinct element method, and the discontinuous deformation analysis method.

Progressive Failure

Progressive failure of natural slopes, embankment dams and excavated slopes is a consequence of non-uniform stress and strain distribution and the strain-softening behavior of earth masses. Thus shear strength of a soil element, or the shear resistance along a discontinuity within a soil or rock mass, may decrease from a peak to a residual value with increasing strain or increasing deformation. Analysis and simulation of progressive failure requires that strain-softening behavior be taken into consideration within the context of changing stress or strain fields. This may be done by using advanced methods such as an initial stress approach or a sophisticated stress-deformation approach. Of the many historical landslides in which progressive failure is known to have played an important part, perhaps the most widely studied is the catastrophic Vaiont slide which occurred in Italy in 1964. The causes and mechanisms have not been fully explained by any one study and there are still uncertainties concerning both the statics and dynamics of the slide. For further details and a list of some relevant references, the reader may refer to Chowdhury et al (2010).

Probabilistic approaches and simulation of progressive failure

A probabilistic approach should not be seen simply as the replacement of a calculated 'factor of safety' as a performance index by a calculated 'probability of failure'. It is important to consider the broader perspective and greater insight offered by adopting a probabilistic framework. It enables a better analysis of observational data and enables the modeling of the reliability of a system. Updating of reliability on the basis of observation becomes feasible and innovative approaches can be used for the modeling of progressive failure probability and for back-analysis of failed slopes. Other innovative applications of a probabilistic approach with pertinent details and references are discussed by Chowdhury et al (2010).

An interesting approach for probabilistic seismic landslide analysis which incorporates the traditional infinite slope limit equilibrium model as well as the rigid-block displacement model has been demonstrated by Jibson et al (2000).

A probabilistic approach also facilitates the communication of uncertainties concerning hazard assessment and slope performance to a wide range of end-users including planners, owners, clients and the general public.

Geotechnical Slope Analysis in a regional context

Understanding geology, geomorphology and groundwater flow is of key importance. Therefore, judicious use must be made of advanced methods of modeling in order to gain the best possible understanding of the geological framework and to minimize the role of uncertainties on the outcome of analyses (Marker, 2009; Rees et al, 2009).

Variability of ground conditions, spatial and temporal, is important in both regional and site-specific analysis. Consequently, probability concepts are very useful in both cases although they may be applied in quite different ways.

Spatial and temporal variability of triggering factors such as rainfall have a marked influence on the occurrence and distribution of landslides in a region (Chowdhury et al, 2010, Murray, 2001)

This context is important for understanding the uncertainties in the development of critical pore-water pressures. Consequently, it helps in the estimation of rainfall threshold for on-set of landsliding. Regional and local factors both would have a strong influence on the combinations of rainfall magnitude and duration leading to critical conditions.

Since earthquakes trigger many landslides which can have a devastating impact, it is important to understand the causative and influencing factors. The occurrence, reach, volume and distribution of earthquake-induced landslides are related to earthquake magnitude and other regional factors. For further details and a list of some relevant references, the reader may refer to Chowdhury et al (2010).

Regional Slope Stability Assessments

Basic Requirements

Regional slope stability studies are often carried out within the framework of a Geographical Information System (GIS) and are facilitated by the preparation of relevant data-sets relating to the main influencing factors such as geology, topography, drainage characteristics and by developing a comprehensive inventory of existing landslides. The development of a digital elevation model (DEM) facilitates GIS based modeling of landslide susceptibility, hazard and risk within a GIS framework. Regional slope stability and hazard studies facilitate the development of effective landslide risk management strategies in an urban area. The next section of this paper is devoted to a brief discussion of GIS as a versatile and powerful system for spatial and even temporal analysis. This is followed by a section providing a brief overview of sources and methods for obtaining accurate spatial data. The data may relate to areas ranging from relatively limited zones to very large regions. Some of these resources and methods have a global reach and applicability. Such data are very valuable for developing digital elevation models (DEMs) of increasing accuracy.

For regional analysis, a DEM is, of course, a very important and powerful tool.

Landslide Inventory

The development of comprehensive databases including a landslide inventory is most desirable if not essential especially for the assessment of slope stability in a regional context. It is important to study the occurrence and spatial distribution of first-time slope failures as well as reactivated landslides.

Identifying the location of existing landslides is just the beginning of a systematic and sustained process with the aim of developing a comprehensive landslide inventory. Among other features, it should include the nature, size, mechanism, triggering factors and date of occurrence of existing landslides. While some old landslide areas may be dormant, others may be reactivated by one or more regional triggering factors such as heavy rainfall and earthquakes.

One comprehensive study of this type has been discussed in some detail in Chapter 11 of Chowdhury et al (2010). This study was made for the Greater Wollongong region, New South Wales, Australia by the University of Wollongong (UOW) Landslide Research Team (LRT). In this paper this study is also referred to as the WOLLONGONG REGIONAL STUDY.

A small segment for the Wollongong Landslide Inventory for the Wollongong Regional Study is shown as Figure 1. The elements of a Landslide Relational Database are shown as Figure 2. Some details of the same are shown in Figures 3 and 4. A successful knowledge-based approach for assessment of landslide susceptibility and hazard has been described by Flentje (2009) and is covered in some detail in a separate section of this paper.

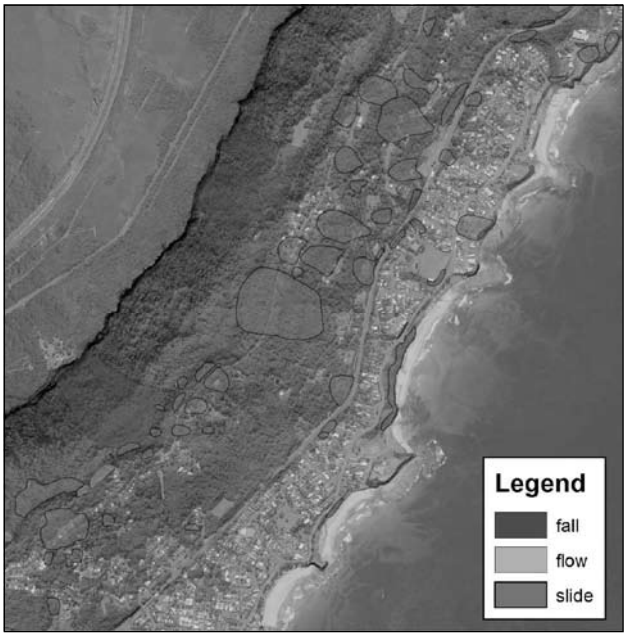


Fig. 1: Segment of the University of Wollongong Landslide Inventory.

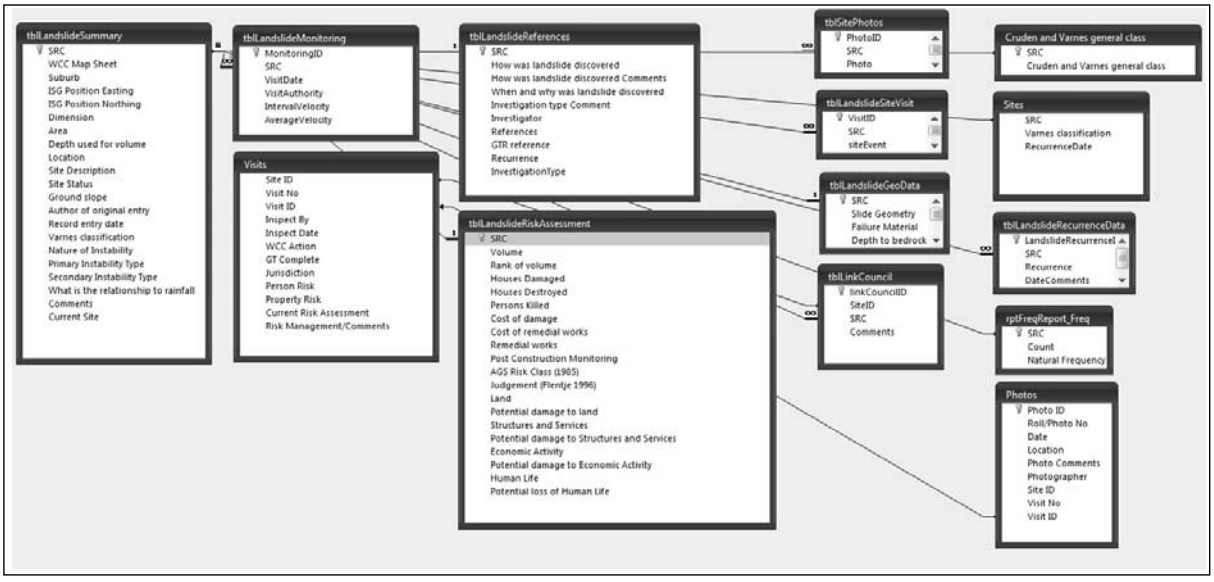


Fig. 2: Elements of a Landslide Relational Database.

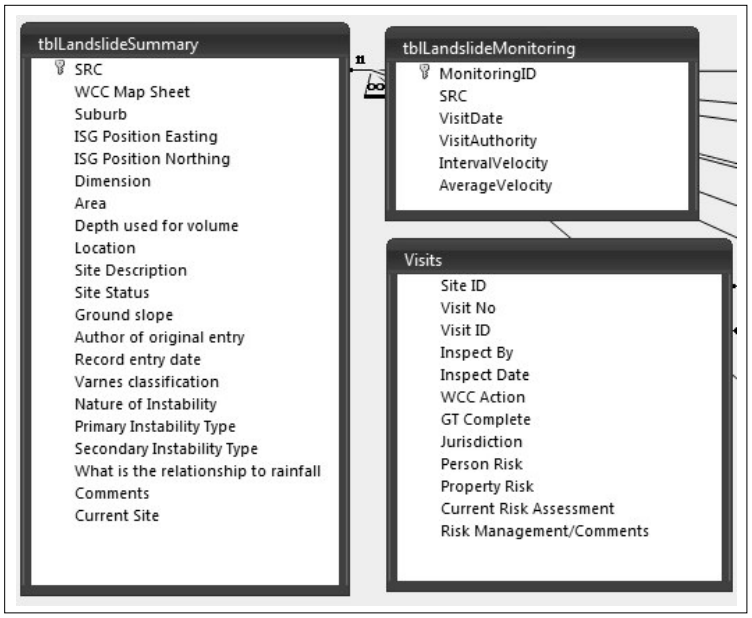


Fig. 3: Details of main tables of Relational Database shown above.

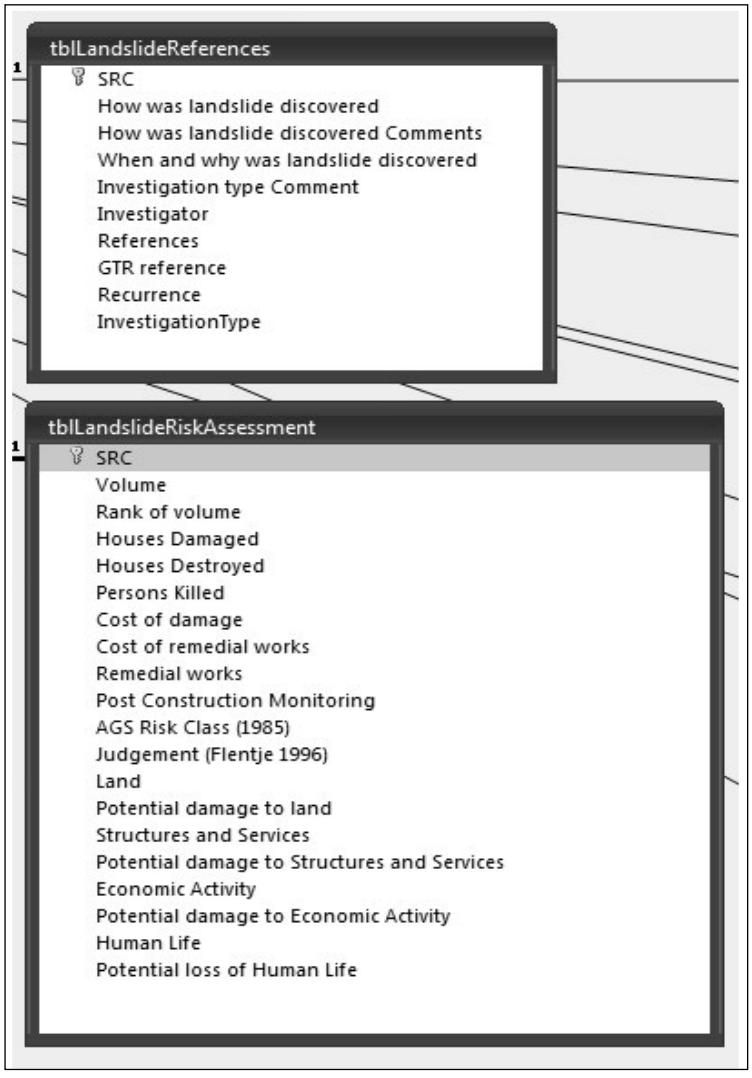


Fig. 4: Details of selected tables of Relational Database shown above.

Role of Geographical Information Systems (GIS)

GIS enables the collection, organization, processing, managing and updating of spatial and temporal information concerning geological, geotechnical, topographical, and other key parameters. The information can be accessed and applied by a range of professionals such as geotechnical engineers, engineering geologists, civil engineers and planners for assessing hazard of landsliding as well as for risk management. Traditional slope analysis must, therefore, be used within the context of a modern framework which includes GIS. Amongst the other advantages of GIS are the ability to deal with multiple hazards, the joining of disparate data and the ability to include decision support and warning systems (Gibson and Chowdhury, 2009).

Papers concerning the application of basic, widely available, GIS systems as well as about the development of advanced GIS systems continue to be published. For instance, Reeves and West (2009), covering a conference session on 'Geodata for the urban environment', found that 11 out of 30 papers were about the 'Development of Geographic Information Systems' while Gibson and Chowdhury (2009) pointed out that the input of engineering geologists (and, by implication, geotechnical engineers) to urban geohazards management is increasingly through the medium of GIS.

Consequently, 3D geological models have been discussed by a number of authors such as Rees et al (2009) who envisage that such models should be the basis for 4D process modeling in which temporal changes and factors can be taken into consideration. They refer, in particular, to time-series data concerning precipitation, groundwater, sea level and

temperature. Such data, if and when available, can be integrated with 3D spatial modeling.

Sources of Accurate Spatial Data Relevant to the Development of Digital Elevation Models

Over the last decade, Airborne Laser Scan (ALS) or Light Detection and Ranging (LiDAR) techniques are increasingly being applied across Australia to collect high resolution terrain point datasets. When processed and used to develop Geographic Information System (GIS) Digital Elevation Models (DEMs), the data provides high resolution contemporary terrain models that form fundamental GIS datasets. Prior to the advent of this technology, DEMs were typically derived from 10 to 50 year old photogrammetric contour datasets. When processed, ALS datasets can comprise point clouds of many millions of ground reflected points covering large areas up to hundreds of square kilometers, with average point densities exceeding one point per square meter. Collection, processing and delivery of these data types are being enhanced and formalized over time. Increasingly, this data is also being collected in tandem with high resolution geo-referenced imagery.

Airborne and Satellite derived Synthetic Aperture Radar (SAR) techniques are also being increasingly developed and applied internationally to develop terrain models, and specifically differential models between return visits over the same area in order to highlight the changes in ground surfaces with time. This is being used to monitor landslide movement, ground subsidence and other environmental change.

NASA and the Japan Aerospace Exploration Society have just recently (mid-October 2011) and freely released via the internet the Advanced Space-borne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) - ASTER GDEM v2 global 30m Digital Elevation Model as an update to the year 2000 vintage NASA SRTM Global DEM at 90m and 30m pixel resolutions. This global data release means moderately high resolution global Digital Elevation Model data are available to all.

The development of ALS terrain models and the free release of the global ASTER GDEM v2 have important implications for the development of high resolution Landslide Inventories and Zoning Maps world wide. These datasets mean one of the main barriers in the development of this work has been eliminated.

Observational Approach – Monitoring and Alert Systems

Geotechnical analysis should not be considered in isolation since a good understanding of site conditions and field performance is essential. This is particularly important for site-specific as well as regional studies of slopes and landslides. Observation and monitoring of slopes are very important for understanding all aspects of performance; from increases in pore-water pressures to the evidence of excessive stress and strain, from the development of tension cracks and small shear movements to initiation of progressive failure, and from the development of a complete landslide to the post-failure displacement of the landslide mass.

Observation and monitoring also facilitate an understanding of the occurrence of multiple slope failures or widespread landsliding within a region after a significant triggering event such as rainfall of high magnitude and intensity (Flentje et al, 2007; Flentje, 2009). Observational approaches facilitate accurate back-analyses of slope failures and landslides. Moreover, geotechnical analysis and the assessment of hazard and risk can be updated with the availability of additional observational data on different parameters such as pore-water pressure and shear strength. The availability of continuous monitoring data obtained in near-real time will also contribute to more accurate assessments and back-analyses. Consequently, such continuous monitoring will lead to further advancement in the understanding of slope behavior.

One part of the Wollongong Regional Study is the development of rainfall intensity - duration curves for the triggering of landslides overlaid with historical rainfall average recurrence interval (ARI) curves as shown in Figure 5. From the very beginning of this research, the potential

use of such curves for alert and warning systems was recognized. In fact, this research facilitated risk management in the Wollongong Study Area during intense rainfalls of August 1998 when widespread landsliding occurred.

More recent improvement and extension of this work involves the use of data from our growing network of continuous real-time monitoring stations where we are also introducing the magnitude of displacement as an additional parameter. Aspects of this research are shown in Figure.5 and, as more data become available from continuous monitoring, additional displacement (magnitude)-based curves can be added to such a plot.

Two examples of continuous landslide performance monitoring are shown in Figures 6 and 7. Figure 6 relates to a coastal urban landslide site (43,000m^3) with limited trench drains installed. The relationship between rainfall, pore water pressure rise and displacement is clearly evident at two different time intervals in this figure. Figure 7 shows data from a complex translational landslide system (720,000m^3) which is located on a major highway in NSW Australia. In the 1970's landsliding severed this artery in several locations resulting in road closures and significant losses arising from damage to infrastructure and from traffic disruptions.

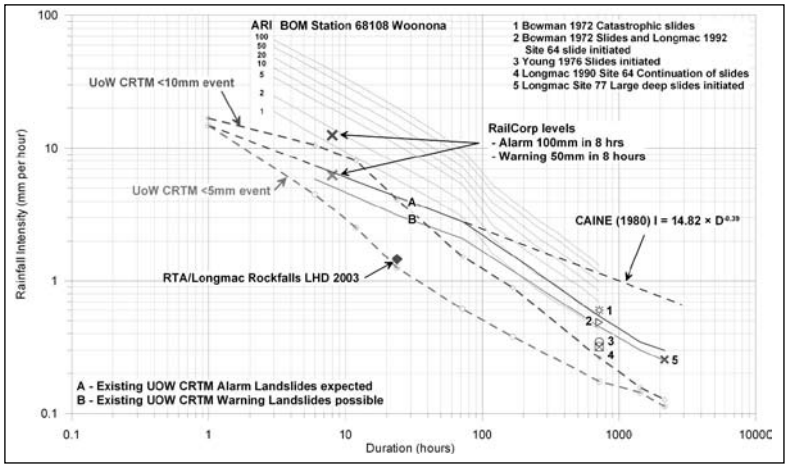


Fig. 5: Interpreted threshold curves for landsliding in Wollongong, superimposed on Annual Recurrence Interval curves for a selected rainfall station.

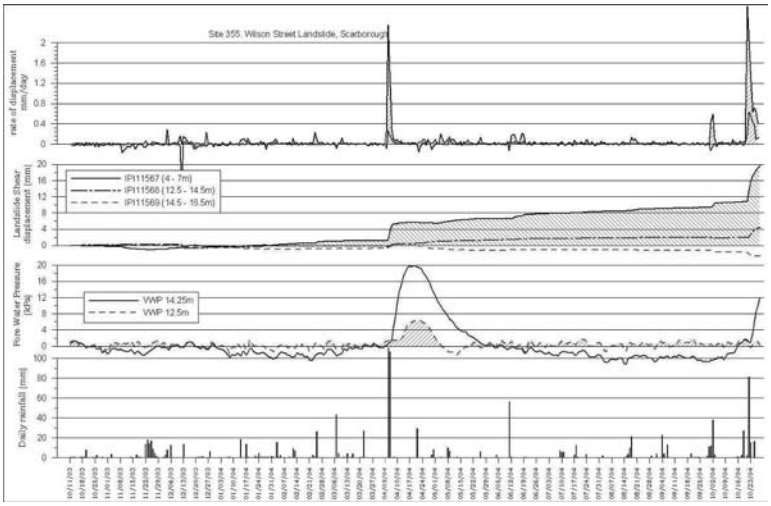


Fig. 6: Hourly logged continuously recorded rainfall, pore water pressure, landslide displacement and rate of displacement data for a 43,000m^3 urban landslide site in Wollongong.

After comprehensive investigations, remedial measures were installed. At this site, a dewatering pump system was installed, which continues to operate to this day. However, this drainage system has been reviewed and upgraded from time to time. Since 2004, this site has been connected to the Continuous Monitoring Network of the University of Wollongong Landslide Research Team. Interpretation of the monitoring data shows that movement has been limited to less than 10 mm since the continuous monitoring commenced as shown in Figure 7 (Flentje et al 2010). However, the occurrence of events of this small movement was considered unacceptable by the authorities. Hence, pump and monitoring system upgrades commenced in 2006 and have been completed in 2011.

Susceptibility and Hazard Assessment (Wollongong Regional Study)

The susceptibility model area and the data-sets

The area chosen within the Wollongong Region for modeling

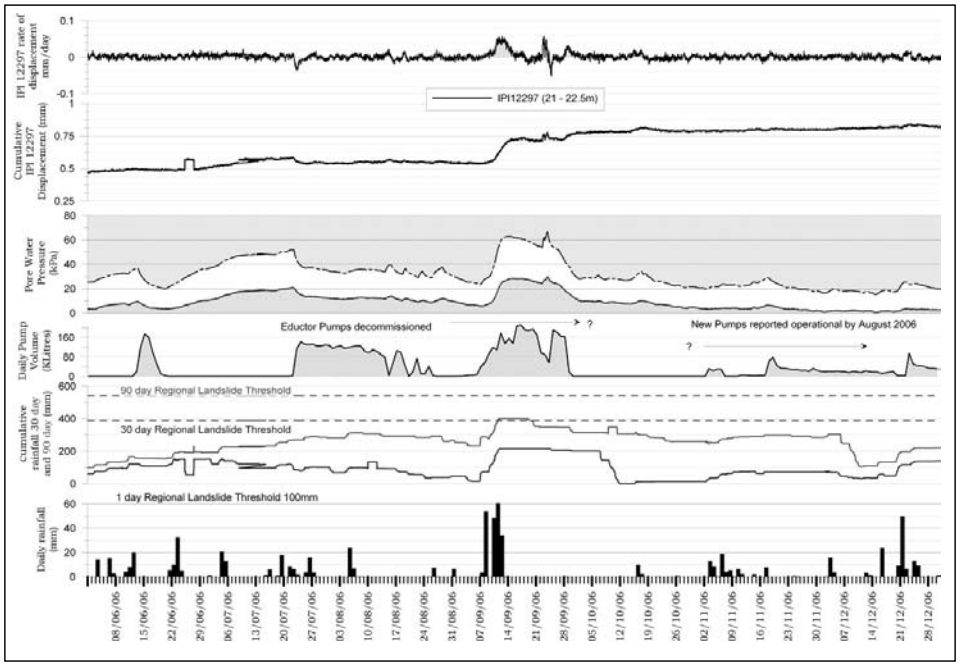


Fig. 7: Hourly logged, continuously recorded rainfall, groundwater pump volumes, pore water pressure, landslide displacement and rate of displacement data for a 720,000m³ landslide affecting a major transport artery in Wollongong.

landslide susceptibility (Susceptibility Model Area) is 188 square km in extent and contains 426 Slide category landslides.

The data sets used for this study include:

- Geology (mapped geological formations, 21 variables)
- Vegetation (mapped vegetation categories, 15 variables)
- Slope inclination (continuous floating point distribution)
- Slope aspect (continuous floating point distribution)
- Terrain units (buffered water courses, spur lines and other intermediate slopes)
- Curvature (continuous floating point distribution)
- Profile curvature (continuous floating point distribution)
- Plan curvature (continuous floating point distribution)
- Flow accumulation (continuous integer), and
- Wetness index (continuous floating point distribution)

Landslide inventory

The landslide inventory for this study has been developed over a fifteen year period and comprises a relational MS Access and ESRI ArcGIS Geodatabase with 75 available fields of information for each landslide site. It contains information on a total of 614 landslides (Falls, Flows and Slides) including 480 slides. Amongst the 426 landslides within the Susceptibility Model Area, landslide volumes have been calculated for 378 of these sites. The average volume is 21800 m^3 and the maximum volume is 720,000 m^3.

***Knowledge-based approach based on Data Mining* model**

The specific knowledge-based approach used for analysis and synthesis of the data sets for this study is the Data Mining (DM) process or model. The DM learning process is facilitated by the software "See 5" which is a fully developed application of "C4.5" (Quinlan, 1993). The DM learning process helps extract patterns from the databases related to the study. Known landslide areas are used for one half of the model training, the other half comprising randomly selected points from within the model area but outside the known landslide boundaries. Several rules are

generated during the process of modeling. Rules which indicate potential landsliding are assigned positive confidence values and those which indicate potential stability (no-landsliding) are assigned negative confidence values. The rule set is then re-applied within the GIS software using the ESRI Model Builder extension to produce the susceptibility grid. The complete process of susceptibility and hazard zoning is described in Flentje (2009) and in Chapter 11 of Chowdhury et al (2010).

Susceptibility and Hazard zones

On the basis of the analysis and synthesis using the knowledge-based approach, it has been possible to demarcate zones of susceptibility and hazard into four categories:

1. Very Low Susceptibility (or Hazard) of landsliding (VL)
2. Low Susceptibility (or Hazard) of landsliding (L)
3. Moderate Susceptibility (or Hazard) of landsliding (M), and
4. High Susceptibility (or Hazard) of landsliding (H)

A segment of the landslide Susceptibility map is shown in Figure 8 below. A segment of the landslide hazard map, an enlarged portion from the bottom left of Figure 8, is reproduced as Figure 9. Relative likelihoods of failure in different zones, estimated from the proportion of total landslides which occurred in each zone over a period of 126 years, are presented in columns 1 and 2 of Table 1 below. This information is only a part of the full table presented as Table 11.3 in Chowdhury et al (2010).

Table 1: Failure Likelihood and Reliability Index for each Hazard Zone (after Chowdhury et al, 2010)

Hazard Zone Description	Failure Likelihood	Reliability Index
Very Low	7.36×10^{-3}	2.44
Low	6.46×10^{-2}	1.51
Moderate	3.12×10^{-1}	0.49
High	6.16×10^{-1}	-0.30

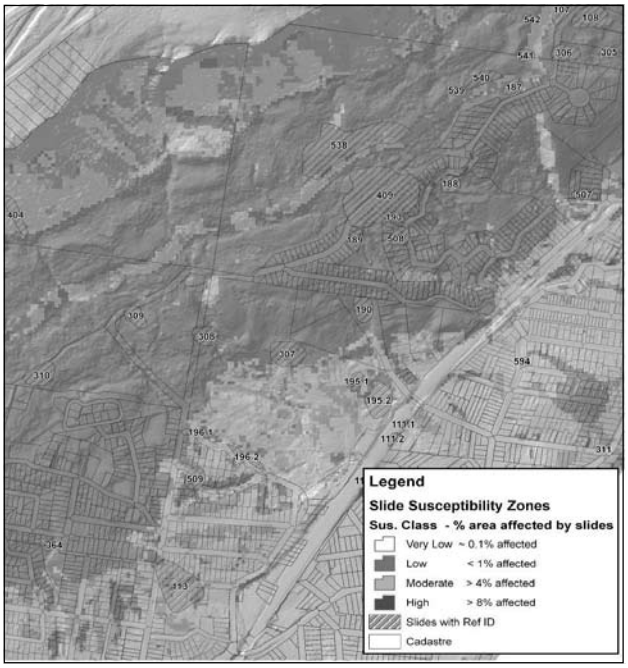


Fig. 8: Segment of Landslide Inventory and Susceptibility Zoning Map, Wollongong Local Government Area, New South Wales, Australia.

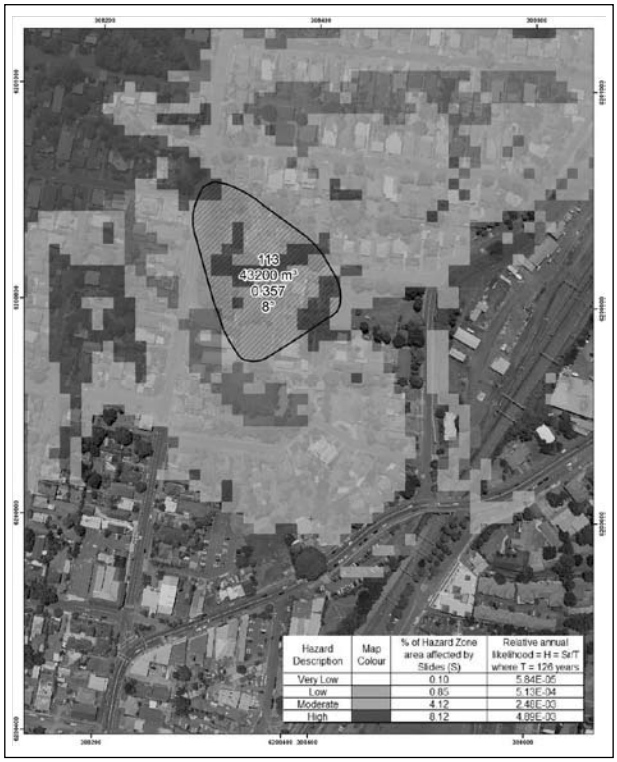


Fig. 9: Segment of Landslide Hazard Zoning Map from the bottom left corner of Fig.8, Wollongong Local Government Area, New South Wales, Australia. Landslide label shows four important particulars of each landslide stacked vertically. These are (1) Site Reference Code,(2) landslide volume,(3) annual frequency of reactivation derived from inventory and(4)landslide profile angle. Hazard zoning in legend shows relative annual likelihood as explained in the text.

Estimated Reliability Indices And Factors Of Safety

An innovative concept has been proposed by Chowdhury & Flentje (2011) for quantifying failure susceptibility from zoning maps developed on the basis of detailed knowledge-based methods and techniques within a GIS framework. The procedure was illustrated with reference to the results of the Wollongong Regional Study and the relevant Tables are reproduced here. Assuming that the factor of safety has a normal distribution, the reliability index was calculated for each zone based on the associated failure likelihood which is assumed to represent the probability of failure. These results are presented in the third or last column of Table 1.

Table 2: Typical Mean value of Factor of Safety (F) for each Hazard Zone considering coefficient of variation to be 10 %.(after Chowdhury& Flentje, 2011)

Hazard Zone Description	Reliability Index	Mean of Factor of Safety, F (V_F = 10%)
Very Low	2.44	1.32
Low	1.51	1.18
Moderate	0.49	1.05
High	-0.3	0.97

Assuming that the coefficient of variation of the factor of safety is 10%, the typical values of mean factor of safety for each zone are shown in Table 2. The results were also obtained for other values of the coefficient of variation of the factor of safety (5%, 10%, 15% and 20%). These results are shown in Table 3.

Table 3: Typical Mean values of Factor of Safety for different values of coefficient of variation (after Chowdhury& Flentje, 2011)

V_F%	Mean of F for different Hazard Zones			
	Very Low	Low	Moderate	High
5	1.14	1.08	1.02	0.98
10	1.32	1.18	1.05	0.97
15	1.57	1.29	1.08	0.96
20	1.95	1.43	1.11	0.94

Most of the landslides have occurred during very high rainfall events. It is assumed here, in the first instance, that most failures are associated with a pore water pressure ratio of about 0.5 (full seepage condition in a natural slope). Furthermore, assuming that the 'infinite slope' model applies to most natural slopes and that cohesion intercept is close to zero, the values of factor of safety can be calculated for other values of the pore pressure ratio (0.2, 0.3 and 0.4) for any assumed value of the slope inclination. The results shown below in Table 4 are for a slope with an inclination of 12 degrees for pore pressure ratios in the range 0.2 - 0.5.

Table 4: Typical Mean Factor of Safety with different values of pore pressure ratio (slope inclination i = 12˚, V_F = 10%). (after Chowdhury & Flentje, 2011)

Pore water pressure ratio	Mean of F for different Hazard Zones			
	Very Low	Low	Moderate	High
0.5	1.32	1.18	1.05	0.97
0.4	1.61	1.44	1.28	1.18
0.3	1.90	1.70	1.51	1.40
0.2	2.19	1.95	1.74	1.61

Discussion on the proposed concept and procedure

The above results were obtained as a typical F value or a set of F values referring to each hazard zone. However, taking into consideration the spatial variation of slope angle, shear strength and other factors, this approach may facilitate the calculation F at individual locations. Well-documented case studies of site-specific analysis would be required for such an extension of the procedure. Other possibilities include estimation of the variation of local probability of failure. The approach may also be used for scenario modeling relating to the effects of climate change. If reliable data concerning pore pressure changes become available, failure susceptibility under those conditions can be modeled and the likelihood and impact of potential catastrophic slope failures can be investigated.

Appendix I — Selected Figures from Power Point Slide Set Entitled "Understanding Risk And Risk Reduction" (Hays 2011)

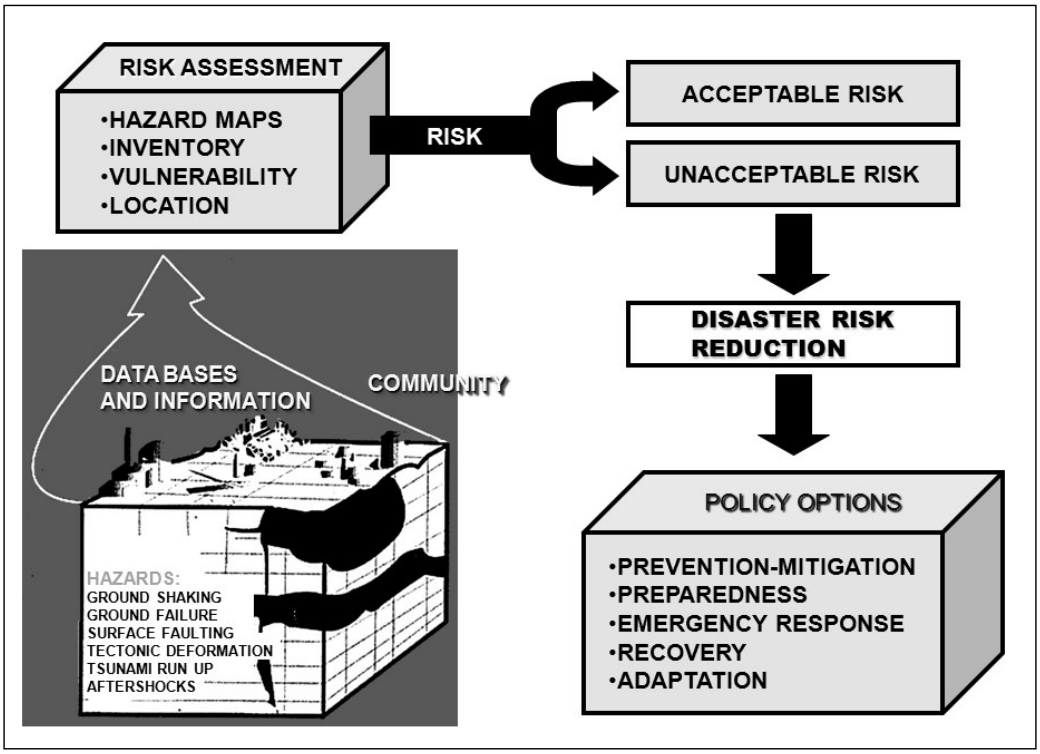


Fig. A-1: Elements of Risk Assessment and Management for Natural Disasters courtesy of Walter Hays, 2011.

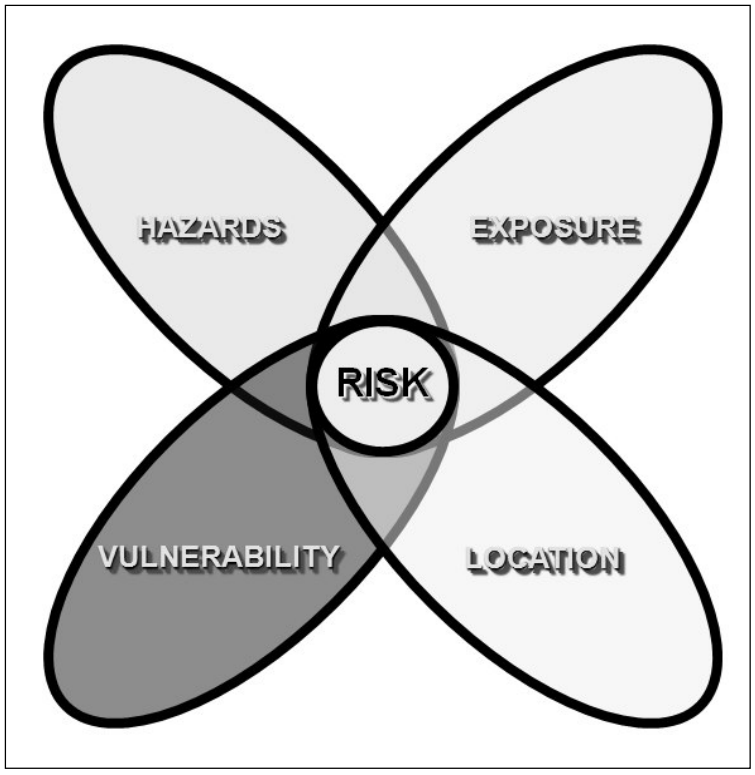


Fig. A-2: Components of Risk courtesy of Walter Hays, 2011.

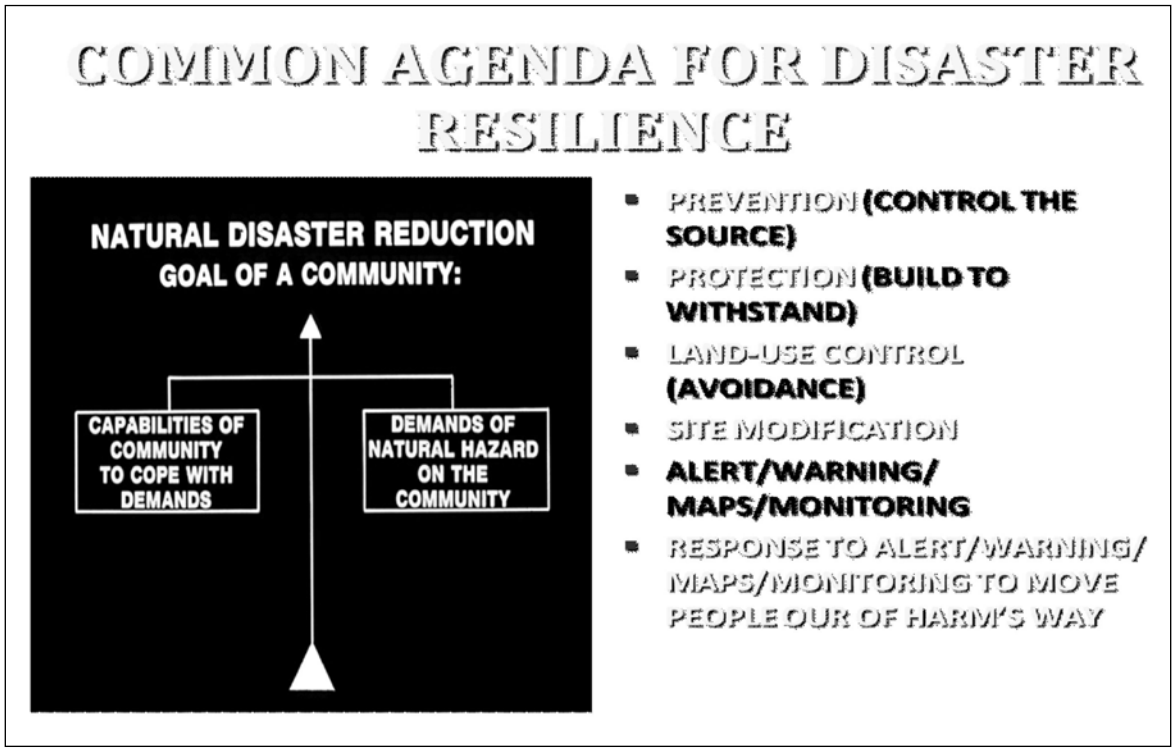


Fig. A-3: Common Agenda for Natural Disaster Resilience, courtesy of Walter Hays, 2011.

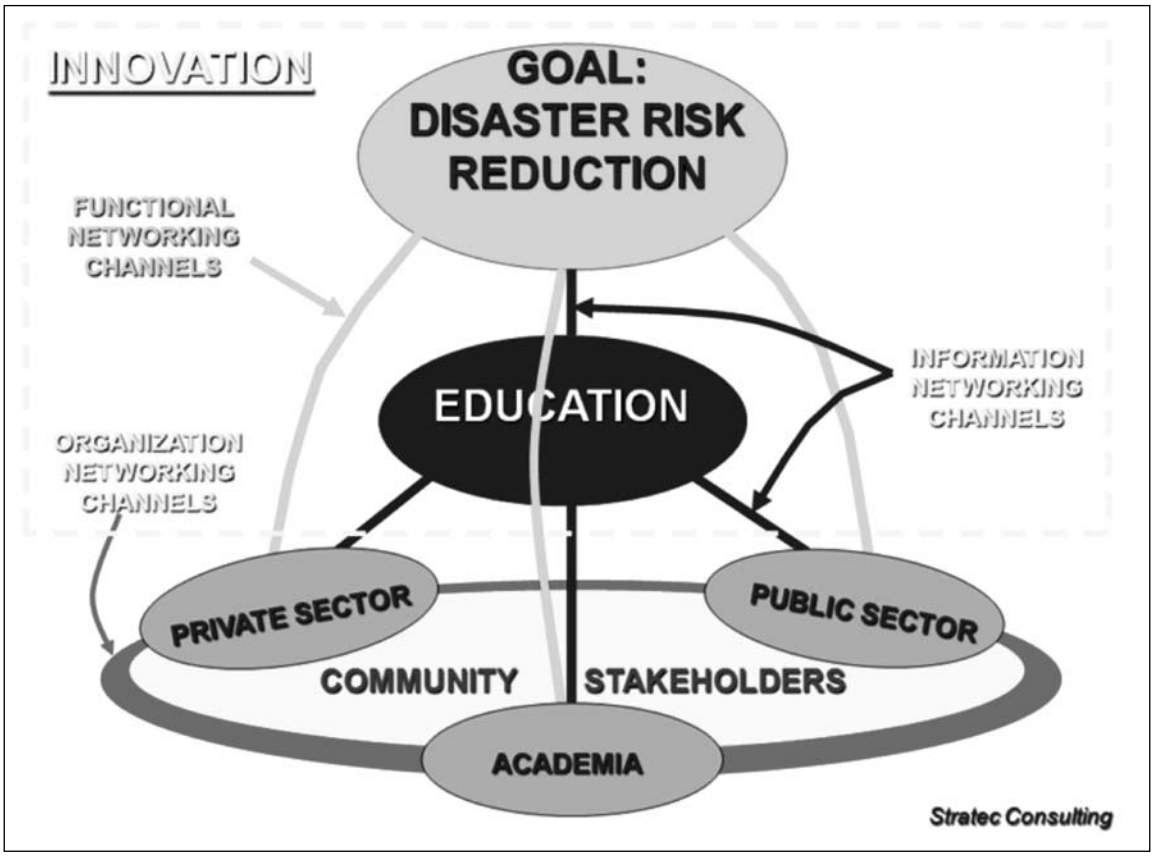


Fig. A-4: The overall context for Innovation in Disaster Management and Reduction, courtesy of Walter Hays, 2011

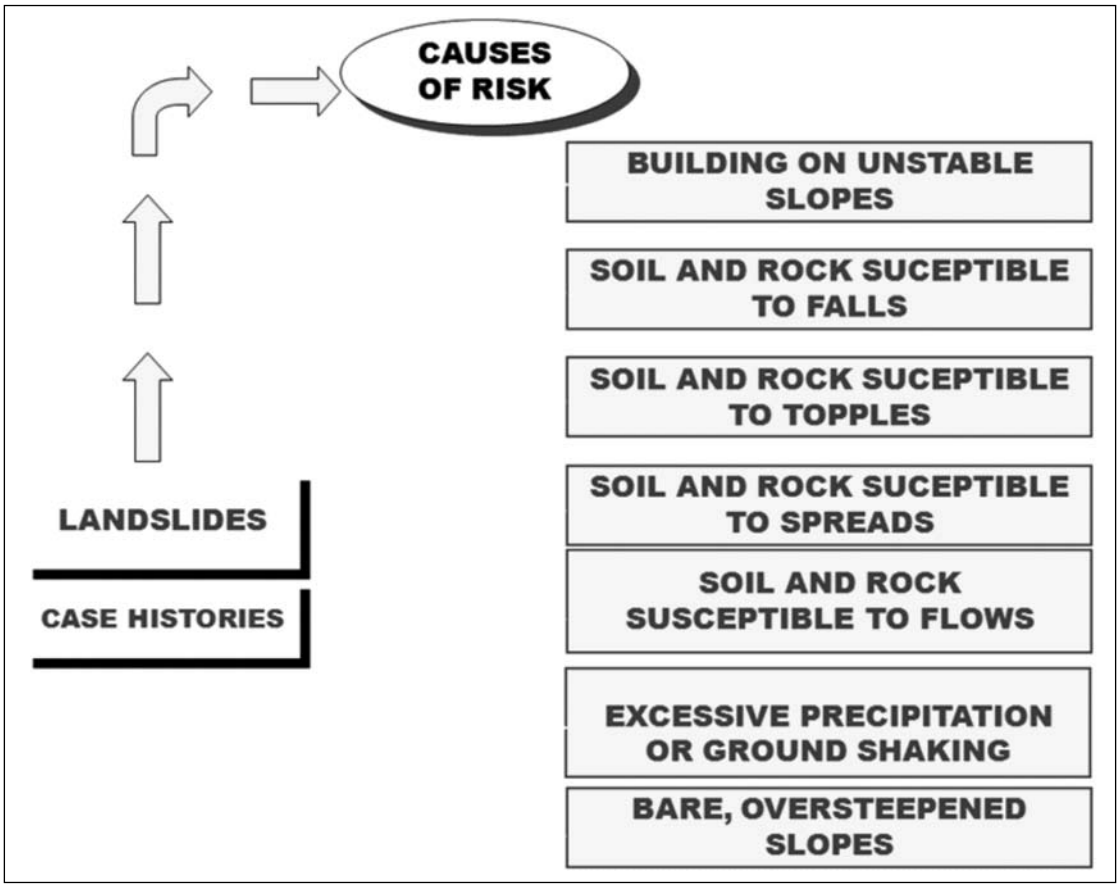


Fig. A-5: Some causes of risk for landslides, courtesy of Walter Hays, 2011

Discussion, Specific Lessons Or Challenges

The focus of this paper has been on hazard and risk assessment in geotechnical engineering. Advancing geotechnical engineering requires the development and use of knowledge which facilitates increasingly reliable assessments even when the budgets are relatively limited. Because of a variety of uncertainties, progress requires an astute combination of site-specific and regional assessments. For some projects, qualitative assessments within the framework of a regional study may be sufficient. In other projects quantitative assessments, deterministic and probabilistic, may be essential.

In this paper, different cases have been discussed in relation to the Wollongong Regional Study. Firstly reference was made to the basis of an alert and warning system for rainfall-induced landsliding based on rainfall-intensity-duration plots supplemented by continuous monitoring. The challenges here are obvious. How do we use the continuous pore pressure data from monitoring to greater advantage? How do we integrate all the continuous monitoring data to provide better alert and warning systems? This research has applications in geotechnical projects generally well beyond slopes and landslides.

The examples concerning continuous monitoring of two case studies discussed in this paper illustrate the potential of such research for assessing remedial and preventive measures. The lesson from the case studies is that, depending on the importance of a project, even very low hazard levels may be unacceptable. As emphasized earlier, the decision to upgrade subsurface drainage at the cost of hundreds of thousands of dollars over several years was taken and implemented despite the shear movements being far below disruptive magnitudes as revealed by continuous monitoring. The challenge in such problems is to consolidate this experience for future applications so that costs and benefits can be rationalized further.

The last example from the Wollongong Regional Study concerned the preparation of zoning maps for landslide susceptibility and hazard. Reference was made to an innovative approach for quantitative interpretation of such maps in terms of well known performance indicators such as 'factor of safety' under a variety of pore pressure conditions. The challenge here is to develop this methodology further to take into consideration the spatial and temporal variability within the study region.

Challenges Due To External Factors

Beyond the scope of this paper, what are the broad challenges in geotechnical hazard and risk assessment? How do we deal with the increasing numbers of geotechnical failures occurring globally including many disasters and how do we mitigate the increasingly adverse consequences of such events? What strategies, preventive, remedial and other, are necessary?

Often catastrophic landslides are caused by high magnitude natural events such as rainstorms and earthquakes. It is also important to consider the contribution of human activities such as indiscriminate deforestation and rapid urbanization to landslide hazard. There is an increasing realization that poor planning of land and infrastructure development has increased the potential for slope instability in many regions of the world.

Issues concerned with increasing hazard and vulnerability are very complex and cannot be tackled by geotechnical engineers alone. Therefore, the importance of working in interdisciplinary teams must again be emphasized. Reference has already been made to the use of geological modeling (2D, 3D and potentially 4D) and to powerful tools such as GIS which can be used in combination with geotechnical and geological models.

At the level of analysis methods and techniques, one of the important challenges for the future is to use slope deformation (or slip movement) as a performance indicator rather than the conventional factor of safety. Also, at the level of analysis, attention needs to be given to better description of uncertainties related to construction of slopes including the quality of supervision.

Research into the effects of climate change and, in particular, its implications for geotechnical engineering is urgently needed (Rees et al 2009; Nathanail and Banks, 2009). The variability of influencing factors such as rainfall and pore-water pressure can be expected to increase. However, there will be significant uncertainties associated with estimates of variability in geotechnical parameters and other temporal and spatial factors. Consequently geotechnical engineers need to be equipped with better tools for dealing with variability and uncertainty. There may also be other changes in the rate at which natural processes like weathering and erosion occur. Sea level rise is another important projected consequence of global warming and climate change and it would have adverse effects on the stability of coastal slopes.

Concluding Remarks

A wide range of methods, from the simplest to the most sophisticated, are available for the geotechnical analysis of slopes. This includes both static and dynamic conditions and a variety of conditions relating to the infiltration, seepage and drainage of water. Considering regional slope stability, comprehensive databases and powerful geological models can be combined within a GIS framework to assess and use information and data relevant to the analysis of slopes and the assessment of the hazard of landsliding. The use of knowledge-based systems for assessment of failure susceptibility, hazard, or performance can be facilitated by these powerful tools. However, this must all be based on a thorough field work ethic.

It is important to understand the changes in geohazards with time. In particular, geotechnical engineers and engineering geologists will face long-term challenges due to climate change. Research is required to learn about the effects of climate change in greater detail so that methods of analysis and interpretation can be improved and extended. Exploration of such issues will be facilitated by a proper understanding of the basic

concepts of geotechnical slope analysis and the fundamental principles on which the available methods of analysis are based.

REFERENCES

1. Bhattacharya, G. and Chowdhury, R. (2011), "Continuing Research Concerning the Residual Factor as a Random Variable", Progress Report, September 2011.
2. Bowles, D., Rutherford, M. and Anderson, L. (2011), "Risk Assessment of Success Dam, California: Evaluating of Operating Restrictions as an Interim Measure to Mitigate Earthquake Risk", Geotechnical Risk Assessment and Management, Proc. GeoRisk 2011, Editors: C H Juang, K K Phoon, A J Puppala, R A Green and G A Fenton, Geo-Institute, A.S.C.E.
3. Brumund, F. (2011), "Geo –Risks in the Business Environment, Geotechnical Risk Assessment and Management", Proc. GeoRisk 2011,Editors:C H Juang, K K Phoon, A J Puppala, R A Green and G A Fenton, Geo-Institute, A.S.C.E.
4. Chowdhury, R. and Bhattacharya, G. (2011) , "Reliability Analysis of Strain-softening Slopes" , Proc. of the 13th International Conference of IACMAG, Vol. II, pp. 1169-1174, Melbourne, Australia, May, 2011.
5. Chowdhury, R. and Flentje, P. (2008), "Strategic Approaches for the Management of Risk in Geomechanics", Theme Paper, Proc. 12 IACMAG conference, Goa, India, CD-ROM, 3031-3042.
6. Chowdhury, R. and Flentje, P. (2010), "Geotechnical Analysis of Slopes and Landslides: Achievements and Challenges" , Paper Number 10054, Proc. 11th IAEG Congress of the International Association of Engineering Geology and the Environment, Auckland, New Zealand, 6 pp.
7. Chowdhury, R. and Flentje, P.(2011), "Practical Reliability Approach to Urban Slope Stability", Proc. ICASP11,the 11th Int. Conf. on Application of Statistics and Probability in Civil Engineering, August 1-4,ETH,Zurich,Switzerland,5 pp .
8. Chowdhury, R., Flentje, P. and Bhattacharya, G., (2010), "Geotechnical Slope Analysis", CRC Press, Balkema, Taylor and Francis Group, 746 pp.
9. Christian, J. T. and Baecher, G. B. (2011), "Unresolved Problems in Geotechnical Risk and Reliability", Geotechnical Risk Assessment and Management, Proc. GeoRisk 2011,Editors:C H Juang, K K Phoon, A J Puppala, R A Green and G A Fenton, Geo-Institute, A.S.C.E.
10. Flentje, P. (2009), "Landslide Inventory Development and Landslide Susceptibility Zoning in The Wollongong City Council Local Government Area", Unpublished Report to Industry Partners-Wollongong City Council, RailCorp and the Roads and Traffic Authority, University of Wollongong, Australia, 73pp.
11. Flentje, P., Chowdhury, R, Miner, A. S. and Mazengarb, C. (2010), "Periodic and Continuous Monitoring to Assess Landslide Frequency-Selected Australian Examples", Proc. 11th IAEG Congress of the International Association of Engineering Geology and the Environment, Auckland, New Zealand, 6 pp.
12. Flentje, P., Stirling, D. and Chowdhury, R. (2007), " Landslide Susceptibility and Hazard derived from a Landslide Inventory using Data Mining – An Australian Case Study", Proceedings of the First North American Landslide Conference, Landslides and Society: Integrated Science, Engineering, Management, and Mitigation. Vail, Colorado June 3-8, 2007. CD, Paper number 17823-024, 10 pages.
13. Gibson, A. D. and Chowdhury, R. (2009), Planning and geohazards, In Engineering Geology For Tomorrow's Cities, Culshaw, M.G, Reeves, H.J., Jefferson, I., and Spink, T.W., (eds.), Geological Society, London, Engineering Geology Special Publication ,Vol.22,113-123.
14. Hays, W., (2011), "Understanding Risk and Risk Reduction-a set of power point slides" , Global Alliance for Disaster Reduction (GADR).
15. Jibson, R. W., Harp, E. L. and Michael, J. A., 2000. A method for producing digital probabilistic seismic landslide hazard maps, Engineering Geology 58 (3-4): 271-289 Dec 2000.
16. Lacasse, S. and Nadim, F., 2011. Learning to Live with Geohazards: From Research to Practice Geotechnical Risk Assessment and Management, Proc.GeoRisk 2011,Editors:C H Juang, K K Phoon, A J Puppala, R A Green and G A Fenton, Geo-Institute, A.S.C.E.
17. Marker, B. R., 2009. Geology of mega-cities and urban areas, In, Engineering Geology For Tomorrow's Cities, Culshaw, M.G, Reeves, H.J., Jefferson ,I.,and Spink,T.W.,(eds.), Geological Society, London, Engineering Geology Special Publication ,Vol.22,33-48.
18. Murray, E., 2001. Rainfall Thresholds for Landslide Initiation in the Wollongong Region, Internal report to Australian Geological Survey Organisation and SPIRT Project Team at the University of Wollongong.
19. Nathanail, J.and Banks, V. 2009 Climate Change: implications for engineering geology practice 2009, In, Engineering Geology For Tomorrow's Cities, Culshaw, M.G, Reeves, H.J., Jefferson, I.,and Spink, T.W.,(eds.), Geological Society, London, Engineering Geology Special Publication ,Vol.22,65-82.
20. Quilan, R. 1993. C 4.5: Programs for Machine Learning, San Mateo, CA: Morgan.
21. Rees, J. G., Gibson, A. D., Harrison, M., Hughes, A. and Walsby, J.C., 2009. Regional modeling of geohazards change, In ,Engineering Geology For Tomorrow's Cities, Culshaw, M.G., Reeves, H.J., Jefferson, I.,and Spink, T.W., (eds.), Geological Society, London, Engineering Geology Special Publication ,Vol.22,49-64.
22. Reeves, H. J. and West, T. R., 2009. Geodata for the urban environment, In ,Engineering Geology For Tomorrow's Cities, Culshaw, M.G., Reeves, H.J., Jefferson, ,I., and Spink, T.W., (eds.), Geological Society, London, Engineering Geology Special Publication ,Vol.22,209-213.
23. Scott, G. A., 2011.The Practical Application of Risk Assessment to Dam Safety, Geotechnical Risk Assessment and Management, Proc.GeoRisk 2011,Editors:C H Juang, K K Phoon, A J Puppala, R A Green and G A Fenton, Geo-Institute, A.S.C.E.
24. Tang, W. H. and Zhang L. M., 2011. Development of a Risk-based Landslide Warning System Geotechnical Risk Assessment and Management, Proc.GeoRisk 2011,Editors:C H Juang, K K Phoon, A J Puppala, R A Green and G A Fenton, Geo-Institute, A.S.C.E.

Life Cycle Reliability and Safety Engineering
Vol. 2 Issue 2 (2012) 44-60

Two Dimensional Software Reliability Modeling and Related Allocation Problems using Genetic Algorithm

P.K.Kapur[*], Ompal Singh, Adarsh Anand
[*]Amity International Business School, Amity University, Noida, Uttar Pradesh, India
Department of Operational Research, University of Delhi, Delhi, India
(Corresponding author: pkkapur1@gmail.com)

Abstract

Modeling is a strong tool to plan the steps in development of a system. A model explains the system at different levels of abstraction. And in software reliability a most important tool that can evaluate the software quantitatively, develops test status, schedules status and monitors the changes in reliability performance is Software Reliability Growth Model (SRGM). During the last three decades, a large number of SRGMs have been proposed in literature. However, almost all of the SRGMs are developed under the assumption that software reliability growth process depends only on testing-time. Later some testing resource dependent SRGMs were also developed. Also, there exists testing coverage based SRGMs in the literature. But all these models do not take into account the simultaneous effect of time and resources or fails to consider the concurrent effect of time and coverage on cumulative number of faults removed from software. Therefore such models can be termed as one-dimensional software reliability growth models. And in order to capture the mutual effect of testing time and resources or simultaneous effect of testing time and coverage two dimensional software reliability growth models (2-D SRGM) are needed. In this paper we develop three different 2-D SRGMs, two of which take into consideration the simultaneous effect of testing time and resources and the third one is developed for concurrent effect of testing time and coverage. Further, this paper also focuses on optimal allocation decisions at unit testing level. Optimization problems which simultaneously allocate testing time and resource are also formulated and solved using Genetic Algorithm.

Introduction

21st century demands sky-scraping hardware performance and high quality software programming in order to accomplish new breakthrough in quality and productivity. Because of the integrating prospective of software, designers can bring revolutionary innovations. However the multidisciplinary scope of software has increased the overall complexities of many systems. As a result, with the increased dependency on software use, the likelihood of crisis from computer failures has also increased.

The major role for the success or failure of software depends on the condition of its testing process. The testing phase of software development life cycle is an utmost essential phenomenon for developing highly reliable software system. Software development firms that fail to execute quality control standards and effectively define the range of tests for an application can wipe out brand credibility, disrupt the overall project and generate a cost blowout. Various incidents have been stated in history where inefficiency of software testing resulted in social problems and financial losses. For example:

- News reports in December of 2007 indicated that significant software problems were continuing to occur in a new ERP payroll system for a large urban school system. It was believed that more than one third of employees had received incorrect paychecks at various times since the new system

went live the preceding January, resulting in overpayments of $53 million, as well as underpayments. An employees' union brought a lawsuit against the school system, the cost of the ERP system was expected to rise by 40%, and the non-payroll part of the ERP system was delayed. Inadequate testing reportedly contributed to the problems.

- A September 2006 news report indicated problems with software utilized in a state government's primary election, resulting in periodic unexpected rebooting of voter checking machines, which were separate from the electronic voting machines, and resulted in confusion and delays at voting sites. The problem was reportedly due to insufficient testing [5].

Hence, for optimizing software use, it becomes obligatory to concentrate on matters such as the reliability of the software products. To model software reliability models, various probabilistic and statistical approaches have been developed. By means of these tools/techniques/methods, software developers can design numerous testing programs or automate testing tools to meet the client's technical requirements, schedule and budget. These techniques can make it easier to test and correct software, detect more bugs, save more time and reduce expenses significantly [6].

There has been much effort expended in quantifying the reliability of a software system through the development of models [32]. These models are collectively called Software Reliability Growth Models (SRGMs). The main goal of these models is to fit a theoretical distribution to time between- failure data, to estimate the time-to-failure based on software test data, to estimate software system's reliability and to design a rule for determining the appropriate time to terminate testing.

During the last three decades, a large number of SRGMs have been proposed in the literature [3,19,20,21,23,26]. However, almost all of the SRGMs are developed under the conjecture that software reliability growth process depends only on testing-time as the software reliability growth factor essentially. An alternative approach based on the NHPP was proposed by Yamada et al. [34], [35], Huang and Kuo [9]. They developed some testing resource dependent SRGMs. Also, in the literature of software reliability modeling the impact of testing coverage behind the growth development during testing phase is also considered. A testing coverage based SRGM was proposed by Malaiya [37]. Inoue and Yamada [13] also developed SRGM with coverage which used a testing-coverage function to describe a time-dependent behavior or of a testing-coverage attainment process with the testing-skill of test-case designer. But all the above models do not take into account the simultaneous effect of time and resources or concurrent effect of time and coverage. Therefore such models can be termed as one-dimensional software reliability growth models. In order to capture the mutual effect of testing time and resources or simultaneous effect of testing time and coverage, two dimensional software reliability growth model (2-D SRGM) is needed. In this paper we develop three different 2-D SRGMs, two of which takes into consideration the simultaneous effect of testing time and resources and the third is developed for concurrent effect of testing time and coverage. Further, this paper also focuses on optimal allocation decisions at unit testing level. Optimization problems which simultaneously allocate testing time and resource are also formulated and solved using Genetic Algorithm.

1.1 Testing Resources in Software Reliability Modeling

In today's world the two foremost possessions after which every company, every organization and every individual is struggling for optimization are Time and Resources. The rationale being, both of them are limited and precious too. And during the testing of software these resources are one of the major factors that have an influence on determining cumulative number of faults removed from the software. It is infact a vital aspect which makes testing phase important because it is this phase only which consumes a chief portion of the total resources

available for the software development. Basically, testing activities account for 30 to 90 percent of labor expended to produce a working program [1]. Therefore, it is necessary to capture the effect of resources along with time in modeling software reliability growth models.

1.2 Testing Coverage in Software Reliability Modeling

The level of testing required usually depends on the potential consequences of undetected bugs. For evaluating the coverage of code being tested for identifying the errors a significant measure of Testing Coverage is used by software developers as it helps in evaluating the quality of the tested software. Testing Coverage determines how much additional effort is needed to improve the reliability of the software besides providing customers with a quantitative confidence criterion while planning to use a software product. Testing Coverage is a structural testing technique in which the software performance is judged with respect to specification of the source code and the extent or the degree to which software is executed by the test cases [12, 13]. Technically, it is defined as the ratio of the number of potential fault-sites sensitized by the test divided by the total number of potential fault-sites under consideration. And as mentioned above, in literature software reliability growth models relating testing coverage to software reliability have been proposed, but they fail to incorporate the concurrent effect of time and Testing Coverage in determining the cumulative number of faults removed.

There have been plenty of coverage measures proposed in literature, such as function coverage, statement coverage, branch coverage, data flow coverage, and so on. Different measures have their advantages and disadvantages when analyzing coverage [29, 22]. However, in this paper we have restricted ourselves to statement coverage only. Statement coverage (also known as line coverage and basic block coverage) tells about each execution. It reports the total number of statements (blocks) that have been executed by the test data. Its great advantage is that it is insensitive to some control structures.

2. Two Dimensional Modeling Framework

It is important to understand that the testing time as the only metric cannot give the complete picture of the reliability of the product. Therefore, in this section, we develop a two-dimensional modeling framework which incorporates the combined effect of testing time and resources/ coverage to remove the faults lying dormant in the software. In recent years, Ishii and Dohi[14] proposed a two dimensional software reliability growth model and its applications. They investigated the dependence of test-execution time as a testing effort on the software reliability assessment, and validate quantitatively the software reliability models with two-time scales. Inoue and Yamada[11] also proposed two dimensional software reliability growth model. However their modeling framework was not a direct representative of using mean value functions to represent of fault removal process. They discussed software reliability assessment method by using two dimensional Weibull-type SRGM.

The models developed in the paper are based on the Cobb Douglas production function. The Cobb–Douglas functional form [12] of production functions is extensively used to characterize the rapport of an output to inputs. It was proposed by Knut Wicksell (1851–1926), and tested against statistical evidence by Charles Cobb and Paul Douglas in 1900–1928. The function of Cobb-Douglas presents a simplified outlook of the economy in which production output is obtained by the amount of labor occupied and the amount of capital invested. While there are many factors influencing economic performance, their model demonstrated remarkable accuracy. The mathematical form of the production function is specified as:

$Y = AL^{v}K^{1-v}$

where: Y = total production (the monetary value of all goods produced in a year)

L = labor input

K = capital input

A = total factor productivity

v is elasticity of labor. This value is constant and is determined by available technology. Figure 2.1 shows graphically how the total production is influenced due to change in the proportion of labor and capital.

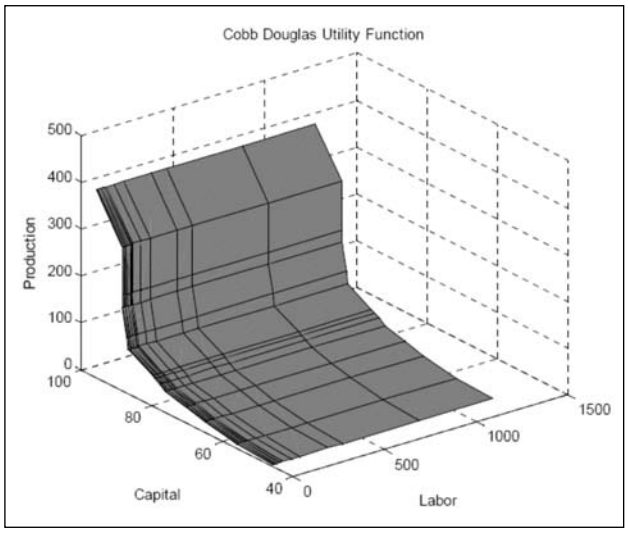


Figure 2.1: A two-input Cobb–Douglas production function

The 2-D SRGM proposed in our work incorporating testing time and resource usage is based on Non Homogeneous Poisson Process (NHPP). Let $\{N(s,u),\ s \geq 0,\ u \geq 0\}$ be a two-dimensional stochastic process representing the cumulative number of software failures by time s and with the usage of resources u. A two-dimensional NHPP with a mean value function m(s, u) is formulated as:-.

$$\Pr(N(s,u)=n)=\frac{(m(s,u))^n}{n!}\exp((-m(s,u)) \text{ , n= 0, 1,2...}$$

3. Time and Resource Dependent Two Dimensional Exponential SRGM

3.1 Model Notations

a	Initial number of faults.
b	Fault detection rate per remaining fault.
s	Testing time.
u	Resources.
α	Resource Elasticity to Testing Time
m (s, u)	Cumulative number of faults removed by time s and with the usage of resources u

3.2 Model Assumptions and Model Development

The basic assumptions of the model are as follows:

1. Failure /fault removal phenomenon is modeled by NHPP.
2. Software is subject to failures during execution caused by faults remaining in the software.
3. Failure rate is equally affected by all the faults remaining in the software.
4. On a failure, the fault causing that failure is immediately removed and no new faults are introduced.
5. To cater for the combined effect of testing time and resources, we use Cobb-Douglas production function of the following form:

$$\tau \cong s^{\alpha}u^{1-\alpha} \qquad 0 \leq \alpha \leq 1 \tag{3.1}$$

Under the above assumptions the differential equation representing the rate of change of cumulative number of faults detected w.r.t. to time and resources is given as:

$$m'(\tau)=b\left(a-m(\tau)\right) \tag{3.2}$$

Integrating above equation with initial condition m(τ =0)=0, and using Eq. (3.1) we get

$$m(s,u)=a(1-\exp(-b.s^{\alpha}.u^{(1-\alpha)})) \tag{3.3}$$

In the above two-dimensional mean value function, if α = 1 we get traditional time dependent GO model [3] and if α = 0 it becomes a resource dependent SRGM.

3.3 Parameter Estimation and Model Validation

To measure the performance of the proposed model we have carried out the parameter estimation on two data sets. First data set (DS-1) is cited from Brooks and Mootely [2]. The data set consists of 1239 faults count observed during 13 months with cumulative manpower resources of 247. Second data set (DS-2) is cited from Wood [30].The software was tested for 20 weeks which consumed 10000 manpower resources and 100

faults were removed. The estimation results for both data sets are given in Table 3.1. To check the performance of the model estimates we have compared the results with the traditional time dependent GO model [3]. The goodness of fit measures used are Mean Square Error (MSE) and Coefficient of multiple determination (R^2). The results are tabulated in table 3.2. Goodness of fit curves is shown in Figure 3.1 and Figure 3.2.

Table 3.1: Parameter Estimates

	a	b	α
DS-1			
Proposed Two dimensional model (eqn. 3.3)	2700	0.006892	0.342473
One dimensional GO model [3]	2733	0.049432	-
DS-2			
Proposed Two dimensional model (eqn. 3.3)	132	0.011457	0.687663
One dimensional GO model[3]	131	0.083166	

Table 3.2: Goodness of Fit Measures

	R^2	MSE
DS-1		
Proposed Two dimensional model (eqn. 3.3)	0.989	1655.75
One dimensional GO model [3]	0.983	2717.90
DS-2		
Proposed Two dimensional model (eqn. 3.3)	0.990	9.01
One dimensional GO model [3]	0.986	11.83

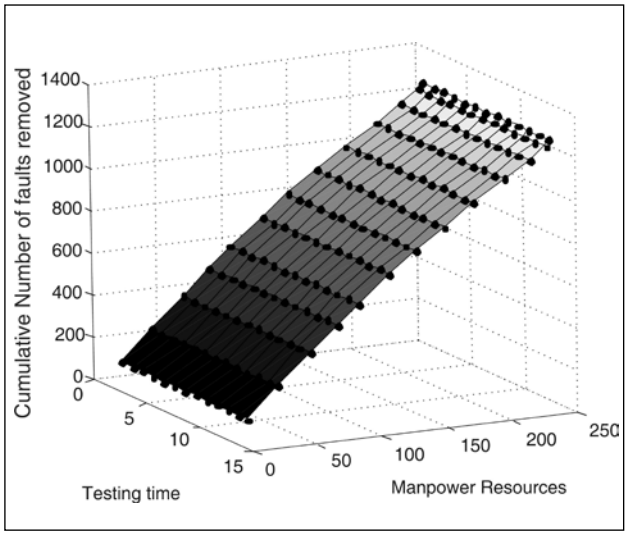


Figure 3.1: Goodness of Fit Curve (DS-1)

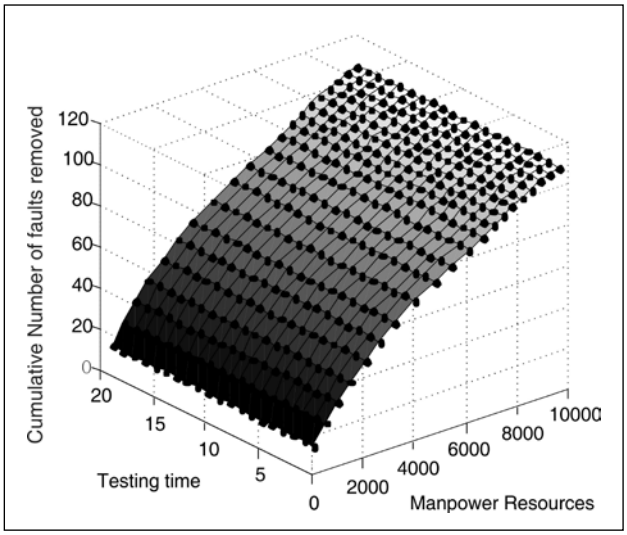


Figure 3.2: Goodness of Fit Curve (DS-2)

4. Time and Resource Dependent Two-Dimensional Model for Faults of Different Severity

4.1 Notations

a_1, a_2, a_3 Constants, representing number of simple, hard, complex faults lying dormant in the software at the beginning of testing phase .

a	Constant, representing number total faults lying dormant in the software at the beginning of testing phase(=a_1+a_2+a_3)
b_1, b_2 ,b_3	Fault detection rate for simple, hard and complex fault.
p	Proportion of simple faults lying in the software
q	Proportion of hard faults lying in the software
s/t	Testing time.
u	Resources.
α	Time Elasticity to fault removal.
m(t)	Cumulative Number of Faults Removed by time t.
m (s, u)	Cumulative number of faults removed by time s and with the usage of resources u.

4.2 Time Dependent Model with Faults of Different Severity [19]: A Review

These SRGMs take into consideration the modeling framework only with respect to time. The assumption that governs these models is, 'the software failure occurs at random times during testing caused by faults lying dormant in software.' And, for modeling the software fault detection phenomenon, counting process $\{N(t); \geq 0\}$ is defined which represents the cumulative number of software faults detected by testing time t. The SRGM based on NHPP is formulated as:

$$Pr\{N(t)=n\}=\frac{m(t)\cdot\exp(-m(t))}{n!},\quad n=0,1,2,...$$

Where m (t) is the mean value function of the counting process N (t).

Modeling Simple Faults

Simple faults are the faults which can be removed instantly as soon as they are observed. Hence Fault removal is modeled as one-stage process:

$$\frac{d}{dt}m_{11}(t)=b_1(a_1-m_{11}(t)) \tag{4.1}$$

Solving differential equation (4.1) with initial condition $m_{11}(t=0)=0$, we get:

$$m_1(t)\equiv m_{11}(t)=a_1(1-\exp(-b_1.t)) \tag{4.2}$$

Modeling Simple Faults

The hard faults consume more testing time for the removal. This means that the testing team will have to spend more time to analyze the cause of the failure and therefore requires greater time to remove them. Hence the removal process for hard faults is modeled as a two-stage process:

$$\frac{d}{dt}m_{21}(t)=b_2(a_2-m_{21}(t)) \tag{4.3}$$

$$\frac{d}{dt}m_{22}(t)=b_2(m_{21}(t)-m_{22}(t)) \tag{4.4}$$

Solving the above differential equations with initial conditions $m_{21}(t=0)=0$, $m_{22}(t=0)=0$, we get

$$m_2(t)\equiv m_{22}(t)=a_2(1-(1+b_2.t).\exp(-b_2.t)) \tag{4.5}$$

Modeling the Complex Faults

These faults require more testing time for removal after isolation as compared to hard faults. Hence they need to be modeled with greater time lag between failure observation and removal. Thus, the removal process for complex faults is modeled as a three-stage process:

$$\frac{d}{dt}m_{31}(t)=b_3(a_3-m_{31}(t)) \tag{4.6}$$

$$\frac{d}{dt}m_{32}(t)=b_3(m_{31}(t)-m_{32}(t)) \tag{4.7}$$

$$\frac{d}{dt}m_{33}(t)=b_3(m_{32}(t)-m_{33}(t)) \tag{4.8}$$

Solving the above differential equation with initial condition $m_{31}(t=0)=0$, $m_{32}(t=0)=0$, $m_{33}(t=0)=0$, we get:

$$m_3(t) \equiv m_{33}(t) = a_3(1-(1+b_3.t+\frac{b^2t^2}{2}).\exp(-b_3.t)) \quad (4.9)$$

Modeling the Total Faults

The total fault removal phenomenon is the superimposition of the simple, hard and complex faults, and is therefore given as:

$$m(t) = m_1(t) + m_2(t) + m_3(t)$$

$$m(t) = \quad a_1(1-\exp(-b_1.t)) + a_2(1-(1+b_2.t).\exp(-b_2.t)) +$$

$$a_3(1-(1+b_3.t+\frac{b_3{}^2t^2}{2}).\exp(-b_3.t)) \quad \textbf{(4.10)}$$

4.3 Time and Resource dependent Model with fault Severity: Assumptions and Development

Apart from the postulation of NHPP, the other assumptions of the time and resource dependent two dimensional model are:

1. The faults existing in the software are of three types: simple, hard and complex.
2. Fault removal process is prefect and failure observation/fault isolation/fault removal rate is constant.
3. Each time a failure occurs, an immediate (delayed) effort takes place to decide the cause of the failure in order to remove it. The time delay and more resource usage between the failure observation and its subsequent fault removal is assumed to represent the severity of the faults.
4. To cater for the combined effect of testing time and resources we use Cobb-Douglas production function of the following form:

$$\tau \cong s^{\alpha}u^{1-\alpha} \qquad 0 \le \tau \le 1 \quad (4.11)$$

Under the above assumptions and extending the dimensions from the time dependent SRGM stated in previous sub-section, to time and resource dimensions; the cumulative number of simple, hard and complex faults removed is given as:

Simple Faults

Such faults get detected and removed with no time delay and require the least resource consumption for their removal. Hence Fault removal is modeled as one-stage process:

Therefore under the above assumptions the differential equation representing the rate of change of cumulative number of faults detected w.r.t. to time and resources for simple faults is given as:

$$m_1'(\tau) = b_1\left(a - m_1(\tau)\right) \quad (4.12)$$

On solving the above equation with initial condition $m(\tau = 0) = 0$; we get

$$m_1(\tau) = a_1(1-\exp(-b_1.\tau)) \quad (4.13)$$

Using equation (4.11) in (4.13) we get the cumulative number of simple faults removed with s time and u usage of resources as:

$$m_1(s,u) = a_1(1-\exp(-b_1.s^{\alpha}u^{1-\alpha})) \quad (4.14)$$

Hard Faults

Apart from time delay in fault detection and removal, these faults need more resource usage than simple faults. Therefore on the similar basis of modeling hard faults of time dependent SRGM, we get the cumulative number of hard faults removed as:

$$m_2(\tau) = a_2(1-(1+b_2.\tau).\exp(-b_2.\tau)) \quad (4.15)$$

Substituting (4.11) in above equation we get

$$m_2(s,u) = a_2(1-(1+b_2.s^{\alpha}u^{1-\alpha}).\exp(-b_2.s^{\alpha}u^{1-\alpha})) \quad (4.16)$$

Complex Faults

In our proposed model, these faults will take highest time for removal and will consume maximum resources as compared to simple and hard faults. The mean value function giving the complex fault removal will be given by:

$$m_3(\tau) = a_3(1-(1+b_3.\tau+\frac{b_3{}^2\tau^2}{2}).\exp(-b_3.\tau)) \quad (4.17)$$

From (4.11) ; (4.17) becomes

$$m_3(s,u) = a_3(1-(1+b_3.s^{\alpha}u^{1-\alpha}+ \cdot \frac{b_3{}^2(s^{\alpha}u^{1-\alpha})^2}{2}).\exp(-b_3.s^{\alpha}u^{1-\alpha})) \quad (4.18)$$

Total Faults

The mean value function with respect to testing time and resources for total faults is:

$$m(s,u) = ap\left(1-\exp(-b_1 s^{\alpha}u^{1-\alpha})\right) + aq\left(1-(1+b_2 s^{\alpha}u^{1-\alpha})\exp(-b_2 s^{\alpha}u^{1-\alpha})\right) + a(1-p-q)\ \ 1-(1+b_3 s^{\alpha}u^{1-\alpha}+\frac{b_3{}^2\left(s^{\alpha}u^{1-\alpha}\right)^2}{2})\exp(-b_3 s^{\alpha}u^{1-\alpha}) \quad (4.19)$$

4.4 Model Validation

To measure the performance of the proposed model we have carried out the parameter estimation on the data set (DS-3) cited from Brooks and Motley [2]. The fault data set is for a radar system of size 124 KLOC (kilo lines of code) tested for 35 months in which 1301 faults were identified and resources of 1846 were consumed. The estimation result for the data set is given in Table 4.1. To check the performance of the model estimates we have compared the results with the traditional time dependent model with fault severity [19]. The goodness of fit measures used are Mean Square Error (MSE) and Coefficient of multiple determination (R^2). Based on the comparison results (large R^2, and small MSE) it is observed that the proposed model is better than its counterpart in one dimension (Table 4.2). This is due to the fact that the estimates are simultaneously based on testing time and manpower resources. The 3-D graph of goodness of fit depicting the cumulative number of faults with respect to time and resources for the data is shown in figure 4.1.

Table 4.1: Parameter Estimates for DS-3

	a	b_1	b_2	b_3	p	q	A
Time Dependent SRGM for faults of different severity (eqn. 4.10)	2416	0.001	0.06383	0.16664	0.41504	0.001	
Time And Resource Dependent SRGM for faults of different severity (eqn. 4.19)	1366	0.12457	0.01680	0.02471	0.09578	0.12402	0.48494

Table 4.2: Goodness of Fit Measures for DS-3

	R^2	MSE
Time Dependent SRGM for faults of different severity (eqn. 4.10)	0.994	43.28571
Time And Resource Dependent SRGM for faults of different severity (eqn. 4.19)	0.997	20.05714

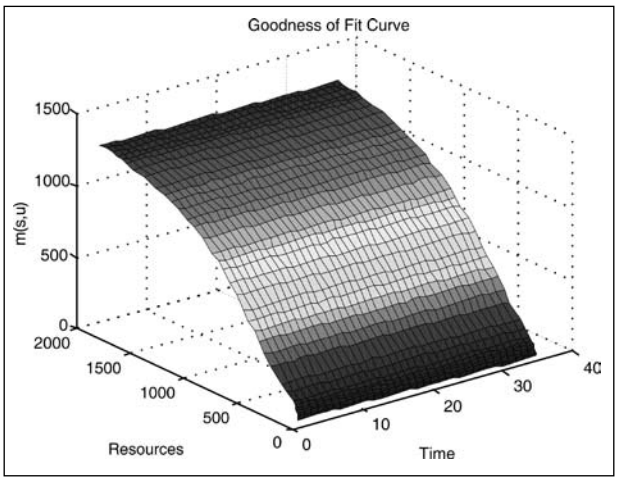


Figure 4.1 Goodness of Fit Curve for DS-3

5. Time and Coverage Dependent Two-Dimensional Model with Change Point

5.1 Change Point in Software Reliability Modeling

One of the most important factors while modeling SRGMs is the fault detection rate (FDR). FDR helps in measuring the effectiveness of fault detection by test techniques and test cases. Many SRGMs assume that this detection rate remains same throughout the testing phase. However, in practical situations it is dubious that the stability of the factor can be guaranteed during the whole process of software testing. Indeed, the characteristic of the software failure-occurrence or the fault-detection phenomenon is notably changed. When this change occurs that point is termed as change point. This would result in a software failure intensity function either increasing or decreasing monotonically [7]. The position of the Change Point can be judged by the graph of actual failure data. The work in this area started with Zhao [39] who introduced the Change Point analysis in Hardware and Software reliability. Shyur [27], Huang [7], Wang [28] also made their contributions in this area.

In this section we develop a two-dimensional model which measures the concurrent effect of time and testing coverage to remove the faults lying dormant in the software in which the FDR is changed at change point.

5.2 Notations

a	Initial number of faults in software
s	Testing time.
u	Testing Coverage.
α	Time Elasticity to fault removal
m (s, u)	Cumulative number of faults removed by time s and with coverage u
λ	$s^{\alpha}u^{1-\alpha} \quad 0 \le \alpha \le 1$
λ_0	Change Point; $s_0^{\alpha}u_0^{1-\alpha} \quad 0 \le \alpha \le 1$
b_1	Fault detection rate per remaining fault before change point.
b_2	Fault detection rate per remaining fault after change point.
β_1	Constant before change point
β_2	Constant after change point

5.3 Model Assumptions

1. On a failure, the fault causing that failure is immediately removed and no new faults are introduced.
2. Fault detection rate changes at λ0.
3. To cater the combined effect of testing time and coverage, we use Cobb-Douglas production function of the following form:
$$\lambda \cong s^{\alpha}u^{1-\alpha} \qquad 0 \le \alpha \le 1 \qquad (5.1)$$
4. The economy of scale (i.e. α) remains same before and after change point.

5.4 Model Development

Under the above assumptions the differential equation representing the rate of change of cumulative number of faults detected w.r.t. to time and usage is given as:

$$m'(\lambda) = b(\lambda)\left(a - m(\lambda)\right) \qquad (5.2)$$

Where

$$b(\lambda) = \begin{cases} \dfrac{b_1}{\left(1+\beta_1 \exp\left(-b_1\lambda\right)\right)} & for\ \lambda \le \lambda_0 \\ \dfrac{b_2}{\left(1+\beta_2 \exp\left(-b_2\lambda\right)\right)} & for\ \lambda > \lambda_0 \end{cases} \qquad (5.3)$$

Case 1: For $\lambda \le \lambda_0$

Solving (5.2) with the initial condition $m(\lambda = 0) = 0$ and using equation (5.1) we get

$$m(s,u) = a\left[1 - \frac{1+\beta_1}{1+\beta_1 \exp(-b_1.(s^{\alpha}u^{1-\alpha}))}\exp(-b_1.(s^{\alpha}u^{1-\alpha}))\right] \qquad (5.4)$$

Case 2: For $\lambda > \lambda_0$

Solving (5.2) with the initial condition $m(\lambda = \lambda_0) = m(\lambda_0)$ and using equation (5.1) we get

$$m(s,u) = a\left[1 - \left(\frac{(1+\beta_1)}{\left(1+\beta_1 \exp\left(-b_1.(s_0^{\alpha}u_0^{1-\alpha})\right)\right)}\right)\left(\frac{\left(1+\beta_2 \exp\left(-b_2.(s_0^{\alpha}u_0^{1-\alpha})\right)\right)}{\left(1+\beta_2 \exp\left(-b_2.(s^{\alpha}u^{1-\alpha})\right)\right)}\right) e \exp\left(-b_1.(s_0^{\alpha}u_0^{1-\alpha}) - b_2(.(s^{\alpha}u^{1-\alpha}) - .(s_0^{\alpha}u_0^{1-\alpha})\right)\right] \qquad (5.5)$$

Combining Case1 and Case 2 we get m(s,u) as:

$$m(s,u)=\begin{cases} a\left[1-\dfrac{1+\beta_1}{1+\beta_1\exp(-b_1.(s^{\alpha}u^{1-\alpha}))}\exp(-b_1.(s^{\alpha}u^{1-\alpha}))\right] & for\ \lambda\le\lambda_0 \\ a\left[1-\left(\dfrac{(1+\beta_1)}{\left(1+\beta_1\exp\left(-b_1.(s_0^{\alpha}u_0^{1-\alpha})\right)\right)}\right)\left(\dfrac{\left(1+\beta_2\exp\left(-b_2.(s_0^{\alpha}u_0^{1-\alpha})\right)\right)}{\left(1+\beta_2\exp\left(-b_2.(s^{\alpha}u^{1-\alpha})\right)\right)}\right)\right] \\ \left.\begin{array}{l} \\ \\ \end{array}\right\} e\exp\left(-b_1.(s_0^{\alpha}u_0^{1-\alpha})-b_2(.(s^{\alpha}u^{1-\alpha})-.(s_0^{\alpha}u_0^{1-\alpha})\right)\Bigg] \end{cases} \quad (5.6)$$

5.5 Model Validation

We have carried out the parameter estimation on two data sets. First data set (DS-4) is cited in Malaiya et al [37]. The data set is Coverage Data set with 796 test cases and 9 cumulative numbers of faults removed with block coverage as 95.99%. Second data set (DS-5) is also from Malaiya et al [37]. The overage data set consists of 9 cumulative faults removal with 1196 test case covering 95.97% of the block coverage. To check the performance of the model estimates we have compared the results with the traditional time dependent flexible SRGM by Kapur and Garg[15]. The goodness of fit measures used are Mean Square Error (MSE), and Coefficient of multiple determination (R^2).

DS-4

The Change point for DS-4 is obtained at (44, 87) (see figure 5.1).

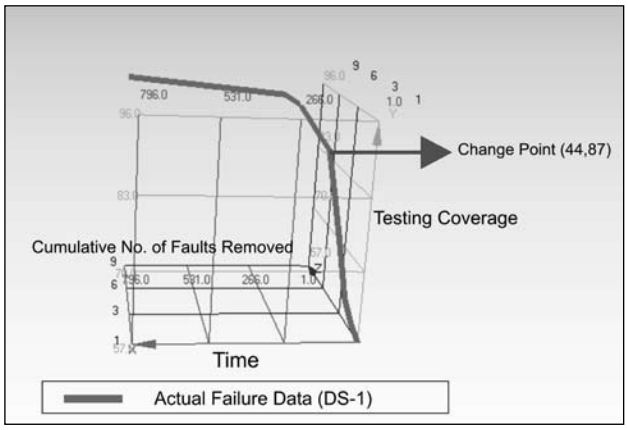


Figure 5.1: Actual Failure Data Set (DS-4)

The parameter estimates results for DS-4 is tabulated in Table 5.1. The comparison result for the data is shown in Table 5.2. The 3-D graph of goodness of fit depicting the estimated cumulative number of faults with respect to time and coverage is shown in figure 5.2

Table 5.1: Parameter Estimates for DS-4

	Proposed Two Dimensional Model (eqn. 5.6)	One Dimensional KG Model [15]
a	10	9
b_1/b	0.031884	0.052937
b_2	0.02248	-
β_1/ β	1.028487	0.001
β_2	0.001	-
α	0.523349	-

Table 5.2: Goodness of Fit Measures for DS-4

	Proposed Two Dimensional Model (eqn.5.6)	One Dimensional KG Model[15]
R^2	0.991	0.922
MSE	0.06847	0.5823

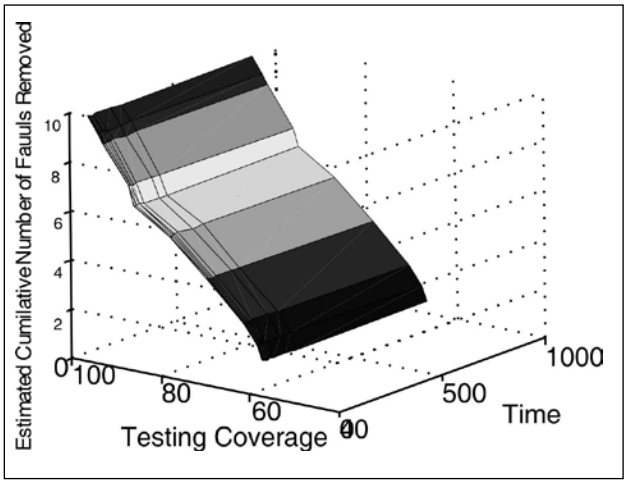


Figure 5.2: Goodness of Fit Curve(DS-4)

DS-5

The Change point for DS-1 is obtained at (20, 70.5) (see figure 5.3).

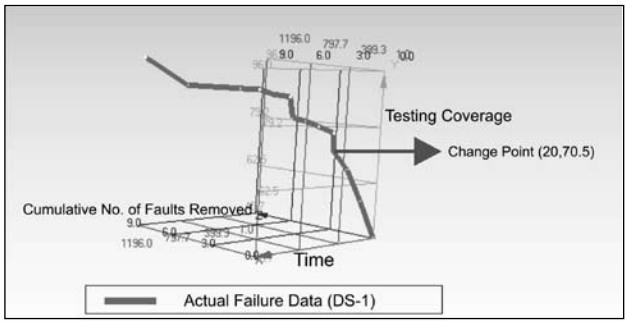


Figure 5.3: Actual Failure Data Set (DS-5)

The parameter estimates results and goodness of fit measures for DS-5 is tabulated in Table 5.3 and 5.4 respectively. The 3-D graph of goodness of fit depicting the estimated cumulative number of faults with respect to time and coverage is shown in figure 5.4.

Table 5.3: Parameter Estimates for DS-5

	Proposed Two Dimensional Model (eqn.5.6)	One Dimensional KG Model [15]
a	10	9
b_1/b	0.016475	0.018364
b_2	0.022992	-
β_1/ β	1.576795	0.001
β_2	2.074583	-
α	0.288297	-

Table 5.4: Goodness of Fit Measures for DS-5

	Proposed Two Dimensional Model (eqn.5.6)	One Dimensional KG Model [15]
R^2	0.973	0.947
MSE	0.2094714	0.4177143

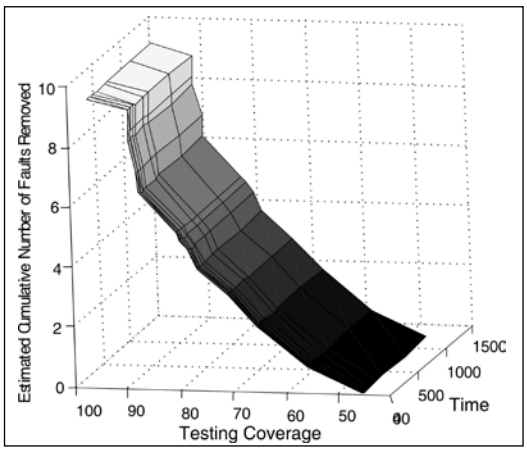


Figure 5.4: Goodness of Fit Curve (DS-5)

Based on the comparison results (large R^2, large adjusted R^2, small bias and small MSE) it is observed that the proposed model is better than its counterpart in one dimension. This is due to the fact that the estimates are concurrently based on testing time and testing coverage.

6. Simultaneous Optimal Allocation of Time and Resources for Modular Software

Allocation decisions are critical for the software systems. Moreover, the objective behind such critical decisions can vary from firms to firms. The motive of the firm can be maximization of software reliability or maximization of number of faults to be removed from each module or it can be minimization of number of faults remaining in the software or minimization of software development cost. Taking into consideration these different aims, various authors have investigated the optimal time or optimal resource allocation problems. Xie and Yang [31] studied problem of optimal testing-time allocation for modular software systems with the aim to maximize the operational reliability of a simple software system. Ohetera and Yamada [24] developed and solved two resource allocation problems for modular software, minimizing the mean number of remaining faults in the software modules when the amount of available testing resources is previously specified and vice versa. Yamada et al. [33] further studied the allocation problem minimizing the mean number of remaining faults in the software modules with a reliability aspiration and budget constraint. Huo et al. [10] determined the optimal amount of resources needed for software module testing using the hyper-geometric software reliability growth model. Kapur et al. [16] discussed the testing resource allocation problem to maximize the total fault removal from software consisting of several independent components. And for the resulting optimization problem, they defined marginal testing effort function (MTEF), where the testing resource consumption was represented in terms of fault removal. Further, Kapur et al.[17, 18] studied various resource allocation problems maximizing the number of faults removed form each module under constraint on budget and

management aspirations on reliability using exponential and S-shaped SRGMs [3, 19]. They have discussed dynamic, mathematical and goal programming approaches to yield solutions of such class of optimization problems. Huang et al. [8] investigated an optimal resource allocation problem in modular software systems where the failure phenomenon of each of the modules was described by an exponential curve during testing phase. The main purpose was to minimize the cost of software development when the number of remaining faults and a desired reliability objective are given.

The above mentioned literature takes into consideration only one aspect of real life situation problems-either time or resource. However, for a software development firm it would be more advantageous if they could concurrently allocate time and resources optimally to modular software product. Our proposed mathematical optimization problem offers this advantage to the development organizations. In this paper we investigate two dimensional optimization problems which assigns testing time and manpower resources among the modules simultaneously.

The resulting problems are solved using genetic algorithm. Genetic Algorithms (GA) stand up a powerful tool for solving search and optimization problems. GA always considers a population of solutions that offers a lot of advantages. GA has been used to solve many difficult engineering problems and is particularly effective for combinatorial optimization problems with large, complex search spaces Zaki et.al [38]. The GA has been applied to the reliability allocation problem of a typical pressurized water reactor in nuclear power plants in Yang et.al [36] and has also been used in system reliability by Painton and Campbell [25].

6.1. Genetic Algorithm

Genetic algorithms are described by the natural process of evolution, which is a rapidly growing field of artificial intelligence. GA is inspired by the theory of Charles Darwin about the natural evolution in the origin of species. Goldberg [4] gave the introduction of GA. For implementing the GA in solving the two dimensional simultaneous time and resource allocation problems, the following steps are considered.

Step 1: Chromosome Representation

Genetic Algorithm starts with the initial population of solutions represented as chromosomes. A chromosome consists of genes where each gene represents a specific attribute of the solution. In our problem each chromosome of length 2N is taken (Figure 5.1). It is divided into two parts. First N genes corresponds to testing time (s_i, i=1,2,…N) and last N to resources(u_i, i=1,2,…N).

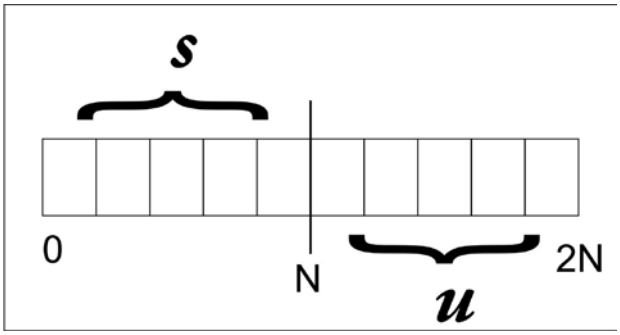


Figure 6.1: Chromosome Representation for two dimensional problem

Step 2: Initial Population

For a given total testing time S and total manpower resources U, GA generates the initial population randomly. It initialize to random values within the limits of each variable. Here, we for the first N variables the limit is with respect to testing time and for the last N variables the limit is with respect to manpower resources

Step 3: Fitness of a Chromosome

The fitness is a measure of the quality of the solution it represents in terms of various optimization parameters of the solution. A fit chromosome suggests a better solution. In the allocation problem (P1), the fitness function is the objective of optimization problem along with the penalties of the constraints that are not met.

Step 4: Selection

Selection is the process of choosing two parents from the population for crossover. The

higher the fitness function, the more chance an individual has to be selected. The selection pressure drives the GA to improve the population fitness over the successive generations. Selection has to be balanced with variation form crossover and mutation. Too strong selection means sub optimal highly fit individuals, will take over the population, reducing the diversity needed for change and progress; too weak selection will result in too slow evolution. We use Tournament selection without replacement here.

Step 5: Crossover

Crossover is the process of taking two parent solutions and producing two similar chromosomes by swapping sets of genes, hoping that at least one child will have genes that improve its fitness. In our two dimensional problem the first N genes of a chromosome are crossed over with the first N genes of other selected chromosome. The crossover of the last N genes of the same chromosome takes place with the last N genes of other selected same chromosome.

Step 6: Mutation

Mutation prevents the algorithm to be trapped in a local minimum. Mutation plays the role of recovering the lost genetic materials as well as for randomly disturbing genetic information. Mutation in our case is done for first N testing time genes and then of the last N genes but on the same selected chromosome.

The steps 3 to 6 are then repeated till the stopping criteria of maximum generation number is reached.

6.2 Simultaneous Time Resource Allocation Formulation for Modular Software

In this sub section we formulate two allocation problems for modular software systems. The first allocation problem uses time and resource dependent exponential SRGM developed in section 3. The second optimization problem optimally allocates time and resources when severity of faults is also taken into consideration.

6.2.1 Concurrent Time and Allocation -Problem 1

This optimization problem aims at minimization of the software development cost under the constraints of limited total testing time available and restricted manpower resources in the hands of the software development firm. Further, it is aspired that at least a pre-defined proportion of faults from each module should be removed during module testing phase.

Modeling the Cost Function

The cost function modeled in this research includes the following cost:

(i) The cost of removing a fault during testing phase (C_1).

(ii) The cost of removing a fault after during operational phase (C_2).

(iii) The unit cost of testing (C_3).

Using (i)-(iii) the cost expression for the modular software system is modeled as:

$$C(S,U) = \sum_{i=1}^{N} C_{1i} m_i(s_i,u_i) + \sum_{i=1}^{N} C_{2i}(a_i - m_i(s_i,u_i)) + C_3 \sum_{i=1}^{N} s_i^{\alpha_i}.u_i^{(1-\alpha_i)} \quad (6.1)$$

The mean value function of fault removal process for the i^{th} module, i=1,2,…N is obtained using the modeling framework given in section 3, i.e.

$$m_i(s_i,u_i) = a_i(1-\exp(-b_i.s_i^{\alpha_i}.u_i^{(1-\alpha_i)})) \quad \forall i = 1,2,...N \quad (6.2)$$

Substituting Eq.(6.2) in (6.1) we get

$$C(S,U) = \sum_{i=1}^{N} C_{1i} a_i(1-\exp(-b_i.s_i^{\alpha_i}.u_i^{(1-\alpha_i)})) + \sum_{i=1}^{N} C_{2i}(a_i - a_i(1-\exp(-b_i.s_i^{\alpha_i}.u_i^{(1-\alpha_i)}))) + C_3 \sum_{i=1}^{N} s_i^{\alpha_i}.u_i^{(1-\alpha_i)}$$

The problem of optimal allocation of testing time and manpower resources to each of the

N independent modules with the objective of minimizing the total expected software cost given by Eq. (6.3), such that at least p_i proportion of the faults are removed from each module with S to be the total testing time and U to be total manpower resources expenditure available is formulated as:

$$Min\ C(S,U) = \sum_{i=1}^{N} C_{1i} a_i (1-\exp(-b_i . s_i^{\alpha_i} . u_i^{(1-\alpha_i)})) + \sum_{i=1}^{N} C_{2i} (a_i - a_i (1-\exp(-b_i . s_i^{\alpha_i} . u_i^{(1-\alpha_i)}))) + C_3 \sum_{i=1}^{N} s_i^{\alpha_i} . u_i^{(1-\alpha_i)}$$

$$\sum_{i=1}^{N} s_i \le S$$

$$\sum_{i=1}^{N} u_i \le U$$

$$m_i(s_i,u_i) = a_i(1-\exp(-b_i . s_i^{\alpha_i} . u_i^{(1-\alpha_i)})) \ge p_i a_i \quad \forall i = 1,2,...N$$

6.2.2 Concurrent Time and Allocation -Problem 2

The problem of optimal allocation of testing time and manpower resources to each of the N independent modules with the objective of removing maximum number of faults such that at least μ_i proportion of the faults are removed from each module with S to be the total testing time and U to be total manpower resources expenditure available is formulated as:

$$Max\ m(s,u) = \sum_{i=1}^{N} m_i(s_i,u_i)$$

$$= \sum_{i=1}^{N} a_i p_i \left(1-\exp(-b_{1i} s_i^{\alpha_i} u_i^{1-\alpha_i})\right) + \sum_{i=1}^{N} a_i q_i \left(1-(1+b_{2i} s_i^{\alpha_i} u_i^{1-\alpha_i})\exp(-b_{2i} s_i^{\alpha_i} u_i^{1-\alpha_i})\right) +$$

$$\sum_{i=1}^{N} a_i (1-p_i-q_i)\ 1-(1+b_{3i} s_i^{\alpha_i} u_i^{1-\alpha_i} + \frac{b_{3i}^2 \left(s_i^{\alpha} u_i^{1-\alpha_i}\right)^2}{2})\exp(-b_{3i} s_i^{\alpha_i} u_i^{1-\alpha_i})$$

Subject to:

$$\sum_{i=1}^{N} s_i \le S$$

$$\sum_{i=1}^{N} u_i \le U$$

$$m_i(s_i,u_i) \ge \mu_i a_i \quad ; \quad \forall i = 1,2,...N$$

$$s_i \ge 0; \quad u_i \ge 0 \ \ i = 1,2...N$$

6.3 Numerical Illustration

6.3.1 Numerical example of the Simultaneous Allocation of Testing Time and Manpower Resources for Allocation Problem 1

Consider modular software system having three modules; each is modeled under the modeling framework given in section 3. The parameters of each module are assumed to be estimated. These estimates are given in table 6.1. The total testing time is assumed to be 1000 hours. The total resources available are assumed to be 60000. The cost parameters for each module is taken to be same only for the sake of numerical example, i.e. $C_{1i}=C_1=5$; $C_{2i}=C_2=10$ and $C_3=2$, i=1,2,3. Also, it is desired that at least 90% of the faults are to be removed from each module.

Table 6.1: Parameter Estimates used in Simultaneous Allocation of Testing time and Resources Problem

Module	a	b	α
M1	218	0.000647	0.479
M2	105	0.001055	0.443
M3	361	0.000574	0.398
	Total Number of faults: 684		

The problem is solved using genetic algorithm proposed using Microsoft Visual C++ 6.0 programming software. The parameters used in GA evaluation are:

Population Size: 250

Number of Generations: 100

Selection Method: Tournament Selection without Replacement

Crossover: Simulated Binary Crossover (SBX) with crossover probability as 0.9.

Mutation: Polynomial mutation with probability of mutation as 0.1.

The optimal allocation of testing time and resources among the modules based upon the above information is shown in Table 6.2. The total testing time spent during module testing came to be 998.24 hours and the total resources allocated were obtained as 59464.55. The optimal

(minimum) cost of testing for the numerical illustration of (P1) came out to be 23308.08. The number of faults removed and remaining and the proportion of the faults removed from each module within the available time and resources is given in Table 6.3

Table. 6.2: Optimal allocation of Testing Time and Resources for modules

Module	Testing Time	Manpower Resources
M1	457.28	23553.58
M2	284.98	11035.67
M3	255.98	24875.30

Table 6.3: Number of faults removed and remaining and the proportion of the faults removed from each module

Module	a	m	a-m	Proportion of faults removed
M1	218	197	21	0.90367
M2	105	95	10	0.904762
M3	361	325	36	0.900277
Total	684	617		

6.3.2 Numerical example of the Simultaneous Allocation of Testing Time and Manpower Resources for Allocation Problem 2

Consider modular software system having three modules; each is modeled under the modeling framework given in section 4. The parameters of each module are assumed to be estimated. These estimates are given in Table 6.4. The total testing time is assumed to be 2000 hours. The total resources available are assumed to be 75000. Also, it is desired that at least 90% of the faults are to be removed from each module.

The parameters used in GA evaluation are:

Population Size: 250

Number of Generations: 120

Selection Method: Tournament Selection without Replacement

Crossover: Simulated Binary Crossover (SBX) with crossover probability as 0.95.

Mutation: Polynomial mutation with probability of mutation as 0.1.

The optimal allocation of testing time and resources among the modules based upon the above information is shown in Table 6.5. The total testing time spent during module testing came to be 1999.068 hours and the total resources allocated were obtained as 74812.293.

Table 6.4: Parameter Estimates used in Simultaneous Allocation of Testing time and Resources Problem

Module	a	b_1	b_2	b_3	p	Q	α
M1	266	0.001245	0.001680	0.002471	0.09578	0.02402	0.4879
M2	105	0.001450	0.001055	0.003432	0.07175	0.05431	0.5143
M3	200	0.002171	0.008055	0.001547	0.02934	0.05908	0.3298
	Total Number of faults: 571						

Table 6.5: Optimal Simultaneous Allocation of Testing time and Resources

Module	Time	Resource	Simple Faults Removed	Hard Faults Removed	Complex Faults Removed	Total Faults Removed	Proportion of faults Removed
M1	974.9669	27325.142	25	6	234	265	0.99
M2	806.7409	33603.24	7	5	91	103	0.98
M3	217.3606	13883.91	5	11	165	181	0.91
	1999.068	74812.293				549	

7. Conclusion and Future Scope

This paper proposes three two dimensional software reliability growth models two of which cater the simultaneous effect of testing time and resources and the third one is developed under the assumption that the fault removal rate is dependent concurrently on testing time and testing coverage. All the models proved to be better when compared with their one dimensional testing time dependent models. Two simultaneous allocation problems of testing time and resources have also been formulated and solved in the paper. For solving the complex non linear two dimensional optimization problems meta heuristic technique of Genetic Algorithm has been employed.

In this paper it is assumed that the model is developed under the perfect debugging environment. The overcoming of this limitation in modeling forms a scope of future research. We have focused only on two dimensional frameworks in this work. In future we can explore the possibility of including multi dimensional software reliability growth modeling so as to take care the effect of not only testing coverage but also other testing factors like testing effort, testing time/number of test cases on the fault removal process simultaneously. One of the functional forms that can be used for this purpose is the multi dimension extension of Cobb Douglas function given by:

$$\tau \cong {s_1}^{\alpha_1} {s_2}^{\alpha_2} \ldots . {s_n}^{\alpha_n}$$

$$where \sum_{i=1}^{n} \alpha_i = 1 \; ; \; \alpha_i \geq 0 \; \forall i$$

8. Acknowledgement

The research work presented in this paper is supported by grants to the first author from Department of Science and Technology (DST) Grant No SR/S4/MS: 600/09, India.

References:

1. Beizer B., Software Testing Techniques. Boston: International Thomson Computer Press, 1990
2. Brooks W.D. and Motley R.W, ''Analysis of Discrete Software Reliability Models'', Technical Report, RADC-TR-80-84, Rome Air Development Center, New York, USA, 1980.
3. Goel A.L. and Okumoto, K., "Time dependent error detection rate model for software reliability and other performance measures", IEEE Transactions on Reliability, 1979, R-28, 206–211.
4. Goldberg, D.E., "Genetic Algorithms in Search of Optimization and Machine Learning". Addison-Wesley, 1989.
5. http://www.sereferences.com/software-failure-list.php
6. Huang C., Kuo S., and Lyu M. R.. Optimal Software Release Policy Based on Cost and Reliability with Testing Efficiency. International Computer Software and Applications Conference, (COMPSAC), pages 468–473, 1999.
7. Huang C.Y., "Performance Analysis of Software Reliability Growth Models With Test Efforts And Change-Point" Journal of Systems and Software Vol. 76, pp. 181-194, 2005
8. Huang C.Y., Lo J.H., Kuo S.K. and Lyu M.R., "Optimal allocation of testing resources considering cost, reliability, and testing –effort", Proceedings of the 10th IEEE Pacific International Symposium on dependable Computing, 2004.
9. Huang, C.Y. and Kuo, S.Y., Analysis of incorporating logistic testing-effort function into software reliability modeling, IEEE Transactions on Reliability, 2002, 51 (3), 261–270.
10. Huo R.H., Kuo S.K., Chang Y.P., "Needed resources for software module test, using the hyper-geometric software reliability growth model", IEEE Trans. on Reliability, 1996, 45(4), 541-549.
11. Inoue S, Yamada S "Two-Dimensional Software Reliability Assessment with Testing-Coverage" Second International Conference on Secure System Integration and Reliability Improvement,2008, July 14-July 17.
12. Inoue S, Yamada S., "Two-Dimensional Software Reliability Measurement Technologies", In the proceedings of IEEE IEEM, 2009.
13. Inoue S., Yamada S., "Testing-Coverage Dependent Software Reliability Growth Modeling" International Journal of Reliability, Quality and Safety Engineering, Vol. 11, No. 4, 303-312, 2004.
14. Ishii T. and Dohi T., "Two-dimensional software reliability models and their application," Proc. 12th Pacific Rim Intern. Symp. Dependa. Comput., 2006, 3–10
15. Kapur P. K. and. Garg R. B., "Software reliability growth model for an error-removal phenomenon," Software Engineering Journal, Vol. 7, No. 4, pp. 291–294, 1992.
16. Kapur P.K., Bardhan A., Yadavalli V., "On allocation of resources during testing phase of a modular software", International Journal of Systems Science, 2007, 38(6), 493-499.
17. Kapur P.K., Jha P., Bardhan A., "Dynamic programming approach to testing resource allocation problem for modular software", Ratio Mathematica, Journal of Applied Mathematics, 2003, 14, 27-40.
18. Kapur P.K., Jha P., Bardhan A., "Optimal allocation of testing resource for a modular software", Asia Pacific Journal of Operational Research, 2004, 21(3), 333-354
19. Kapur, P.K., Garg, R.B. and Kumar, S., Contributions to Hardware and Software Reliability, World Scientific: Singapore 1999.
20. Lyu M.R., Handbook of Software Reliability Engineering, McGraw-Hill, New York, 1996.
21. Musa J.D., Iannino A., Okumoto K., Software Reliability:

Measurement, Prediction, Application, McGraw-Hill, New York, 1987.

22. Ntafos S.C., "A Comparison of Some Structural Testing Strategies", IEEE Trans. Software Engineering, Vol. 14, No. 6, pp. 868-874, 1988.
23. Obha, M., "Software reliability analysis models",IBM Journal of Research and Development, 1984, 28, 428–443.
24. Ohetera H. and Yamada S., "Optimal allocation and control problems for software testing resources", IEEE Transactions on Reliability, 1990, 39 (2), 171-176.
25. Painton, L., Campbell, J., "Genetic algorithms in optimization of system reliability". IEEE Transactions on Reliability, 1995, 44 (2), 172–178.
26. Pham H., Software Reliability, Springer-Verlag, Singapore, 2000.
27. Shyur H.J., "A stochastic software reliability model with imperfect-debugging and change-point," Journal of Systems and Software, vol. 66, pp. 135–141, 2003
28. Wang Z. and Wang J. "Parameter Estimation of Some NHPP Software Reliability Models With Change-Point," Communications in Statistics- Simulation and Computation, vol. 34, pp. 121-134, 2005
29. Weiser M., Gannon J.D., McMullin P.R., "Comparison of Structural Test Coverage Metrics", IEEE Software, Vol. 2, No. 2, pp 80-85, 1985.
30. Wood A., ''Predicting Software Reliability'', IEEE Computers, 1996, 11, 69–77.
31. Xie M., and Yang B., "Optimal Testing time Allocation for Modular Systems", International Journal of Quality and Reliability Management, 2001, 18(4), 854-863.
32. Xie M.. Software Reliability Modelling. World Scientific, Singapore, 1991.
33. Yamada S. Ichitmori T. Nishiwaki M., "Optimal allocation policies for testing-resource based on a software reliability growth model", Mathematical and Computer Modelling, 1995, 22(10-12), 295-301.
34. Yamada, S., Hishitani, J. and Osaki, S., Software-reliability growth with a Weibull test-effort, IEEE Transactions on Reliability, 1993, 42 (1), 100–106.
35. Yamada, S., Ohtera, H. and Narihisa, H., Software reliability growth models with testing effort, IEEE Transactions on Reliability,1986, 35 (1), 19–23.
36. Yang, J.E., Hwang, M.J., Sung, T.Y., and Jin, Y., "Application of genetic algorithm for reliability allocation in nuclear power plants". Reliability Engineering Systems Safety, 1999, 65 (3), pp. 229–238.
37. Yashwant K. Malaiya, Michael Naixin, M Bieman. James and Rick Karcich, "Software Reliability Growth with Testing Coverage", IEEE Transactions on Reliability, Vol. 51, No. 4, pp. 420-426, 2002
38. Zaki, M., El-Ramsisi, A., and Omran, R., "A soft computing approach for recognition of occluded shapes". J. Systems. Software, 2000, 51(1), 73–83.
39. Zhao M., "Change-Point Problems In Software And Hardware Reliability", Communications in Statistics-Theory and Methods, Vol. 22, No. 3, pp. 757-768, 1993.

Life Cycle Reliability and Safety Engineering
Vol.2 Issue 2 (2012) 61-70

Reliability and Imbalance Modeling of a Low Pressure Turbine Rotor

V. M. S. Hussain and V. N. A. Naikan
Reliability Engineering Centre, Indian Institute of Technology, Kharagpur, India
Corresponding author: naikan@hijli.iitkgp.ernet.in, Phone: +919434722566

Abstract:

Rotary systems are subjected to several dynamic forces such as imbalance, misalignment, oil whirl, bend shaft, mechanical looseness etc. The focus of this paper is on development of a new model for reliability assessment of a rotating system having imbalance and its effect on reliability. The objective of this work is also to establish the safe and critical limits of imbalance as well as rotational speed for achieving the target reliability. A simulation based solution methodology is used for solving the problem. A case study of a turbine main shaft made of an alloy steel 2.5 Ni-Cr-Mo-V is provided for demonstrating the practical application.
Key words: ***Rotary system, reliability, imbalance, stress-strength interference, simulation***

1. Introduction:

The focus of this research work is on the reliability studies of mechanical systems, with special emphasis on rotary systems. Mechanical systems can be classified as stationary, reciprocating, sliding and rotational. The type of forces and the failure mechanisms acting on these systems are also of different nature. Therefore, reliability studies on mechanical system require different approaches depending on the system.

In this paper a mathematical model based on the simulation methodology has been developed for the assessment of reliability of a low pressure turbine shaft made of alloy steel 2.5 Ni-Cr-Mo-V. This section provides a brief literature review on the failure mechanisms of mechanical systems, failure models, stress-strength models, reliability and imbalance in rotary systems.

1.1 Reliability of Mechanical Systems

The approaches for modeling mechanical system reliability have been discussed by several authors. Carter [1] opined that the mechanical system reliability mostly relies on the factor of safety or safety margins based on the empirical methods and experience, rather than on the scientific and statistical approach. Yoshikawa [3] also has pointed out that due to complexities of mechanical systems; conventional reliability theory is not adequate. Carter [4, 5] did extensive research on the reliability of the components under fatigue and wear-out (degradation) failures and shown that due to degradation in strength the failure rate steadily increases. The author has shown that a modified Weibull distribution can be used to represent the early life failure patterns in mechanical systems.

Application of stress-strength interference technique using moments and Monte-Carlo simulation by Keceoglu [2], mechanical reliability review by Martin *et al* [6], a method to evaluate the reliability of the components by vibration testing by Chegdaev and Samsonov [7], a multi-state fault tree method by Charlesworth and Rao [8], a model based on the Esscher method by Radhakrishna [9], a system-reliability model for mechanical systems based on the graph theory and Boolean function by Tang [10], a method to implement the reliability analysis at the conceptual design phase by Avontuur and Weff [11], a digraph method to find the reliability of the interacting tribo pair by Sharma and Gandhi [12] are several recent approaches for modeling mechanical systems reliability.

Modelling of degradation failure mechanisms using stress-strength technique by Srinath [36],

Dasgupta and Pecht [37], modelling of fatigue, diffusion, inter-diffusion, creep, corrosion, wear, Radiation damage, large elastic deformation, yield, buckling, fracture by Dasgupta and Pecht [37, 38], inter-diffusion and creep failures by Li and Dasgupta [39, 40], wear modeling by Engel [41], large elastic deformation by Dasgupta and Hu [42], buckling problem by Dasgupta and Haslach [43] fracture processes by Dasgupta and Hu [44] are significant contributions in modeling of mechanical failure mechanisms. Dasgupta and Pecht [37] have classified the failure mechanisms into four types of failure models such as Stress – Strength interference (SSI) models, Damage – Endurance models, Challenge – Response model, Tolerance – Requirement models. They have presented several examples for these models.

From the afore discussed research works of Dasgupta and other authors [36 – 44], it is identified that while listing the conceptual models for failure states they stated that the Stress–Strength interference (SSI) model is suitable for the failure analysis of mechanical systems and components where strength and stress applied are often treated as the random variables. A critical analysis of the above literature shows that technical issues arising out of the several dynamic problems of mechanical systems are not considered for reliability modeling studies. The literature on imbalance in rotary systems is presented in the next sub-section.

1.2 Rotating systems

Othman *et al* [13] have stated that rotating machinery are vulnerable to varieties of problems originating from rotating shafts, gears, pumps etc. and a little defect such as a crack or chip interference could lead to a major disaster. The authors stress that the mass imbalance, shaft misalignment, and the improper surface finish need to be addressed to avoid disasters. The mass imbalance in the rotating machines causes vibrations which in turn lead to unwanted motion of the rotors. Rotor-to-stator rub is one such problem arises due to the vibration in rotating machinery. Muszynska [14, 15], Goldman and Muszynska [16], Chu and Lu [17], Jiang *et al* [18] studied about rotor-to-stator rub and published their research findings. Collacot [19] has correctly stated that the forces generated due to the imbalance of rotor can be very large especially at high rpm of rotor. The author also suggested the need of further studies on this.

Lum *et al* [20], Zhou and Shi [21], Hredzak and Guo [22], Lee *et al* [23] and many other authors such as DeSmidt [24], Oladejo *et al* [25], Horvarth *et al* [26], Ohtomi *et al* [27], Bansal *et al* [28], Lum *et al* [29 – 32], Manchala *et al* [33, 34], Kang [35] have done extensive work on balancing of rotors and published their research findings regarding the balancing of rotating systems. A critical review of this literature reveals that most of the works carried in this area are of two types. The first one is on effect of rotor-to stator-rub due to imbalance, and the second one is on balancing the rotary systems. No publication is found on reliability modelling of rotary systems with mass imbalance. However several authors have suggested the need of such work. The authors of the present paper felt that SSI models can help to develop such models. Therefore, the literature on stress-strength models is reviewed in the next section.

1.3 Stress – Strength interference models

As mentioned by Kotz et al. [45] the germ of the stress-strength inference theory was introduced by Birnbaum [46]. The theory is further developed by Birnbaum and McCarty [47]. Owen *et al* [48] constructed confidence limits for $P(X<Y)$, when X and Y are dependent or independent normally distributed random variables. Many authors such as Kapur and Lamberson [49], Dhillon [50], Chung [51], Reiser and Guttman [52, 53], Dargahi-Nourbary [54], Murty and Naikan [55, 56], Murthy and Verma [57], Wang and Liu [58, 59], English *et al.* [60], Aminzadeh [61], Alam and Roohi [62], Khan and Islam [63], Zong-wen *et al* [64] have discussed and published their research findings about SSI models and their applications.

On the application side of SSI models, Boehm and Lewis [65] applied the SSI approach to the reliability analysis of the ceramics, where the fracture caused by volume embedded cracks. Miller and Freivalds [66] used the SSI model to

study the incident rate of Carpal Tunnel Syndrome. A critical analysis of the literature shows that the principles of stress-strength interference are currently used for modeling failure mechanisms of mechanical systems. However, finding out the joint probability distributions of stresses acting due to several failure mechanisms and the associated difficulties in solving complicated problems are the challenges. Simulation seems to be a possible tool for simplifying the solution procedures.

Based on the survey of literature on mechanical systems reliability with emphasis on rotary systems, it is observed that the research work on mechanical systems reliability have been the focus of several researches in recent times. Rotor balancing has been researched well by several authors, but no research is done to establish the relationship between reliability and imbalance. In the view of these research gaps, an attempt has been made in this paper to close the gap by presenting a reliability model for the rotary systems with imbalance. The effect of imbalance is included the SSI model for reliability studies. Due to the complexity of finding analytical closed form solution, simulation approach is used.

The remaining part of the paper is organized as follows: Section 2 presents the rotary system and their failures with specific emphasis on imbalance, section 3 presents the details of the proposed model, section 4 illustrates the proposed methodology with a case study on low pressure turbine shaft and section 5 presents the conclusions with scope of future work. The paper ends with references and notations are listed in the appendix.

2 Rotor imbalance:

If the center of rotation of the rotor is not coinciding with the axis of the rotation of the shaft, the rotor becomes unbalanced. This generates large forces when the system rotates, especially at higher speeds. The quantity of the unbalanced force depends both on the unbalanced mass, its distance from the center of rotational axis and the rotational speed [19].

The force due to imbalance is given by

$$F_{im} = mr\,\omega^2 \tag{1}$$

It can be seen that even a little imbalance in the rotor generates huge force resulting in considerable stress. All the rotary systems are vulnerable to imbalances. Steam and gas turbines, rotary compressors, electrical motors, industrial draft fans and several other systems are examples. There are also other rotary systems for transferring power or torque. These systems will have components such as gears, pulleys, rollers, blades etc. mounted on the shaft. Over a period of operation, these are likely to create imbalance of the system. An attempt is made to model the effect of imbalance on rotor reliability.

Literature shows several theories for designing the shafts. Important amoung these are Rankine theory, Tresca or Guest's theory, Saint Venant theory, Haigh's theory, and Von Mises theory [67]. During the transmission of power as stand alone transmitter or with the help of mounted gears, pulleys etc., the shaft is subjected to both twisting moment and bending moment. Therefore, when the shaft is designed, both moments are considered simultaneously, to compensate the elastic failures due to combined stresses. The Rankine's (maximum normal stress theory) and Guest's (maximum shear stress theory) theories are widely used [67].

The Maximum shear stress in the shaft, according to Guest theory [67] is given by,

$$\tau_{\max} = \frac{16\sqrt{M^2 + T^2}}{\Pi d^3} \tag{2}$$

The Maximum normal stress in the shaft, according to Rankine's theory [67] is given by,

$$\sigma_{\max} = \frac{32\left(\dfrac{(M) + \sqrt{M^2 + T^2}}{2}\right)}{\Pi d^3} \tag{3}$$

The computed dimensions are then multiplied with factor of safety (*FS*) based on whether the load acting on the component is steady load, live (dynamic) load or shock load [64]. Some times for dynamic loads, the factor of safety will be as high as 12 to 15 depending on the material used and the application. These *FS* are

empirical assessment based on the observations and experience, and not the exact one [1]. The ill effects of *FS* are large dimensions, space, weight and cost.

The stress caused by the imbalance force will produce considerable bending moment and a negligible amount of torque in the shaft [68]. The bending moment due to force caused by mass imbalance together with the bending moment of the rotor due to usual operational forces, act on the shaft. The combined effect of these stresses when exceeds the ultimate strength the rotor shaft, it fails.

Effective stress on the shaft:

To find the effective stress, the bending moment due to the force of imbalance bending moment and torque due to imbalance force should be added to the actual bending moment and actual torque respectively, in the Maximum stress equations (2 and 3) given by Rankine and Guest theories. Therefore the Effective maximum shear stress is given by,

$$\tau_{eff} = \frac{16\sqrt{(M_{im} + M)^2 + T^2}}{\Pi d^3} \quad (4)$$

The Effective maximum normal stress is given by,

$$\sigma_{eff} = \frac{32\left(\frac{(M_{im} + M) + \sqrt{(M_{im} + M)^2 + T^2}}{2}\right)}{\Pi d^3} \quad (5)$$

The equations (4 and 5) give total stress acting on the shaft, including the stress produced due to force of imbalance.

3. *Reliability modeling:*

The effective stress equation proposed above consists of several variables which are probabilistic in nature. The variables such as, radial distance between axis of the shaft and the location of unbalance mass, mass causing imbalance, length of the shaft, ultimate tensile stress and ultimate normal stresses are probabilistic in nature. The stochastic nature of the aforementioned variables are due to many reasons such as, manufacturing variations while metal working or casting etc., variation in heat treatment, dimensional variations due to human or machine accuracy, power fluctuations, environmental effects, etc..

According to the SSI technique the reliability can be computed by [49]

$$R = \int_{-\infty}^{\infty} f_{\alpha}(\alpha)\left[\int_{-\infty}^{\alpha} f_{\tau_{eff}}(\tau_{eff})\, d\tau\right] d\alpha \quad (6)$$

Due to stochastic nature of some variables the effective stress (equation 4 and 5) will also have properties of randomness. This will follow a probabilistic distribution. We need to find the distribution of the effective stress from the distributions of the variables. Due to the complex nature of these equations analytical methods can not be used for finding the probability distribution of these stresses. Therefore we propose to use simulation method for developing the probability distribution of the effective stress distribution, given the probability distribution of the constituent variables.

The ultimate material strength of the shaft will also have stochastic nature due to variations in raw material quality, manufacturing process, heat treatment, surface defects, and several other causes. Therefore we also propose to generate the ultimate stress values using the simulation method. Thereafter the two distributions (stress and strength) are combined using the well known time independent Stress-Strength interference (SSI) technique to compute the reliability using equation (6). In several literatures (for example Murthy and Naikan [55, 56]) this reliability is termed as reliability strength or static reliability.

For example, if both effective shear stress and ultimate shear strength values obtained by simulation follow normal distributions, following equation (7) can be used to compute the reliability.

$$R = 1 - \Phi(z) \quad (7)$$

where, z is given by

$$z = -\frac{(\mu_\alpha - \mu_{\tau_{eff}})}{\sqrt{\varsigma_\alpha{}^2 + \varsigma_{\tau_{eff}}{}^2}}$$

Similar methods can be applied for other distributions also. For more details Kapoor and Lamberson [49] may be referred.

Flow chart of the proposed model:

In this sub-section we presented the proposed model in the form of flow chart for better clarity.

The proposed method is illustrated by

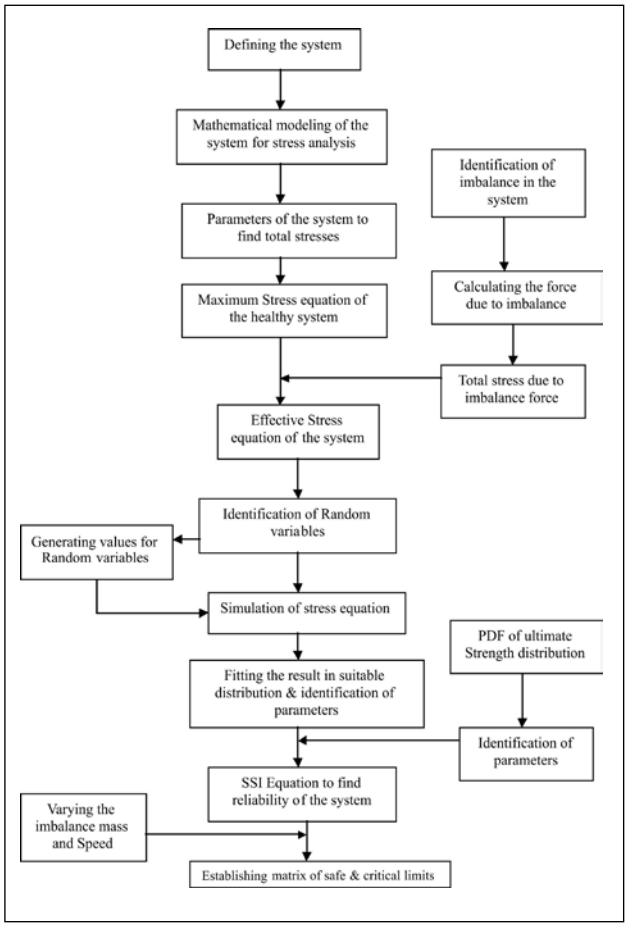


Figure 1: Flow chart for Reliability and imbalance modeling.

numerical examples in the following section.

4. Case study

In this case study, we have taken a low pressure turbine shaft made of 2.5 Ni-Cr-Mo-V steel for its reliability assessment through afore discussed model. The steel is heat treated for a temperature of 950°, for one hour and oil quenched. The steel is having an ultimate tensile strength of 700MPa [70, 71].

Table 1: Chemical composition of 2.5 Ni-Cr-Mo-V steel (weight %)

C	Si	Mn	Cr	Ni	Mo	V	S	P
0.24	0.24	0.34	0.40	2.60	0.28	0.10	0.01	0.01

The turbine shaft which is considered for this case study is of diameter 0.064 m. This shaft is attached with an inlet fan assembly at end A, which exerts a vertical force of 7200 N and the shaft is supported by a bearing at B. Figure 2 gives all the dimensional details.

In this case study, on the basis of experimental

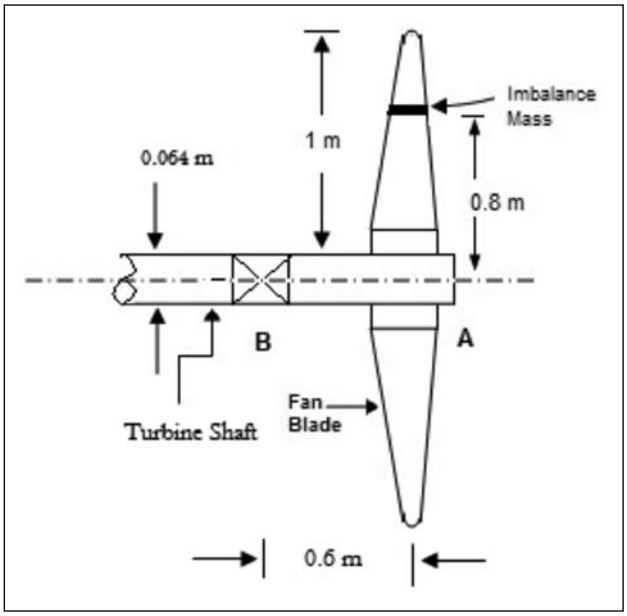


Figure 2: Shaft and four blade fan assembly with imbalance mass

results an attempt has been made to establish the ranges of speed safer for the system and imbalance mass for a given target reliability. The effective total stress is computed using the proposed model discussed in afore discussed section, and the reliability of the shaft is computed by simulation technique.

In this illustration, the simulation is first done starting with a unbalance mass of 0.1 kg (and for different imbalance masses up to 1 kg) and for the various speeds. The effective

shear and tensile stresses are computed by the equations (4 and 5). These effective stress values are fitted to various distributions and it is found that the data can be best fitted by a normal distribution. The ultimate strength is assumed to be following normal distribution. The reliability is then computed using equation (7).

The simulation cycles are replicated for 50000 times. The results are presented below.

Results:

The results for the above example are presented in the figure 3 in graphical form. Variations of reliability with respect to imbalance mass and the rotational speed are shown in this figure.

The ranges of safer and critical rotational speeds are presented in the table 2.

Table 2: Safe and critical RPM of rotor

Imbalance mass (Kg)	Safe RPM (R ≥ 0.99)	Unsafe RPM (0.95 ≤ R ≤ 0.99)	Critical RPM (R < 0.95)
0.1	≤ 4502	4502 – 4525	> 4525
0.2	≤ 3150	3150 – 3205	> 3205
0.3	≤ 2605	2605– 2615	> 2615
0.4	≤ 2255	2255 – 2265	> 2265
0.5	≤ 2010	2010 – 2025	> 2025
0.6	≤ 1843	1843– 1847	> 1847
0.7	≤ 1702	1702 – 1712	> 1712
0.8	≤ 1592	1592 - 1602	> 1602
1	≤ 1423	1423 – 1431	> 1431

The analysis and discussion of experimental results for the illustration of the 2.5 Ni-Cr-Mo-V low pressure turbine alloy steel shows that for the given dimensions, the imbalance mass is having negligible impact on the reliability up to the speed of around 1420 rpm. If the rotor system with imbalance mass used for higher speed applications its reliability decreases sharply. If the imbalance can be limited by regular monitoring and correction, then the system can be highly reliable and safe to operate at higher speeds. Therefore it is suggested that this system needs periodic monitoring and balancing.

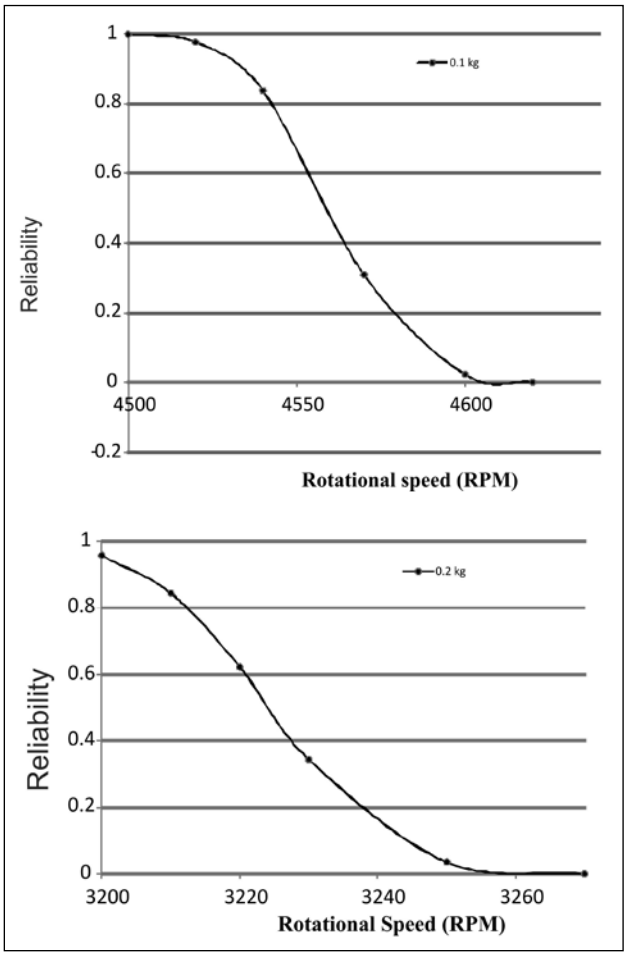


Figure 3: Reliability of the shaft at different rotational speeds (RPM).

5. Conclusions:

The rotary systems which are subjected to imbalance produce enormous dynamic forces especially at higher rotational speeds. These forces may be several times higher than actual static load or force acting on the system. Because of these huge forces produced, it is required to develop models to study and analyze the effect of imbalance in rotary system reliability. These models can also help us to find the safe, unsafe and critical levels of speed for a known imbalance mass. It can also help us to take necessary actions, including maintenance and balancing, to keep the imbalance mass within the threshold limit for a given or required rotational speed for the safety of the rotary system.

In this paper, a reliability model for the generalized rotary system subjected to imbalance

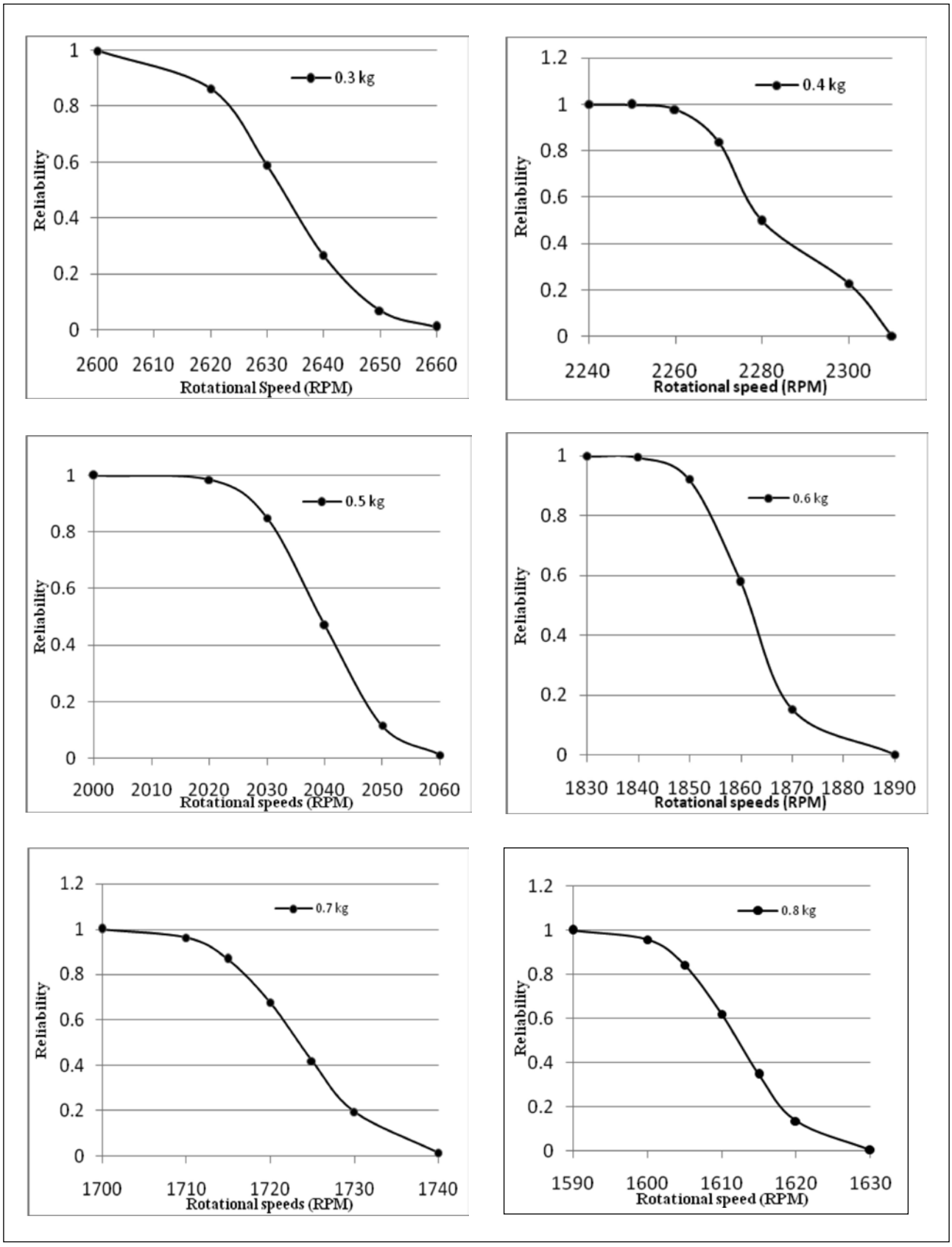


Figure 3: Reliability of the shaft at different rotational speeds (RPM). (Continued…)

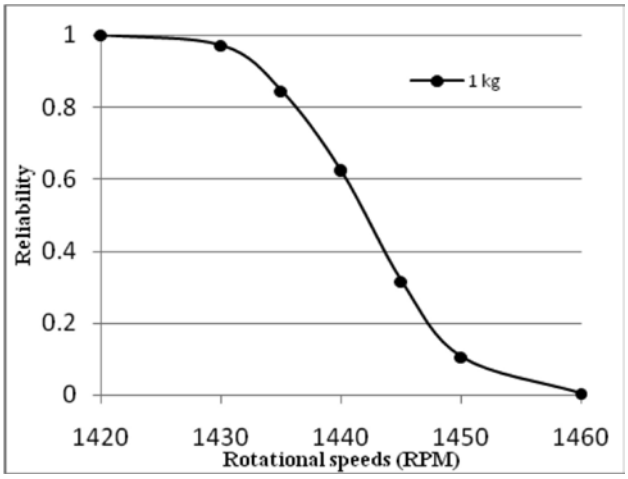


Figure 3: Reliability of the shaft at different rotational speeds (RPM). (Continued…)

is presented and a case study for turbine alloy steel 2.5 Ni-Cr-Mo-V is presented. The model can be effectively used along with simulation methodology to predict the operational reliability of a rotor system given the imbalance present and its rotational speed. If the reliability target is specified, safe ranges of operational speed and allowable imbalance mass (imbalance tolerance) can be established using this model.

The discussed model considered only the effect of static imbalance on rotary system. There is a scope for improvement of the model by considering other types of dynamic forces on the rotor system which includes dynamic imbalance. Furthermore, models for the other problems such as oil whirl, misalignment, and wear effects, which are also the source of dynamic forces can be developed and integrated with this model for the complete reliability analysis of the system.

References:

1. Carter, A.D.S., "Mechanical Reliability & Design", 1997, MacMillan, Hong Kong.
2. Keceoglu, D., "Reliability analysis of Mechanical systems and components", Nuclear engineering and Design, 1972, Vol. 19, pp 259 – 290.
3. Yoshikawa, H., "Fundamentals of Mechanical Reliability and its Application to computer aided machine design", Annals of CIRP, 1975, Vol. 24, No. 1, pp 297-302.
4. Carter, A. D. S., "Wear out failure patterns and their interpretation", Journal of Mechanical Engineering science, 1980, vol. 22, pp 143-151.
5. Carter, A.D.S., "Early life Failures", Reliability Engineering, 1983, vol. 4, pp 41-59.
6. Martin, P., Strutt, J. E., and Kinkead, N., "A review of mechanical reliability modeling in relation to failure mechanisms", Reliability Engineering, 1983, vol. 6, pp 13-42.
7. Chegodaev, D. E., and Samsonov, V. N., "Evaluating the reliability of mechanical systems", Problemy prochnosti, 1987, 1720-1723.
8. Charlesworth, W. W., and Rao, S. S., "Reliability analysis of continuous mechanical systems using multistate fault trees", Reliability Engineering and System Safety, 1992, vol. 37, pp195-206.
9. Murthy, A. S. R., and Krishna, A. R., "Mechanical / Structural reliability evaluation through Esscher's method of approximation", Reliability engineering and system safety, 1997, vol. 57, pp. 171-176.
10. Tang, J., "Mechanical system reliability analysis using a combination of graph theory and Boolean function", Reliability engineering and system safety, 2001, vol. 72, pp. 21-30.
11. Avontuur, G. C., and Weff, K. V., "An implementation of reliability analysis in the conceptual design phase of drive trains", Reliability engineering and system safety, 2001, vol. 73, pp. 155-165.
12. Sharma, B., and Gandhi, O. P., "Digraph-based reliability assessment of a tribo-pair", Industrial Lubrication and Tribology, 2008, vol. 60/3, pp. 153-163.
13. Othman, N. A., Damanhuri, N. S., and Kadirkamanathan, V., "The study of fault Diagnosis in rotating machinery", IEEE - Proceedings of 5th International colloquium on signal processing & its application, Kuala Lumpur, Malaysia , 2009, pp. 69-74.
14. Muszynska, A., "Partial lateral rotor to stator rubs", Proceedings of the Institution of Mechanical Engineers, 1984, **C281** (84), pp. 327–335.
15. Muszynska, A., "Rotor-to-stationary element rub-related vibration phenomena in rotating machinery-literature survey", Shock and Vibration Digest, 1989, vol. **21** (3), pp. 3–11.
16. Goldman, P., and Muszynska, A., "Rotor-to-Stator rub related thermal / mechanical effects in rotating machinery", Chaos solitons and fractals, 1995, vol. 5(9), pp. 1579-1601.
17. Chu, F., and Lu, W., "Determination of rubbing location in multi-disk rotor system by means of dynamic stiffness identification", Journal of sound and vibration, 2001, vol. 248 (2), pp. 235-246.
18. Jiang, J., Ulbrich, H., and Chavez, A., "Improvement of rotor performance under rubbing conditions through active auxiliary bearings" International journal of non-linear mechanics, 2006, vol. 41, pp. 949-957.
19. Collacott, R. A., "Vibration Monitoring and diagnosis", John Willey, New York, 1979.
20. Lum, K., Coppola, V. T., and Bernstein, D. S., "Adaptive autocentering control for an active magnetic bearing supporting a rotor with unknown mass imbalance", IEEE Trans. on control system technology, Sep 1996, vol 4, no.5, pp. 587-597.
21. Zhou, S., and shi, J., "Active Balancing and vibration control of rotating machinery: A survey", The shock and vibration digest, 2001, vol.33, no. 4, pp 361-371.
22. Hredzak, B., and Guo, G., "New electromechanical balancing device for active imbalance compensation", Journal of sound and vibration, 2006, vol. 294, pp. 737-751.
23. Lee, J.K., Ihn, Y.S., Koo, J.C., Dongho O., and Lee, H. S., "Active correction of mass imbalance for precise rotor", proceedings of IEEE- Asia pacific magnetic recording

conference, Jan 2009, pp. AB-2.
24. DeSmidt, H.A., "Imbalance vibration suppression of a supercritical shaft via an automatic balancing device", Journal of Vibration and Acoustics, Transactions of the ASME, 2009, vol. 131, no. 4, pp. 0410011-04100113.
25. Oladejo, K.A., Koya, O.A. and Adekoya, L.O., "A computational model for static and dynamic balancing of masses on rotating shafts", Computer Assisted Mechanics and Engineering Sciences, 2008, vol. 15, no. 1, pp. 23-35.
26. Horvath, R., Flowers, G.T. and Fausz, J., "Passive balancing for rotor systems using pendulum balancers", Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference - DETC2005, 2005, pp. 967.
27. Ohtomi, K., Otsuki, F., Uematsu, H., Nakamura, Y., Chida, Y., Nishimura, O. and Okamura, R., "Active mass auto-balancing system for centrifuge rotor providing an artificial gravity in space", Proceedings of the ASME Design Engineering Technical Conference, 2001, pp. 691.
28. Bansal, A.S., Rao, V.R.B. & Singh, S., "Dynamic Balancing of a Rigid Multi-mass Thresher Rotor and Vibration Control", Agricultural Mechanization in Asia, Africa and Latin America, 1998, vol. 29, no. 4, pp. 35-42.
29. Lum, K., Bhat, S.P., Bernstein, D.S. and Coppola, V.T., "Adaptive virtual autobalancing for a magnetic rotor with unknown mass imbalance", Proceedings of the American Control Conference, 1995, pp. 3796.
30. Lum, K., Bhat, S.P., Bernstein, D.S. and Coppola, V.T., "Adaptive virtual autobalancing for a magnetic rotor with unknown mass imbalance Part I: static balancing", ASME, Design Engineering Division (Publication) DE, 1995, pp. 1419.
31. Lum, K., Bhat, S.P., Bernstein, D.S. and Coppola, V.T., "Adaptive virtual autobalancing for a magnetic rotor with unknown mass imbalance Part II: dynamic balancing", ASME, Design Engineering Division (Publication) DE, 1995, pp. 1427.
32. Lum, K., Coppola, V.T. and Bernstein, D.S., "Adaptive virtual autobalancing for a rigid rotor with unknown mass imbalance supported by magnetic bearings", Journal of Vibration and Acoustics, Transactions of the ASME, 1998, vol. 120, no. 2, pp. 557-570.
33. Manchala, D.W., Palazzolo, A.B., Kascak, A.F., Montague, G.T. and Brown, G.V., "Active vibration control of sudden mass imbalance in rotating machinery", ASME, Design Engineering Division (Publication) DE, 1994, pp. 133.
34. Manchala, D.W., Palazzolo, A.B., Kascak, A.F., Montague, G.T. and Brown, G.V., "Constrained quadratic programming, active control of rotating mass imbalance", Journal of Sound and Vibration, 1997, vol. 205, no. 5, pp. 561-578.
35. Kang, Y., Liu, C., and Sheen, G., "A modified influence coefficient method for balancing unsymmetrical rotor-bearing systems", Journal of Sound and Vibration, 1996, vol. 194, no. 2, pp. 199-218.
36. Srinath, L.S., "Mechanical reliability", East West Press, New Delhi, 2002.
37. Dasgupta, A., and Pecht, M., "Material failure mechanisms and damage models", IEEE trans. Reliability, Dec 1991, vol. 40, pp 531-536.
38. Dasgupta, A., "Failure mechanisms models for cyclic fatigue", IEEE trans. Reliability, Dec 1993, vol. 42, no. 4, pp 548-555.
39. Li, J., and Dasgupta, A., "Failure mechanism models for material ageing due to interdiffusion", IEEE trans. Reliability, Mar 1994, vol. 43, no. 1, pp 1 - 10.
40. Li, J., and Dasgupta, A., "Failure mechanism models for creep and creep rupture", IEEE trans. Reliability, Sep 1993, vol. 42, no. 3, pp 339 – 353.
41. Engel, P. A., "Failure models for mechanical wear modes & Mechanisms", IEEE trans. Reliability, June 1993, vol. 42, no. 2, pp 262 – 267.
42. Dasgupta, A., and Hu, J. M., "Failure-mechanisms models for cyclic excessive elastic deformation", IEEE trans. Reliability, Mar 1992, vol. 41, no. 1, pp 149-154.
43. Dasgupta, A., and Haslach Jr, H. W., "Mechanical design failure models for buckling", IEEE trans. Reliability, Mar 1993, vol. 42, no. 1, pp 9-16.
44. Dasgupta, A., and Hu, J. M., "Failure-mechanisms models for ductile fracture", IEEE trans. Reliability, Dec 1992, vol. 41, no. 4, pp 489-495.
45. Kotz, S., Lumelskii, Y., and Pensky, M., "The stress-Strength model and its generalizations", World Scientific, 2003, Singapore.
46. Birnbaum, Z. W., "On a use of the Mann-Whitney statistic", Proceedings of Third Berkeley Symposium on Math. Statist. Prob., Vol. 1, Contributions to the Theory of Statistics and Probability, pp. 13–17, University of California Press, Berkeley
47. Birnbaum, Z. W., and McCarty, B. C., "A distribution-free upper confidence bounds for Pr(Y<X) based on independent samples of X and Y". Annals of mathematical statistics, 1958, vol. 29, pp. 558-562.
48. Owen, D. B., Craswell, K. J., Hanson, D. L., "Nonparametric upper confidence bounds for P(Y<X) when X and Y are normal", Journal of American statistical association, 1964, vol. 59, pp. 906-924.
49. Kapoor, K. C., and Lamberson, L. R., "Reliability in Engineering Design", 1997, John Willey, New York.
50. Dhillon, B. S., "Stress-Strength Reliability models", Microelectronics reliability, 1980, vol.20, pp. 513-516.
51. Chung, W. K., "Some stress-strength reliability models", Microelectronics reliability, 1982, vol. 22, pp. 277-280.
52. Reiser, B., and Guttman, I., "Sample size choice for reliability verification in stress-strength models", The Canadian Journal of Statistics, Sep 1989, vol. 17, no. 3, pp. 253-259.
53. Reiser, B., and Guttman, I., "Statistical interference for Pr(Y<X): The normal case", Technometrics, 1986, vol.28, no. 3, pp. 253-257.
54. Dargahi-Nourbary, G. R., "A parametric solution for simple Stress-strength model of failure with an application", Journal of computational and applied Mathematics, 1988, vol.23, pp. 185-197.
55. Murthy, A. S. R., and Naikan, V. N. A., "Condition monitoring strategy-a risk based interval selection", Reliability engineering and system safety, 1994, Vol.34, No.1, pp 285-296.
56. Murthy, A. S. R., and Naikan, V. N. A., "Inverse distribution in reliability design – a revisit", Reliability Engineering and System Safety, 1994, vol. 44, No. 2, pp 167-171.
57. Murthy, A. S. R., and Verma, A. K., "Inverse normal and lognormal distributions for reliability design", Reliability Engineering, 1986, Vol. 15. pp. 55-60.
58. Wang, J. D., and Liu, T. S., "A discrete stress-strength interference model for unreliability bounds", Reliability engineering and system safety, 1994, vol. 44, no. 1, pp. 125-130.

59. Wang, J. D., and Liu, T. S., "Fuzzy reliability using a discrete stress-strength interference model", IEEE Tran. on Reliability, Mar 1996, Vol. 45, no. 1.
60. English, J. R., Sargent, T., and Landers, T. L., "A discretizing approach for stress / strength analysis", IEEE tran. on Reliability, Mar 1996, Vol. 45, no.1, pp. 84-89.
61. Aminzadeh, M. S., "Estimation of reliability for experimental stress-strength models with explanatory variables", Applied mathematics and computation, 1997, vol. 84, pp. 269-274.
62. Alam, S.N., and Roohi, "On facing an exponential stress and strength having power function distribution", *Aligarh Journal of Statistics*, 2003, Vol. 23 pp.57-63.
63. Khan, M. A., and Islam, H. M., "On facing Rayleigh stress with strength having power function distribution", Journal of applied statistical science, 2008, vol. 16, no. 2, pp 9-18.
64. Zong-Wen, A., Huang, H., and Liu, Y., "A discrete stress-strength interference model based on universal generating function", Reliability Engineering and System Safety, 2008, vol. 93, pp 1485-1490.
65. Boehm, F., and Lewis, E. E., "A stress-strength interference approach to reliability analysis of ceramics: Part-1 Fast fracture", Probabilistic Engineering Mechanics, 1992, Vol. 7, pp. 1-8.
66. Miller, S. A., and Freivalds, A., "A stress-strength interference model for predicting CTD probabilities", Int. Journal of industrial ergonomics, 1995, Vol. 15, pp. 447-457.
67. Khurmi, R. S., and Gupta, J. K., "Machine Design", 2009, S.Chand, New Delhi.
68. Beer, F. P., Jhonston, F.R., and Dewolf, J. T., "Mechanics of Materials", 2006, McGraw-Hill, Singapore.
69. Haldar, A., and Mahadevan, S., "Reliability assessment using Stochastic finite element analysis", 2000, John Wiley, New York.
70. Spink, G. M., "Fretting fatigue of a 2.5 % NiCrMoV low pressure turbine shaft steel – The effect of different contact pad material and of variable slip amplitude" Wear, 1990, (136), pp 281-297
71. Bhambri, S. K., Singh, V., Jayaraman, G., "The effect of microstructure on stage-II fatigue crack growth rates in 2.5 Ni-Cr-Mo-V steel", Int. Journal of Fatigue, 1989, (1) pp 51 – 54.

Appendix I:

Notation:

d	diameter of the shaft
F_{im}	force due to imbalance
$f_{\alpha}(\alpha)$	density function of ultimate shear strength
$f_{\tau_{eff}}(\tau_{eff})$	density function of effective maximum shear stress
M	bending moment
M_{im}	bending moment due to imbalance force
m	imbalance mass
R	reliability
RPM	rotations per minute
r	distance between center of rotational axis and imbalance mass
T	torque
α	ultimate shear strength
μ_{α}	mean of the normally distributed ultimate shear strength
$\mu_{\tau_{eff}}$	mean of the normally distributed effective maximum shear strength
ς_{α}	standard deviation of the normally distributed shear strength
$\varsigma_{\tau_{eff}}$	standard deviation of the normally distributed effective maximum shear strength
σ_{max}	maximum normal stress
σ_{eff}	effective maximum normal stress
τ_{max}	maximum shear stress
τ_{eff}	effective maximum shear stress
z	standardized normal variate
ω	angular velocity
$\Phi(z)$	standard normal cumulative probability value of z

Life Cycle Reliability and Safety Engineering
Vol.2 Issue 2 (2012) 71-80

Uncertainty Quantification of Contaminant Transport Model Using Dempster-Shafer Evidence Theory

D. Datta
Health Physics Division, Bhabha Atomic Research Centre, Mumbai – 400085
Email: ddatta@barc.gov.in

Abstract

Uncertainty is an attribute of information. It is particularly crucial in safety studies where misleading representations of uncertainties can lead to incautious and potentially hazardous decisions. Uncertainty associated with environmental parameters is of epistemic in nature due to their insufficient measurements and hence imprecise probabilistic approach is applied for quantification of the uncertainty. The present approach of quantifying uncertainty is based on Dempster-Shafer evidence theory. Evidence theory based various measures of uncertainty such as belief, plausibility, expectation value, 5^{th} and 95^{th} percentile, dissonance and non-specificity is computed. Overall computation strategy is illustrated using a standard one dimensional imprecise probability based model of contaminant transport. Expert's opinions are adopted to figure out the evidences of the uncertain parameters of the model.
Keywords: Uncertainty, Dempster-Shafer, Percentile, Belief, Plausibility, Expert's opinion

Introduction

Uncertainties are one of the most challenging problems in safety related environmental studies of complex systems [1-3]. They are present in any environmental parameters due to either randomness (inherent variability) or lack of information (subjectivity). Generally uncertainty analysis consists of modeling the uncertainty of the uncertain parameters and propagating parameter uncertainties to the output of the model. Efficient and accurate uncertainty quantification can provide an assessment of risk or confidence in the design and is essential for safety analysts to obtain a real safe and optimal solution.

Uncertainty can be classified into two subtypes: (a) aleatory and (b) epistemic uncertainty. Aleatory uncertainty is also called irreducible and inherent uncertainty. It is the inherent variation associated with the physical system or the environment under consideration [4]. Epistemic uncertainty is subjective and reducible because it arises from lack of knowledge or data. It represents uncertainty of the outcome due to lack of knowledge or information in any phase or activity of the modeling process [4].

Classical framework of probability theory for uncertainty quantification has been performed in many engineering disciplines for decades. The probabilistic analysis for uncertainty quantification uses probabilistic method to represent sources of uncertainty and then propagates the uncertainty through a model (response function) by Monte Carlo simulation (sampling technique) [5-7]. In general, probability theory is very effective if sufficient data are available to characterize the parameters by specific probability distribution. But, if the available data are not sufficient probability theory will fail to quantify the uncertainty. For this reason, even though the probability theory is reasonably well founded conceptual framework for uncertainty management, a number of alternative theories have also found to carry out the uncertainty especially for that variety where information is insufficient. These theories include possibility theory and the so-called Dempster-Shafer (D-S) theory of belief functions, or Evidence theory.

The evidence theory (D-S theory) is an especially interesting methodology because of its applicability in areas where information

(evidence) must be combined and can be considered as generalization of classical probability theory and also as a generalization of possibility theory. Possibility theory [8], first introduced by Zadeh [9] is usually chosen to handle epistemic uncertainty. Possibility theory uses fuzzy measures to describe the possibility or membership grade by which a certain event can be plausible or believable. Contrary to the classical probability theory, possibility theory is usually used to quantify only epistemic uncertainty. Besides possibility theory, interval analysis can be applied when the information is available in the form of an interval (lower bound, upper bound).

The D-S theory (evidence theory) is more general than probability and possibility theories. It uses plausibility and belief to measure the likelihood of event, without making additional assumptions. When the belief and plausibility measures are equal, the general evidence theory reduces to the classical probability theory. Therefore, the classical probability theory is special case of evidence theory. Moreover, evidence theory can combine empirical evidence from different experts to construct coherent picture of reality. The D-S theory has several interpretations such as the Transferable Belief Model (TBM) [9]. The TBM is completely dissociated from any model based on probability functions. The possibility theory is usually employed to quantify only epistemic uncertainty. Hence, in this work epistemic uncertainty associated with the transport of contaminant for assessing the safety issues is quantified using Dempster-Shafer theory [9].

These new theoretical developments have been motivated by the growing recognition that all forms of partial information are not easily amenable to representation in the probabilistic framework. This applies, for instance, to vague linguistic statements often used by experts to express their knowledge, such as "if x is small, then y is very likely to be large". An extreme situation is that of complete ignorance: a quantity "x" of interest may be totally unknown, and the representation of this lack of information by a probability distribution (even a uniform one) may be shown to lead to paradoxes. Among existing tools for uncertainty representation, belief functions appear to play a pivotal role as they generalize both probability and possibility distributions.

The uncertainty analysis of a model (such as a computer code used for accident management procedures in nuclear industry) is usually performed in four steps depicted below:

a. Identification of uncertain parameters

All important factors affecting the model results must be identified. These factors are generally referred to as the "uncertainty sources" or as the "uncertain parameters".

b. Quantification of the knowledge about uncertain parameters

The available information about uncertain parameters is formalized. The uncertainty of each uncertain parameter is quantified. If dependencies are known between uncertain parameters and judged to be potentially important, they also need to be specified.

c. Propagation of uncertainties through the computer code

The propagation requires, except for very simple computer codes, a coupling between the code and statistical software.

d. Treatment and interpretation of the code responses

The paper presents a short description of Dempster-Shafer theory and its implementation towards the quantification of epistemic uncertainty. Evidence of the parameters of the model under investigation is collected from expert's knowledge. Uncertainty bounds are expressed in terms of belief and plausibility. Computations of various measures of imprecision such as expectation value, 5^{th} percentile, 95^{th} percentile, dissonance, non-specificity and aggregated uncertainty are presented in this paper. Expectation value, 5^{th} and 95^{th} percentiles are expressed in terms of corresponding coherent lower and upper bounds, whereas dissonance, non-specificity and aggregated uncertainty are profiled for various downstream distances.

2.0 Evidence Theory (Dempster-Shafer Theory)

Evidence theory is based on two dual semi continuous non additive measures (fuzzy measures): belief measure and plausibility measure. It is a theory of evidence because it deals with weights of evidence and numerical degrees of support based upon evidence. Further, it contains a viewpoint on the representation of uncertainty and ignorance. It is also a theory of plausible reasoning because it focuses on the fundamental operation of plausible reasoning, namely the combination of evidences. Alternatively, evidence theory also is known as Dempster-Shafer theory. The Dempster-Shafer theory is based on the idea of placing a number between zero and one to indicate the degree of evidence for a proposition. The theory also includes reasoning based on the rule of combination of degrees of belief according to different evidences. The addition axiom P(A) + P(not A) = 1 for any proposition A in classical Bayesian theory does not necessarily correspond to the description of the real world because ignorance was not taken into account. Therefore, without enough evidences for proposition A, it is appropriate to assume that the sum of both degrees of belief are not equal to one, i.e., P(A) + P(not A) < 1.

2.1 Basic Probability Assignment

One of the basic concepts of the Dempster-Shafer theory is that of basic probability assignment, that is to assign a function m: $2^{\theta} \rightarrow [0,1]$, such that

$$m(0) = 0 \quad (1) \text{ and } \quad \sum_{A \in \theta} m(A) = 1 \quad (2)$$

The number m(A) is called basic probability assignment of A. The equation (1) states that no belief is committed to the empty set and the equation (2) states that the total belief is equal to one.

2.2 Formulation of the Representation of Evidence

Evidence theory is based on two dual nonadditive measures: belief measure and plausibility measure. Given a universal set X, assumed here to be finite, a *belief* measure is a function, $bel:P(X) \rightarrow [0,1]$ such that $bel(\phi) = 0$ and $bel(X) = 1$. Let 6 be a set of propositions about the exclusive and exhaustive possibilities in a domain. For example, if we are rolling a die, G contains the six propositions of the form 'the number showing is 'k', where $1 < k < 6$. The θ is called the frame of discernment and 2^{θ} is the set of all subsets of θ. Elements of 2^{θ} are the class of general propositions in the domain; for example, the proposition the number showing is even corresponds to the set of the three elements of G that assert the die shows either a 2,4, or 6.

The following are the key assumptions of the evidence theory approach:

1. If some of the evidence is imprecise we can quantify uncertainty about an event by the maximum and minimum probabilities of that event. Maximum (minimum) probability of an event is the maximum (minimum) of all probabilities that are consistent with the available evidence.
2. The process of asking an expert about an uncertain variable is a random experiment whose outcome can be precise or imprecise. There is randomness because every time we ask a different expert about the variable we get a different answer. The expert can be precise and give a single value or imprecise and provide an interval. Therefore, if the information about uncertainty consists of intervals from multiple experts, then we have uncertainty due to both imprecision and randomness.

If all experts are precise they give us pieces of evidence pointing precisely to specific values. In this case, we can build a probability distribution of the variable. But if the experts provide intervals, we cannot build such a probability distribution because we do not know what specific values of the random variables each piece of evidence supports. In this case, we can use second order probability, or we can calculate the maximum and minimum values of the probabilities of events. The latter approach does not require any additional information beyond what is already

available. However, if experts provide intervals as well as the knowledge of the probability distribution of the uncertain parameters, then we can build the Dempster-Shafer structure using an inverse sampling technique.

2.2.1 Belief and Plausibility Function

Let m be a given basic probability assignment. A function *bel:* $2^{\theta} \rightarrow [0, 1]$ is called a belief function over θ if and only if

$$bel(A) = \sum_{B \subseteq A} m(B) \tag{3}$$

and, for any collection $A_1, \ldots, A_n$ of subsets θ

$$Bel(A_1 \cup A_2 \cup \ldots \cup A_n) \geq \sum_j Bel(A_j) - \sum_{j<k} Bel(A_j \cap A_k) + \ldots + (-1)^{(n+1)} Bel(A_1 \cap A_2 \cap \ldots \cap A_n) \tag{4}$$

The plausibility function, denoted by *pls*, is represented as a function *pls:* $2^{\theta} \rightarrow [0,1]$ such that

$$pls(A) = 1 - bel(\overline{A}) \tag{5}$$

The plausibility function *pls(A)* may be expressed in terms of the basic probability assignment m of *bel* as

$$pls(A) = 1 - bel(\overline{A}) = \sum_{B \subseteq \theta} m(B) - \sum_{B \subseteq A} m(B) = \sum_{B \cap A = \phi} m(B) \tag{6}$$

Suppose, the B_i intervals are nested according to the following structure:

$$B_i \subseteq B_{i+1} \quad \text{for i = 1, 2, ..., n-1} \tag{7}$$

Since the B_i intervals are equally credible, they can be given a basic assignment m = 1/n. The *belief* and *plausibility* measures can be computed as follows:

$$bel(B_i) = \sum_{all\ B_j \subseteq B_i} m(B_j) \tag{8}$$

$$pls(B_i) = \sum_{all\ B_j \cap B_i \neq \Phi} m(B_j) \tag{9}$$

2.2.2 Dempster Rule of Combination

In the evidence theory, evidences pooled from more than one resource are aggregated by using Dempster rule of combination, which is written mathematically as

$$m_1 \oplus m_2(C) = \begin{cases} 0 & \text{if } C = \phi \\ \dfrac{\sum_{A \cap B = C} m_1(A)\, m_2(B)}{1 - \sum_{A \cap B = \phi} m_1(A)\, m_2(B)}, & \text{otherwise} \end{cases}$$

, where (10)

$m_1(A)$ = probability mass based on evidence *A*

$m_2(B)$ = probability mass based on evidence *B*

m(C)= probability mass for pooled evidence, ϕ - null set, *m*(φ) = mass assigned to ϕ

2.2.3 Uncertainty Measures of Evidence Theory

Evidence theory provides two types of uncertainty measures. One is due to the imprecision in the evidence; the other is due to the conflict. Non-specificity and Strife measure the uncertainty due to imprecision and conflict, respectively. Both measures are expressed in bits of information. Only non-specificity measure is presented in this paper. A detailed presentation of non-specificity and strife can be found elsewhere in [10]. The larger the focal elements of a body of evidence, the more imprecise are the evidence and, consequently, the higher is Non-specificity. When the evidence is precise (all of the focal elements consist of a single member), Non-specificity is zero. Strife measures the degree to which pieces of evidence contradict each other. Consonant (nested) focal elements imply little or no conflict. Disjoint elements imply high conflict in the evidence.

2.2.4 Entropy Measures of Evidence Theory

2.2.4.1 Measure of Dissonance

Dissonance is a state of contradiction between claims, beliefs, or interests [11]. The measure of dissonance, D, can be defined based on evidence theory as follows:

$$D(m) = -\sum_{i=1}^{n} m(A_i) \log_2(pls(A_i)) \tag{11}$$

where $m(A_i) > 0$; $\{A_1, A_2, \ldots, A_n\}$ = a family of subsets, i.e., focal points, that contains some or all elements of the universal set X; $m(A_i)$ = a basic assignment that is interpreted as the degree of evidence supporting the claim that a specific element belongs to the subset, A_i, but not any special subset of A_i. Thus dissonance measures the information that has no overlap with the set A.

2.2.4.2 Measure of Confusion

The measure of confusion characterizes the multitude of subsets supported by evidence as well as the uniformity of the distribution of strength of evidence among the subsets. The greater the number of subsets involved and the more uniform the distribution, the more confusing the presentation of evidence. The measure of confusion, C, is defined as

$$C(m) = -\sum_{i=1}^{n} m(A_i)\log_2(bel(A_i)) \quad (12)$$

Where $bel(A_i)$ = belief measure, which represents the total evidence or belief that the element of concern belongs to the subset A_i as well as to the various special subsets of A_i. confusion measures the information that has some elements outside of the set A.

2.2.5 Aggregate Uncertainty in Evidence Theory

An aggregate uncertainty (AU) in evidence theory measures the combined non-specificity and conflict provided by a given body of evidence. The function AU is defined as a mapping from the set of all belief measures (B) to the nonnegative real line (R_+) as follows:

$$AU : B \rightarrow R_+ \quad (13)$$

The measure is given by

$$AU(bel) = \max_{P_{bel}}[-\sum_{x \in X} p_x \log_2(p_x)] \quad (14)$$

where P_{bel} is the set of all probability distributions (p_x) defined over X that satisfy the following two constraints:

$$p_x \in [0,1] \quad \text{for all } x \in X \quad \text{and} \quad \sum_{x \in X} p_x = 1 \quad (15)$$

$$bel(A) \leq \sum_{x \in X} p_x \leq 1 - bel(\overline{A}) \quad \text{for all } A \subseteq X \quad (16)$$

3.0 Methodology of Uncertainty Quantification

Evidence theory has been applied to quantify the epistemic uncertainty associated with any model. Basically when information on model parameters is insufficient or vague and not able to characterize their uncertainty using traditional probability theory, one can apply this evidence theory to characterize their uncertainty. Generally, evidences on the uncertainty of the representative parameters of a model are collected using expert's opinion. The domain expert knows the range value (interval) of the uncertain parameter and its continuity character within that defined range. Hence, domain expert discretize the complete range into small sub intervals and assigns a basic probability (basic mass assignment) to each sub intervals. These sub intervals are known as focal element and the complete structure {Focal element, Basic probability mass} is called as Dempster-Shafer (D-S) structure. On the basis of this D-S structure created from the knowledge of domain expert and the probabilistic behavior (e.g., Normal, Uniform, Lognormal, Weibull, etc) of the model parameter, uncertainty of the parameter is simulated using standard Monte Carlo simulation. Simulated results are nothing but the belief (lower bound) and plausibility (upper bound) of the parameter of interest. Finally, propagation of uncertainty is carried out using the model of interest. Final result also comes out as Dempster-Shafer structure constituting the belief and plausibility. Therefore, uncertainty of the model is expressed as the range of belief and plausibility.

4.0 Case Study – Contaminant Transport Model

One dimensional solute transport in saturated porous media is considered for assessing the uncertainty analysis using evidence theory. The model computes the concentration of a dissolved chemical species (contaminant)

in an aquifer at any time and at any specified distance from the point of release of the chemical. Measured parameters associated with the present model are flow velocity (u m/day) and longitudinal dispersivity ($\in_L$ m) and these parameters are considered as uncertain due to the insufficient measurement. Uncertainty of both these parameters is addressed using the knowledge of domain expert. The governing equation of one dimension solute transport in saturated porous media [12] is given by

$$\frac{\partial C}{\partial t} = D_L \frac{\partial^2 C}{\partial x^2} - u \frac{\partial C}{\partial x} \quad (17)$$

where C represents the concentration of the chemical species (solute) in mg/l, u is the flow velocity in m/day, D_L is the longitudinal dispersion coefficient (m^2/day), x is the downstream distance (m) and t is the time of observation (days). Initial and boundary conditions of the solute transport problem are $C(x,0) = 0$, $x \geq 0$, $C(0,t) = C_0$, $t \geq 0$, and $C(\infty, t) = 0$, $t \geq 0$. Longitudinal dispersion coefficient by definition is given by $D_L = \in_L u$. Analytical solution of the model described can be written as $C(x,t) = 0.5C_0 (\varphi + \xi\delta)$, where C_0 = initial concentration of the contaminant poured into the aquatic media, $\varphi = \text{erfc}[(x - ut) / \sqrt{(4D_L t)}]$, $\xi = \exp(ux/D_L)$, $\delta = \text{erfc}[(x + ut)/\sqrt{(4D_L t)}]$, and $\text{erfc}(\beta) = 1 - \text{erf}(\beta)$. The error function is given by $\text{erf}(\beta) = \frac{2}{\sqrt{\pi}} \int_0^{\beta} \exp(-t^2) dt$. Epistemic uncertainty of concentration is quantified at a fixed time of observation for various downstream distances (100 m -1500 m). The initial concentration, C_0 = 100 mg/l and fixed time of observation, t = 400 days are used for computation. Downstream distances ranging from 100 m to 1500 m with a step size of 50 m are taken into consideration for computation. Uncertain parameters of the model are chosen as flow velocity, u m/day and longitudinal dispersivity, $\in_L$ (m).

4.1 Results and Discussions

Uncertain parameters for the problem chosen for illustration of the evidence theory based uncertainty quantification are groundwater velocity (u m/day) and longitudinal dispersivity ($\in$ m). Evidence related to uncertainty for these parameters are collected from expert's knowledge and are expressed in terms of an interval (focal element) and corresponding basic probability assignment (BPA). The table 1 presents this D-S or BPA structure. It is also known from both the experts that the uncertain parameters are continuous and their uncertainty follows the uniform distribution. This knowledge helps for sampling the uncertainty of the parameter of interest. An inverse sampling technique of uniform distribution is applied on this BPA structure of each expert to generate the belief and plausibility of each uncertain parameter. Evidences come from more than one expert can be aggregated either by using weight mixing or Dempster rule of combination. In the present paper, Dempster rule of combination is applied to aggregate the evidence of uncertainty of a specific parameter from more than one expert. The present model contains the uncertain parameter as ground water velocity and longitudinal dispersivity. The belief and plausibility plot

Table 1: BPA Structure of Input Uncertain parameters

Para-meter	Expert 1				Expert 2				Expert 3			
	LB	Mass	UB	Mass	LB	Mass	UB	Mass	LB	Mass	UB	Mass
V (m/d)	[1.6, 1.8]	0.3	[2.9, 3.0]	0.4	[1.6, 1.7]	0.3	[2.9, 3.1]	0.5	[1.6,1.8]	0.4	[2.9,3.1]	0.5
	[1.7, 1.9]	0.3	[3.0, 3.2]	0.4	[1.7, 1.9]	0.4	[3.0, 3.2]	0.3	[1.7,2.0]	0.5	[3.0,3.2]	0.3
	[1.8,2.0]	0.4	[3.1,3.2]	0.2	[1.8,2.0]	0.3	[3.1,3.2]	0.2	[1.8,2.0]	0.1	[3.1,3.2]	0.2
α (m)	[8.0,9.0]	0.4	[19,20]	0.5	[8,9]	0.3	[118,20]	0.4	[8,8.5]	0.3	[18,18.5]	0.3
	[8.5,9.5]	0.3	[19, 20]	0.4	[8.5,9.5]	0.5	[19, 21]	0.4	[8,9]	0.5	[18,19]	0.3
	[8.5,10]	0.3	[19,21]	0.1	[9,10]	0.2	[20,21]	0.2	[8.5,9.5]	0.2	[18.5,20]	0.4

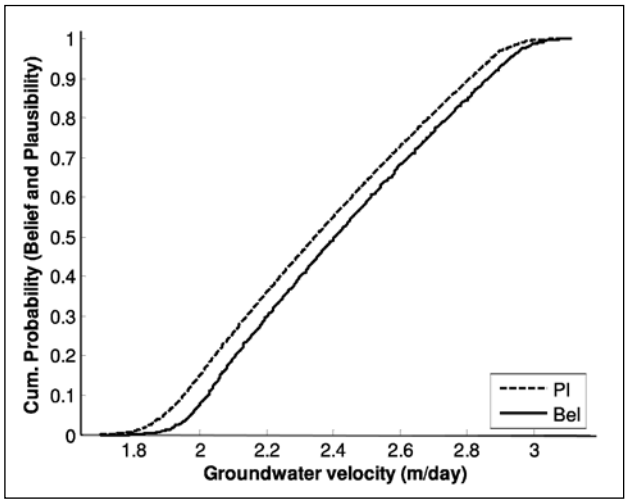


Figure 1 Belief and Plausibility of groundwater velocity

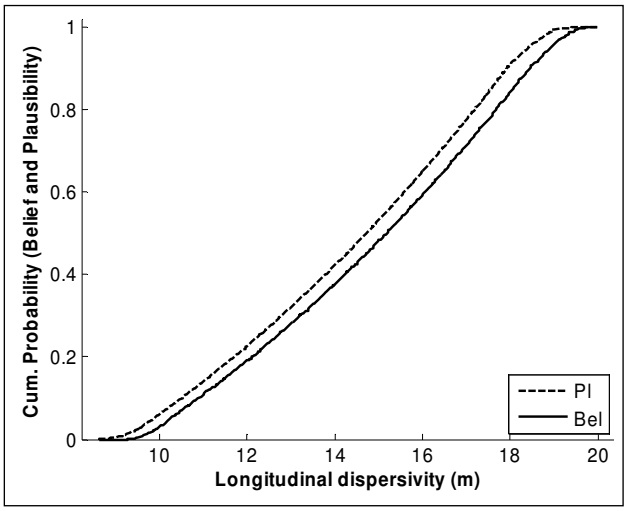


Figure 2 Belief and Plausibility of longitudinal dispersivity

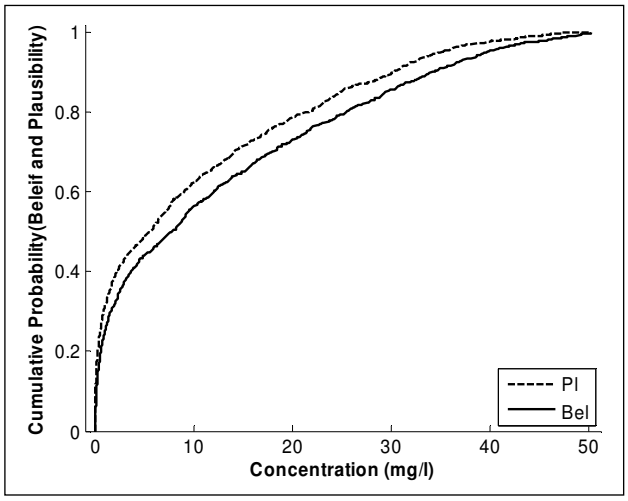


Figure 3 Belief and Plausibility of Concentration profile at x = 1220 m and at t = 400 days

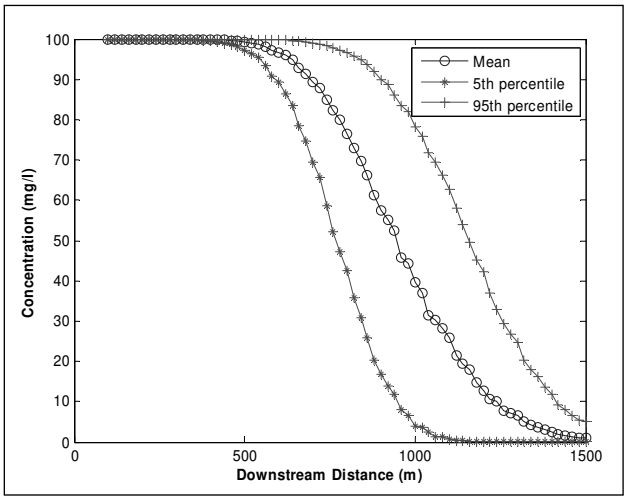


Figure 4 Plot of the concentration profile with uncertainty

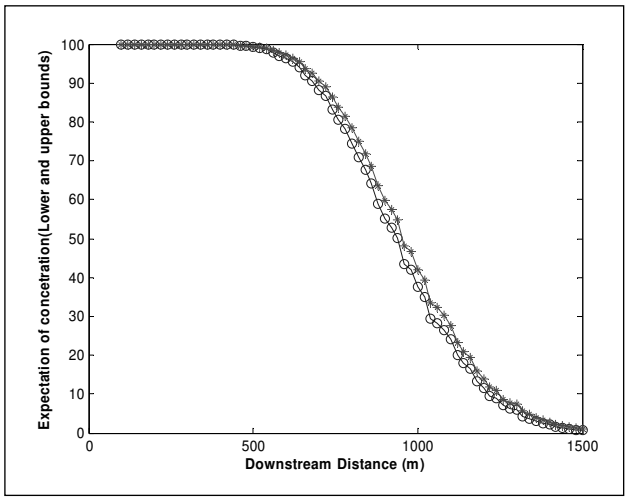


Figure 5 Uncertainty of expected concentration

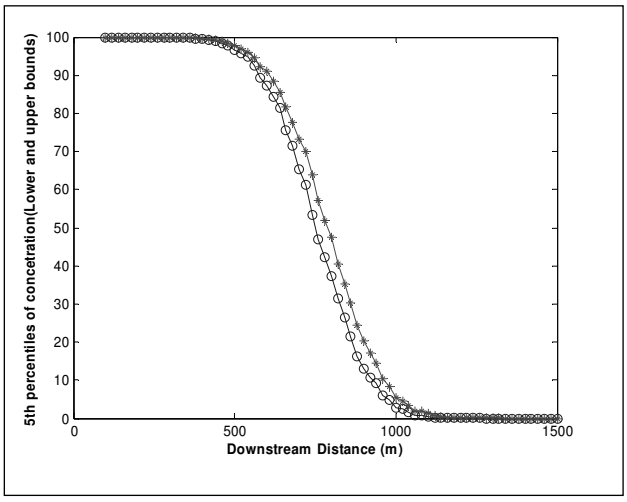


Figure 6 Uncertainty of 5th percentile of concentration

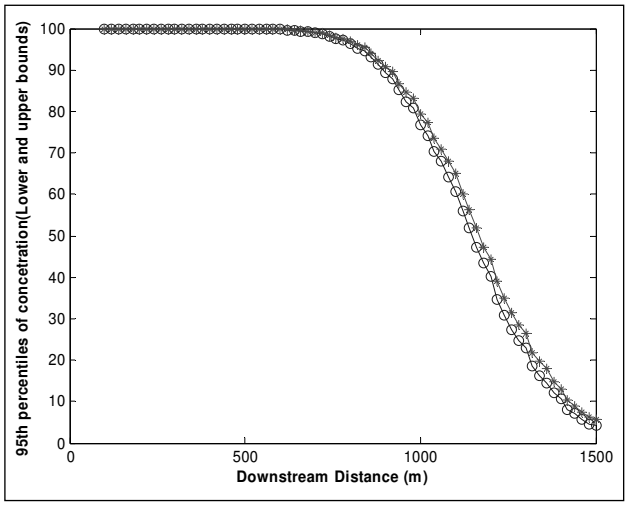


Figure 7 Uncertainty of 95th percentile of concentration

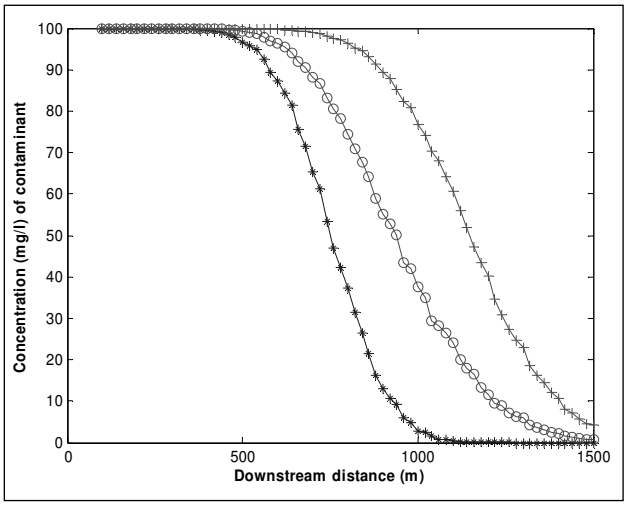


Figure 8 Lower bound of expected concentration with lower bounds of 5th and 95th percentiles

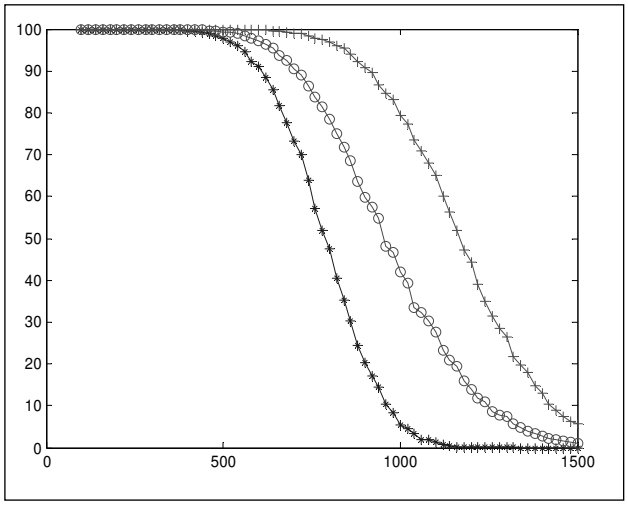


Figure 9 Upper bound of expected concentration with lower bounds of 5th and 95th percentiles

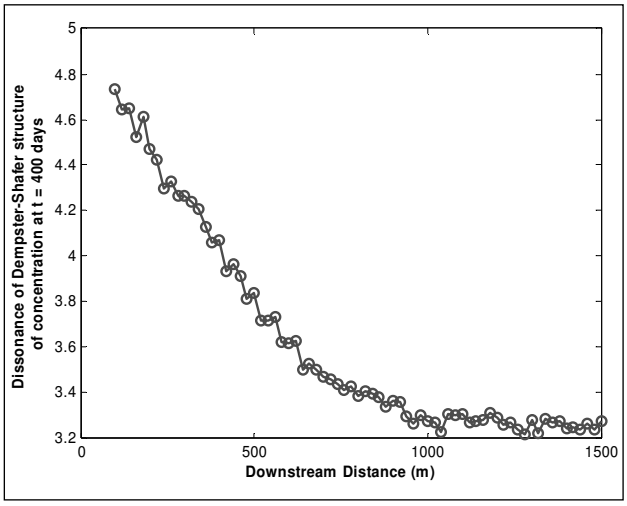


Figure 10 Dissonance measure of uncertainty of concentration profile

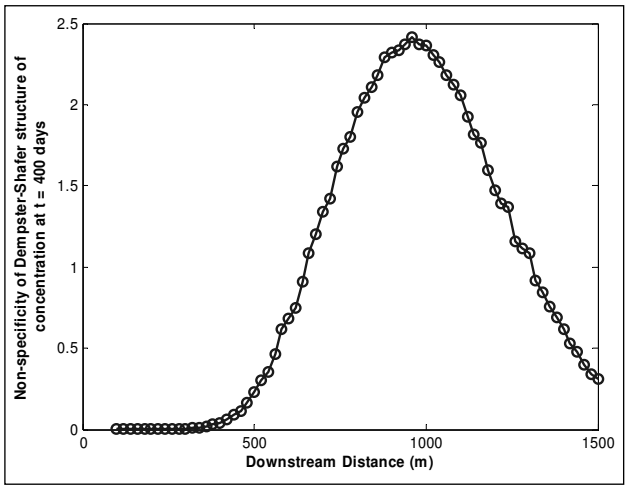


Figure 11 Non-specificity measure of uncertainty of concentration profile

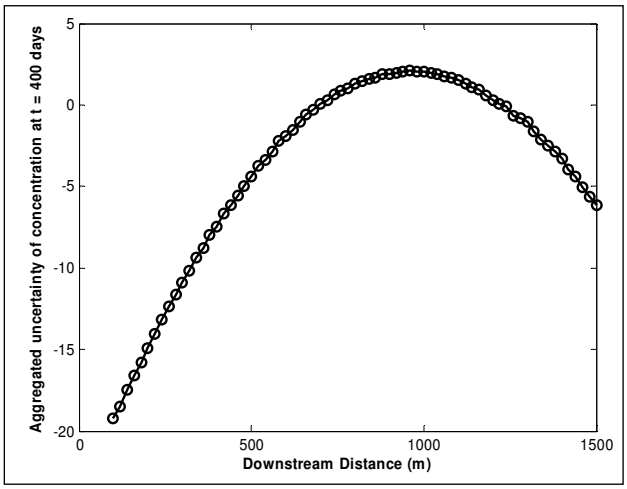


Figure 12 Aggregate measure of uncertainty of concentration profile

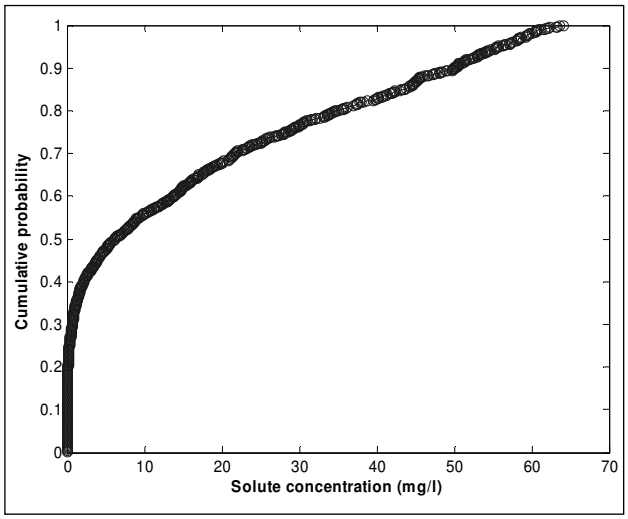


Figure 13 Cumulative probability plot of concentration profile at (x = 1220m, t = 400 days)

of groundwater velocity and longitudinal dispersivity are as shown in figures 1 and 2. In order to compute the belief and plausibility of concentration of contaminant, time of observation is kept fixed at t = 400 days and downstream distances are varied from 100 m to 1500 from the point of discharge of the contaminant with a step size of 20 m. Figure 3 presents the plot of belief and plausibility of the concentration of contaminant at downstream distance, x = 1220 m at time t = 400 days. The belief of estimated concentration 10.71mg/l at x = 1220 m and at t = 400 days is seen from figure 3 as 0.5760 whereas the plausibility is computed as 0.6420. Therefore, the uncertainty of estimated concentration 10.71 mg/l is expressed as [*bel* = 0.5760, *pls* = 0.6420]. Computation of various other uncertainty measures of concentration of the contaminant such as expectation value, 5th percentile and 95th percentile at each downstream location and at the same time, t = 400 days are carried out. Figure 4 presents the mean concentration profile with mean 5th and 95th percentiles. Mean concentration here is signified as the average value of the lower and upper bound of the concentration and the same is true for mean 5th and 95th percentiles also. Figures 5, 6 and 7 show the lower and upper bound of expectation of concentration profile and the corresponding profiles of 5th and 95th percentiles. Interpretation of these figures can be framed in this way that the average width of the bounds of each individual quantity (expectation, 5th and 95th percentiles) is very narrow signifying further that the spread at each level of measures is not very significant. The separate profiles of lower and upper bound of the expectation value, 5th and 95th percentile of concentration of contaminant for various downstream distances are as shown in figures 8 and 9. It can be easily notified from figure 8 and 9 that, at downstream distance of 700 m, the lower and upper bound of the mean value of concentration of the contaminant is [88.19, 90.51] mg/l, that of 5th percentile is [65.42, 73.29] mg/l and the same of 95th percentile is [98.94, 99.08] mg/l. Computation of dissonance, non-specificity and aggregated uncertainty of concentration of contaminant at time t = 400 days for same downstream distances are also carried out using evidence theory. Results of these computations are presented in figures 10, 11 and 12. It can be interpreted from figure 10 that dissonance measure of uncertainty of concentration over downstream distances (100 m – 1500 m) decays exponentially with the increase of downstream distance. Therefore, at minimum dissonance, information on overlap attains nil. The non-specificity measure also follows the same philosophy. Aggregated uncertainty (AU = Non-specificity + Conflict) provides the estimation of the conflict by subtracting it from the measured non-specificity. Measures of non-specificity and aggregated uncertainty of concentration of contaminant at downstream location, x= 960 m at time 400 days from figures 11 and 12 are 2.419 and 2.086 respectively. Therefore, conflict is computed as -0.33, signifying that there is no conflict on the evidence obtained from the experts. In order to prove the fact that probability of occurrence of the value of random variable lies between belief and plausibility, computation of the contaminant transport problem is carried out further by taking into account the model parameters as uniformly distributed (probabilistic feature). Cumulative probability distribution plot of the concentration profile for the same location and at the same time is as shown in figure 13. The probability of occurrence of mean concentration of 10.71 mg/l can be easily observed from figure 13 as 0.59. Therefore, it is proved that the probability always

lies within the interval of belief and plausibility. All the computations are carried out using in-house software RAUDSET (Risk Analysis Using Dempster Shafer Evidence Theory) developed using Visual Basic and C++ (Microsoft Visual Studio).

5.0 Conclusions

Impreciseness of data or knowledge leads to uncertainty known as epistemic uncertainty. Epistemic uncertainty can be modeled using evidence theory. Evidences are collected using expert's knowledge. The procedures of modeling and propagation of epistemic uncertainty using Dempster-Shafer evidence theory is illustrated through contaminant transport model. Evidence theory shows promising to handle epistemic uncertainty. Uncertainty of this variety is expressed in terms of belief and plausibility. Computation of various other measures such as aggregated uncertainty, dissonance, non-specificity, expectation value, 5th percentile and 95th percentile of the model output proves the power of evidence theory. However, if the model is complex (multi-component), it requires extensive computational effort to evaluate belief and plausibility of each component and joined them using a suitable network. So, the future work will focus to develop an algorithm to compute belief networks.

References

1. INTERNATIONAL ATOMIC ENERGY AGENCY, *Generic Models for Use in Assessing the Impact of Discharges of Radioactive Substances to the Environment*, Safety Series No. 19, IAEA, Vienna, 2001.
2. INTERNATIONAL ATOMIC ENERGY AGENCY, *Generic Models and Parameters for Assessing the Environmental Transfer of Radionuclides from Routine Releases: Exposures of Critical Groups*, Safety Series No. 57, IAEA, Vienna, 1982.
3. INTERNATIONAL ATOMIC ENERGY AGENCY, *Hydrological Dispersion of Radionuclide Material in Relation to Nuclear Power Plant Siting*, Safety Series No. 50-SG-S6, IAEA, Vienna 1985.
4. Datta, D., and H.S. Kushwaha, Fundamental Statistics for Uncertainty Analysis, *Uncertainty Modeling and Analysis*, Bhabha Atomic Research Centre, Trombay, Mumbai, 400085, (Editor: H.S. Kushwaha), pp. 1-48, 2009, ISBN: 978-81-907216-0-8.
5. Ayyub, B.M and Klir, G.J., *Uncertainty Modeling and Analysts in Engineering and the Sciences,* 2006, Chapman & Hall/CRC Press, Boca Raton, FL.
6. Datta, D., Statistics of Monte Carlo Methods Used in Radiation Transport Calculation*, Applications of Monte Carlo Methods in Nuclear Science and Engineering*, Bhabha Atomic Research Centre, Trombay, Mumbai, 400085, (Editor: H.S. Kushwaha), April, 2009, ISBN: 978-81-8372-047-2
7. Christos E. Papadopoulos, Hoi Yeung, Uncertainty estimation and Monte Carlo simulation method, Flow Measurement and Instrumentation, 2001, vol. 12, pp. 291–298.
8. Dubois, D., Nguyen H.T. and Prade H., Possibility Theory, Probability and Fuzzy Sets: Misunderstandings, Bridges and Gaps, *Fundamentals of Fuzzy Sets*, Kluwer Academic Publishers, Boston, 2000, pp. 343-438.
9. Shafer G., *A Mathematical Theory of Evidence*, 1976, University Press, Princeton.
10. Klir, G. J., Wierman, M. J., 1998, *Uncertainty-Based Information*, A Springer-Verlag Company.
11. Yager, R.R., Entropy and specificity in a mathematical theory of evidence, Int. J. Gen. Syst/., 1983, vol. 9, pp. 249-260.
12. Dou, C., Woldt W., Bogardi I., and Dahab M., Numerical solute transport simulation using fuzzy sets approach, Journal of contaminant hydrology, 1997, vol. 27, pp. 107-126.

Life Cycle Reliability and Safety Engineering
Vol.2 Issue 2 (2012) 81-87

Search For Optimal Preventive Maintenance Policy Of Equipment Under An Uncertainty Of Detection Of Its Condition

Anil Rana*, Professor Ajit Kumar Verma, *Professor A Srividya******
Indian Institute of Technology, Powaii
**ranaanil13@hotmail.com*
***akv@ee.iitb.ac.in*
****asvidya@civil.iitb.ac.in*

Abstract

Preventive maintenance of equipment is generally chosen over the corrective maintenance policy in order to preclude the chances of sudden failure that incurs high opportunity and repair costs. However, the choice between a time based preventive maintenance and a condition based preventive maintenance is generally carried out under an assumption that the probability of detection of the deteriorating condition of the equipment is 1. This assumption may be far fetched, as most of the condition monitoring techniques have a probability of correct detection of equipment condition less than 1. In some other cases, even the deteriorating condition that is being measured may not have a perfect correlation with the equipment state. The uncertainties involved in use of a condition monitoring system may result in making an improper choice of the PM policy resulting in sub-optimal use of resources. This paper presents a method of selection of a suitable preventive maintenance policy under the uncertainty involved in correct detection of the deteriorating condition of a equipment. A non-stationary Gamma wear process has been used to model the deteriorating condition of the equipment and the wear thresholds for alarm and time for monitoring the condition have been included as decision variables for deciding the optimal CBPM.
Key words : Gamma wear process, TBPM, CBPM

Introduction

For an equipment or a machinery, there are only two kinds of maintenance actions: Preventive maintenance and Corrective maintenance. Preventive maintenance can either be time based or condition based. A time based maintenance is understood to be a maintenance action where in, no condition monitoring is undertaken, instead the equipment is replaced or maintained at periodic (or fixed time or age) time intervals. Condition based maintenance on the other hand involves monitoring of the condition of the equipment. When a specified level of deterioration or wear of the subject equipment is surpassed, the equipment may be replaced or repaired. There will always be a chance of breakdown of machinery under both the above preventive maintenance policies that will give rise to a corrective maintenance incurring high cost of repair and opportunity and at times this may have some safety implications too. The optimal choice of the preventive maintenance policy will therefore be guided by the degree to which the chance of corrective maintenance is minimized.

From Barlow and Hunter [1] in 1960, till date, there have been many models and case studies on preventive maintenance policies. References from [1] to [5] are few such examples. Wang[6] provided a thorough review of time based preventive maintenance approaches in the literature. The author has discussed age dependent preventive maintenance policies, periodic preventive maintenance policies, failure rate limit policies, sequential preventive maintenance policies, repair limit polices, opportunistic maintenance polices and optimization approaches for maintenance policies. Blischke and Murthy[7]

have also provided a broader view of many of the maintenance policies available in practice. Mann et al [8] provided a review of time based PM models and condition based models. Endrenyi et al [9] proposed use of RCM (Reliability Centre Maintenance) to determine the most cost effective maintenance policy for a given system. Whereas Saranga[10] proposed a structured approach method called the RCP or Relevant Condition Parameter which selects the maintenance significant items according to a risk priority number. Condition based maintenance has been explored by many researchers such as Grall et al [11], Fouadirad et al[12] and Barata et al [13] and many others. Grall et al [11] has proposed a varying time based monitoring interval based on the extent of deterioration. The other authors have considered continuous monitoring and few others have considered joint effect of shock and deterioration as the failure process of the equipment. In what follows, we consider the comparison between time based PM and condition based PM with the choice of monitoring time intervals, wear threshold and probability of detection.

TBPM or CBPM

Time based maintenance are generally proposed as an effective strategy for less critical systems which also have a comparatively smaller degree of variability in the failure time distributions. Condition based maintenance techniques on the other hand are justified for highly critical systems that require effective maintenance planning and execution. However, before a CBPM based policy can even be applied on a equipment, availability of a particular parameter that can accurately detect the deteriorating condition of a particular failure mode of the equipment need to be analyzed. For a CBPM policy to be applied on an equipment, it is important that the wear or deterioration progress with respect to time be completely defined in terms of a continuous stochastic process. The process can then be used to define two different levels of deterioration, one the alarm level and the other the failure level. The time for the equipment deterioration to reach the failure level from the alarm level becomes an important consideration in deciding whether enough time is available for the maintainers to act before the equipment fails catastrophically.

The question, therefore, that would often arise during a condition based maintenance decision making is that what should be the interval of monitoring of the equipment condition ? or at what probability of detection of the equipment condition, should one consider the CBPM to be economically viable ? or given a probability of detection if one has to shift away from the optimal time interval of monitoring to another time schedule, would the CBPM sill be advantageous over the time based maintenance? In the above arguments we have safely assumed that the corrective maintenance actions are the most cost and safety prohibitive and therefore need not be considered.

Modeling of Wear/Deterioration of Equipment Condition Using Gamma Process

The gamma process was applied in a series of papers in the fifties to model water flow into a dam, Moran [14,15,16]. However, it was proposed to model deterioration occurring random in time only in 1975. Since then it has been satisfactorily fitted to data on creep of concrete Cinlar et al[17], fatigue crack growth Lawless et al[18], corroded steel gates Frangpool et al[19], thinning due to corrosion Kallenet al [20] etc. A method for estimating a gamma process by means of expert judgment is proposed in Nicolai et al[21]. Gamma wear process which is non-stationary has been shown as the most suitable process in Pandey et al [22] that can take care of the temporal variability of the wear process. In this process the system failure behavior might be described by a damage accumulation model or shock model. The system state at any time 't' can be summarized by a random ageing variable/deterioration W_t. In the absence of repair or replacement actions, W_t is an increasing stochastic process, with W_0=0. The system will fail when the ageing variable or deterioration exceeds a predetermine threshold level W_f. The gamma process is also a reasonable extension of a deterioration process with exponential jumps. The gamma process is parameterized by α and β which can be estimated

from the deterioration data . If W_t (deteriorating state) is a gamma process then for all 0 < = s < t the random variable $W_t - W_s$ (increments of deterioration between s and t) has a gamma pdf with shape parameter α(t-s) and a scale parameter β, given by :

$$f_{\alpha(t-s),\beta}(w) = \frac{\beta^{\alpha(t-s)}}{\Gamma(\alpha(t-s))}.w^{\alpha(t-s)-1}.e^{-w.\beta}I_{\{x\geq 0\}} \quad (1)$$

The gamma process has a non-negative independent increment property. The mean and variance of its degradation rate can be expressed as α(t-s)/β and α(t-s)/β^2. For such a process the deteriorating state starting from w_0, the associated failure time distribution, CDF for a given failure threshold, W_f can be expressed as

$$F_{\alpha,\beta}(w) = 1 - \frac{1}{\Gamma(\alpha.t)}.\int_0^{(W_f - w_0).\beta} e^{-u}.u^{(\alpha t-1)}.du \quad (2)$$

Consider the process of wear or deterioration of a equipment shown in figure 1 below (Grall [23]). As time progresses, the equipment deteriorates. The equipment is monitored at regular intervals for its deterioration or wear. There are two wear levels which are of some consequence. A wear threshold level is an alarm level. If on an inspection it is observed that the wear of deterioration of the equipment has crossed the threshold level, it is preventively replaced with a new one or maintained so that its condition becomes as good as new. If however, the wear crosses the 'wear limit', it is considered to have failed and the equipment needs to be correctively replaced. A gamma wear process helps include the wear levels of alarm and failure into the model calculations.

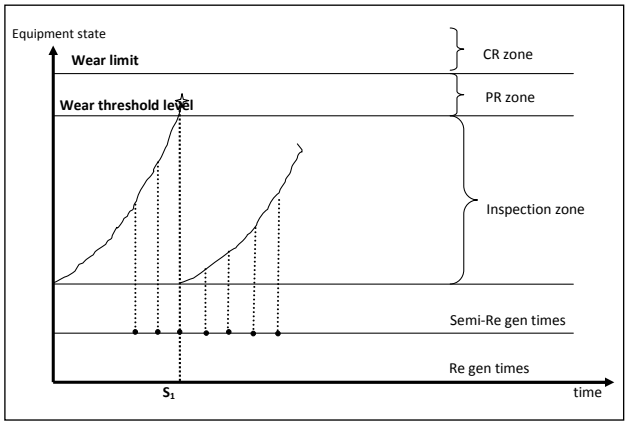


Fig. 1 Schematic evolution of the maintained system state

Choice Between TBPM and CBPM Using An Example

Using a renewal cycle method we show evaluation of an optimal preventive maintenance policy based on time based PM schedule and a condition based, probability of detection based and monitoring time interval based CBPM. We consider two main probabilities of maintenance : a probability of carrying out preventive maintenance (which can be either time based or a condition based) and a probability of corrective maintenance. When the TBPM is in force, a corrective maintenance is possible only when the equipment fails before the scheduled time 'T' for PM is clocked. When the CBPM is in force, a corrective maintenance is possible only when the wear reaches the alarm level 'W_{th}' and the failure level 'W_{lim}' between two consecutive monitoring time schedules. It can also take place when the alarm threshold level is present during a monitoring time however the correct detection does not take place during the monitoring.

Using a non-stationary gamma wear process to map the wear or deterioration we evaluate the costrate of maintenance of a equipment under consideration. The model is placed at Appendix 'A'

Assumptions

The following assumptions are being made in the model being discussed

- The equipment is preventively replaced when the condition being monitored reaches a wear threshold level. When following the time based PM, the equipment is replaced at regular time based intervals
- If the probability of detecting the correct condition of the equipment is 'P', there is 0.5*(1-P) chance of making a wrong detection on the safer side. That is raising an alarm, when the condition is still normal
- Though 'P' is the probability of correct detection of the condition of the equipment,

there is a perfect correlation between the parameter being monitored and the actual condition of the equipment

- The equipment follows a non-stationary gamma wear process with shape parameter '$\alpha.t^{\zeta}$' and scale parameter 'β' where α = 0.02278; β =1.2; ζ=1 . The wear threshold for failure is a non-dimensional number = 10
- Cost of setting up a comprehensive condition monitoring system has not been included in the example

Results

Cost rate drawn up for various wear threshold levels have been shown in figure 2. Similarly the probability of carrying out PM or preventive maintenance under various probabilities of detection is shown in figure 4 . Figure 2 has been drawn up with probability of detection of 1 and replacement at monitoring time interval or 'T'. Figure 3 shows a similar plot for a time based PM with renewal cycle= T; the time for maintenance interval. The plots in figure 2 display the optimal wear threshold level and the optimal time for monitoring the condition of a given equipment. The plots also display the alternatives available with the maintenance engineer in deciding the wear threshold level he would like to choose for his equipment. Since the time between the wear threshold level and the wear limit level is crucial in making the logistics arrangement ready for the upcoming maintenance actions, a maintenance engineer may like to choose the wear threshold level that may not be an optimal solution. The time available to the maintenance engineer, once the wear threshold level has been reached can be given in accordance with the approximation formula Rana et al [24]

$$\left(\frac{\text{wearlimit}.\beta}{\lambda}+\frac{0.479}{\lambda}.e^{\left(\frac{1-\zeta}{\Gamma(\beta+\zeta)}\right)}\right)^{\frac{1}{\zeta}}-\left(\frac{\text{wearthreshold}.\beta}{\lambda}+\frac{0.479}{\lambda}.e^{\left(\frac{1-\zeta}{\Gamma(\beta+\zeta)}\right)}\right)^{\frac{1}{\zeta}} \quad (3)$$

The maintenance engineer may also not be able to provide regular monitoring at every optimum time interval because of various constraints, instead he can choose the time interval suitable to him and know the consequences in terms of cost rate as per the plots displayed in figure 2. It may be noted in figure 2 that as the wear threshold level rises from 5.0 to 9.0 the optimal time interval in that particular wear threshold moves to the left. This is because the time available for the equipment to reach the wear limit for failure (assumed to be 10 in this case) becomes shorter and therefore the monitoring becomes more frequent.

Going by the solutions in figure 2 and 3, one can see that the CBPM is a better preventive maintenance policy than the TBPM, when the wear alarm threshold level is maintained at 8 and the monitoring is carried out every 12 days. However, if the logistics time delay (time in arranging resources for replacement of the equipment) does not allow for this wear

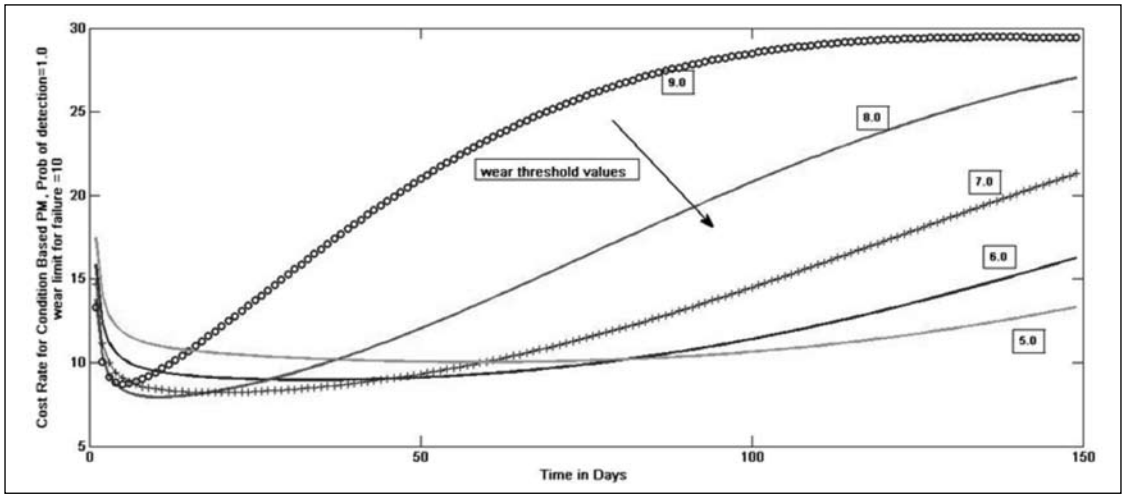


Fig.2 Cost rate for various wear threshold levels and condition monitoring intervals

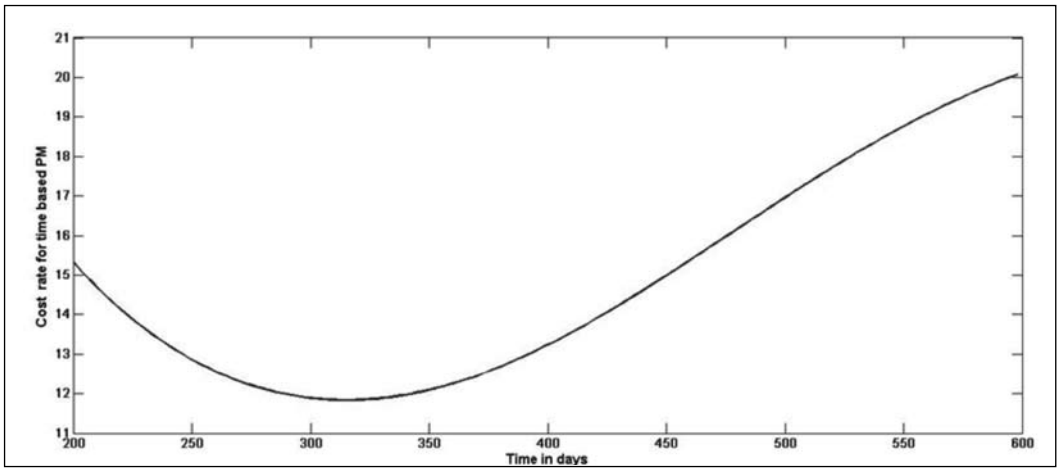


Fig. 3 Cost rate for time based PM

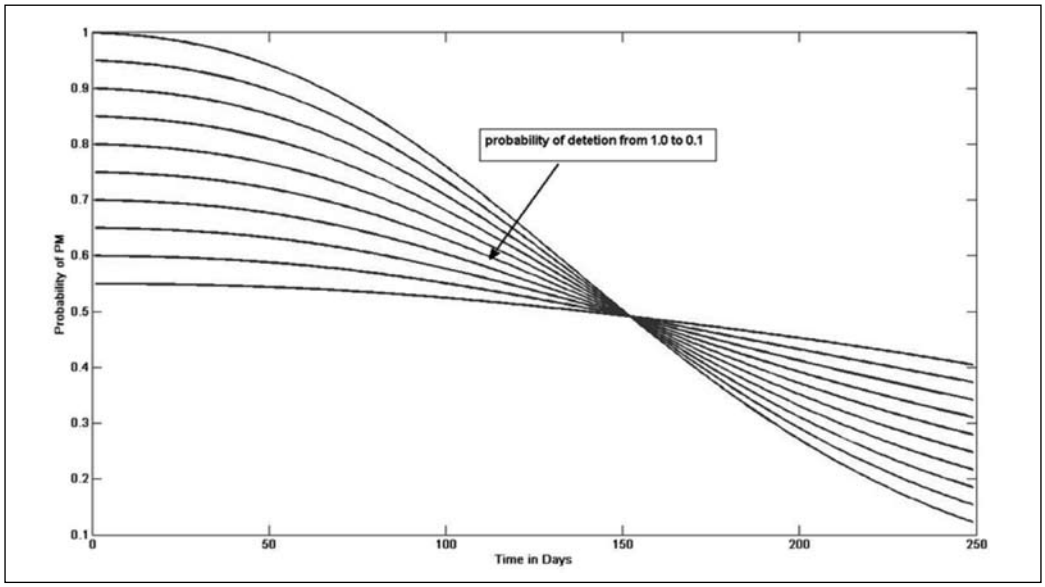


Figure 4 Probability of carrying out PM for various probabilities of detection and monitoring interval a=0.02278;β=1.2;ζ=1;wear threshold =7; wear limit=10

alarm threshold, and if this threshold has to be maintained at 5, the TBPM seems to be a better PM policy. It may be noted that these results are drawn up for probability of detection of 1. For values less than 1 a different set of graphs will need to be drawn up. The effect of the probability of detection on the probability of PM has been clearly shown at figure 4.

REFERENCES

1. R Barlow and Larry Hunter, "Optimum preventive maintenance policies", Operational Research. Vol.8, pp 90-100 feb 1960

APPENDIX 'A'

$$\text{Probabilty of PM} = \sum_{n=0}^{\infty} \text{Probability of detection X(Probability of the detrioration crosses threshold level between } n+1.\Delta t \text{ and } (n)\Delta t)$$
$$\text{X(Probability that deterioration doesn't cross wear limit between } n+1.\Delta t \text{ and } (n)\Delta t))$$
$$+0.5(1-\text{Probability of detection})\text{X}$$
$$\text{(Probability that deterioration crosses threshold level between only after } n.\Delta t\text{))} \qquad \text{(A1)}$$

$$P_{PM} = \sum_{n=0}^{\infty} P_d \cdot \left[\int_0^{(W_{th})} \frac{\beta^{\alpha.(n+1)\Delta t^{\zeta}}}{\Gamma(\alpha.(n+1)\Delta t^{\zeta})} x^{\alpha.(n+1)\Delta t^{\zeta}-1} .e^{-\beta x} .dx - \int_0^{(W_{th})} \frac{\beta^{\alpha.(n)\Delta t^{\zeta}}}{\Gamma(\alpha.(n)\Delta t^{\zeta})} x^{\alpha.(n)\Delta t^{\zeta}-1} .e^{-\beta x} .dx \right] .$$

$$\left[1 - \int_0^{(W_{\lim} - W_{th})} \frac{\beta^{\alpha.\Delta t^{\zeta}}}{\Gamma(\alpha.\Delta t^{\zeta})} x^{\alpha.\Delta t^{\zeta}-1} .e^{-\beta x} .dx \right] + 0.5.(1 - P_d).$$

$$\left[\int_0^{(W_{th})} \frac{\beta^{\alpha.(n+2)\Delta t^{\zeta}}}{\Gamma(\alpha.(n+2)\Delta t^{\zeta})} x^{\alpha.(n+2)\Delta t^{\zeta}-1} .e^{-\beta x} .dx - \int_0^{(W_{th})} \frac{\beta^{\alpha.(n+1)\Delta t^{\zeta}}}{\Gamma(\alpha.(n+1)\Delta t^{\zeta})} x^{\alpha.(n+1)\Delta t^{\zeta}-1} .e^{-\beta x} .dx \right] \quad \text{(A2)}$$

$$\text{Cost Rate} = \frac{P_{PM}(\text{PM cost}) + P_{CM} * (\text{CM cost}) + CI * \text{Renewal cycle}/\Delta t}{\text{Renewal cycle}}$$

where

we have $P_{CM} = 1 - P_{PM}$

$$\text{Renewal Cycle} = \sum_{n=0}^{n=\infty} (n+1).\Delta t.P_d \left(\begin{matrix} (\text{Prob of wear threshold being reached between } n\Delta t \text{ and } (n+1)\Delta t). \\ \text{prob of wear limit not being reached between } n\Delta t \text{ and } (n+1)\Delta t) \end{matrix} \right) +$$

$$\sum_{n=0}^{\infty} \left(\begin{matrix} (\text{Prob of wear threshold being reached between } n\Delta t \text{ and } (n+1)\Delta t). \\ \text{mean time of wear limit being reached between } n\Delta t \text{ and } (n+1)\Delta t) \end{matrix} \right) +$$

$$\sum_{n=0}^{n=\infty} (n+1).\Delta t \left(\begin{matrix} (\text{Prob of wear threshold being reached between } (n+1)\Delta t \text{ and } (n+2)\Delta t).0.5(1 - P_d). \\ (\text{This represents time due to wrong alarm}) \end{matrix} \right)$$

$$\sum_{n=0}^{\infty} \left(\begin{matrix} 0.5(1 - P_d).(\text{Prob of wear threshold being reached between } (n-1)\Delta t \text{ and } (n)\Delta t). \\ \text{mean time of wear limit being reached between } (n)\Delta t \text{ and } (n+1)\Delta t) \end{matrix} \right)_{n>1} \quad \text{(A3)}$$

meantime to reach the wear limit between $n\Delta t$ and $(n+1)\Delta t$ can be given as $\int_{n\Delta t}^{(n+1)\Delta t} x.f(x).dx$

P_{PM} = probability of preventive maintenance

P_{CM} = probability of corrective maintenance

P_d = Probability of detection

W_{th} = wear threshold level

W_{lim} = wear limit for failure

Δt = time interval for condition monitoring

$\alpha.t^{\zeta}$ = shape parameter of the gamma wear process

β = scale parameter of the gamma wear process

n = number of cycles

CI = cost of inspection; PM ost = cost of PM and CM cost = cost of CM

2. H Mine and H Kawai, "Preventive replacement of a 1 unit system with a wear out state", IEEE Transactions on reliability, vol R-23, no.1, April 1974
3. T. Nakagawa, "Optimum preventive maintenance policies for repairable systems", IEEE Transactions on Reliability, vol R-26,1977
4. KS Park, "Optimal continuous wear limit replacement under periodic inspections", IEEE Transactions on Reliability , 37(1), 97-102. 1988
5. IB Gertsbakh, "Models of Preventive Maintenance", 1977, North Holland Publishing company
6. Wang H. A survey of maintenance policies of deteriorating systems. Eurpoean Journal of Operational Research, 139, 469-489, 2002
7. Blischke WR and Murthy DN, " Case Studies in Reliability and Maintenance", New Jersey, John Wiley and Sons. (2003)
8. Mann L, Saxena A, Knapp G, " Statistical Based or Condition Based Maintenance? Journal of Quality in Maintenance, 1, 46-59, 1995.
9. Endrenyi J, Aboresheid S, Allan N, Anders GJ, Asgarpoor S, Billinton R, The Present Status of Maintenance Strategies and the Impact of Maintenance on Reliability. IEEE Transactions on Power Systems, 16, 638-646, 2001
10. Saranga, H. Relevant Condition Parameter Strategy for an Effective Condition Based Maintenance. Journal of Quality

in Maintenance Engineering, 8, 92-105, 2002.

11. Grall A, Berenguer C and Dieulle L. A condition Based Maintenance Policy for Stochastically Deteriorating Systems. Reliability Engineering and System Safety 76, pp 167-180, 2002
12. Fouladirad, Grall A, Dieulle L. On the Use of Oi-Line Detection for Maintenance of Gradually Deteriorating Systems. Reliability Engineering and System Safety. 93, pp 1814-1820, 2008
13. Barata J, Soares CG, Marseguerra M, Zio E. Reliability Engineering and System Safety, 76, pp 255-264, 2002
14. Moran PAP. A probability theory of dams and storage systems. Australian Journal of applied science 1954;5(2):116-24
15. Moran PAP. A probability theory of dams and storage systems: modifications of the release rules. Australian Journal of applied science 1955;6(2):117-30
16. Moran PAP. The theory of storage. London:Methuen: 1959
17. Cinlar E, Bazant ZP, Osman E., Stochastic process for extrapolating concrete creep. J Engg Mech Div 1977; 103 (EM6):1069-88
18. Lawless J, CrowderM, "Covariates and random effects in a gamma process model with application to degradation and failure", Lifetime data Analysis 2004;103(3):213-27
19. Frangpool DM, Kallen MJ, Van Noortwijk JM, "Probabilistic models for life cycle performance of deteriorating structures: review and future directions", Prog Struct Eng Mater 2004:6(4):197-212
20. Kallen MJ, Van Noortwijk JM. Optimal maintenance decisions under imperfect inspection. Reliability Engineering and System Safety 2005;90(2-3) 177-85
21. Nicolai RP, Dekker R, Van Noortwijk JM, "A comparison of models for measurable deterioration: An application to coatings on steel structures.", Reliability and Engg System Saftey 2007, in press, doi:10.1016/j.ress. 2006.09.021
22. Pandey MD, Yuan X-X, Van Noortwijk JM. A comparison of probabilistic deterioration models for life cycle management of structures. Structure Infrastructure Engineering, 2007.
23. Grall A, Laurence D, B Christophe and R Michael," Continuous time predictive maintenance scheduling for a deteriorating system", IEEE Transactions on Reliability, vol 51, no. 2, June 2002
24. Rana Anil, Verma AK, Srividya A, "Approximation of MTTF calculation of a non-stationary Gamma wear process", (Under review) International Journal of System Assurance Engineering and Management, submitted 18 Jul 2011.

Life Cycle Reliability and Safety Engineering
Vol. 2 Issue 2 (2012) 88-96

A Multi-Agent based control scheme for Accelerator pre-injector and transport line for enhancement of accelerator operations

R. P. Yadav, P. Fatnani, *RRCAT, Indore, India,* **P. V. Varde**, BARC, *Mumbai, India,*
P.S.V. Nataraj, IIT, *Mumbai, India*

Abstract:

Reliable accelerator operation requires control system with higher level of automation, flexibility, robustness, and optimisation. In this paper a multi-agent system based control scheme is presented for optimal control of accelerator system that improves the plant performance in wide-range of operations. The multi-agent based control schemes for accelerators have been reported in literature. But the scheme proposed in this paper differs significantly form the existing schemes. In this work the agent architecture is formulated based on the control requirements of pre-injector accelerator subsystem (Microtron in particular) and transport line of synchrotron radiation sources. The scheme consists of two software agents at supervisory level that work in an autonomous manner for the optimized control of dynamic system. The Microtron agent architecture augments model assisted adaptive controller for realizing feedback control action at lower layer and goal based logic controller with pre-structure model identifier along with the pattern recognizer at supervisory layer. The TL-1 agent has a model-based, goal-based modular architecture and optimizes the TL-1 control using differential evolution based algorithm. The simulation results of applying this scheme to model of Microtron and Transport Line-1 of INDUS complex shows that this approach is very effective in optimizing the Microtron and TL-1 tuning.

Introduction:

Multi-agent based control schemes has been proposed by many researchers for exercising the robust control for large scale industrial system like power plants, power distribution systems, cement industry, ship board automation systems and accelerator control systems [1-5]. The existing multi-agent based control architectures for accelerator control [5- 12] could not be used directly for exercising the intelligent agent for controlling Microtron like accelerators, which exhibits dynamic nonlinear input-output behavior. Agent architecture for such systems requires augmenting functionality for adaptive feedback controller, dynamic and static model identifiers, and system state predictors based on historical data along with the supervisory level optimisation, communication, coordination and planning functionalities.

For controlling dynamic nonlinear systems using multi-agent based approach researchers have proposed different single agent architectures and organisation for multiple agents. J. D, Head et.al. [13] and S. Jin. et.al. [14] has proposed a three level based agent organisation with high level agents provide the man-machine-interface functionalities, at middle level the data base handling, task delegation and monitoring functionalities are handled, the lover most level implements the feedback and feedforward controllers with PI gain optimizer. The gain optimizer considers the current output of the plant and simulates the plant's response to feedback controllers using candidate gain values. Based on the response of the plant model, new candidate gain values are generated and tested. This process continues until the current set of candidate gain values meet all criteria deemed necessary in order to be considered acceptable. Once a set of optimal gains has been found, they are sent to the Feedback agent for immediate implementation. For generating the model of

subsystems neural network based off-line and on-line identifiers are proposed.

S. Kamalasadan [15, 16] has shown that the multi-agent based approach can be effectively used for controlling the dynamic systems showing multiple modes and drastic parametric jumps. Single link flexible robotic manipulator was controlled using three agents. The first agent is a heuristic based multiple fuzzy reference model generator that moves the reference model, mapping the system auxiliary state when it shows multi modality. This agent generates suitable reference model structure at every time instant. The second agent is a radial basis function neural network based controller that is used to augment the traditional model reference adaptive control in the presence of system functional uncertainty. Main emphasis was given to the use of neural network to approximate inverse dynamics of the plant working in parallel with a linear adaptive control law. The third agent is the traditional model reference adaptive controller which adaptively controls the system, linearises the parameters over a specific domain and forces the output or other plant variables to a suitable reference model structure.

Ben Nasr et.al.[17] proposed a model predictive control of a non liner fast dynamic system based on the multi-agent concept. The global system was first decomposed into sub-systems independent of one another. For each sub-system a model predictive control unit was made constituting the agent controller. Based on the analytical solution corresponding to the solution of the local receding horizon sub-problems, a logic unit was designed which by switching tries to find the best sequence of actions sent to the nonlinear system that gives the desired trajectory. In this way the sequences of actions were identified that bring the global system in a desired trajectory avoid any violation constraints on actions. A fuzzy controller was also made with an objective to handle the results of the actions on the global system and monitor the closed-loop system.

Agent-based control offers the ability to learn the patterns in system dynamics and use this information in determining the optimal, or near optimal control schema. Further to this by propagating this information among different agents in a multi-agent environment the global goals and global constraints could be easily handled. Such learning capabilities have not been sufficiently addressed in the literature. The current approach limits the learning ability more or less to online and offline system model identification only. Further to this almost all of the strategies rely upon using the neural network based models for modeling the nonlinear system. This requires a large amount of data set to be generated for offline identification and validation there by increasing the offline identification time. This can degrade the system performance in case such identification is needed more often for example for dynamic systems where the system model is needed to predict the faraway operating points. A less data driven approach could be the use of predefined model structure in the static model identification block where the best suited predefined model structure can be found according to the problem domain separately.

This paper presents an agent-based methodology for controlling the pre-injector and transport line operations. The agents learn the patterns observed in the system dynamics for both short-term and long-term basis and optimize there individual operations as well as there joint goals based on the learned patterns. The Microtron agent architecture augments model assisted adaptive controller for realizing feedback control action at lower layer and goal based logic controller with pre-structure model identifier along with the pattern recognizer at supervisory layer. The TL-1 agent has a model-based, goal-based modular architecture and optimizes the TL-1 control using differential evolution based algorithm. The rest of the paper is organized as follows. Section 2 describes the framework for proposed multi-agent based accelerator control. Section 3 describes the coordination scenario under normal conditions and section 4 gives the coordination scenario under constrain on actuator scenario. Section 5 gives the simulation results followed by conclusion.

The Accelerator System

The accelerator system comprises of three main parts; Microtron: it is a small accelerator which accelerates the electron beam upto 20MeV. It acts as pre-injector to synchrotron accelerator named Booster; Booster: it is another accelerator which accelerates the electron beam from 20MeV to 450MeV and 550MeV for injection to INDUS-1 and INDUS-2 respectively; TL-1: It is the transport line between Microtron and Booster accelerators which transfers the electron beam from one accelerator to other accelerator and serves the purpose of matching the parameters of beam available from Microtron to that of beam acceptance parameters at the Booster injection septum. The flow of beam between three parts is first beam is produced by the Microtron accelerator it then enters to the TL-1 which transports it to the Booster injection point.

The Microtron Model

The model of Microtron currently used in the development of the multi-agent system is based on the experimental identification of interdependence between different parameters. It is modeled as a four-input five-output nonlinear Simulink model shown in figure 1. The inputs into the system are Cathode current (I_{ca} in A) that controls the temperature of LaB_6 cathode inside Microtron RF cavity, RF frequency (f_{rf} in GHz) that provides the basic RF signal which is amplified by presiding amplifier stage and fed to the RF cavity for producing the required electric field in the cavity, acceleration start point in terms of I_{ca} ($ASPI_{ca}$ in A) which defines the system state and depends of various factors, Cavity resonant frequency (f_{cav} in GHz) it is the resonant frequency of the cavity at the particular time and depends primarily of the cavity temperature and electron

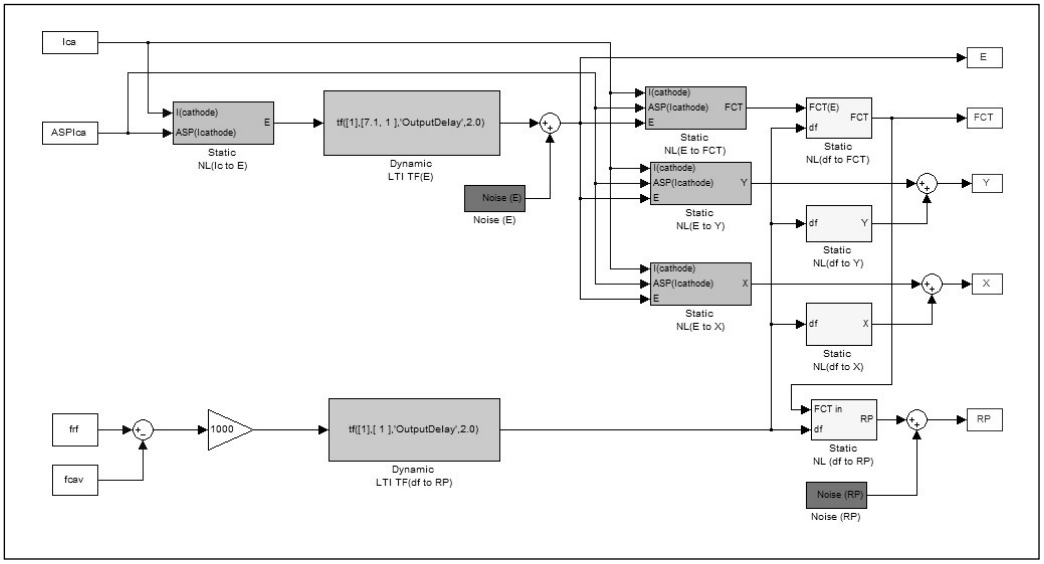


Figure 1. Simulink block diagram for Microtron model

emission level in cavity. The outputs of the model are Emission (E in V) which gives the measure of electrons emitted from cathode, Fast current transformer signal (FCT in V) which gives the measure of electrons actually accelerated to 20MeV level, beam position (X and Y in mm) at the extraction point, reflected power signal (RP in V) that gives the measure of power reflected by the cavity. The Eq. 1 to 9 gives the interdependence between different parameters used for Microtron modeling.

$$E = \begin{cases} 4.35 I_{ca} - 126.6 & \text{for } I_{ca} \le ASPI_{ca} \\ 1.324 I_{ca} - 35.11 & \text{for } I_{ca} > ASPI_{ca} \end{cases} \quad (1)$$

$$FCT_E = \begin{cases} 0.2 & \text{for } I_{ca} \le ASPI_{ca} \\ -0.1318E^2 + 1.389E - 3.036 & \text{for } I_{ca} > ASPI_{ca} \\ 0 & \text{for } E < 3.1 \\ 0.1 & \text{for } E > 7.4 \end{cases} \quad (2)$$

$$FCT = \begin{cases} 0.01 \\ -5.007\delta f^5 - 2.889\delta f^4 + 0.03362\delta f^3 - \\ 0 \end{cases}$$

$$\text{for } \delta f < -0.6 \text{ or } \delta f > 0.6$$
$$0.3748\delta f^2 + 0.3415\delta f + FCT_E \quad \text{for } -0.6 < \delta f < 0.6$$
$$\text{for } FCT < 0 \tag{3}$$

$$X_E = \begin{cases} 0.08935E & \text{for } I_{ca} \le ASPI_{ca} \\ 0.2935E & \text{for } I_{ca} > ASPI_{ca} \end{cases} \tag{4}$$

$$Y_E = \begin{cases} 1.391E & \text{for } I_{ca} \le ASPI_{ca} \\ 4.571E & \text{for } I_{ca} > ASPI_{ca} \end{cases} \tag{5}$$

$$X_{\delta f} = 0.9196\left(1 - e^{-\frac{(\delta f)^2}{0.2507}}\right) \tag{6}$$

$$Y_{\delta f} = 1300\left(1 - e^{-\frac{(\delta f)^2}{60.16}}\right) \tag{7}$$

$$RP = 0.115\left(1 - e^{\frac{(\delta f)^2}{0.5758}}\right) + C_0 \quad \text{Where } C_0 =$$

$$0.0672\frac{(FCT_{\max} - FCT)}{FCT_{\max}} \tag{8}$$

where $\delta f = (f_{rf} - f_{cav})$ is the deviation of RF generator frequency from the cavity resonant frequency expressed in MHz, and the beam position X and Y using Eq. 4 to 7 are calculated as below.

$$Beam\ Position\ (X,Y) = \begin{cases} X = X_E + X_{\delta f} - 1.6 \\ Y = Y_E + Y_{\delta f} - 24.8 \end{cases} \tag{9}$$

Noise at emission signal *Noise(E)* and noise at the reflected power signal *Noise(RP)* are modeled by autoregressive model given by Eq. 10 and 11.

$$Noise(E) = \frac{e(t)}{1 - 0.9982q^{-1} - 0.0007436q^{-2}}$$

$$Where\ \ e(t)\ =\ white\ noise \tag{10}$$

$$Noise(RP) = \frac{e(t)}{1 - 0.9983q^{-1} - 0.000969q^{-2}}$$

$$Where\ \ e(t)\ =\ white\ noise \tag{11}$$

And the dynamic response transfer function (TF) for emission signal (E) and reflected power (RP) are modeled as given by Eq. 12 and 13.

$$TF(E) = \frac{e^{-2.0s}}{7.0s + 1} \tag{12}$$

$$TF(RP) = e^{-2.0s} \tag{13}$$

For calculating the different settings the Microtron agent uses the static model thus bypassing the dynamic TF for E and RP.

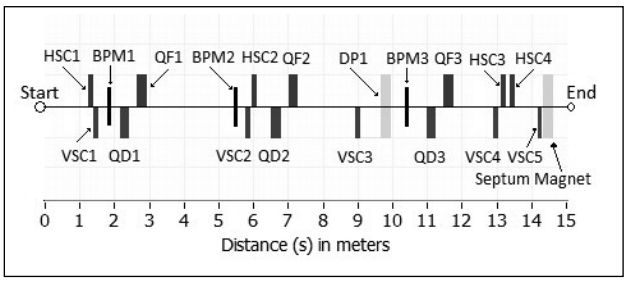


Figure 2. Layout of TL-1

The TL-1 Model

The layout of TL-1 is shown in Figure 2 using which the model of TL-1 is constructed by multiplying the transfer matrixes of individual elements. This model accepts the macro particle beam with attributes (x, y, x',y') and magnet settings $S=[I_1, I_2, \ldots, I_n]$ to produces the attributes (x, y, x',y') for macro particles at BPM1, BPM2, BPM3 locations and at the end of TL-1, The particle at the start of TL-1 with $X=[x_0, x_0']$ and $Y=[y_0, y_0']$ are transferred to the end of TL-1 using

Eq. 14 and 15. Similarly the transformations from TL-1 start to the respective BPM are given by Eq. 16 to 21

$$X_{End}^{T} = M_3^{x} C_2^{x} M_2^{x} C_1^{x} M_1^{x} X_0^{T} \quad (14)$$

$$Y_{End}^{T} = M_3^{y} C_2^{y} M_2^{y} C_1^{y} M_1^{y} Y_0^{T} \quad (15)$$

$$X_{BPM1}^{T} = M_{BP1}^{x} C_1^{x} M_1^{x} X_0^{T} \quad (16)$$

$$Y_{BPM1}^{T} = M_{BP1}^{y} C_1^{y} M_1^{y} Y_0^{T} \quad (17)$$

$$X_{BPM2}^{T} = M_{BP2}^{x} C_1^{x} M_1^{x} X_0^{T} \quad (18)$$

$$Y_{BPM2}^{T} = M_{BP2}^{y} C_1^{y} M_1^{y} Y_0^{T} \quad (19)$$

$$X_{BPM3}^{T} = M_{BP3}^{x} C_2^{x} M_2^{x} C_1^{x} M_1^{x} X_0^{T} \quad (20)$$

$$Y_{BPM3}^{T} = M_{BP3}^{y} C_2^{y} M_2^{y} C_1^{y} M_1^{y} Y_0^{T} \quad (21)$$

Where the Matrix M^x_1, M^x_2, M^x_3, M^y_1, M^y_2, M^y_3, M^x_{BPM1}, M^y_{BPM1}, M^x_{BPM2}, M^y_{BPM2}, M^x_{BPM3}, M^y_{BPM3} are calculated for the typical values of magnet settings at which the TL-1 is normally operated is given as below

$$M_1^x = \begin{bmatrix} 1 & 1.3280 \\ 0 & 1 \end{bmatrix} \quad M_2^x = \begin{bmatrix} -0.9556 & -0.5919 \\ 0.8418 & -0.5250 \end{bmatrix} \quad M_3^x = \begin{bmatrix} -0.7580 & -0.9286 \\ 0.6410 & -0.5340 \end{bmatrix}$$

$$M_1^y = \begin{bmatrix} 1 & 1.4680 \\ 0 & 1 \end{bmatrix} \quad M_2^y = \begin{bmatrix} -2.1105 & 3.8592 \\ -0.8149 & 1.0163 \end{bmatrix} \quad M_3^y = \begin{bmatrix} -0.8901 & 0.3016 \\ -0.0284 & -1.1138 \end{bmatrix}$$

$$M_{BP1}^x = \begin{bmatrix} 1 & 0.6840 \\ 0 & 1 \end{bmatrix} \quad M_{BP2}^x = \begin{bmatrix} -0.30238 & 0.99932 \\ -0.84182 & -0.52496 \end{bmatrix} \quad M_{BP3}^x = \begin{bmatrix} -0.87364 & 0.8775 \\ -0.6402 & -0.50159 \end{bmatrix}$$

$$M_{BP1}^y = \begin{bmatrix} 1 & 0.5440 \\ 0 & 1 \end{bmatrix} \quad M_{BP2}^y = \begin{bmatrix} -1.62313 & 3.25144 \\ -0.81491 & 1.01634 \end{bmatrix} \quad M_{BP3}^y = \begin{bmatrix} -0.97504 & 5.0733 \\ -0.44863 & 1.30872 \end{bmatrix}$$

And the operations $C_1^x, C_2^x, C_1^y, C_2^y$, are given as below

$$C_1^x = \begin{Bmatrix} x_1 = x_0 \\ x_1' = x_0' + (I_{HSC1}/13.0) \end{Bmatrix} \qquad C_2^x = \begin{Bmatrix} x_1 = x_0 \\ x_1' = x_0' + (I_{HSC2}/13.0) \end{Bmatrix}$$

$$C_1^y = \begin{Bmatrix} y_1 = y_0 \\ y_1' = y_0' + (I_{VSC1}/13.0) \end{Bmatrix} \qquad C_2^y = \begin{Bmatrix} y_1 = y_0 \\ y_1' = y_0' + (I_{VSC2}/14.0) \end{Bmatrix}$$

C. The Booster Model

The booster model accepts the beam composed of *n* number of macro particle and calculates the normalized booster current successfully injected into the booster using Eq. 22 by evaluating the pass/lost condition for each macro particle. The pass/ lost condition for each particle is evaluated using Eq. 23 and 24. These equations are calculated by obtaining the acceptance in phase space for Booster using the MAD [18] based Booster model for the typical magnet settings at which Booster is normally operated.

$$I_{BR} = \frac{1}{n}\sum_{i=1}^{n} f_B^x(i) \times f_B^y(i) \quad (22)$$

$$f_B^x(i) = \begin{cases} 1 & for\ 0.363x_i^2 - 0.663x_i x'_i + 3.328x'^2_i \le 33.765 \\ 0 & for\ 0.363x_i^2 - 0.663x_i x'_i + 3.328x'^2_i > 33.765 \end{cases} \quad (23)$$

$$f_B^y(i) = \begin{cases} 1 & for\ 1.83y_i^2 + 0.364y_i y'_i + 0.568y'^2_i \le 51.359 \\ 0 & for\ 1.83y_i^2 + 0.364y_i y'_i + 0.568y'^2_i > 51.359 \end{cases} \quad (24)$$

Where x and y are in millimeter and x' and y' are in millirad the TL-1 agent uses this model for predicting the current injected into the booster under different operating conditions.

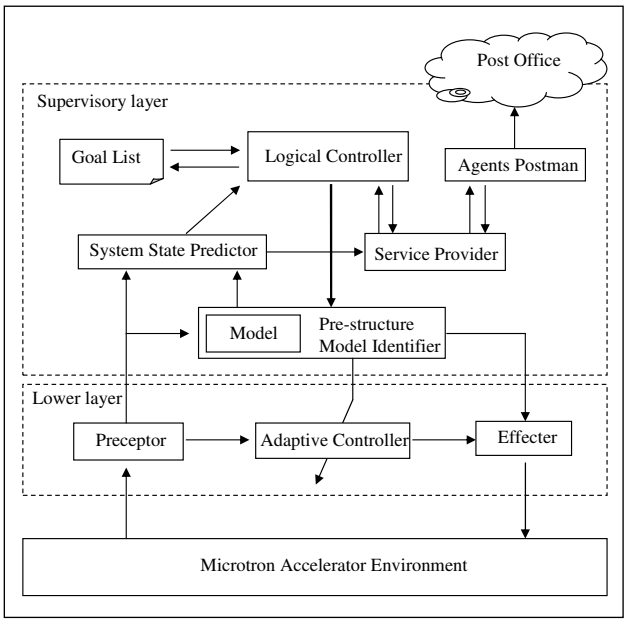


Figure 3. Microtron Agent architecture

Multi-agent based accelerator control

A. Microtron Agent:

The Microtron agent is made with the architecture shown in figure 3. The agent architecture comprises of two loops, the first loop: comprised of "preceptor", "adaptive controller" and "effecter" blocks. This loop is responsible for continuously maintaining

the machine operating point under dynamic conditions. The loop works on the principle of sense-think-control cycle where the accelerator environment is continuously sensed and if some drift in the operating point is observed the corresponding corrective action is calculated by the adaptive controller and applied to the accelerator environment through effecter. The second loop is the supervisory loop responsible for autonomously controlling the agent actions and the interaction with other agents. The "pre-structure model identifier" block when required /asked by the "logical controller" identifies the plant model in the pre-structured model form by directly taking the control of "effecter" and "preceptors" and using the predefined action recipe. This block also provided this identified model to other blocks like "system state predictor" block, "adaptive controller" block and "logical controller" block for their functions. The "system state predictor" block continuously tries to learn the system dynamics and predicts the system dynamics for future *n* steps using the auto-regressive moving average with exogenous (ARMAX) algorithm. This block also provides the functionalities of predicting the future machine states/parameters under the influence of dynamics using the currently identified Microtron model. "Service provider" block is the communication interface of the agent with the other agents. This block is responsible for serving the requests obtained from different agents and from "logic controller" which requires some data from other agents. The "postman" is the communication medium between the agent and the post office for exchange of messages between different agents. The "logic controller" is the brain of the agent and is responsible for managing and synchronizing all the activities of the agents towards the achievement of goals.

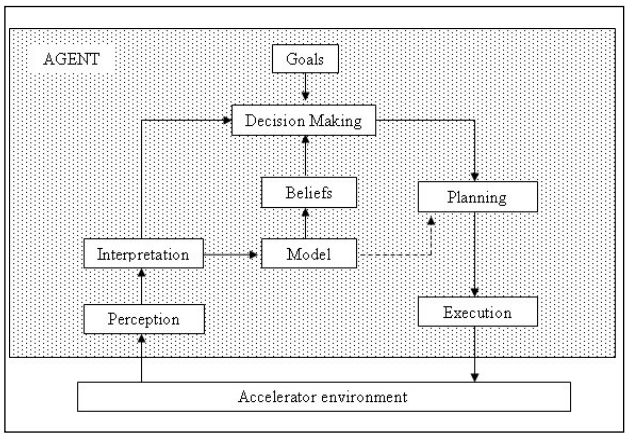


Figure 4: TL-1 agent architecture

B. TL-1 Agent:

The TL-1 agent is developed with a model-based, goal-based modular architecture shown in figure 3. [19] and optimizes the TL-1 control using differential evolution based algorithms. The "Perception" and "Execution" blocks directly interact with the accelerator environment. In TL-1 case it will interact with the TL-1 power supplies and beam diagnostic devices (Fast Current Transformer (FCT) through Oscilloscope, Fluorescent Beam Position Monitor (BPM) screens). Function of the "Perception" block is to read different P/S settings and read-back values, FCT & Oscilloscope traces and BPM images. Depending upon the read data it then generates the appropriate event. Events are passed directly to the respective blocks in the form of messages along with the required data. The "Interpretation" block serves the purpose of processing the raw data acquired by the "perception" block to convert it to the required form in TL-1 this block extracts the beam position (x , y) and beam sizes (σ_x, σ_y) from the BPM images and the injection current value from FCT and CRO traces. "Beliefs" block is the agent's data storage. This stores the system state and other meta data required in the processing / decision making steps. "Goal" block contains the definition for all the goals and provision for enabling / disabling of goals. Definition of goal comprises of plan list. Plans in the list are the alternate plans by which the goal could be achieved in different system conditions. The position of the plan in the plan list decides its priority. The plan at higher level in the list has higher priority. The "Decision Making" block depending upon the current events and the agent beliefs decide the plans to be executed to achieve all the active goals. It does this by evaluating the plan applicability function and selecting the highest priority applicable plan from the list for each active goal. The "Planning" block serves

the purpose of executing the selected plan in synchronised/coordinated way and updating the active goal list. Each plan body comprises of sequence of actions i.e. steps to be followed to attain the desired goal. The "Execution" block sends the commands obtained from different blocks in the form of messages to machine components after checking them for the validity. The "Model" block in itself is an agent comprising of the TL-1 model and serves the purpose of providing the information about the probable outcome of the stated actions on the machine.

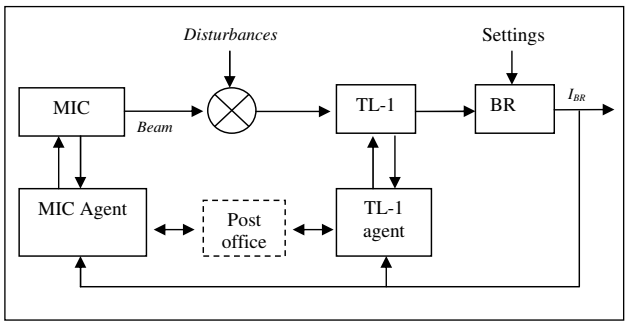


Figure 5: Block diagram of multi-agent based control of Microtron and TL-1

C. Multi-agent based control:

Figure 4 shows the multi-agent based control system block diagram for controlling Microtron and TL-1. The multi-agent based control of Microtron and TL-1 towards the cooperative tuning requires that both of the agents should try to maintain there individual operation to the optimum according to there local priorities on one hand and cooperatively decides their operating points such that their joint goal of increasing the overall injection current in the booster is achieved. This is achieved by jointly identifying the operating points which maximises the cost function J_1 given by Eq. 25. Subject to the conditions that the demand of change in the TL-1 magnet settings is to be reduced while always maintaining the required level of injection current in booster.

$$\max \quad J_1 = I_{MIC}\left(x_t, x_t^{'}, y_t, y_t^{'}, OP_{MIC}\right)\times I_{TL1}\left(x_t, x_t^{'}, y_t, y_t^{'}, OP_{TL1}\right) \tag{25}$$

For the case of cooperative optimisation based on the dynamics learning the Microtron agent at the time of deciding the new operating point for optimisation cooperatively maximises the cost function J_2 given by Eq. 26 considering the n steps ahead future disturbances based on the past movement history provided by the "system state predictor" block.

$$\max \quad J2 = \sum_{i=1}^{n} I_{MIC}\left(x_i, x_i^{'}, y_i, y_i^{'}, OP_{MIC}\right)\times I_{TL1}\left(x_i, x_i^{'}, y_i, y_i^{'}, OP_{TL1}\right) \tag{26}$$

Simulation results

For checking the effectiveness of this scheme the system comprised of accelerator model, Microtron agent and TL-1 agent as shown in figure 5 is simulated. The results of the agent based control when beam coming out of Microtron is subjected to the disturbance shown in figure 6 for three different scenarios, scenario1: when both the TL-1 and Microtron agent works independently to achieve their individual goal, scenario 2: when both of the agents work cooperatively to maximize the booster injection current, and scenario 3: when both of the agents cooperatively with dynamics learning capability works to maximize the booster injection current with 20 steps ahead predicted beam dynamic behavior were calculated for the beam disturbance shown in figure 6 . Figure 7 shows the injection current in booster for the three scenarios. Figure 8 shows the different operating points for which the TL-1 was adjusted by the TL-1 agent for the three scenarios. Figure 9 shows the beam current provided by the Microtron for the three different scenarios.

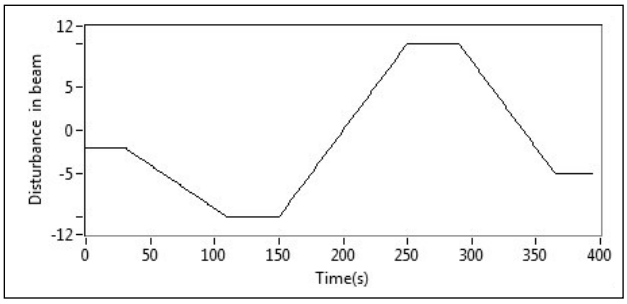


Figure 6 Disturbance in beam added at Microtron output

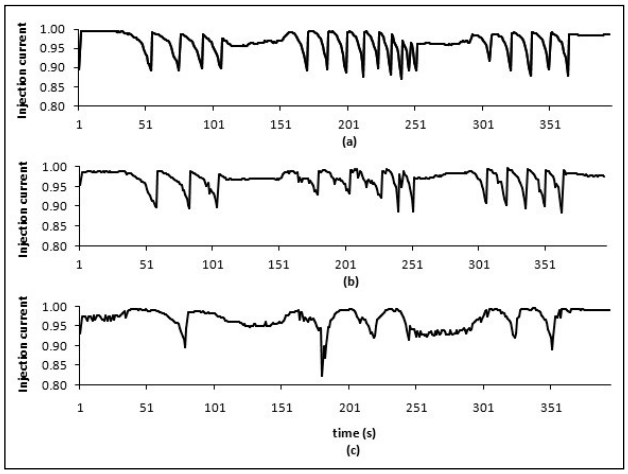


Figure 7: Normalized injection current in booster for (a) when Microtron and TL-1 agents optimizes there operations individually (b) when both of the agents work cooperatively for optimizing the injection current as well as reducing the no of changes in TL-1 settings (c) When both of the agents work cooperatively with dynamics learning case.

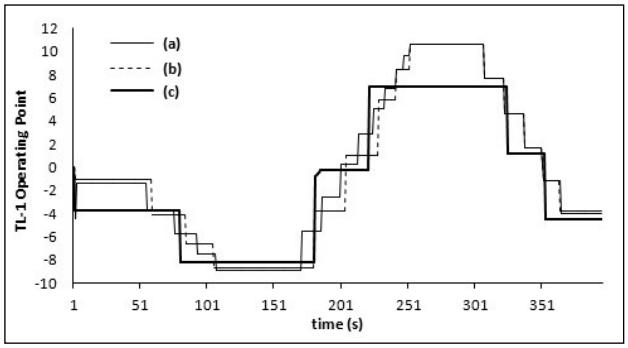


Figure 8 TL-1 operating points for which TL-1 was adjusted by TL-1 agent for (a) scenario 1 (b) scenario 2 (c) scenario 3.

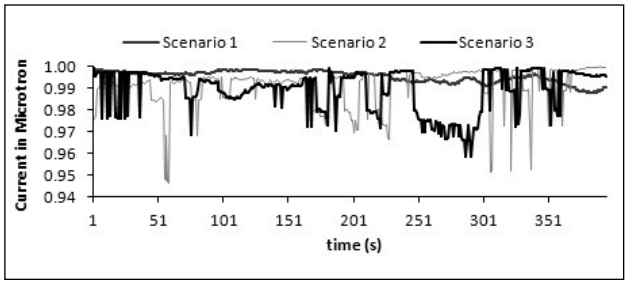


Figure 9 Beam current at Microtron output when Microtron is operated at different operating points by Microtron agent under different scenarios.

From Figure 7 (b) it can be seen that the booster current in case of cooperative optimization is showing lesser number of variation in Booster current with respect to the case when agents work individual (figure 7(a)). This variation further reduces with the application of the dynamics based learning algorithm (figure 7(c)). From figure 8 it can be seen that for the case of dynamics based learning the changes in operating point of TL-1 is minimum as compared to other two cases. From Figure 9 the effect of cooperative optimization in choosing the operating points by Microtron agent can be seen clearly. Where for the scenario 1 the Microtron current remains always at its best operating value but for the other two scenarios the agent gives priority to the common goals and thus opted for slightly sub optimal operating points.

Conclusion

In this paper the application of a multi-agent based approach in control of pre-injector and transport line at synchrotron accelerator facilities was discussed. The novel concept of cooperative optimization with system dynamics learning capability for multi-agent based control approach was presented. The individual agent architecture for controlling Microtron and Transport line and there organization as multi-agent for cooperative control was designed. The simulation results of the presented concept for controlling the pre-injector accelerator Microtron and Transport line under the influence of disturbance on beam shows that this scheme can be used successfully for their optimal control without operator interventions.

Reference:

1. J. S. Heo and K. Y. Lee, "A Multi-Agent System-Based Intelligent Heuristic Optimal Control System for A Large-Scale Power Plant," *Proc. of the IEEE World Congress on Computational Intelligence*, pp. 5693-5700, Vancouver, Canada, July 16-21, 2006.
2. Sheng Gehao, Jiang Xiuceng, Zeng Yi. , "Optimal Coordination For Multi-Agent Based Secondary Voltage Control In Power System," *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, pp.1-6, 2005.
3. Linlin Qian, Yinong Zhang, Xiuli Su, "Cement rotary kiln control system realized by PROFIBUS based on multi-agent," *Fuzzy Systems and Knowledge Discovery (FSKD), 2011 Eighth International Conference on* , vol.3, pp.1364-1367, 26-28 July 2011
4. Francisco P. Maturana, Pavel Tichý, Petr Slechta, Fred Discenzo, Raymond J. Staron, Kenwood H. Hall, "Distributed multi-agent architecture for automation systems". *Expert Syst. Appl.* Vol.26, no.1, pp. 49-56, 2004.
5. Jennings N. R. , Corera J. M. , Laresgoiti I. , Mamdani E. H., Perriollat F., Skarek P. , Varga L. Z., "Using, ARCHON

to develop real-world DAI applications for electricity transportation management and particle accelerator control", *IEEE Expert.* no.11, pp.64-70, 1996.

6. Klein W. , Stern C., Luger G. and Olsson E. , "Designing a portable architecture for intelligent particle accelerator control", *Proceedings of Particle Accelerator Conference*. 12-16 May 1997, Vancouver, BC, Canada, Vol. 2, pp.2422-2424, 1997.
7. Stern C. , Olsson E., Kroupa M., Westervelt R., Luger G. and Klein W., "A Control System for Accelerator Tuning Combining Adaptive Plan Execution with Online Learning ", *International Conference on Accelerator and Large Experimental Physics Control Systems* 1997, 3-7 Nov 1997, Beijing, China, 1997.
8. Klein W.B., Westervelt R.T., and Luger G.F., "An Architecture for intelligent control of particle accelerators" available from http://www-bd.fnal.gov/icalepcs/abstracts/PDF/wpo63.pdf , 2011.
9. Klein W.B., Stern C.R., Luger G.F., and Olsson E.T., "An Intelligent Control Architecture for Accelerator Beamline Tuning", *in Proceedings of AAAI / IAAI*, pp.1019-1025, 1997.
10. Klein W.B., Westervelt R.T., and Luger G.F., "A general purpose intelligent control system for particle accelerators". *In Journal of Intelligent and Fuzzy Systems*, New York: John Wiley, 1999.
11. Klein W.B., Stern C.R., Luger G.F., and Pless D., "Teleo-reactive Control for Accelerator Beamline Tuning", *in Proceedings of Arti_cial Intelligence and Soft Computing*, 24-26 July, 2000, Ban_, Alberta, Canada. pp.372-379, 2000.
12. Pugliese R., Bille F., Abrami A. and Svensson O., "Applying Intelligent System Concepts to Automatic Beamline Alignment", *International Conference on Accelerator and Large Experimental Physics Control Systems* 1997, 3-7 Nov 1997, Beijing, China, 1997
13. J.D Head, J.R Gomes, C.S Williams, K.Y Lee, "Implementation of a multi-agent system for optimized multiobjective power plant control," *North American Power Symposium (NAPS), 2010* , pp.1-7, 26-28 Sept. 2010.
14. Heo, J.S., Lee, K.Y., "A multi-agent system-based intelligent identification system for power plant control and fault-diagnosis," *Power Engineering Society General Meeting, 2006. IEEE*, 2006.
15. S. Kamalasadan, "An Organizational Coordinated Control Paradigm for Complex Systems based on Intelligent Agent Supervisory Loops", *Proceedings of International Joint Conference on Neural Networks (IJCNN),* Orlando, Florida, USA, August 12-17, 2007, August 2007
16. Sukumar Kamalasadan, "A Novel Multi-Agent Controller for Dynamic Systems based on Supervisory Loop Approach", *Engineering Letters* vol.14,no. 2,pp.81-88,2007.
17. Hichem, B.N., Faouzi, M. , "A multi-agent predictive control approach based on fuzzy supervisory loop for fast dynamic systems,", *5th International Multi-Conference on Systems, Signals and Devices 2008. IEEE SSD 2008.*, pp.1-6, 20-22 July 2008.
18. H. Grote, F.C. Iselin, The MAD Program (Methodical Accelerator Design) version 8.16, User's Reference Manual, CERN/SL/90-13(AP), (rev. 4), March 27, 1995.
19. R. P. Yadav, P. Fatnani, P. V. Varde, Intelligent agent based control of TL-1, In proceedings of *Indian particle accelerator Conference, (InPAC2011),* February 15-18, Delhi, India, 2011.

SRESA JOURNAL SUBSCRIPTION FORM

Subscriber Information (Individual)

_____ _____ _____ _____

Title First Name Middle Name Last Name

_____ _____

Street Address Line 1 Street Address line 2

_____ _____ _____ _____ _____

City State/Province Postal Code Country

_____ _____ _____

Work Phone Home Phone E-mail address

Subscriber Information (Institution)

Name of Institution/ Library _____

Name and Designation of Authority for Correspondence _____

Address of the Institution/Library _____

Subscription Rates

	Subscription Quantity	**Rate**	**Total**
Annual Subscription (in India)	_____	Rs 10,000	_____
(Abroad)	_____	$ 500	_____
	_____		_____
	_____		_____

Payment mode (please mark)

Cheque ☐ Credit Card ☐ Master Card ☐ Visa ☐ Online Banking ☐ Cash ☐ De mand Draft ☐

Credit card Number _____

Credit Card Holders Name _____

Credit Card Holde _____

Guidelines for Preparing the Manuscript

A softcopy of the complete manuscript should be sent to the Chief-Editors by email at the address: editor@sresa.org.in. The manuscript should be prepared using 'Times New Roman' 12 font size in double spacing, on an A-4 size paper. The illustrations and tables should not be embedded in the text. Only the location of the illustrations and tables should be indicated in the text by giving the illustration / table number and caption.

The broad structure of the paper should be as follows: a) Title of the paper – preferably crisp and such that it can be accommodated in one or maximum two lines with font size of 14 b) Name and affiliation of the author(s), an abstract of the paper in ~ 100 words giving brief overview of the paper and d) Five key words which indicates broad subject category of the paper. The second page of the paper should start with the title followed by the Introduction

A complete postal address should be given for all the authors along with their email addresses. By default the first author will be assumed to be the corresponding author. However, if the first author is not the corresponding author it will be indicated specifically by putting a star superscript at the end of surname of the author.

The authors should note that the final manuscript will be having double column formatting, hence, the size of the illustration, mathematical equations and figures should be prepared accordingly.

All the figures and tables should be supplied in separate files along with the manuscript giving the figure / table captions. The figure and table should be legible and should have minimum formatting. The text used in the figures and tables should be such that after 30% reduction also it should be legible and should not reduced to less than font 9.

Last section of the paper should be on list of references. The reference should be quoted in the text using square bracket like '[1]'in a chronological order. The reference style should be as follow:

1. Pecht M., Das D, and Varde P.V., "Physics-of-Failure Method for Reliability Prediction of Electronic Components", Reliability Engineering and System Safety, Vol 35, No. 2, pp. 232- 234, 2011.

After submitting the manuscript, it is expected that reviews will take about three months; hence, no communication is necessary to check the status of the manuscript during this period. Once, the review work is completed, comments, will be communicated to the author.

After receipt of the revised manuscript the author will be communicated of the final decision regarding final acceptance. For the accepted manuscript the author will be required to fill the copy right form. The copy right form and other support documents can be down loaded from the SRESA website: http://www.sresa.org.in

Authors interested in submitting the manuscript for publication in the journal may send their manuscripts to the following address:

Society for Reliability and Safety
RN 68, Dhruva Complex
Bhabha Atomic Research Centre,
Mumbai – 400 085 (India)
e-mail : editor@sresa.org.in

The Journal is published on quarterly basis, i.e. Four Issues per annum. Annual Institutional Subscription Rate for SAARC countries is Indian Rupees Ten Thousand (Rs. 10,000/-) inclusive of all taxes. Price includes postage and insurance and subject to change without notice. For All other countries the annual subscription rate is US dollar 500 ($500). This includes all taxes, insurance and postage.

Subscription Request can be sent to SRESA Secretariat (please visit the SRESA website for details)

SRESA Journal of

Life Cycle Reliability and Safety Engineering

Contents

Vol. 2 Issue No. 2 April – June 2012 ISSN – 2250 0820