A Novel Framework for Calculating the Maintenance Improvement Factor based on Human Error Factors and Unbiased Expert Judgment

Rogelio Emmanuel Jáuregui Miramontes
Centre for Management of Technology and Entrepreneurship, Chemical Engineering & Applied Chemistry, University of Toronto, Toronto, Canada;
rogelio.jaureguimiramontes@mail.utoronto.ca, rogelio_mir@hotmail.com

Pasi Petteri Luukka,
School of business, Lappeenranta University of Technology, Lappeenranta, Finland
pasi.luukka@lut.fi

Yuri A. Lawryshyn
Centre for Management of Technology and Entrepreneurship, Chemical Engineering & Applied Chemistry, University of Toronto, Toronto, Canada
yuri.lawryshyn@utoronto.ca

Abstract

Preventive maintenance aims to keep components in good operating condition and to reduce the probability of failure. Optimizing the frequency of preventive maintenance continues to be an important research topic because it can significantly impact operating costs in industrial settings. A key factor in optimizing preventive maintenance is the calculation of the probability of component failure. Estimates of this probability have commonly been based on historical data and the maintenance quality. The improvement factor (IF) was introduced in the literature as a way of measuring maintenance quality so managers and planners could better estimate the probability of component failure, post maintenance. The IF estimate can be based on either historical failure data or expert judgment. It has been demonstrated in previous works that relying on expert judgment is the better of the two approaches when estimating the IF. Such judgment is, however, inherently biased. We propose a framework for estimating the IF based on expert judgment that aims to circumvent such biases. To demonstrate the effectiveness of the proposed framework, we estimate the optimum maintenance interval of a system at a real-world offshore oil and gas installation. The results show our framework yielding an improved maintenance interval.

Keywords: Maintenance; Fuzzy Logic; Human Error Factors; Simulation Model; Optimization

1. Introduction

In a production facility, component failures can lead to downtime which can significantly increase a company's maintenance costs and negatively impact revenues. For this reason, executives and managers need to focus on evolving strategies for reducing component failure. One well-known strategy is based on the expected probability of component failure (Tezuka et al. 2015). Inadequate probability estimates can, however, lead to excessive maintenance which increases overall production costs in the short term (Ding & Kamaruddin 2015).

A key aspect for calculating the probability of component failure is having a good understanding of the effectiveness of the maintenance being performed. Much work has been published attempted to quantify maintenance effectiveness by introducing a parameter referred to as the maintenance improvement factor (IF) (Pham & Wang 1996; Wang 2006; Letot et al. 2015; and Hosseini & Ghadimi 2016). These papers estimate the IF based on probabilistic approaches by using failure data. Maintenance quality is, however, directly related to the expertise of maintenance personnel (Khatab

& Aghezzaf 2016). Consequently, the personnel who perform maintenance activities can often provide a better understanding of the maintenance effectiveness than can data-centered probabilistic approaches, especially given that it is often difficult to quantify the maintenance effectiveness of a given maintenance event (Noortwijk et al. 1992; Hauge et al. 2016).

Multiple studies have found that employees, both experts and non-experts, have a tendency to overrate their own job performance (Doyle et al.2009; Kirkeb¢en 2009; and Hester 2012). Ecken et al. (2011) consider this self-overrating to be a form of desirability bias. This bias means that employees may base their judgments on any possible gain they can get from a study's outcome. Walls & Quigley (2001), Kynn (2008), and Tredger et al. (2016) further show how employees can intentionally or unintentionally make unrealistic judgments. Thus, one can expect an inherent bias in interview or survey results coming from expert and non-expert maintenance personnel who are asked about the effectiveness of maintenance performance. This bias impacts the maintenance performance estimates, yielding an unrealistic IF. A method for incorporating de-biased expert judgment into the IF estimates is clearly needed.

The main objective of this paper is to introduce an approach for estimating the improvement factor that aims to circumvent the inherently biased opinions of maintenance personnel. This is achieved by integrating the results of a survey administered to non-expert maintenance personnel with estimates of maintenance performance provided by recognized maintenance experts, typically lead supervisors. Hoffman et al. (1995) acknowledge that the nature of a question can itself bias survey results. Tredger et al. (2016) indicate that managing bias begins with managing the way questions are structured and designed. The proposed survey does not simply include questions directly pertaining to the maintenance performance. The survey incorporates, rather, 18 multiple choice questions related to factors that can lead to human maintenance error. These factors are referred to as Human Error Identification (HEI) factors. Given that both the survey results and expert estimates are in linguistic form, an approach based on fuzzy set theory is used to develop a unique Expert Improvement Factor (EIF).

The EIF output is within the unit interval: zero being 'as good as new' maintenance performance, and one being 'as bad as old' maintenance performance. From this, an estimate can be made of the probability of component failure post maintenance. A simple simulation model is used to optimize the maintenance interval of two components operating in parallel. The results show that the accuracy of the EIF is a key driver of the total maintenance cost, proving that the EIF based on unbiased expert judgment increases the accuracy of optimum maintenance interval estimates.

2. Literature Review

Maintenance is meant to either preserve or improve components, and its effectiveness impacts the probability of component failure (Al-Najjar & Alsyouf 2003). The degree to which maintenance activities impact this probability is referred to as maintenance performance. This paper uses four maintenance performance levels as follows:

<u>Worst maintenance</u>: restores a component in a way that makes it fail immediately. An example of worst maintenance is the application of an incorrect maintenance procedure (Pham & Wang 1996).
<u>Minimal maintenance</u>: partially restores a component, providing a state 'as bad as old.' The component behaviour remains the same as before it failed. An example of minimal maintenance is the lubrication of bearings (Jardine & Tsang 2013).
<u>Imperfect maintenance</u>: lies between minimal maintenance and perfect maintenance and restores the component to a state between 'as good as new' and 'as bad as old' (Brown & Proschan 1980).
<u>Perfect maintenance</u>: restores a component completely, providing a state 'as good as new.' Completely replacing a component with a new one is also considered perfect maintenance (Pham & Wang 1996).

Maintenance performance impacts the probability of component failure and is well documented in the literature as the maintenance improvement factor (IF). Nakagawa (1979), Block et al. (1985) and Makis & Jardine (1992) have mathematically modelled the effect of maintenance on the probability of component failure and provide foundational methods for calculating the IF. Others have developed enhanced approaches based on these methods, such as: Zhou et al. (2016), Kim et al. (2016), and Hosseini & Ghadimi (2016).

The above authors proposed probabilistic models, and historical maintenance and repair data sets to determine the IF. These data sets do not, however, incorporate subjective information about maintenance and repair activities.

Subjective information includes things like detailed component-specific failure analysis, quality of the spare replacement, maintenance procedure compliance and restricted time for finishing a repair (Hauge et al. 2016). This lack of information can result in sub-optimal maintenance strategies (Garg & Sharma 2012; Panchal & Kumar 2016). Moreover, Liu et al. (2011) acknowledge that incorporating expert judgment into maintenance optimization models yields robust results. Accordingly, when there is a lack of data, subjective information provided by surveys and experts can be incorporated to evaluate maintenance effectiveness, with the end goal of enhancing maintenance strategies (Said et al. 2016).

Expert judgment provides subjective information with respect to maintenance performance, which typically cannot be extracted from an operational or maintenance data set (Peres et al. 2007; Doyle et al. 2009). The value of this expert judgment is acknowledged by a number of authors and has been incorporated into models such as those for improving maintenance downtime optimizing human safety, optimizing condition-based maintenance, calculating optimal preventive maintenance policies, optimizing maintenance management and identifying latent maintenance errors (Hussin et al. 2012; Azadeh et al. 2014; Asadsadeh & Azadeh 2014; Hosseini & Ghadimi 2016; Qi et al. 2015; Chiu & Hsieh 2016). Furthermore, expert judgment has been included in reliability models associated with the gas chilling and liquefaction unit of a hydrocarbon processing facility, a manufacturing system and a safety instrumented system (Hennequin & Arango 2009; Rahimi & Rausand 2013; Hammed et al. 2016). The above mentioned works recognize expert judgment as a tool for enhancing reliability estimates. They do not, however, address the issue of countering bias in expert judgment. Tredger et al. (2016) break out the most common biases associated with expert judgment into six types: anchoring, availability, framing, small sample, over-confidence, and substitution.

Anchoring occurs when an expert uses initial or prior information to make a judgment. Availability refers to the use of the first thought as a judgment. Framing takes place when the questions posed to an expert are poorly designed. Small sample bias results from having too few experts. Overconfidence occurs when an expert is overly confident in his or her knowledge. Substitution takes place when an expert answers only the easiest questions. These bias types, do not, however, include the respondent's desirability of the outcome. That is to say, an expert's desire to achieve or influence a certain outcome, either positively or negatively (Ecken et al. 2011). For this reason, the works of Tredger et al. (2016) and Ecken et al. (2011) point to a seventh type of bias. Maintenance performance based on expert judgment can be biased by one or all of the above types.

Numerous studies have been published that aim to reduce expert bias in different fields, including organizational evaluation, auditing, health sciences, foresight and actuarial work (Kromrei 2015; Nelson & Tan 2005; Croskerry et al. 2013; Ecken et al. 2011; Tredger et al. 2016) These studies aim to de-bias judgments by training to increase their awareness of possible biases. Few studies are found in the maintenance field that are aimed at reducing expert bias. Doyle et al. (2009) rely on coaching and training as core techniques for reducing bias when estimating component lifetime distributions. Hussin et al. (2012) rely on a trained external assessor for reducing bias in offshore gas compression downtime estimates. But training does not eliminate the desirability bias because experts may still prefer one alternative over another based on the expected gain resulting from their analysis.

Despite the importance of subjective information for optimizing maintenance models, there does not appear to be a method for calculating the IF based on unbiased expert judgment. This paper presents a framework that aims to fill this gap, by calculating the IF based on the Expert Improvement Factor (EIF). The proposed framework aims to reduce all seven bias types mentioned above by asking four recognized experts, typically lead supervisors, for their appraisal of the following maintenance performance levels, based on the degree of component restoration: as bad as old; imperfect low; imperfect medium; imperfect high; and as good as new. These performance levels will help to lessen the influence of desirability bias and also reduce anchoring, availability, and substitution biases. This framework further includes the administering of a survey to all non-expert maintenance personnel. This survey aims to reduce the framing, overconfidence and small sample biases by incorporating HEI factors that determine the conditions that can lead to human error in maintenance.

It should be noted that HEI factors are referred to by different names in the literature, including performance indicator factors, performance shaping factors, and error producing conditions (Rausand 2011). Maintenance activities involve human actions and, therefore, bring with them the possibility of human error (Alireza et al. 2013). It has been shown that human error lead to 25% of unexpected nuclear industry shutdowns and up to 90% of all industrial accidents (Heo & Park 2010; Rausand 2011). The consequences of human error vary from the insignificant to the catastrophic. Human error in maintenance can impact the equipment failure rate, the occurrence of accidents, and the IF (Davies 2000; Dhillon 2009; Kumar & Gandhi 2011; Antonovsky et al. 2016).

HEI factors have been discussed in a number of studies (Antonovsky et al. 2016; Asadsadeh & Azadeh 2014 and, in the fuzzy logic context, Kumar & Gandhi 2011). Generally, HEI factors are classified into three groups: external, internal and stressor. Working conditions comprise the external factors. Internal factors recognize the individual characteristics of workers, such as experience, skill and motivation. Stressor factors encompass conditions affecting the cognitive performance of the maintenance personnel (Abbassi et al. 2015).

Antonovsky et al. (2016) interviewed maintenance personnel at a real-world petroleum installation in an effort to identify the HEI factors most likely to lead to maintenance failure. While they do not classify the factors, a total of 26 factors were identified, of which 16 can be considered external and 10 internal. None of the factors identified in the study fall into the HEI stressor group. Specifically, workload and fatigue were not singled out as factors. (Wang & Tsai 2014) found, however, that workload and fatigue play an important role in maintenance error. The survey that will be conducted as part of this framework incorporates external, internal, and stressor HEI factors.

As noted above, Kumar & Gandhi (2011) use fuzzy logic to transform linguistic terms supplied by experts into numerical data in an effort to quantify HEI factors. Fuzzy set theory, introduced by Zadeh (1965), provides a mathematical method for transforming linguistic terms into numerical data. Fuzzy set theory has a rich history as documented in Zadeh (1965); Zimmermann (1985); Klir (1995), Benitez et al. (2007), and CIM (2011). Since both the survey administered to maintenance crew members and the experts estimates of maintenance performance are expressed in linguistic terms, the proposed framework also uses fuzzy set theory.

3. Methodology

The Expert Improvement Factor (EIF) is introduced with the intent of reducing expert bias to better estimate the probability of component failure. The EIF is incorporated in a simulation model to optimize the maintenance strategy associated with the operation of two components at a real-world offshore oil and gas installation. The following subsections introduce the EIF methodology and the simulation model. Results of the methodology are presented in the subsequent section where the impact of using the biased IF versus the EIF is also presented.

3.1 Novel EIF Methodology

The proposed framework for calculating the EIF involves three key steps: administering a maintenance personnel survey, conducting expert interviews, and calculating the EIF. Figure 1 depicts these steps. Step one involves administering a survey to every member of the maintenance crew at a real-world offshore oil and gas installation to determine the quality of the maintenance. Step two involves interviewing maintenance experts to work out survey question weights, survey multiple choice impact on maintenance performance, the respondent baseline, and the establishment of maintenance performance baselines. Step three integrates the survey and interview results to determine the EIF.

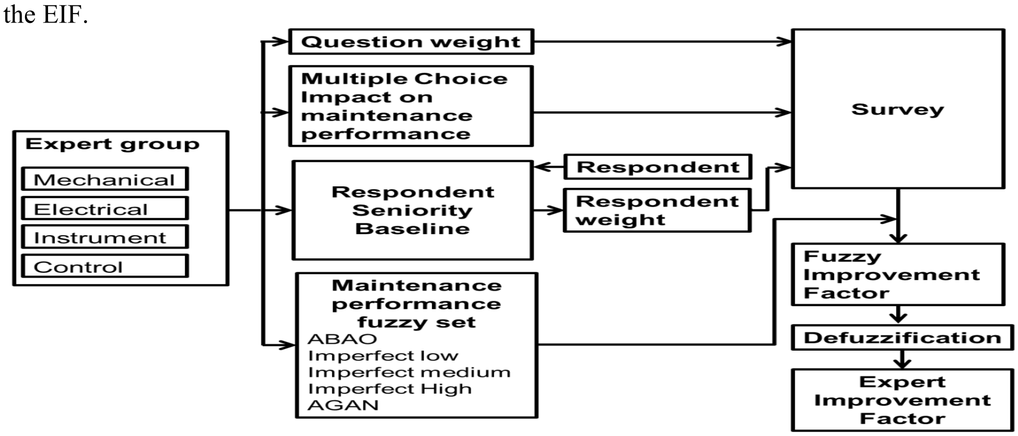

Figure 1. Methodology for calculating the unbiased *EIF*

A main goal of the approach presented here is to account for the seven biases associated with expert judgment. This is accomplished through the integration of the maintenance survey data with data collected from the expert interviews. The proposed framework steps are explained below.

3.1.1 Maintenance Personnel Survey

This paper presents a survey that incorporates HEI factors to best estimate the quality of maintenance. It should be noted that the survey includes only 18 multiple-choice questions (shown in Table 1) because the time allotted for answering the questions is ten minutes, a duration that has proven to be the most effective (Galesic & Bosnjak 2009). The formulation of these questions followed a thorough review of published papers identifying HEI factors in fields such as human error probability data, human reliability analysis, maintenance, and work shift scheduling (Basra & Kirwan 1998; Davies 2000; Kim & Jung 2003; Mannan 2005; Rausand 2011; Antonovsky et al. 2013; Wang & Chuang 2014). These studies duplicate some of the factors. Thus, the consolidation of these HEI factors leads to a total of 116 unique factors, including 55 external, 37 internal and 24 stressor.

Ren et al. (2008) recognize significant variability associated with human error. It is this variability that limits one's ability to identify the most representative HEI factors. The works of: Crichton (2005); DiMattia et al. (2005); Wang et al. (2015); Tabibzadeh & Meshkati (2014); Landucci & Paltrinieri (2016); Vinnem et al (2012); Skogdalen & Vinnem (2011) all make use some of the 116 HEI factors and further show how using HEI factors improves their risk analyses. The integration of the factors used across these works leads to 18 specific HEI factors. These factors form the basis of the maintenance survey questionnaire (see Table 1). Of these, particular consideration was given to the external HEI factor group pertaining to organizational culture. These factors have been acknowledged by Reiman & Oedewal (2006), and Asaad & Yusoff (2013) as playing a significant role in the carrying out of maintenance activities and in the development of maintenance strategies. As a result, eleven questions were designed based on the external HEI factor group, while the remaining questions were divided between the internal and the stressor HEI factor groups.

Table 1. Survey questionnaire.

#	Question	Very Low	Low	Medium	High	Very High
1	Do you know which are your responsibilities and duties ?	o	o	o	o	o
2	How often your supervisor carry out a job performance review	o	o	o	o	o
3	How much are you trained for the job?	o	o	o	o	o
4	How often do you receive training?	o	o	o	o	o
5	How often your coworkers act as a team	o	o	o	o	o
6	How often do you follow a written procedure?	o	o	o	o	o
7	How often the written procedures are available ?	o	o	o	o	o
8	How often the written procedures are actualized?	o	o	o	o	o
9	How often do you receive a clear oral instruction?	o	o	o	o	o
10	How often do you receive a clear written work order	o	o	o	o	o
11	How well do you know the company maintenance policy?	o	o	o	o	o
12	How effective are the tools you actually use in the maintenance?	o	o	o	o	o
13	How often do you work in night shifts?	o	o	o	o	o
14	How often do you work extra time	o	o	o	o	o
15	How often are you overloaded of activities ?	o	o	o	o	o
16	How often are you short of time to finish a job?	o	o	o	o	o
17	How often your workplace space is limited	o	o	o	o	o
18	How often the maintenance is postponed due lack of resources e.g. spare parts, tools.	o	o	o	o	o

The survey questionnaire solicits a respondent's level of agreement with each question. Each question aims to quantify one HEI factor that influences human error associated with maintenance activity and, consequently, is known to impact maintenance performance. In this way, the survey results provide a useful tool for estimating IF. Moreover, the survey aims to lessen the influence of five of the seven expert judgment biases discussed previously—anchoring, framing, small sample, overconfidence and desirability. Anchoring and desirability biases are reduced by polling for the level of agreement to a question not directly tied with the IF. Framing bias is reduced by designing each question based on the HEI factors discussed above. Both overconfidence bias and small sample bias are reduced by administering the survey to all non-expert personnel at the installation. (If one respondent shows overconfidence in his responses, the remaining responses are used to mitigate this bias.). The remaining two biases not reduced in this survey tool (availability and substitution) are addressed by the experts' interviews that are explained next.

3.1.2 Expert Interviews

As discussed earlier (Step 2) the methodology involves interviewing maintenance experts to determine survey question weights, survey multiple-choice impact on maintenance performance, the respondent baseline, and the establishment of maintenance performance baselines. The math involved to interpret the interview results is presented below. Fuzzy set theory is used to transform the interview linguistic results into crisp data.

Survey Questions Weight.

Four recognized experts were asked individually what they thought was the appropriate weight for each of the 18 survey questions. Weights indicate the degree of importance a question has when compared to all other questions, based on their perceived impact on maintenance performance. Accordingly, the i-th weight is provided by the j-th expert as a crisp (non-fuzzy) number between 1 and 100. This results in a $N_E\ x\ N_Q$ matrix $\boldsymbol{W_{ij}}$ that incorporates the weights of the N_Q = 18 survey questions provided by each of the N_E experts. As a result, the weight Q_{W_i} given to the i-th question can be calculated by incorporating the geometric mean (Dong, Zhang, Hong, & Xu 2010),

$$\boldsymbol{Q_{W_i}} = \frac{\left(\prod_{j=1}^{N_E} w_{ij}\right)^{\frac{1}{N_E}}}{\sum_{k=1}^{N_Q}\left(\prod_{j=1}^{N_E} w_{kj}\right)^{\frac{1}{N_E}}}, \qquad (1)$$

Multiple Choice Impacts on Maintenance Performance

Each of the 18 survey questions is presented with an agreement scale consisting of the following levels: very low (VL), low (L), medium (M), high (H), and very high (VH). Each choice is thus associated with a different degree of impact on maintenance performance. For example, the eighth question (How much are you trained for the job?) may elicit a response of VL, indicating the respondent has received little job training. An employee who lacks training is likely more prone to maintenance error than his fully trained counterpart and that can negatively impact on maintenance performance.

The agreement scales are expressed in linguistic terms and a method for transforming these terms into numerical data is required. As mentioned above, fuzzy set theory is the mathematical method for transforming linguistic terms into numerical data; thus, the agreement scales can be expressed as fuzzy numbers. While the chosen experts have an insight into the impact of each agreement scale level on the maintenance performance, some may lack knowledge of fuzzy set theory. Hence, an explanatory introduction of fuzzy set theory was provided to the experts with special attention given to fuzzy sets. A fuzzy set is defined mathematically as follows: $Z = \{(x, \mu_{VL}(x)) | x \epsilon X\}$,where $Z \subset X,\ \mu_{VL}(x)$ is the membership function defined as $\mu_{VL}(x): X \rightarrow [0,1]$, and where zero indicates non-membership and one indicates full membership to a fuzzy set (Zimmermann 1985). A triangular fuzzy set can be defined by a triplet $(x: a, b, c; \mu(x))$ and a trapezoidal fuzzy set can be defined by the quadruplet $(x: a, b_L, b_H, c; \hat{\mu}(x))$ where a, b_L, b_H, and c are the low, medium low, medium high, and high estimates respectively of a fuzzy set. Note, that a triangular fuzzy set can be easily transformed into a trapezoidal fuzzy set, such that $(x: a, b, c; \mu(x)) \rightarrow (x: a, b, b, c; \hat{\mu}(x))$ (see Grzegorzewski & Mrówka 2005; and Bojadziev & Bojadziev 2007).

Following a brief introduction to fuzzy set theory, the experts were asked to provide consensual estimates of the degree of component restoration for the questionnaire scales. Specifically, the experts were tasked with determining the appropriate parameters $a^{(1)}$, $b^{(1)}$, $c^{(1)}$, $d^{(1)}$and $e^{(1)}$as depicted in Figure 2. Based on these parameters, appropriate fuzzy sets can be determined for each of the respondent questionnaire scale choices, namely, VL, L, M, H and VH. Specifically, the consensual fuzzy sets include both trapezoidal and triangular fuzzy sets; thus, the triangular fuzzy sets are transformed into trapezoidal fuzzy sets. The trapezoidal fuzzy sets are as follows: $VL(x) = (x: 0, 0, a^{(1)}, b^{(1)}; \hat{\mu}(x))$; $L(x) = (x: a^{(1)}, b^{(1)}, b^{(1)}, c^{(1)}; \hat{\mu}(x))$; $M(x) = (x: b^{(1)}, c^{(1)}, c^{(1)}, d^{(1)}; \hat{\mu}(x))$; $H(x) = (x: c^{(1)}, d^{(1)}, d^{(1)}, e^{(1)}; \hat{\mu}(x))$; and $VH(x) = (x: d^{(1)}, e^{(1)}, e^{(1)}, 100; \hat{\mu}(x))$.

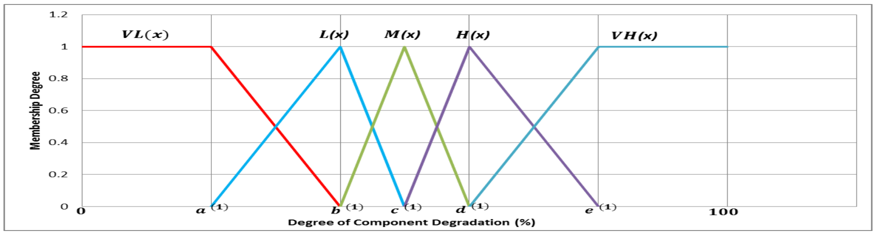

Figure 2**. Agreement scale fuzzy sets.**

Respondent Seniority Baseline

Survey respondents almost certainly have varying degrees of experience. Since the survey responses from a highly experienced employee are likely to be more reliable than those from a less experienced employee, a baseline is proposed in which each respondent can be compared and subsequently weighted based on their seniority. This baseline for the j-th expert is expressed as a trapezoidal fuzzy set $R_j(x) = (x\colon 0, a_j^{(2)}, b_j^{(2)}, \infty; \hat{\mu}(x))$, where $a_j^{(2)}$, and $b_j^{(2)}$ are the seniority levels: medium low, and medium high in years. Figure 3 depicts the respondent baseline. If a respondent has seniority below $a_j^{(2)}$ the respondent is discarded. However, if a respondent holds seniority equal or greater than $b_j^{(2)}$ then this respondent is assigned the highest weight of one; otherwise, the respondent is assigned a weight of a $\hat{\mu}(x)$.

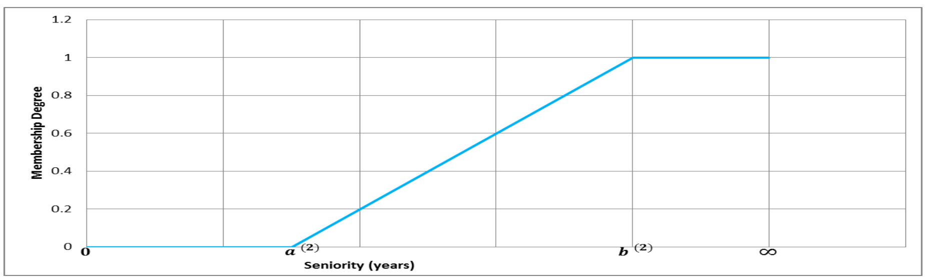

Figure 3. Respondent seniority baseline.

Participating experts were asked to provide the seniority levels individually. These levels were averaged to merge the seniority estimates into a single fuzzy set (Bojadziev & Bojadziev 2007). The respondent baseline $R(x) \in (0,1)$ is considered to be: $R(x) = \left(x\colon 0, \frac{1}{N_E}\sum_{j=1}^{N_E} a_j^{(2)}, \frac{1}{N_E}\sum_{j=1}^{N_E} b_j^{(2)}, \infty; \hat{\mu}(x)\right)$.

Maintenance Performance Baselines

As previously established, expert judgment is inherently biased. The expert appraisal of the installation maintenance performance is therefore also biased, which can inappropriately impact maintenance optimization models. In order to reduce the effects of the remaining two biases (availability and substitution) the four experts were individually asked to provide their appraisal of the maintenance performance levels—as bad as old (ABAO), imperfect low (IL), imperfect medium (IM), imperfect high (IH), and as good as new (AGAN)—based on the low, medium, and high percentage of component restoration associated with each level. By asking the experts to appraise maintenance performance levels instead of the maintenance performance, unbiased appraisals should be produced. Moreover, these appraisals, in combination with the proposed framework, will circumvent the inherent expert biases during the IF assessment.

Each of the j-th experts was asked to individually provide estimates of the parameters $a_j^{(3)}$, $b_j^{(3)}$, $c_j^{(3)}$, $d_j^{(3)}$, and $e_j^{(3)}$ relating the degree of restoration to the maintenance performance levels ABAO, IL, IM, IH and AGAN utilizing Figure 4. Following a similar procedure to that above, the j-th fuzzy number can be written as $ABAO_j(x) = \left(0,0,a_j^{(3)},b_j^{(3)},\hat{\mu}(x)\right)$, $IL_j(x) = \left(a_j^{(3)},b_j^{(3)},b_j^{(3)},c_j^{(3)}\hat{\mu}(x)\right)$, $IM_j(x) = \left(b_j^{(3)},c_j^{(3)},c_j^{(3)},d_j^{(3)}\hat{\mu}(x)\right)$, $IH_j(x) = \left(c_j^{(3)},d_j^{(3)},d_j^{(3)},e_j^{(3)}\hat{\mu}(x)\right)$, $AGAN_j(x) = \left(d_j^{(3)},e_j^{(3)},e_j^{(3)},100,\hat{\mu}(x)\right)$ These estimates are then averaged as follows(Bojadziev & Bojadziev 2007): $ABAO(x) = \left(0,0,\frac{1}{N_E}\sum_{j=1}^{N_E} a_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} b_j^{(3)}; \hat{\mu}(x)\right)$, $IL(x) = \left(\frac{1}{N_E}\sum_{j=1}^{N_E} a_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} b_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} b_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} c_j^{(3)}; \hat{\mu}(x)\right)$, $IM(x) =$

$(\frac{1}{N_E}\sum_{j=1}^{N_E} b_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} c_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} c_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} d_j^{(3)}; \hat{\mu}(x))$, $IH(x) = (\frac{1}{N_E}\sum_{j=1}^{N_E} c_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} d_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} d_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} e_j^{(3)}; \hat{\mu}(x))$, $AGAN(x) = (\frac{1}{N_E}\sum_{j=1}^{N_E} d_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} e_j^{(3)}, \frac{1}{N_E}\sum_{j=1}^{N_E} e_j^{(3)}, 100; \hat{\mu}(x)$.

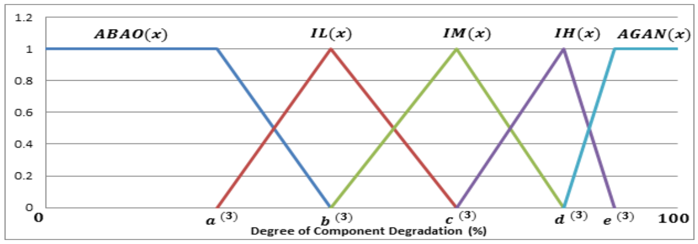

Figure 4. Maintenance performance fuzzy sets

3.2 *EIF* Calculation

The goal of this paper is to present a method for estimating an unbiased Expert Improvement Factor (EIF) This EIF is calculated by integrating the survey results with the maintenance performance baselines. This is achieved with methods based on fuzzy set theory.

3.2.1 Survey Results

As mentioned previously, the survey incorporates 18 questions and five-scale agreement choices, namely, VL, L, M, H and VH. For the sake of simplicity, each questionnaire response for the k-th respondent can be expressed as a matrix $[M]$ with N_Q rows and N_S columns, where N_Q is the total number of questions (18 in this case), and N_S is the total number of scale agreement choices (5 in this case). For the i-th question the k-th respondent selects the l-th scale agreement. Respondents make just one choice for any given question and null answers are not allowed. Thus, if a respondent chooses “Low” for Question 4 then $[M]_{4,2}^k = 1$, and, $[M]_{4,1}^k = [M]_{4,3}^k = [M]_{4,4}^k = [M]_{4,5}^k = 0$.

Each choice indicates either a decline or an increase in maintenance performance. This effect on maintenance performance is based on the HEI factor associated with each question. For example, if a respondent selects ‘Very Low’ for Question 3 (How much are you trained for the job?), the response indicates a decline in maintenance performance. If, however, a respondent selects ‘Very Low’ for Question 12 (How often do you work extra time?), the response indicates an improvement in maintenance performance. Accordingly, the scale agreement choices have a positive and a reverse direction (see Likert (1932)). The positive agreement scale direction is represented by the vector $[F_1]$, and the reversed agreement scale direction by the vector $[F_2]$ as follows:

$$\boldsymbol{F_1} = \begin{bmatrix} \boldsymbol{VL(x)} \\ \boldsymbol{L(x)} \\ \boldsymbol{M(x)} \\ \boldsymbol{H(x)} \\ \boldsymbol{VH(x)} \end{bmatrix}, \quad (2), \quad \boldsymbol{F_2} = \begin{bmatrix} \boldsymbol{VH(x)} \\ \boldsymbol{H(x)} \\ \boldsymbol{M(x)} \\ \boldsymbol{L(x)} \\ \boldsymbol{VL(x)} \end{bmatrix}, \quad (3)$$

The fuzzy result of the k-th questionnaire is calculated as $\boldsymbol{\Psi_k} = \frac{1}{18}(\sum_{i=1}^{11} \boldsymbol{Q_{w_i}}\left[[\boldsymbol{M}]^{\boldsymbol{k}} * [\boldsymbol{F_1}]\right]_{\boldsymbol{i}} + \sum_{i=12}^{18} \boldsymbol{Q_{w_i}}\left[[\boldsymbol{M}]^{\boldsymbol{k}} * [\boldsymbol{F_2}]\right]_{\boldsymbol{i}}$. It should be emphasized that Ψ_k represents a maintenance performance factor as estimated by the k-th respondent.

Fuzzy Improvement Factor

The fuzzy improvement factor is calculated based on the fuzzy subsethood between the estimated Ψ_k and the maintenance performance baselines. The fuzzy subsethood refers to the proportional area of two intersecting fuzzy sets A and B, and is defined as (Klir & Yuan 1995; Takac 2016) $S(A,B) = \frac{card(A \cap B)}{card(A)}$, where $S(A,B) \in (0,1)$ is the subsethood of two fuzzy sets, $A\ and\ B$ are two independent fuzzy sets, and $card$ is the cardinality of a fuzzy set defined as $\int_U m_A(x)dx$, where $m_A(x)$ is all the possible membership degrees of a fuzzy set, $U \in (0,1)$ (Bandemer & Hans 1992; Martinez et al. 2014). The fuzzy subsethood provides a tool for weighting each of the maintenance performance fuzzy sets based on the survey results. These weights provide a tool for merging expert and employee appraisals of maintenance performance. The weighted fuzzy sets are aggregated as follows:

$$\boldsymbol{P_k(\Psi_k)} = \frac{S(ABAO,\Psi_k)*ABAO+S(IL,\Psi_k)*IL+S(IM,\Psi_k)*IM+S(IH,\Psi_k)*IH+S(AGAN,\Psi_k)*AGAN}{S(ABAO,\Psi_k)+S(IL,\Psi_k)+S(IM,\Psi_k)+S(IH,\Psi_k)+S(AGAN,\Psi_k)}, \quad (4)$$

As mentioned above, each respondent (maintenance employee) is weighted according to his or her seniority level, (see respondent seniority baseline chapter). This weight is incorporated into the corresponding $P_k(\Psi_k)$ to calculate the trapezoidal fuzzy expert improvement factor which is considered to be: $FIF = \sum_{k=1}^{N_R} \frac{R_k * P_k(\Psi_k)}{\sum_{k=1}^{N_R} R_k}$.

Deffuzification

The maximizing method is used to transform the trapezoidal FIF set into a crisp number (Bojadziev & Bojadziev 2007). This crisp number is referred to as the expert improvement factor $EIF \in (0,1)$ which is calculated as follows:

$$\boldsymbol{EIF} = \left(\mathbf{1} - \frac{\frac{(FIF_L+FIF_{M_L}+FIF_{M_H}+FIF_H)}{4}}{\mathbf{100}}\right), \quad (5)$$

Utilizing the EIF, it is possible to perform simulations to improve maintenance operations. A simulation model was developed for this study for the purpose of comparing the performance of optimized maintenance based on EIF and IF calculations. The results show that overestimating the improvement factor can lead to costly, sub-optimal maintenance.

3.3 Simulation Model

The simulation model can be used to estimate the impact of the implementation of different maintenance strategies on the cost of operating a system. The system modelled here consists of two identical and brand new repairable components. One component is required for production and the second remains on standby and is ready to work (available to work). If a component requires maintenance or repair, the component that is available to work will start to work without delay. If both components require repair, the system suffers production loss. In this simplified model, the cost of maintenance and (unscheduled) repair and costs associated with lost production, occurring only when both components are down, are all taken into account.

The proposed model was coded in MATLAB. The simulation starts by setting the status of component one (C1) as working and component two (C2) as available to work. The components can have one of the following statuses associated with their state: working, available, maintain, and repair. A condition where both components are in maintenance status is not allowed in this model because the objective of the system is to have at least one component working at any time, if possible, to avoid costly downtime. The age of the working component is compared to the simulated time to failure and the time to maintain. If the age is greater than or equal to the simulated time to failure, the component's status is changed to repair. Similarly, if the age is greater than or equal to the time to maintain, the component's status is set to maintain—unless, of course, the other component is unavailable—and the component continues to operate.

The maintenance and repair states have associated random downtimes. These random downtimes are estimated based on the historical maintenance and repair downtimes probability distribution. The historical maintenance downtimes associated with the simulated system are log normally distributed with parameters μ=1.38, and σ=2.41, and the historical repair downtimes are log normally distributed with parameters μ=2.22, and σ=1.45. The random downtimes are estimated by using the MATLAB random lognormal number generator function. A component being maintained or repaired remains in that state until the associated downtime expires. As mentioned above, when a component is being maintained or repaired, maintenance of the working component is postponed. Note that, when one component is undergoing maintenance or repair, if the second (working) component's age reaches its simulated time to failure, the system goes down, resulting in costly downtime. During system downtime the component whose maintenance or repair is completed first is set to work and the system is no longer in a down state. The simulation proceeds until a user-provided time horizon, T, is reached.

For simplicity, the estimated stochastic time to failure is based on a hypothetical failure rate cumulative distribution function, $F(X)$. The k-th component's virtual age (see Kijima et al. 1988) after the i-th repair or maintenance is denoted by $x_{k_i} = EIF * (x_{k_{i-1}} + t_{F_{k_i}})$, where $t_{F_{k_i}}$ is the simulated time to failure. The conditional probability that $t_{F_{k_i}}$ takes a value between $x_{k_{i-1}}$ and X is determined as follows:

$$\boldsymbol{P}\left\{\boldsymbol{t_{F_{k_i}}} \leq \boldsymbol{X} \middle| \boldsymbol{x_{k_{i-1}}} = \boldsymbol{y}\right\} = \frac{F(X+y)-F(y)}{1-F(y)}, \quad (6)$$

The $\boldsymbol{t_{F_{k_i}}}$ is simulated by assigning a random uniform variable $U_{k_i} \in (0,1)$ to $P\left\{t_{F_{k_i}} \leq X \middle| x_{k_{i-1}} = y\right\}$, and solving Equation 6 in terms of $t_{F_{k_i}}$ as depicted next: $\boldsymbol{t_{F_{k_i}} = F^{-1}\left(U_{k_i} * \left(1 - F(x_{k_{i-1}})\right) + F(x_{k_{i-1}})\right) - x_{k_{i-1}}}$.

The application of preventive maintenance is a well-known method for reducing the probability of component failure (Bentley 1999). Preventive maintenance is administered at fixed periods of time T_M. The maintenance interval renders a variable number of component and system failures. The total cost for the si-th simulation at a given T_M is calculated as follows:

$$\boldsymbol{TC_{si}(T_M) = C_M\left(\sum_{q=1}^{N_{M_1}^{(T_M)}} D_{M_1}^{(q)} + \sum_{q=1}^{N_{M_2}^{(T_M)}} D_{M_2}^{(q)}\right) + C_F\left(\sum_{q=1}^{N_{F_1}^{(T_M)}} D_{F_1}^{(q)} + \sum_{q=1}^{N_{F_2}^{(T_M)}} D_{F_2}^{(q)}\right) + C_{S_F}\left(\sum_{q=1}^{N_{S_F}^{(T_M)}} D_{S_F}^{(q)}\right)}, \quad (7)$$

where C_M, C_F, and C_{S_F} are user-defined cost rates for maintenance, repair, and system downtime respectively; $N_{M_1}^{(T_M)}$ and $N_{M_2}^{(T_M)}$are the total number of maintenance downtime intervals for C1 and C2 respectively; $N_{F_1}^{(T_M)}$ and $N_{F_2}^{(T_M)}$ are the total number of failures for C1 and C2 respectively; $N_{S_F}^{(T_M)}$ is the total number of system failures; $D_{M_1}^{(q)}$.and $D_{M_2}^{(q)}$ are the downtimes associated with the q-th maintenance; $D_{F_1}^{(q)}$ and $D_{F_2}^{(q)}$ are the downtime associated with the q-th repair; $D_{S_F}^{(q)}$ is the downtime related to the q-th system failure. A total of N_{si} simulation runs were conducted for each T_M. Hence, the total cost for a given maintenance period, T_M is the average of the estimated costs, $TC(T_M) = \frac{1}{N_{si}}\sum_{si=1}^{N_{si}} TC_{si}(T_M)$. Note that different maintenance frequencies yield unique average total costs. The optimum maintenance interval is the frequency that yields the minimum average total cost. Thus, the optimum maintenance interval is calculated as follows: $T_{Mopt} = \arg\min_{T_M} TC(T_M)$.

4. Results and Discussions

The ultimate goal of this paper is to present a method for estimating an unbiased expert improvement factor (EIF) from which the probability of component failure following the application of maintenance or repair can be calculated. This probability can be used to estimate the optimum maintenance interval of the proposed system. As discussed previously, to demonstrate the effectiveness of the proposed framework it is assumed that the EIF factor indicates the actual maintenance performance. Thus, a comparison between the estimated optimum maintenance interval based on EIF and the maintenance interval based on an IF, to be discussed below, was performed. This comparison shows the advantages of using an accurate improvement factor when using simulation to estimate an optimal maintenance interval.

Four recognized experts at an oil and gas installation were individually asked for their appraisal of the maintenance quality. Each provided a number between zero and one, where one indicates that the component remains the same as before the maintenance, and zero shows a renewal of the component. The interview results were averaged yielding an optimistic IF =0.30. As previously explained, all such judgments are inherently biased. In order to circumvent the judgment biases of our experts, the above methodology was applied to calculate the EIF. Experts were interviewed and questionnaires were administered to all 39 maintenance crew members at the oil and gas installation. This methodology yielded an EIF factor of 0.49, indicating that the maintenance performance was between 'as bad as old' and 'as good as new' and certainly worse than the IF of 0.30 estimated solely by interviewing experts. The optimum maintenance interval was determined through simulations for both the IF ($T_{Mopt_{IF}}$) and the EIF ($T_{Mopt_{EIF}}$) utilizing the following parameters: virtual component age (x_{k_i})=0, repair cost per hour (C_F)= 53, maintenance cost per hour (C_M)= 20, system failure cost per hour (C_{S_F})= 700, time horizon hr (T)= 57,684, iterations (N_{si})= 10000, and maintenance range (T_M)= [720 ,26280], (Note that these parameters do not show the cost currency, as it cannot be disclosed since the oil and gas installation used in this survey has requested confidentiality). Figure 5 shows the total cost associated with the maintenance interval, T_M, where it is assumed the improvement factor IF=0.30. As can be seen in the figure, the optimal maintenance interval is $T_{Mopt_{IF}} = 5760\ hr$, and the total cost at this value is 171,681.

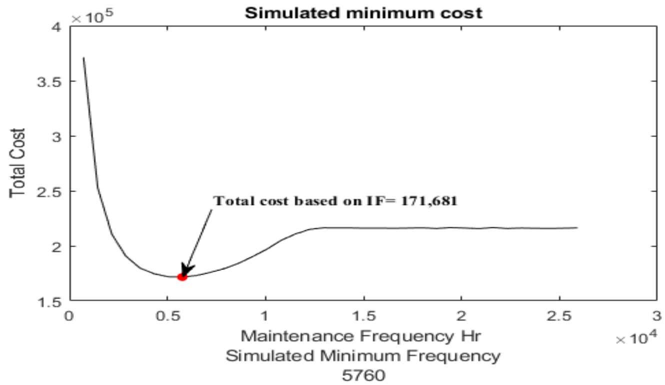

Figure 5. Optimum maintenance interval based on biased *IF* of 0.30.

It is assumed that the EIF factor indicates the maintenance performance at the installation. Thus, the simulation results (total costs) based on this factor are considered to be an accurate representation of the installation. Figure 6 shows the total cost associated with the maintenance interval where the improvement factor EIF = 0.49. The $T_{Mopt_{IF}}$ and the $T_{Mopt_{EIF}}$ values are 5760 hr and 5040 hr respectively. Figure 6 shows that the 5760 hr interval results in a cost of 205,956 and the 5040 hr interval results in a cost of 203,936. Thus, the 720 hr maintenance interval delay allowed by the IF results in an additional expense of 2,020.

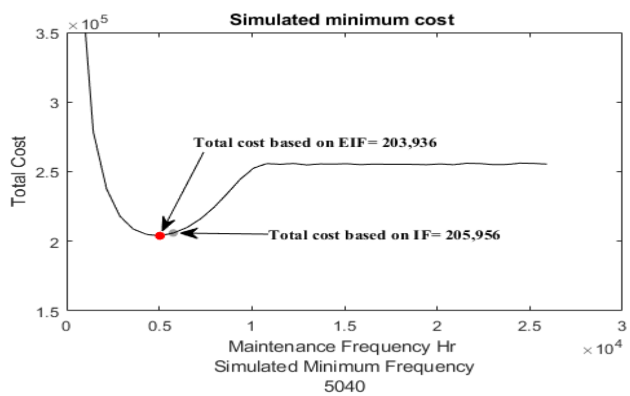

Figure 6. Cost comparison between *IF* and *EIF*

Accordingly, an accurate EIF is the building block for calculating the optimum maintenance interval. The proposed method, in combination with a simulation model, provides a tool for managers to optimize maintenance costs. Based on the EIF estimate of 0.49, the simulation shows that a maintenance interval of 5040 hr is most suitable for the given system of the oil and gas installation.

5. Conclusions

This paper presents a novel methodology for calculating an unbiased EIF based on expert estimates and a survey that incorporates human error identification factors. The proposed methodology aims to overcome the limitations of other methods previously described in the literature, by circumventing the bias embedded in expert opinion. This methodology was tested using a simulation model that estimated the optimum maintenance interval for a given system. The EIF and the IF were calculated for a system located in a real-world offshore oil and gas installation. The IF was estimated by asking recognized experts at the installation for their appraisals of the maintenance performance. Consistent evidence of a bias embedded in expert opinion has been found in the literature; hence, this IF is arguably biased. The simulation results from using the EIF and the IF to estimate the optimum maintenance intervals indicate there is likely a significant cost saving from utilizing the EIF. The EIF methodology can be customized and applied to other industries aiming to reduce total costs. Further work on the simulation model can be done through the design of an appropriate failure rate approach and by increasing the number of components within the system.

6. References

Abbassi, R., Khan, F., Garaniya, V., Chai, S., Chin, C., & Hossain, K. A. (2015). An integrated method for human error probability assessment during the maintenance of offshore facilities. *Process Safety and Environmental Protection. Vol. 94*, 172-179.

Alireza, N., Khakzad, N., Khan, F., MacKinnon, S., & Abbassi, R. (2013). The role of human error in risk analysis: Application to pre- and post-maintenance procedures of process facilities. *Reliability Engineering and System Safety. Vol. 119*, 251-258.

Al-Najjar, B., & Alsyouf, I. (2003). Selecting the most efficient maintenance approach using fuzzy multiple criteria decision making. *International journal of production economics. Vol. 84, Issue 1*, 85-100.

Antonovsky, A., Pollock, C., & Straker, L. (2013). Identification of the Human Factors Contributing to Maintenance Failures in a Petroleum Operation. *Human Factors and Ergonomics Society. Vol. 56, Issue 2*, 306-321.

Antonovsky, A., Pollock, C., & Straker, L. (2016). System reliability as perceived by maintenance personnel on petroleum production facilities. *Reliability Engineering and System Safety. Vol. 152,* , 58-65.

Asaad, M. N., & Yusoff, R. Z. (2013). Organizational culture influence on total productive maintenance (TPM) and operational performance using RASCH model analysis. *The Asian Journal of Technology Management. Vol. 6*, 72-81.

Asadsadeh, S. M., & Azadeh, A. (2014). An integrated systemic model for optimization of condition-based maintenance with human error. *Reliability Engineering & System Safety. Vol. 124*, 117-131.

Azadeh, A., Gaeni, Z., & Moradi, B. (2014). Optimization of HSE in maintenance activities by integration of continuous improvement cycle and fuzzy multivariate approach: A gas refinery. *Journal of Loss Prevention in the Process Industries. Vol. 32*, 415-427.

Bandemer, H., & Hans, W. (1992). *Fuzzy data analysis.* Netherlands: Kluwer academic publishers.

Basra, G., & Kirwan, B. (1998). Collection of offshore human error probability data. *Reliability engineering and system safety. Vol. 61, Issues 1-2*, 77-93.

Benitez, J. M., Martin, J. C., & Roman, C. (2007). Using fuzzy number for measuring quality of service in the hotel industry. *Tourism Management. Vol. 28, Issue 2*, 544-555.

Bentley, J. P. (1999). *Introduction to Reliability & Quality Engineering.* Harlow: Addison Wesley.

Block, H. W., Borges, W. S., & Savits, T. H. (1985). Age-Dependent minimal repair. *Journal of Applied Probability. Vol. 22, Issue 2*, 370-385.

Bojadziev, G., & Bojadziev, M. (2007). *Fuzzy Logic for Business, Finance, and Management.* Singapore: World Scientific Publishing Co.

Brown, M., & Proschan, F. (1980). Imperfect Maintenance. *Journal of Applied Probability*, 179-187.

Chiu, M. C., & Hsieh, M. C. (2016). laten human error analysis and efficient improvement strategies by fuzzy TOPSIS in aviation maintenance tasks. *Applied Ergonomics. Vol. 54*, 136-147.

CIM. (2011, October 20). Interview with Professor Lotfi A. Zadeh. (CIM, Interviewer)

Crichton, M. (2005). Attitudes to teamwork, leadership, and stress in oil industry drilling teams. *Safety Science. Vol. 43, Issue 9*, 679-696.

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ quality & safety. Vol. 22, Issue Suppl 2*, 58-64.

Davies, P. (2000). *Improving maintenance a guide to reducing human error.* Sudbury: Health and Safety Executive.

Dhillon, B. (2009). *Human Reliability, Error, and Human Factors in Engineering Maintenance.* Boca Raton : Taylor & Francis group.

DiMattia, D. G., Khan, F. I., & Amyotte, P. R. (2005). Determination of human error probabilities for offshore platform musters. *Journal of Loss Prevention in the Process Industries. Vol. 18, Issues 4-6*, 488-501.

Ding, S.-H., & Kamaruddin, S. (2015). Maintenance policy optimization- literature review and directions. *International Journal of Advanced Manufacturing Technology. 76, issue 5-8*, 1263-1283.

Dong, Y., Zhang, G., Hong, W.-C., & Xu, Y. (2010). Consensus models for AHP group decision making under row geometric mean prioritization method. *Decision Support Systems. Vol. 58, Issue 6*, 281-289.

Doyle, E., Lee, C.-G., & Cho, D. (2009). Justification for the next generation of maintenance modelling techniques. *Journal of Operational Research Society. Vol.60, Issue 4*, 461-470.

Ecken, P., Gnatzy, T., & Gracht, H. A. (2011). Desirability bias in foresight: consequences for decision quality based on Delphi results. *Technological Forecasting & Social Change. Vol. 78, Issue 9*, 1654-1670.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire lenght on participation and indicators of response quality in a web survey. *Public Opinion Quarterly. Vol. 73, Issue 2*, 349-360.

Garg, H., & Sharma, S. (2012). Behavior analysis of synthesis unit in fertilizer plant. *International Journal of Quality and Reliability Management. Vol. 29, Issue 2*, 217-232.

Grzegorzewski, P., & Mrówka, E. (2005). Trapezoidal approximations of fuzzy numbers. *Fuzzy sets and Systems. Vol. 153, Issue 1*, 115-135.

Hammed, A., Khan, F., & Ahmed, S. (2016). A risk based shutdown inspection and maintenance interval estimation considering human error. *Process Safety and Environmental Protection. Vol. 100*, 9-21.

Hauge, S., Hokstad, P., Habrekke, S., & Lundteigen, M. (2016). Common cause failure in safety instrumented systems using field experience from the petroleum industry. *Reliability Engineering and System Safety. Vol. 151*, 34-45.

Hennequin, S., & Arango, G. (2009). Optimization of imperfect maintenance based on fuzzy logic for a single stage single product production system. *Journal of Quality in Maintenance Engineering. Vol. 15, Issue 4*, 412-429.

Heo, G., & Park, J. (2010). A framework for evaluating the effects of maintenance-related human errors in nuclear power plants. *Reliability Engineering and System Safety. Vol. 95, Issue 7*, 797-805.

Hester, P. (2012). Epistemic Uncertainty Analysis: An Approach Using Expert Judgment and Evidential Credibility. *International Journal of Quality, Statistics, and Reliability. Vol. 2012*, 1-8.

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting Knowledge from Experts: A methodological Analysis. *Organizational Behavior and Human Decision Processes. Vol. 62, Issue 2*, 129-158.

Hosseini, F. M., & Ghadimi, N. (2016). Optimal preventive maintenance policy for electric power distribution systems based on the fuzzy AHP methods. *Complexity. Vol. 21, Issue 6*, 70-88.

Hussin, H., Ashim, F. M., Ramli, O. H., & Ghazali, S. M. (2012). Maintainability analysis of an offshore gas compression train system, a case study. *International Journal of Quality & Reliability Management. Vol. 30, Issue 5*, 495-510.

Jardine, A. K., & Tsang, A. H. (2013). *Maintenance, Replacement, and Reliability.* New York: CRC Press.

Khatab, A., & Aghezzaf, E.-H. (2016). Selective maintenance optimization when quality of imperfect maintenance actions are stochastic. *Reliability Engineering and System Safety. Vol. 150*, 182-189.

Kijima, M., Morimura, H., & Susuki, Y. (1988). Periodical replacement problem without assuming minimal repair. *European Journal of Operational Research. Vol. 37, Issue 2*, 194-203.

Kim, J. W., & Jung, W. (2003). A taxonomy of performance influencing factors for human reliability analysis of emergency tasks. *Journal of Loss Prevention in the Process Industries. Vol. 16, Issue 6*, 479-495.

Kim, J. W., Choi, G. B., Suh, J. C., & Lee, J. M. (2016). Dynamic optimization of maintenance and improvement planning for water main system: Periodic replacement approach. *Korean Journal of Chemical Engineering. Vol. 33, Issue 1*, 25-32.

Kirkeb¢en, G. (2009). Decision Behaviour - Improving Expert Judgment. In K. S. Terry M Williams, *Making Esential Choices with Scant Information* (pp. 169-194). London: Palgrave Mcmillan.

Klir, G. J. (1995). Fuzzy Logic Unearthing its meaning and significance. *IEEE Potentials*, 10-15.

Klir, G. J., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic.* New Jersey: Prentice Hall.

Kromrei, H. (2015). Enhancing the Annual Performance Appraisal Process: Reducing Biases and Engaging Employees Through Self-Assessment. *Performance Improvement Quarterly. Vol. 28, Issue 2*, 53-64.

Kumar, V. A., & Gandhi, O. (2011). Quantification of human error in maintenance using graph theory and matrix approach. *Quality and Reliability Engineering International. Vol. 27, Issue 8*, 1145-1172.

Kynn, M. (2008). The heuristics and biases bias in expert elicitation. *Royal Statistical Society. Vol. 171, Issue 1*, 239-264.

Landucci, G., & Paltrinieri, N. (2016). A methodology for frequency tailorization dedicated to the Oil & Gas sector. *Process Safety and Environmental Protection. Vol. 104, Issue 1*, 123-141.

Letot, C., Dehombreux, P., Rivière-Lorphèvre, E., Fleurquin, G., & Lesage, A. (2015). A degradation model for maintenance improvement in respect of cost and availability. *Journal of Quality in Maintenance. Vol. 21 Issue: 1,*, 55-69.

Likert, R. (1932). *A technique for the measurement of attitudes.* New York: Achives of psychology No.140.

Liu, Y., Li, Y., Huang, H.-Z., & Kuang, Y. (2011). An optimal sequential preventive maintenance policy under stochastic maintenance quality. *Structure and Infrastructure Engineering. Vol. 7, Issue 4*, 315-322.

Makis, V., & Jardine, A. K. (1992). Optimal replacement policy for a general model with imperfect repair. *The journal of the Operational Research Society. Vol. 43, Issue 2*, 111-120.

Mannan, S. (2005). Human factors and human error. In S. Mannan, *Lee's loss prevention in the process industries* (pp. 14/1-14/91). Burlington: Butterworth-Heinemann.

Martinez, J. C., Sanchez, D., Hidalgo, J. M., & Jimenez, P. M. (2014). A discussion on fuzzy cardinality and quantification. Some applications in image processing. *Fuzzy Sets and Systems. Vol. 257*, 85-101.

Nakagawa, T. (1979). Optimum policies when preventive maintenance is imperfect. *IEEE Transactions on Reliability. Vol. R-28, Issue 4*, 331-332.

Nelson, M., & Tan, H.-T. (2005). Judgement and decision making research in auditing: a task, person, and interpersonal interaction perspective. *Auditing a Journal of Practice & Theory*, 41-71.

Noortwijk, J. M., Dekker, R., Cooke, R. M., & Mazzuchi, T. A. (1992). Expert judgment in maintenance optimization. *IEEE Transactions on Reliability. Vol. 41. Issue: 3*, 427-432.

Panchal, D., & Kumar, D. (2016). Stochastic behaviour analysis of power generating unit in thermal power plant using fuzzy methodologies. *Operational Research of India. Vol. 53, Issue 1*, 16-40.

Peres, F., Bouzaiene, L., Bocquet, J. C., Billy, F., lannoy, A., & Haik, P. (2007). Anticipating aging failure using feedback data and expert judgment. *Reliability Engineering & System Safety. Vol. 92, Issue 2*, 200-210.

Pham, H., & Wang, H. (1996). Imperfect maintenance. *European Journal of Operational Research. Vol. 94*, 425-438.

Qi, H. S., Alzaabi, R. N., Wood, A. S., & Jani, M. (2015). A fuzzy criticality assessment system of process equipment for optimised maintenance management. *International Journal of Computer Integrated Manufacturing. Vol. 28, Issue 1*, 112-125.

Rahimi, M., & Rausand, M. (2013). Monitoring human and organizational factors influencing common cause failures of safety instrumented system during the operational phase. *Reliability Engineering and System Safety. Vol. 120*, 10-17.

Rausand, M. (2011). Human reliability analysis. In M. Rausand, *Risk Assessment Theory, Methods, and Applications* (pp. 409-456). New Jersey: WILEY.

Reiman, T., & Oedewal, P. (2006). Assessing the maintenance unit of a nuclear power plant, identifiying the cultural conceptions concerning the maintenance work and the maintenance organization. *Safety Science. Vol. 44, Issue 9*, 821-850.

Ren, J., Jenkinson, I., Wang, J., Xu, D., & Yang, J. (2008). A methodology to model causal relationships on offshore safety assessment focusing on human and organizational factors. *Journal of Safety Research. Vol. 39, Issue 1*, 87-100.

Said, A. B., Shahzad, M. K., Zamai, E., Hubac, S., & Tollenaere, M. (2016). Experts' knowledge renewal and maintenance actions effectiveness in high mix low volume industries, using Bayesian approach. *Cognition, Technology & Work. Vol. 18, Issue 1*, 193-213.

Skogdalen, J. E., & Vinnem, J. E. (2011). Quantitative risk analysis offshore- human and organizational factors. *Reliability Engineering and System Safety. Vol. 96, Issue 4*, 468-479.

Tabibzadeh, M., & Meshkati, N. (2014). Learning from the BP Deepwater Horizon accident: risk analysis of human and organizational factors in negative pressure test. *Environment System Decision. Vol. 34, Issue 2*, 194-207.

Takac, Z. (2016). Subsethood measures for interval valued fuzzy sets based on the aggregation of interval fuzzy implications. *Fuzzy Sets and Systems. Vol. 283, Issue 15*, 120-139.

Tezuka, M., Munakata, S., & Sawada, M. (2015). Maintenance schedule optimization based on failure probability distribution. *Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management* (p. 5). Dubai: IEEE.

Tredger, E., J.T.H. Lo, S., Lau, H., Bonello, N., Hlavka, B., & Scullion, C. (2016). Bias, guess and expert judgment in actuarial work. *British Actuarial Journal. Vol. 21, Issue 3*, 545-578.

Vinnem, J., R. Bye, B. G., Kongsvik, T., Nyheim, O., Okstad, E., Seljelid, J., & Vatn, J. (2012). Risk modelling of maintenance work on major process equipment on offshore petroleum installations. *Journal of Loss Prevention in the Process Industries. Vol. 25, Issue 2*, 274-292.

Walls, L., & Quigley, J. (2001). Building prior distributions to support bayesian reliability growth modelling using expert judgment. *Reliability Engineering and System Safety. Vol. 74, Issue 2*, 117-128.

Wang, C.-H., & Tsai, S.-W. (2014). Optimizing bi-objective imperfect preventive maintenance model for series-parallel system using established hybrid genetic algorithm. *Journal of Intelligent Manufacturing. Vol. 25, Issue 3*, 603-616.

Wang, H. (2006). Imperfect maintenance and dependence. In W. H., *Reliability and Optimal Maintenance* (pp. 13-30). London: Springer Series in Reliability Engineering. .

Wang, T.-C., & Chuang, L.-H. (2014). Psychological an physiological fatigue variation and fatigue factors in aircraft line maintenance crews. *International Journal of Industrial Ergonomics. Vol. 44, Issue 1*, 107-113.
Wang, Y. F., Li, Y. L., Zhang, B., Yan, P. N., & Zhang, L. (2015). Quantitative risk analysis of offshore fire and explosion based on the analysis of human and organizational factors. *Mathematical Problems in Engineering. Vol 2015*, 1-10.
Zadeh, L. (1965). Fuzzy Sets. *Information and Control. Vol. 8, Issue 3*, 338-353.
Zhou, X., Wu, C., Li, Y., & Xi, L. (2016). A preventive maintenance model for leased equipment subject to internal degradation and external shock damage. *Reliability Engineering and System Safety. Vol. 154*, 1-7.
Zimmermann, H. (1985). *Fuzzy Set Theory and its Applications.* Boston: Kluwer-Nijhoff.

Biographies

Rogelio Emmanuel Jauregui Miramontes received a PhD at the University of Toronto, a BEng and MASc degrees from the Monterrey Institute of Technology and Higher Education Mexico, and a maintenance administration diploma at the National Autonomous University of Mexico. He has 13 years of experience in industry. His research interests include simulation, optimization, reliability, maintenance scheduling, and human factors.

Yuri A. Lawryshyn is currently a fulltime Associate Professor. Yuri Lawryshyn received BASc and MASc degrees from the University of Toronto in Mechanical Engineering, a PhD from the Department of Chemical Engineering and Applied Chemistry at the University of Toronto, an MBA from the Richard Ivey School of Business (University of Western Ontario) and a Financial Engineering Diploma from the Schulich School of Business (York University). Yuri specializes in the area of numerical modelling, including financial modelling, and real options analysis. Yuri has supervised over 60 projects related to financial modelling, trading, econometrics, customer analytics, operational risk, cyber security and FinTech.

Pasi Petteri Luukka is currently a fulltime Professor. Pasi Luukka received the M.Sc. degree from the Department of Information Technology, Lappeenranta University of Technology, and a D.Sc. degree in applied mathematics from the Department of Mathematics and Physics, Lappeenranta University of Technology. His research interests include fuzzy data analysis, classification, feature selection, and fuzzy decision making. He has authored over 50 journal papers.