Ocean Engineering 58 (2013) 293–303

Contents lists available at SciVerse ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

# A modified CREAM to human reliability quantification in marine engineering

Z.L. Yang [a,*], S. Bonsall [a], A. Wall [a], J. Wang [a], M. Usman [b]

[a] Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK
[b] Marine Engineering Consulting, Lloyd's Register, London, UK

ARTICLE INFO

*Article history:*
Received 30 May 2012
Accepted 11 November 2012
Available online 20 December 2012

*Keywords:*
Marine engineering
Safety engineering
Maritime human reliability
Human reliability quantification
CREAM
Fuzzy logic
Bayesian network

ABSTRACT

Modern shipping activities are carried out via a highly sophisticated man–machine system within which technological, social and environmental factors often contribute to the occurrence of human failures. Due to the high risks caused by such failures, human reliability analysis (HRA) has always been a serious concern in marine engineering safety. However, the problem of lack of data, together with the complexity of marine engineers' behaviour, has weakened the applicability of well-established HRA methods (i.e., cognitive reliability and error analysis method (CREAM)) in the maritime context. This paper proposes a modified CREAM to facilitate human reliability quantification in marine engineering by incorporating fuzzy evidential reasoning and Bayesian inference logic. The core of the new method is to use evidential reasoning to establish fuzzy IF–THEN rule bases with belief structures, and to employ a Bayesian inference mechanism to aggregate all the rules associated with a marine engineer's task for estimating its failure probability. Consequently, the outcomes of this work can also provide safety engineers with a transparent tool to realise the instant estimation of human reliability performance for a specific scenario/task.



## 1. Introduction

During the past decades, human activity has been the source of ecological disasters, evidenced by accidents such as the "Three Mile Island", "Chernobyl" and "Prestige" tragedies. The shipping industry shares part of this responsibility with the fact that 80–85% of all the recorded maritime accidents are directly due to human error or associated with human error (Harati-Mokhtari et al., 2007). The shipping community has therefore taken a series of measures in order to minimise such events via international bodies like the International Maritime Organization (IMO). They include the introduction of Human Reliability Analysis (HRA) to the Formal Safety Assessment (FSA). FSA was developed as a risk-based goal-setting maritime safety rule-making methodology to address the inappropriate effectiveness of purely using a prescriptive approach to make the regulations of attempting to reduce risks in ship design and operation. However, effective implementation of such a measure is challenging given the difficulties of directly adopting the mature HRA methods available in literature to the maritime domain. For example, lack of historical data in general and human failure statistics in particular is a well-recognised problem in maritime safety research (Ren et al., 2008; Yang et al., 2010). More importantly, traditional tasks onboard ships requiring much 'doing' are being replaced by more 'thinking' jobs with a nature of higher contextual dependency where technological, social and environmental factors often contribute to a dynamic sea working environment in an interactive way. 'Second generation' HRA methods have therefore been developed to overcome such difficulties by appropriately taking into account the contextual influence of a task and by being equipped with the more powerful ability of incorporating expert judgments to deliver quantitative human failure analysis results. Although attractive, these methods have still exposed some shortcomings in their practical application. For instance, the prospective quantification process of a Cognitive Reliability and Error Analysis Method (CREAM), which normally produces an interval approximation analysis result, cannot provide an appraisal of the consequences of human performance on maritime system safety which can be further used in a FSA framework.

This paper therefore develops a human failure quantification methodology by incorporating fuzzy logic, evidential reasoning and Bayesian network techniques into the prospective analysis of CREAM. It can be used either as a stand-alone method for human failure screening or as part of an integrated method for the support of using FSA in a large maritime system safety analysis. More specifically, the new fuzzy Bayesian CREAM approach can be used by maritime safety analysts (a) to identify the tasks that require human cognition and depend on cognitive reliability, (b) to determine the conditions where cognitive reliability may be reduced and therefore constitute a source of risk (Hollnagel, 1998), (c) to provide a precise/point (instead of interval) human failure rate estimation which can be used as input in

* Corresponding author. Tel.: +44 151 231 2531; fax: +44 151 207 3460.
*E-mail address:* z.yang@ljmu.ac.uk (Z.L. Yang).

0029-8018/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.oceaneng.2012.11.003

FSA (Konstandinidou et al., 2006), (d) to improve the accuracy of human failure probabilities through avoiding the loss of useful information in classical fuzzy Max–Min inference operations and (e) to realise the instant estimation of human failure probability for a specific task. Given that (a) and (b) above have been addressed through the well-known retrospective and prospective analysis in CREAM, this study will focus on the investigation of (c) using fuzzy logic, (d) using evidential reasoning and (e) using Bayesian inference.

To achieve the above aim, the research problem is formulated by identifying the weaknesses of the traditional HRA approaches in general and CREAM in specific in the ensuing section. In Section 3, a generic human failure quantification methodology is developed to overcome the problems of applying the traditional CREAM to maritime HRA by integrating fuzzy evidential reasoning with Bayesian logic. Its applicability and feasibility are demonstrated by studying the case of seafarers' performances in a maritime emergency situation due to engine failures in Section 4. A sensitivity analysis is also conducted in this section to examine the logicality of the new method and to show its superiority over the classical fuzzy CREAM approaches. Section 5 concludes the research with a discussion of potential future work.

## 2. Problems of using CREAM in quantifying maritime human failures

Human work onboard ships can be characterised by a change moving from 'Tell me what to do' to 'Show me how to do it' and then to 'Involve me in it'. The development of modern technology has changed the nature of human work from being manual skills to being mostly knowledge intensive functions and cognitive tasks such as diagnosis, planning and problem solving. HRA methods have correspondingly evolved, starting from the 'first generation' Technique for Human Error Rate Prediction (THERP), passing through the 'second generation' CREAM and arriving to A Technique for Human Error ANAlysis (ATHEANA) (under development) (Konstandinidou et al., 2006).

### 2.1. First generation HRA methods

The 'first generation' HRA methods are developed based on the pivotal concept of human error mainly resulting from the inherent deficiencies of humans (Marseguerra et al., 2006), including task-based THERP, response time-based Operator Action Tree (OAT), Human Cognitive Reliability (HCR) analysis, expert judgement-based TESEO (Tecnica Empirica Stima Errori Operatori), Human Error Assessment and Reduction Technique (HEART) and Success Likelihood Index Methodology (SLIM), etc. (Kim and Bishu, 2006). Human error probability (HEP) is treated as the conventional probability which is often used to describe the failure likelihood of pure hardware like mechanical and structural systems. Such an argument has however been challenged over the past 20 years because extensive studies of human performance in accidents have shown the importance of the contextual conditions in which the task is performed is greater than the characteristics of the task itself (Marseguerra et al., 2006). From this evidence, the criticism of the 'first generation' approaches are that it has generally recognised shortcomings in a scarcity of data, lack of consistency in treating error of commission, inadequate proof of accuracy, inadequate psychological realism, insufficient treatment of performance shaping factors (PSFs), inadequate treatment of dynamic situations, a mechanical view of human, high level of uncertainty, lack of systematic task analysis structure and inadequate error reduction strategies (Hollnagel, 1998; Kim and Bishu, 2006). Quantitative comparisons among some classical HRA methods are also analysed by He et al. (2005), which indicates the superiority of the 'second generation' HRA methods over the 'first generation' ones, when the behavioural rather than cognitive focus is mainly concerned.

### 2.2. Second generation HRA methods

For developing a new HRA paradigm, issues such as the probabilistic approach of human behaviour to risk analysis, cognitive model complexity, integration of PSF and model validation have been raised as critical points. In order to address the needs identified, a considerable amount of effort has been devoted to propose alternative 'second generation' HRA methods including the COGnitive evENT tree system (COGENT), Human Interaction TimeLINE (HITLINE), ATHEANA, Connectionism Assessment of Human Reliability (CAHR), CREAM, etc. As one of the best known 'second generation' HRA methods, CREAM presents a consistent error classification system that integrates individual, technological and organisational factors. The classification describes the relations between causes and effects by defining a number of sub-groups and tables, which are provided for the error modes on the one hand and the organisational causes on the other. To model the causal relations, the CREAM methodology has been derived from its core, the Contextual Control Model (COCOM). COCOM focuses on the principle that human performance is the outcome of the purposive use of competence adjusted to specific working conditions rather than of the pre-determined sequence of response to given events. The model therefore defines four characteristic control modes according to the human cognition and action context, which are determined by nine Common Performance Conditions (CPCs). The four control modes, namely "Scrambled", "Opportunistic", "Tactical" and "Strategic", are linked with different failure probability intervals representing human action failure probabilities. The nine CPCs are "adequacy of organisation (#1)", "working conditions (#2)", "adequacy of man–machine interface and operational support (#3)", "availability of procedures and plans (#4)", "number of simultaneous goals (#5)", "available time (#6)", "time of day (#7)", "adequacy of training and experience (#8)" and "crew collaboration quality (#9)". Appropriate linguistic variables chosen to describe the nine CPCs are categorised into three sets in terms of their effects (positive, negative or neutral) on human performance reliability (Table 1). Having given the classification and COCOM, CREAM can be used to conduct its bi-directional HRA inference, retrospective analysis and prospective analysis (Hollnagel, 1998).

### 2.3. Problem analysis of CREAM

One of the purposes of prospective analysis is to provide quantification of human performance reliability. CREAM approaches the quantification in two steps by providing a basic and an extended method (He et al., 2008). The basic CREAM method corresponds to initial screening of the task and its major segments through calculating the distinct sums of the positive and negative influencing CPCs to determine the relevant control mode and failure rate interval. However, the failure rate intervals from the basic method appear to be too wide even for the use in screening (Fujita and Hollnagel, 2004). It is also difficult to further use such failure rate intervals in the FSA framework. The extended method uses the output from the basic CREAM and appropriate data sources to calculate the probability of each cognitive function failure and to deal with the tasks needing more precision and detail (Hollnagel, 1998). Lack of the critical mass in statistical failure data, however, proves the tasks of adapting the extended method in the maritime area to be challenging and the generation of novel methods to be urgent. Recently, new quantification approaches for CREAM have been proposed using fuzzy logic (Konstandinidou et al., 2006; Kim and Bishu, 2006). While the methods provide a systematic procedure to account for the ambiguity in the quantification of CPCs for calculating specific numerical