DEFENCE S&T TECHNICAL BULLETIN

VOL. 11 NUM. 1 YEAR 2018 ISSN 1985-6571

CONTENTS

Ministry of Defence
Malaysia

SCIENCE & TECHNOLOGY RESEARCH INSTITUTE
FOR DEFENCE (STRIDE)

AIMS AND SCOPE

The Defence S&T Technical Bulletin is the official technical bulletin of the Science & Technology Research Institute for Defence (STRIDE). The bulletin, which is indexed in, among others, Scopus, Index Corpenicus, ProQuest and EBSCO, contains manuscripts on research findings in various fields of defence science & technology. The primary purpose of this bulletin is to act as a channel for the publication of defence-based research work undertaken by researchers both within and outside the country.

WRITING FOR THE DEFENCE S&T TECHNICAL BULLETIN

Contributions to the bulletin should be based on original research in areas related to defence science & technology. All contributions should be in English.

PUBLICATION

The editors' decision with regard to publication of any item is final. A manuscript is accepted on the understanding that it is an original piece of work that has not been accepted for publication elsewhere.

PRESENTATION OF MANUSCRIPTS

The format of the manuscript is as follows:

a) Page size A4
b) MS Word format
c) Single space
d) Justified
e) In Times New Roman ,11-point font
f) Should not exceed 20 pages, including references
g) Texts in charts and tables should be in 10-point font.

Please e-mail the manuscript to:

1) Gs. Dr. Dinesh Sathyamoorthy (dinesh.sathyamoorthy@stride.gov.my)
2) Dr. Mahdi bin Che Isa (mahdi.cheisa@stride.gov.my)

The next edition of the bulletin (Vol. 11, Num. 2) is expected to be published in November 2018. The due date for submissions is 15 August 2018. **It is strongly iterated that authors are solely responsible for taking the necessary steps to ensure that the submitted manuscripts do not contain confidential or sensitive material.**

The template of the manuscript is as follows:

TITLE OF MANUSCRIPT

Name(s) of author(s)

Affiliation(s)

Email:

ABSTRACT

Contents of abstract.

Keywords: *Keyword 1; keyword 2; keyword 3; keyword 4; keyword 5.*

1. TOPIC 1

Paragraph 1.

Paragraph 2.

1.1 Sub Topic 1

Paragraph 1.

Paragraph 2.

2. TOPIC 2

Paragraph 1.

Paragraph 2.

Figure 1: Title of figure.

Table 1: Title of table.

Content	**Content**	**Content**
Content	Content	Content
Content	Content	Content
Content	Content	Content

Equation 1 (1)
Equation 2 (2)

REFERENCES

Long lists of notes of bibliographical references are generally not required. The method of citing references in the text is 'name date' style, e.g. 'Hanis (1993) claimed that...', or '…including the lack of interoperability (Bohara *et al.*, 2003)'. End references should be in alphabetical order. The following reference style is to be adhered to:

Books

Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.

Book Chapters

Goodchild, M.F. & Quattrochi, D.A. (1997). Scale, multiscaling, remote sensing and GIS. *In* Quattrochi, D.A. & Goodchild, M.F. (Eds.), *Scale in Remote Sensing and GIS*. Lewis Publishers, Boca Raton, Florida, pp. 1-11.

Journals / Serials

Jang, B.K. & Chin, R.T. (1990). Analysis of thinning algorithms using mathematical morphology. *IEEE T. Pattern Anal.*, **12**: 541-550.

Online Sources

GTOPO30 (1996). *GTOPO30: Global 30 Arc Second Elevation Data Set*. Available online at: http://edcwww.cr.usgs.gov/landdaac/gtopo30/gtopo30.html (Last access date: 1 June 2009).

Unpublished Materials (e.g. theses, reports and documents)

Wood, J. (1996). *The Geomorphological Characterization of Digital Elevation Models*. PhD Thesis, Department of Geography, University of Leicester, Leicester.

COMBATING CORROSION: RISK IDENTIFICATION, MITIGATION AND MANAGEMENT

Mahdi Che Isa*, Abdul Rauf Abdul Manaf & Mohd Hambali Anuar

Magnetic Research & Treatment Centre, Science & Technology Research Institute for Defence (STRIDE), Malaysia

*Email: mahdi.cheisa@stride.gov.my

ABSTRACT

Corrosion can be found everywhere, it occurs all the time and appears in different forms at different environments. Even though corrosion is well known to everybody, we sometimes do not realise that corrosion attack is one of the threats that has jeopardised safety, security and economy of countries in the world, and the consequences of its attack are often troublesome and very costly. Corrosion problems can be evaluated from different point of views, either using technical or non-technical knowledge. Methods based on engineering and based on management are two approaches used in mitigating corrosion impact. Risk assessment is the first step to be applied in the corrosion management process. Given an assessment of risk, a strategy of corrosion management can be constructed, implemented and improved. The corrosion management strategy can be integrated into the policy system to prevent, analyse and solve the problem caused by corrosion. The corrosion management strategy is the route for the implementation of corrosion management activities to accomplish the targets established by the corrosion management policy. A description of the nature of corrosion attack, a risk assessment methodology, management strategy to fight corrosion, and preventive measure for successful corrosion mitigation are discussed in this article.

Keywords: *Corrosion attack; risk identification; corrosion management; economic loss.*

1. INTRODUCTION

Corrosion is a naturally occurring phenomenon that affects our society on a daily basis, causing degradation and damage to household appliances, automobiles, airplanes, highway bridges, energy production and distribution systems, and much more. Like other threat such as earthquakes or severe weather disturbances, corrosion can cause dangerous and expensive damage to everything and costs associated with the damage is substantial. Corrosion has a serious impact on various infrastructure or equipment. For example, in the Gulf war, a serious problem of rotor blade damage in helicopter was caused by dessert sand (Wood, 1999; Edwards & Davenport, 2006). The storage of equipment is a serious matter for countries with corrosive environments such as our tropics environment with the presence of high humidity. Humidity is the biggest killer of hardware and from the above conditions, it is observed that corrosion attack is everywhere, there is no industry or house where it does not penetrate, and it demands a state of readiness for engineers and scientists to combat this problem (Dehri & Erbil, 2000; Guedes Soares *et al*., 2009; Gil *et al*., 2010). Figure 1 shows photos retrieved from google images of corrosion attack on the plant infrastructures, water distribution pipeline, explosive cartridge and metallic storage tanks.

They are no materials which are resistant to corrosion. They must be matched to the environment which they will encounter in service. The most dangerous environmental impact of corrosion is that it occurs in major industrial plants. Typical examples in this regard would be electrical power plants as well the chemical processing plants. Things can reach extreme consequences that a plant may even shut down. Some other major consequences could be contamination of the product; loss of efficiency as well as damages to the adjacent product placed behind the corrosive material. The impact could also be social such as safety, health as well as depletion of natural resources. On the safety aspect it

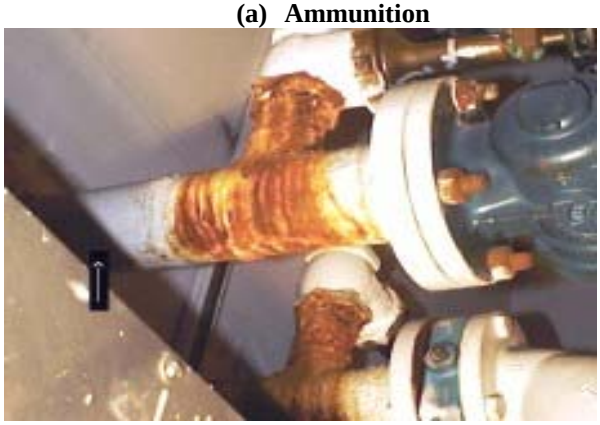
could be sudden release of a toxic product, and on the health perspective it could be pollution from the escaped product (Garverick, 1994; Javaherdashti, 2000; Patil & Ghanendra, 2013).

While corrosion can take many forms, it is generally defined as a chemical or electrochemical reaction between materials and its environments that produce a deterioration of the material and its properties. ISO 8044 defines corrosion as "*physicochemical interaction, which is usually of an electrochemical nature, between a metal and its environment which results in changes in the properties of the metal and which may often lead to impairment of the function of the metal, the environment, or the technical system of which these form a part*" (Mattsson, 1989).

Categorisation of the form of corrosion threat has existed in various schemes for many years. A more focused view would categorise corrosion in various subsections such as uniform corrosion, localised corrosion, high temperature corrosion, metallurgical influenced corrosion, and microbiological influenced corrosion. Almost all corrosion problems and failures encountered in service can be associated with one or more of the eight basic forms of corrosion: general corrosion, galvanic corrosion, concentration-cell (crevice) corrosion, pitting corrosion, intergranular corrosion, stress corrosion cracking, dealloying, and erosion corrosion.

(a) Ammunition

(b) Underground storage tank



(c) Piping system

(d) Processing plant

Figure 1: The severity of corrosion attacks (Retrieved from Google images).

1.1 General Corrosion

General corrosion (sometimes called uniform corrosion), is well distributed and low level attack against the entire metal surface with little or no localized penetration. The corrosion rate is nearly constant at all locations. Microscopic anodes and cathodes are continuously changing their electrochemical behavior from anode to cathode cells for a uniform attack. The general corrosion

rates for metals in a wide variety of environments are known, and common practice is to select materials, with rates that are acceptable for the application.

1.2 Galvanic Corrosion

Galvanic (dissimilar metals) corrosion occurs when two electrochemically dissimilar metals are metallically connected and exposed to a corrosive environment, this is an aggressive and localized form of corrosion due to the electrochemical reaction often found between two or more dissimilar metals in an electrically conductive environment. The less noble metal (anode) suffers accelerated attack and the more noble metal (cathode) is cathodically protected by the galvanic current.

1.3 Concentration-cell Corrosion

Concentration-cell corrosion occurs because of differences in the environment surrounding the metal. This form of corrosion is sometimes referred to as “crevice corrosion”, “gasket corrosion”, and “deposit corrosion” because it commonly occurs in localized areas where small volumes of stagnant solution exist. Normal mechanical construction can create crevices at sharp corners, spot welds, lap joints, fasteners, flanged fittings, couplings, threaded joints, and tube sheet supports. At least five types of concentration cells exist: the most common are the “oxygen” and “metal ion” cells. Areas on a surface in contact with an electrolyte having a high oxygen concentration generally will be cathodic relative to those areas where less oxygen is present (oxygen cell). Areas on a surface where the electrolyte contains an appreciable quantity of the metal’s ions will be cathodic compared to locations where the metal ion concentration is lower (metal ion cell).

1.4 Pitting Corrosion

Pitting is the most common form of corrosion found where there are incomplete chemical protective films and insulating or barrier deposit of dirt, iron oxide, organic, and other foreign substances at the surface. Pitting corrosion is a randomly occurring, highly localized form of attack on a metal surface, characterized by the fact that the depth of penetration is much greater than the diameter of the area affected. Pitting is one of the most destructive forms of corrosion, yet its mechanism is not completely understood. Steel and galvanized steel pipes and storage tanks are susceptible to pitting corrosion and tuberculation by many potable waters. Various grades of stainless steel are susceptible to pitting corrosion when exposed to saline environments.

1.5 Intergranular Corrosion

Intergranular corrosion is a localized condition that occurs at, or in narrow zones immediately adjacent to, the grain boundaries of an alloy. Although a number of alloy systems are susceptible to intergranular corrosion, most problems encountered in service involve austenitic stainless steels (such as 304 and 316) and the 2000 and 7000 series aluminum alloys. Welding, stress relief annealing, improper heat treating, or overheating in service generally establish the microscopic, compositional inhomogeneities that make a material susceptible to intergranular corrosion.

1.6 Stress Corrosion Cracking

Stress corrosion cracking (environmentally induced-delayed failure) describes the phenomenon that can occur when many alloys are subjected to static, surface tensile stresses and are exposed to certain corrosive environments. Cracks are initiated and propagated by the combined effect of a surface tensile stress and the environment. When stress corrosion cracking occurs, the tensile stress involved is often much less than the yield strength of the material; the environment is usually one in which the material exhibits good resistance to general corrosion.

1.7 Dealloying

Dealloying or selective leaching is a corrosion process in which one element is preferentially removed from an alloy. This occurs without appreciable change in the size or shape of the component; but the affected area becomes weak, brittle, and porous. The two most important examples of dealloying are the preferential removal of zinc from copper-zinc alloys (dezincification), and the preferential removal of iron from gray-cast iron (graphitic corrosion). Graphitic corrosion sometimes occurs on underground cast iron water mains and leads to splitting of the pipe when the water pressure is suddenly increased.

1.8 Erosion Corrosion

Erosion corrosion refers to the repetitive formation (a corrosion process) and destruction (a mechanical process) of the metal's protective surface film. This is the gradual and selective deterioration of a metal surface due to mechanical wear and abrasion. It is attributed to entrained air bubbles, suspended matter and particulates under a flow rate of sufficient velocity. This typically occurs in a moving liquid. Erosion is similar to impingement attack, and is primarily found at elbows and tees, or in those areas where the water sharply changes direction. Softer metals such as copper and brasses are inherently more susceptible to erosion corrosion than steel. An example is the erosion corrosion of copper water tubes in a hot, high velocity, soft water environment. Cavitation is a special form of erosion corrosion.

2. THE ECONOMIC LOSS FROM CORROSION ATTACK

Corrosion and its effects have a profound impact on the economy and the integrity of infrastructure and equipment worldwide. This impact is manifested in significant economic loss, maintenance, repair and replacement efforts, reduced access and availability, poor performance and unsafe conditions associated with facilities and equipment. People and organisations involved in corrosion prevention, control, and repair activities in all types of industries have always required reasonably accurate estimates of the costs of corrosion in order to provide persuasive cost/benefit analyses. On a problem-specific or company-specific scale, the current cost of a particular corrosion problem is required to perform life cycle cost (LCC) estimates or return on investment (ROI) calculations to help choose between one or more possible corrective actions. On a much broader industry-wide or national scale, estimates of the cost of corrosion can be used to demonstrate the impact of corrosion on the industry or economy, and the need for investment in facilities, training, research, and policy. The primary metric reflecting this impact is cost. A recent study estimates that the annual cost of corrosion in the U.S. alone is $276 billion (Bhaskaran *et al.*, 2005; Koch *et al*., 2002). Corrosion costs associated with labour, material, and related factors have substantial effects on the economies of industrial nations and more specifically, on the civil/industrial and government sectors of these economies as shown in Table 1 (Koch *et al*., 2016).

According to available data, between 4% to 6% of Gross Domestic Products (GDP) is lost to corrosion – that's USD $1.6 trillion dollars lost every year from the world's economy (Jackson, 2017). In the Malaysian context, 4% of GDP in the year 2016 (RM 600.0 Billions) would mean a loss of around RM 24.0 billion a year – that's more than RM 1200 annually for every man, woman and child in the country and works out to just about the entire Malaysian healthcare budget for 2008 (DOS, 2017; Mukhriz, 2010). These corrosion cost figures come from the National Association of Corrosion Engineers (NACE), the leading global corrosion-control standards organisation whom have recently opened an office in Kuala Lumpur in recognition of the importance of corrosion prevention and control in Malaysia and Asia. NACE estimates that as much as 30% of the cost of corrosion could be saved by using appropriate technology (Mukhriz, 2010). However, it has been estimated that 25% to 30% of annual corrosion costs (RM 24 billions) could be saved if optimum corrosion management practices were employed and therefore we need a comprehensive "National Strategy" for corrosion control.

Table 1: Global corrosion cost by region by sector (Billion USD).

Economic Region	*Corrosion cost by sector*				Total GDP	% GDP
	Agriculture	Industry	Services	Total		
USA	2.0	303.2	146.0	451.3	16,720	2.7
India	17.7	20.3	32.3	70.3	1,670	4.2
European Region	3.5	401	297	701.5	18,331	3.8
Arab World	13.3	34.2	92.6	140.1	2,789	5.0
China	56.2	192.5	146.2	394.9	9,330	4.2
Russia	5.4	37.2	41.9	84.5	3,113	4.0
Japan	0.6	45.9	5.1	51.6	5,002	1.0

For example, most of defence equipment and facilities are composed of materials that are susceptible to oxidation, stress, surface wear and other chemical and environmental mechanisms that cause corrosion. The Malaysian Ministry of Defence (MinDef) maintains billions of Malaysian Ringgit (RM) worth of equipment, systems and infrastructure used in various degrees of corrosive environments around the country. The impact of corrosion on the equipment, systems and infrastructure will cause it to deteriorate, reducing availability and performance capability. The equipment, systems and infrastructure in military services have long recognized the pervasive and insidious effects of corrosion. Therefore, it required corrosion inspection, repair, and replacement and the decreasing availability of critical systems reduces mission readiness. The indirect cost of corrosion to our forces can also be associated with troop safety, weapon system reliability, and overall readiness of the military operation.

To reduce the losses due to corrosion attack, MinDef need to develop a long–term strategy to reduce corrosion threat and its effects. This strategy is to include expanded emphasis on corrosion control, a uniform application of processes for testing and certifying new corrosion prevention technologies, the implementation of programs that ensure a focused and coordinated approach to corrosion-related information distribution, information sharing and a coordinated corrosion control research & development program. In addition, each of the military services tended to develop different approaches to the corrosion problem based on the conditions unique to each service. This has led to stringent standards and processes associated with military corrosion control practices. At the same time, the civil/industrial sector has been driven toward more economic standards and processes because of the inherent profit motive in the competitive marketplace.

In the government sector, the military has been battling corrosion for many years, and the military objective has reliability and readiness as its primary target. Thus, the government sector, led by the MinDef, is the most suitable agency to embark on a major effort to standardise corrosion prevention and control strategies, policies, training, best practices, and research and development across government agencies.

3. CORROSION RISK

Corrosion is so prevalent and takes so many forms that its occurrence and associated costs cannot be eliminated completely. The bottom line is that the use of appropriate corrosion prevention strategies and control methods protects public safety, prevents damage to property and the environment, and saves money. It is widely recognized within any organisation or industry that effective management

of corrosion will contribute towards achieving the following benefits such as reduction in leaks, increased plant availability, reduction in deferment costs, reduction in unplanned maintenance and statutory or corporate compliance with safety, health & environmental policies.

Risk assessment is the key element in the overall corrosion management strategy, identifying critical items requiring high focus in view of inspection, monitoring programs, repair and maintenance. Risk analysis is a new technique and useful tools to determine and to find causes of risk mainly to detect possibility (predict) of failure and damage in an operating system (Pasman *et al.*, 2009). This new technique can be applied in corrosion because it has the capability to identify areas where corrosion may be safely ignored and where it must be attended to. It even provides the pointer to where resources will be spent with greatest reward. Thus, it provides the evidence that permits the construction of a cost-effective corrosion management programme. Cost consequences of accidents are an important part of risk assessment and risk management for critical infrastructure. Cost are included in various measurement scales of the seriousness of incidents and are used as inputs to assess the impacts of such incidents, and the methodology used here can be used for those purposes, allowing risk managers to identify factors that determine and quantify risk in a relatively straightforward way (Restrepo *et al*., 2009; Simonoff *et al*., 2010). The corrosion risk analysis is very important in classifying the relative severity of corrosion risk in the 'low risk', 'medium risk' and 'high risk' categories.

The risk analysis procedure starts with a process flow diagram. For example, the condition of facility or plant equipment is then considered with respect to the likelihood of possible corrosion phenomena, and the consequence of any failure of that piece of equipment. The likelihood probabilities are given the ratings 1, 2, 3…etc, 1 being the lowest probability with other numbers indicating increasingly higher probabilities. Similarly the consequence probabilities are given the ratings A, B, C.…etc, A being the lowest probability with other letters indicating increasingly higher probabilities. These probabilities are plotted in an X-Y Risk Matrix as shown in Figure 2 below.

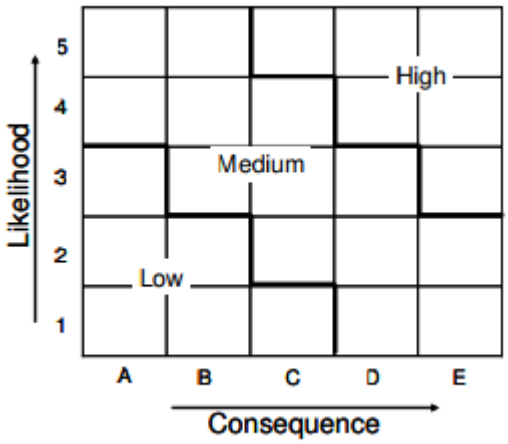


Figure 2: Qualitative risk matrix.

The corrosion risk assessment will have produced a risk ranking for all equipment or any facility. This will enable a strategy for corrosion management to be set down. Table 2 illustrates an example of a risk that might be drawn up for an industrial facility. Accordingly, inspection resources can be planned to allocate greater attention on those areas identified to be in the 'high risk' and 'medium risk' categories (Perumal, 2014).

It should be noted that corrosion prevention, or careful corrosion control, is dictated by high risk classification. By contrast, a low risk classification justifies no corrosion controlling action. A medium risk requires some action. Thus, corrosion management involves a spectrum of activity from no action to considerable action according to the risk. However, taking no action, or taking action, is not corrosion management unless the decision to follow the particular course has been based on an

assessment of risk. Action where it is not needed, like inaction where it is, represents a waste of resources and a tax on profits. Thus, corrosion monitoring is necessary, and it then forms an essential segment of the corrosion management plan. Table 3 lists some monitoring techniques and indicates how they may be used in corrosion management. The key to effective corrosion management is information since it is on the basis of that information that on-going adjustments to corrosion control are made. Information is valid data. Thus, to make effective corrosion management decisions on a day-to-day basis, the monitoring date must be valid. It is important to ensure that the target will involve excessive corrosion control costs, whilst undershooting the target may lead to a situation that cannot economically be recovered.

Table 2: An example of risk and option in corrosion management.

Assessed Risk	*Alternative Corrosion Management Options*
High	Corrosion prevention, or corrosion control for life, or corrosion control to meet planned maintenance or planned replacement
Medium	Corrosion control for life, or planned maintenance
Low	No action, replace if required

Table 3: Corrosion monitoring in corrosion management practices.

Corrosion Control Strategy	*Monitoring Technique*	*Examples of Adjustments and Activities Based on Data Monitoring*
Inhibition of onboard cooling system pipelines	On-line probes (e.g. Coupons, electrical resistance probes)	Adjust inhibitor dosage, change inhibitor type, discontinue imbibition
De-oxygenation of boiler feed-water	O_2 probes	Adjust oxygen scavenger, check pump seals, etc.
Impressed Current Cathodic Protection	Potential	Adjust system output

4. MANAGING CORROSION PROBLEMS

Corrosion threat never stops but its scope and severity can be lessened. To mitigate such a threat, an integrity management system is required. In a sense, methods of corrosion control can be divided into two general types.

i. Methods based on technology or corrosion engineering
ii. Methods based on management or corrosion management

4.1 Corrosion Engineering

Corrosion engineering (CE) is the specialist discipline of applying scientific knowledge, natural laws and physical resources in order to design and implement materials, structures, devices, systems and procedures to manage the materials degradation phenomenon. CE may be defined as the design and application of methods evolved from corrosion science to prevents or minimize corrosion. The main common characteristic of corrosion engineering approaches are using state of the art instruments, analysing figures, evaluating curves and so on without considering human factors. It should begin ideally at the design phase. The corrosion engineering structure can be based on components such as design, material selection and environmental control. At the design phase, corrosion engineering largely depends on the use of the available corrosion data, existing standards and past experience. Proper use of the existing data, along with process and environmental data is the first step in

determining corrosion issues and possibly designing out. Further CE inputs might be necessary post-commissioning and throughout the operation phase because of poor CE input at the design phase or changes in processor operation. Some of the best known methods of corrosion engineering control are material selection, appropriate design, using inhibitors or chemical treatments, cathodic and anodic protection methods, use of coating and linings, also using protective dyes (Cramer & Covino, 2005).

Corrosion problems can be approached from different point of views. Corrosion management (CM) deals with the implementation techniques and methods to control corrosion by keeping the corrosion rate within acceptable limits throughout the operation phase. CM start immediately after post-commissioning, provides early warning signs of impending failures, develops correlations between processes and effects on system and is closely associated with the operation phase (Morshed, 2013; Ghalsasi *et al*., 2016). CM is strongly influenced and affected by both the extent and the quality of the initial CE input during the design phase. The better the quality of the CE input, the more straightforward and simple the CM can be. Corrosion management allows us to use the in-hand resources in a more profitable way, to mitigate corrosion by lowering its threat through modeling, human expertise, strategic planning, education & training, maintenance practices and leadership commitment & policy. Corrosion management is a dynamic approach to control and monitor an asset's technical integrity related material degradation. It is recognised that there are many ways to organise and operate successful corrosion management systems, each of which is asset specific depending on factors such as design, stage in life cycle, process conditions, operational history and visual inspection (Emenike, 1993; Javaherdashti, 2003, 2006a, 2006b).

Corrosion is regarded as a primary threat to the integrity of any asset and platforms, such as radar or any surveillance equipment, aircraft, ground vehicles, battle ships, bridge across the sea, power plants, pipelines and so on. Therefore, corrosion management is required to mitigate and to control economic losses due to corrosion. For any asset, proper and efficient corrosion management is always achieved through an asset corrosion management strategy (CMS). A CMS is defined as "a suite of procedures, strategies, and systems to maintain asset integrity through preventing or mitigating corrosion throughout the asset's operation phase." Any CMS comprises components or stages in the form of a loop, such as developing the strategy, implementing the strategy and learning & improvement as shown in Figure 3. The success of any CMS is reliant upon auditing and measurement of performance. Both activities also contribute feedback ensuring continuous improvement in corrosion management activities. The CMS subjects and benefits of successful CMS are listed in Table 4.

CMS can be improved by using corrosion key performance indicators (KPI) and through regular assessment. Table 5 is a proposed KPI table in an asset monthly corrosion monitoring report where each individual activity along with its corresponding range or threshold, monthly performance and compliance target level are listed. Regular monitoring of their performance will immediately reveal whether an asset CMS has been functioning satisfactorily or not. The benefits of using corrosion KPI in the CMS are listed below (Morshed, 2008):-

i. Corrosion KPIs are an efficient way of capturing, trending, and assessing data related to the most important activities affecting the integrity of the process pressure systems of an asset.
ii. They can help to immediately identify shortcomings or problems during the implementation phase of the asset CMS. This is of great benefit; in particular, to the mature assets suffering from various acute corrosion problems.
iii. They improve the supervision of the responsible corrosion engineer over the most crucial activities (related to the asset integrity) and the individuals who have to regularly carry them out.
iv. help improve motivation among the team as team members constantly endeavor to achieve higher individual and average KPI compliances.
v. Corrosion KPIs are an efficient, quick, and brief way of reporting issues related to asset integrity and asset corrosion management; in particular to the senior management.

Table 4: CMS subjects and benefits.

CMS subjects	*CMS Benefits*
i. **Identification and review of the corrosion threats.**	i. Improved reliability
ii. **Identification of corrosion control measures.**	ii. Less maintenance required and
iii. **Corrosion Risk Assessment / Risk Based Inspections.**	iii. Reduced cost of ownership
iv. **Implementation of corrosion monitoring and inspection, corrosion control measures and their effectiveness.**	
v. **Identification and implementation of corrective actions, repairs, changes.**	
vi. **Auditing and assimilation of lessons learnt from operational experience.**	
vii. **Review of Corrosion Management Process.**	

Table 5: A KPI table within an asset corrosion management monthly report (MIL-A-18001K).

Performance Measured	*System*	*Target Value/Range*	*Compliance*
Ferum (Fe) concentration in Zn anode- impurity	Cathodic protection	0.005%	Max
Copper (Cu) concentration in Zn anode-impurity	Cathodic protection	0.005%	Max
Aluminum (Al) concentration in Zn anode-alloying element	Cathodic protection	0.1 – 0.5	-
Cadmium (Cd) concentration in Zn anode-alloying addition	Cathodic protection	0.025 – 0.07	-

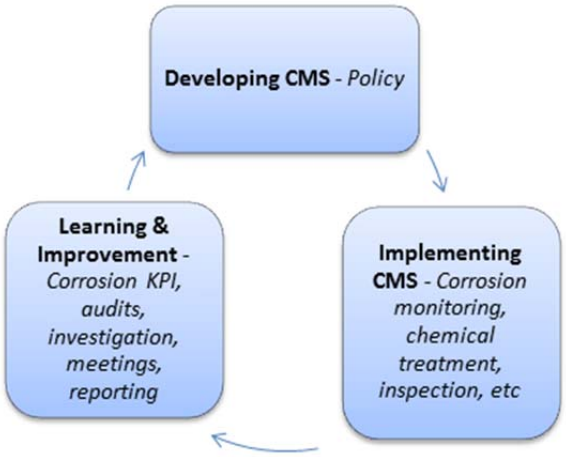


Figure 3: Components of corrosion management strategy.

5. PREVENTIVE STRATEGIES

The current study showed that technological changes have provided many new ways to prevent corrosion and the improved used of available corrosion management techniques. However, better corrosion management can be achieved using preventive strategies in non-technical and technical areas. These preventive strategies include (Koch *et al*., 2002; Payer & Latanision, 2017).

i. Rapidly replace aging assets with new kinds of systems
ii. Increase awareness of significant corrosion costs and potential cost-savings by using effective web-based strategies for communicating and sharing best practices
iii. Change the misconception that nothing can be done about corrosion attack by establishing mechanisms to coordinate and oversee prevention and mitigation plan
iv. Change policies, regulations, standards, and management practices to increase corrosion cost-savings through sound corrosion management
v. Develop clearly defined goals, outcome-oriented objectives, and performance measures that measure progress toward achieving objectives (including return on investment and realized net savings of prevention projects)
vi. Improve education and training of staff in the recognition of corrosion control
vii. Implement advanced design practices for better corrosion management by developing standardized methodologies for collecting and analyzing corrosion related cost, readiness, and safety data.
viii. Develop advanced life prediction and performance assessment methods
ix. Review and update all acquisition-related directives and other documents to reflect policies and requirements concerning corrosion prevention and control.
x. Streamline and standardize application of specification, standards, and qualification processes
xi. Improve corrosion technology through research, development, and implementation

While corrosion management has improved over the past several decades, Malaysia is still far from implementing optimal corrosion control practices. There are significant barriers to both the development of advanced technologies for corrosion control and the implementation of those technological advances. In order to realize the savings from reduces costs of corrosion; changes are required in three areas:

i. The policy and management framework for effective corrosion control
ii. The science and technology of corrosion control, and
iii. The technology transfer and implementation of effective corrosion control

The policy and management framework is crucial because it governs the identification of priorities, the allocation of resources for technology development, and the operation of the system. Incorporating the latest corrosion strategies requires changes in industry management and government policies, as well as advances in science and technology. It is necessary to engage a larger constituency comprised of the primary stakeholders, government and industry leaders, the general public, and consumers. A major challenge involves the dissemination of corrosion awareness and expertise that are currently scattered throughout government and industry organisations. In fact, there is no focal point for the effective development, articulation, and delivery of corrosion cost-savings programs. Therefore, the following recommendations are made:

i. Form a committee on corrosion control and prevention
ii. Develop a national focus on corrosion control and prevention
iii. Improve policies and corrosion management
iv. Accomplish technological advances for corrosion-savings
v. Implement effective corrosion control

By following appropriate strategies and obtaining sufficient resources for corrosion programs, best engineering practices can be achieved. Controlling corrosion requires significant expenditures, but the payoff includes increased public safety, reliable performance, maximised asset life, environmental protection, and more cost-effective operations in the long run.

6. CONCLUSION

Corrosion is a natural phenomenon that cannot be ignored. The consequences of corrosion attack must always be considered. If the consequences are unacceptable, steps must be taken to manage it throughout the facility's life at a level that is acceptable. To manage is not simply to control. Good corrosion management aims to maintain, at a minimum life cycle cost, the levels of corrosion within predetermined acceptable limits. This requires that, where appropriate, corrosion control measures be introduced and their effectiveness ensured by judicious, and not excessive, corrosion monitoring and inspection. Good corrosion management serves to support the general management plan for a facility. Since the later changes as market conditions, for example, change, the corrosion management plan must be responsive to that change. The perceptions of the consequences and risk of a given corrosion failure may change as the management plan changes. Equally, some aspects of the corrosion management strategy may become irrelevant. Changes in the corrosion management plan must, inevitably, follow. Corrosion is everyone's problem and all can contribute to prevention and control.

ACKNOWLEDGEMENT

The author is thankful and grateful for the support of colleagues, technical assistance and reviewers for their instructive comments, inspiring and beneficial discussions.

REFERENCES

Bhaskaran, R., Palaniswamy, N., Rengaswamy, N.S. & Jayachandran, M. (2005). Global cost of corrosion–A historical review. In *ASM Handbook,* Vol. 13B, Corrosion: Materials, pp. 621-628.

Cramer, S.D & Covino Jr. B.S. Eds. (2005). *ASM Handbook, Vol 13A. Corrosion: Fundamental, Testing and Protection.* ASM International, Ohio.

Dehri, I. & Erbil, M. (2000). The effect of relative humidity on the atmospheric corrosion of defective organic coating materials: An EIS study with a new approach. *Corros. Sci.*, **4**: 969-978.

DOS (Department of Statistics Malaysia) (2017). *Gross Domestics Products by States 2016. The Office of the Chief Statistician Malaysia*. Department of Statistics Press Release. Available online at: https://www.dosm.gov.my (Last access date: 27 October 2017).

Edwards, K.L. & Davenport, C. (2006). Materials for rotationally dynamic components: rationale for higher performance rotor-blade design *Mater. & Des.* **27**: 31-35

Emenike, C.O. (1993). The application of knowledge-based systems to corrosion management *Mater. Des.* **14:**331-337

Garverick, L. (1994). *Corrosion in the Petrochemical Industry*. ASM International, Ohio.

Ghalsasi, S., Fultz, B. & Colwell, R. (2016). Primary constituents of a successful corrosion management program. NACE Corrosion Risk Management Conference, paper no. RISK16-8734, NACE Int., Houston, Texas

Gil H., Caldern J.A., Buitrgo CP, Echavarria, A & Echeverria, F. (2010). Indoor atmospheric corrosion of electronic materials in tropical – mounting environments *Corros. Sci.,* **52**: 327-337

Guedes Soares, C., Garbatov, Y., Zayed, A & Wang. G. (2009). Influence of environmental factor to corrosion of ship structures in marine atmosphere. *Corros. Sci,*. **51**: 2014-2026.

Jackson, J.E. (2017). *Cost of Corrosion Annually in the US Over $1.1 Trillion in 2016*. Available online at: http://www.g2mtlabs.com/corrosion/cost-of-corrosion (Last access date: 16 August 2017).

Javaherdashti R. (2000). How corrosion affects industry and life. *Anti-Corros. Meth. Mater.*, **47**: 30-34.

Javaherdashti, R. (2003). Managing corrosion by corrosion management: A guide for industry managers. *Corros. Rev.*, **21**: 311-326.

Javaherdashti, R. (2006a). Corrosion knowledge management: An example of an interface for more efficient managerial communication. *Mater. Corros.*, **57**: 945-950.

Javaherdashti, R. (2006b). Using corrosion management to protect the environment. *Mater. Perform.* **45:** 52-54

Koch, G.H., Brongers, M.P.H., Thompson, N.G., Virmani, Y.P. & Payer, J.H. (2002). *Corrosion Costs and Preventive Strategies in the United States*. U.S. Federal Highway Administration Report FHWA-RD-02-256, March 2002.

Koch, G.H., Thompson, N.G., Moghissi, O., Payer J.H. & Varney, J. (2016). *IMPACT (International Measures of Prevention, Application, and Economics of Corrosion Technologies Study*. Report No. OAPUS310GKOCH (AP110272), NACE Int. 2016, Houston, Texas.

Mattsson, E. (1989). *Basic Corrosion Technology for Scientists and Engineers.* Ellis Horwood, West Sussex, England

Morshed, A. (2008). Improving asset corrosion management using KPIs. *Mater. Perform.*, **47**: 50-54.

Morshed, A. (2013). The evolution of the corrosion management concept. *Mater. Perform.*, **52**:66.

Mukhriz, M. (2010). *Speech: Officiating the Opening of A & E Global Headquaters on 25 Jan 2010.* Available online at: http://www.mukhriz.com/speech (Last access date: 5 February 2010).

Pasman, H.J., Jug, S., Prem, K. Rogers, W.J. & Yang, X. (2009). Is risk analysis a useful tool for improving process safety? *J. Loss Prev. Proc. Ind.*, **22**: 760-777.

Patil, C.V. & Ghanendra, K.B. (2013). An investigation into the environmental impacts of atmospheric corrosion of building materials. *Int. J. Chem. Sci. Appl.*, **4**: 1-6.

Payer, J.H. & Latanision, R. (2017). *Preventive Strategies*. Available online: https://plastibond.com/downloads/ preventative_strategies_CC_06.pdf (Last access date: 23 August 2017).

Perumal, K.E. (2014). Corrosion risk analysis, risk based inspection and a case study concerning a condensate pipeline. *Proc. Eng*. **86:** 597-605.

Restrepo, C.E., Simonoff, J.S. & Zimmerman, R. (2009). Causes, cost consequences and risk implications of accident in US hazardous liquid pipeline infrastructure *Int. J. Crit. Infrastruc. Prot.*, **2**: 38-50.

Simonoff, J.S., Restrepo, C.E. & Zimmerman. R. (2010), Risk management of cost consequences in natural gas transmission and distribution infrastructures *J. Loss Prev. Proc. Ind.* **23**: 269-279.

Wood, R.J.K. (1999). The sand erosion performance of coating. *Mater. Des.*, **20**: 179-191.

MECHANICAL PROPERTIES EXTRACTION OF COMPOSITE MATERIAL USING DIGITAL IMAGE CORRELATION VIA OPEN SOURCE *NCorr*

Ahmad Fuad Ab Ghani[1,2], Jamaluddin Mahmud[2], Saiful Nazran[3] & Norsalim Muhammad[4]

[1]Faculty of Engineering Technology (Mechanical), Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
[2]Faculty of Mechanical Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia
[3]CTRM Aero Composites Sdn. Bhd, Malaysia
[4]Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
[*]Email: ahmadfuad@utem.edu.my

ABSTRACT

This paper describes and provides a comprehensive overview of the use of digital image correlation technique via open source platform Ncorr on composite materials that is widely used in aerospace and defence industries. Deformation displacement, in plane strain xx, in plane strain yy and in plane shear strain xy are extracted from digital image correlation technique using high speed camera that captures during experiment. Data will be used to determine the composite behaviour (properties and parameters for Finite Element Modelling (FEM) and analytical modelling). Tests are conducted on samples in accordance to ASTM standard described later in this section to obtain mechanical properties of composite materials under loading set up of tensile, shear, and flexural. Digital Image Correlation (DIC) is found to be a reliable, consistent and affordable (low cost) non-contact deformation measurement technique which can assist in extracting mechanical properties of composite materials. The use of DIC is proven to be a practical Non Destructive Evaluation (NDE) technique for composite material characterisation as well as Non Destructive Technique (NDT) for structural health monitoring.

Keywords: *Digital Image Correlation (DIC); composite material; Ncorr; non-contact strain; open source DIC.*

1. INTRODUCTION

Nondestructive testing (NDT) is a technique in inspection of components/assemblies for homogenous, defects detection, voids, discontinuities, or differences in characteristics without destroying the physical attributes of the parts under study. In other words, when the inspection or test is completed the part can still be used. Current techniques applied in various industries for NDT inspection includes; Acoustic Emission, Electromagnetic, Guided Wave (GW), Laser Testing, Leak Test, Magnetic Flux Leakage, Microwave, Liquid Penetrant, Magnetic Particle, Radiographic Testing, Thermal Infrared Testing, Ultrasonic Testing, Vibration Analysis and Visual Testing. Apart from the use of NDT technique, DIC can be used to detect non homogenise characterisation and material characterisation or Non Destructive Evaluation (NDE) of materials. The aim of this paper is to enhance understanding of the application of DIC in evaluating mechanical properties of composite materials which is used in aerospace industries that includes Longitudinal Young's Modulus, E_{11}, Transverse Young's Modulus, E_{22}, Shear Modulus, G_{12}, ductility, yield, ultimate tensile strength, etc. The major contribution of FRP composite can be seen in the designs of high performance and light weight solutions in the aerospace and defence industries (Gay & Hua, 2007). The high strength weight ratio of the FRP materials may be customised in order to design optimal structures compared to traditional structures which made using metal alloys. The use of reliable design and prediction methods will ensure superior performance of composite. Composite is a combination of two or more material that differs in properties and composition to form unique properties. Normally, composite provides an increase of the durability or strength over many other materials and at the same time it

may provide additional benefits such as resistance corrosion (Harris, 1987). The stress strain relationship is an essential principle for mechanics of composite study. Composite materials include some of the most advanced engineering materials today. The addition of high strength fibres to a polymer matrix can greatly improve mechanical properties such as ultimate tensile strength, flexural modulus, and temperature resistance. In order to extract the mechanical properties from a composite material, several testing have to be made. ASTM's standard is the reference for determination of the physical, shear, tensile, flexural, and compressive properties of various forms of composite materials used in structural applications. This paper presents the use of an innovative method, DIC as a tool in measuring displacement and strain which are essential in characterising and determining mechanical properties of composite material.

DIC is used to measure the deformation/strain of a specimen under tensile loading. DIC tracks the position of the same physical points shown in a reference image and the deformed image. To achieve this, a square subset of pixels is identified on the speckle pattern around point of interest on a reference image and their corresponding location determined on the deformed image. The combination method of strain gauges and DIC allows in the enhancement of the identification for mechanical properties of composite in testing and contributes to a deeper understanding of deformation and towards the development of optimised systems (Tekieli *et al*., 2016). In another research (Siddiqui *et al*., 2011), incorporating experimental works such as computation of longitudinal and measurement of lateral strains in uniaxial test utilizing DIC as full field strain measurement tool. Strain gauge is normally limited with unsuitable material surface or small size samples where strain gauge mounting is not practical. Costs involved in using strain gauges are quite high. Digital Image Correlation tool has been used to calculate the strains and as well as Poisson's ratio in various selection of metal and composite specimens. The strains computed using DIC method were then compared with strain gauges and extensometer, as shown in Figure 1 for validation of strain measurement. DIC has proved to be inexpensive and consistent technique for strain measurement as well as Poisson's ratio of metallic (homogeneous) and composite (heterogeneous) materials.

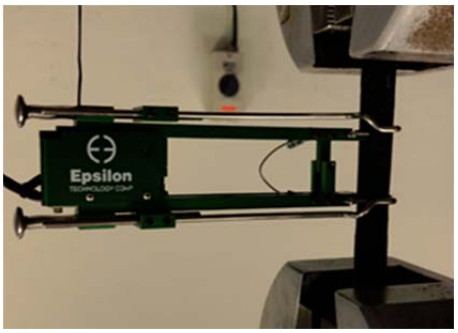


Figure 1: Conventional method of strain measurement using extensometer.

In other research, tests were conducted using UTM, and the load was applied under displacement control mode until tensile rupture of the coupon (Tekieli *et al*.,2016). Textiles of 600mm total length was clamped in the wedges of the testing machine with aluminium tabs to ensure a uniform load distribution and avoiding local damage to the filaments. Digital Image Correlation was used to compute the average strain over the gauge length of the coupon. Both CivEng Vision and *Ncorr* software programs (Blaber *et al*., 2015; Reu *et al*., 2015) were used to process the digital images taken at 5 seconds time interval with the aim of computed the displacements. For surface deformation computation utilizing 2D Digital Image Correlation (DIC) technique, emphasised should be given on positioning of specimen under testing, light intensity and sources as well as camera lens and its capability/resolution/frame rate of camera (Blaber *et al*., 2015). Accurate measurement relies heavily on imaging system configuration. In principle, sample with random speckle pattern sprayed on the surface must be positioned perpendicular to the camera to avoid any out of plane motion. After the entire load applied events, a series of images are taken before and after loading and deformation and finally stored in the computer for post processing images to obtain displacement contour/field using DIC algorithm as shown in Figure 2. Basically from technical perspectives, for 2D DIC, image

resolution plays a vital role in measurement accuracy (Blaber *et al*., 2015). Figure 2 depicts the working principle of DIC which captures images with digital camera during the deformation process to evaluate the changes in surface characteristics and understand the behaviour of the specimen while it is subjected to incremental loads. This technique starts with a reference image (before loading) and followed by a series of pictures taken during the deformation. Deformed images show a different dot pattern relative to the initial non deformed reference image. These patterns difference can be calculated by performing correlation of the pixels of the reference image and any deformed image and a full-field displacement measurement can be computed. The strain distribution can then be obtained by applying the derivatives in the displacement field. To apply this method, the object under study needs to be prepared with random dot pattern speckle pattern to its surface (Reu *et al*., 2015).

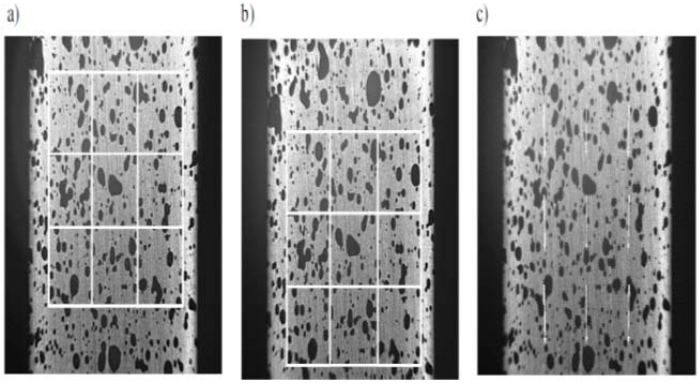


Figure 2: Computation of the displacement vectors using the digital image correlation:a) reference image; b) deformed subset/image; c) displacement field/contour (Blaber *et al*., 2015) .

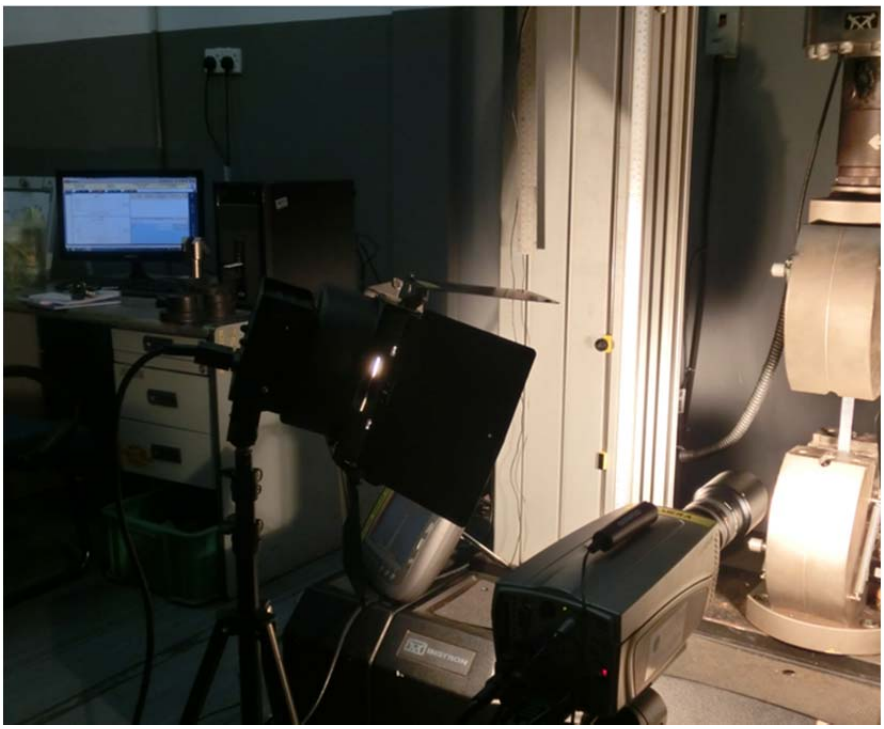


Figure 3: Digital Image Correlation measuring in plane strain.

Olympus I-Speed 2 camera was used to capture images for tensile and bending test. DIC system uses optic method through stereoscopic sensor arrangement to analyse the deformation of object under study. It emphasizes on each point subset based on grey value of digital image captured from specimen under study to compute the strain (Siddiqui *et al*., 2011). The camera is positioned perpendicular to the specimen under testing. In order for the digital image correlation algorithm able to perform the correlation analysis, speckle pattern must be sprayed onto the surface of the coupon as

shown in Figure 3. The pattern must be contrast and small enough to capture the deformation as displayed in Figure 4. The technical specification of the high speed camera are of frequency 60 – 200,000 fps; Shutter minimum of 1 μm, Nikkor 18-55mm lens, an open source software for DIC which is *Ncorr* platform, with installed Matlab version of 2012 and Microsoft Visual C++ as compiler.

Figure 4: View of reference image and current image sprayed with speckle pattern as seen in *Ncorr* platform.

The processing of image in *Ncorr* started with seeding the region where deformation/strain is to be measured as shown in Figure 5.

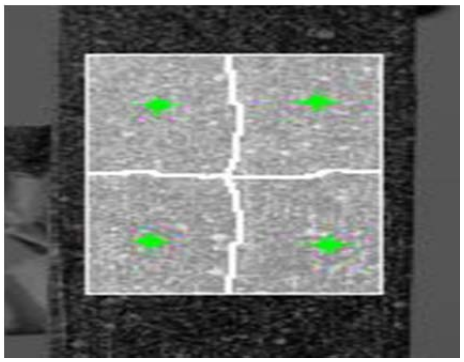


Figure 5: Seeding of region for deformation measurement in *Ncorr* platform.

2. METHODOLOGY

2.1 Tensile Test

Tensile properties such as lamina's Young's Modulus E_{11}(longitudinal) and E_{22}(transverse), Poisson's ratio and lamina longitudinal and transverse tensile strength are measured by static tension testing along 0^o and 90^o with principal direction (fiber direction) according to the ASTM D3039 standard test method (ASTM D3039,2008). The test specimen is given in the figure below. Measurement of E_{11}, E_{22}, v_{12}, and v_{21} and Tensile Strength can be obtained from the uniaxial tensile test using the DIC method.

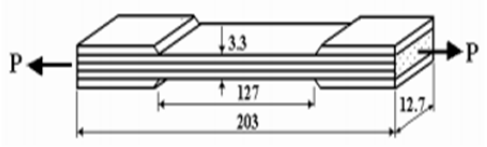


Figure 6: Dimension for 0°(longitudinal) unidirectional ASTM D3039.

The tensile specimen is placed in a testing machine aligning the longitudinal axis of the specimen and pulled at a crosshead speed of 0.5 mm/min. The specimens are loaded step by step till they fail under uniaxial loading. The load and deflection are recorded using the digital data acquisition system. The axial and the transverse strains are obtained by a pair of two strain gauges rosettes, which are attached to the gauge section of the specimen. The stress strain behavior is obtained to be linear and the final failure occurs catastrophically. The values of Young's Modulus, Poisons Ratio and Axial Strength are obtained as follows:

$$\sigma_1 = \frac{P}{A} \quad \nu_{12} = -\frac{\varepsilon_1}{\varepsilon_2} \quad E_1 = \frac{\sigma_1}{\varepsilon_1} \quad X_t = \frac{P_{ult}}{A} \qquad (1)$$

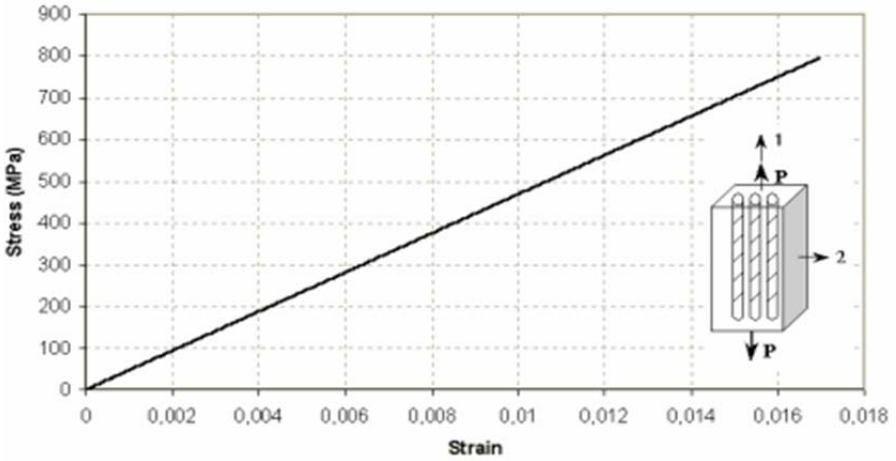


Figure 7: Stress strain plot for 0 °(Longitudinal) unidirectional composite (Nettles, 1994).

The transverse Young's modulus, minor Poisson's ratio and transverse tensile strength are calculated from the tension test data of 90° unidirectional lamina. The tension test is manufactured based on the ASTM 3039 standards and the specimen dimensions are given in the figure below.

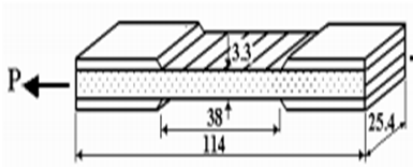


Figure 8: Dimension for 90° (transverse) unidirectional ASTM D3039.

The specimen is loaded gradually until rupture and an indicator measures the strains. The values of Young's modulus, Poisson's ratio and transverse strength are given as follows:

$$\sigma_2 = \frac{P}{A} \quad \nu_{21} = \frac{\nu_{12} E_2}{E_1} \quad E_2 = \frac{\sigma_2}{\varepsilon_2} \quad Y_t = \frac{P_{ult}}{A} \qquad (2)$$

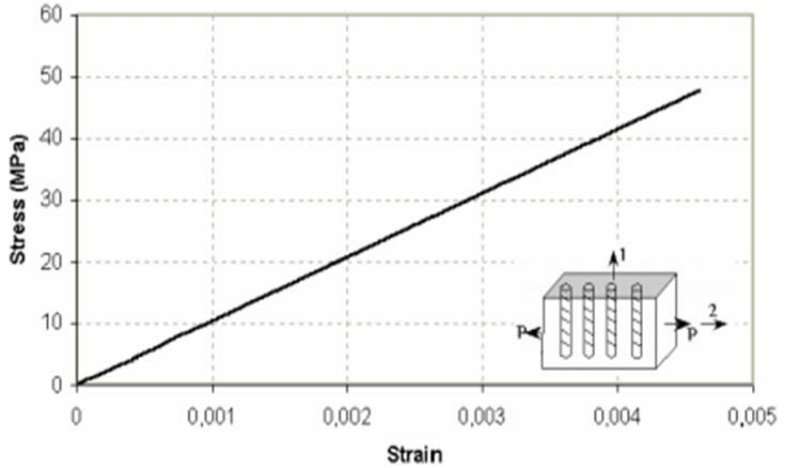


Figure 9: Stress strain plot for 90° (Transverse) unidirectional composite (Nettles, 1994).

2.2 Off Axis Shear Test

In orthotropic material, the shear modulus is a private residence which must be mechanically measured for each different material. The normal procedure for doing so is to create specimen geometry and loading systems that produce pure shear conditions with respect to the direction of the main ingredients. ASTM D 3039-76 is a standard test method for off axis tests on composite materials in general (Xavier *et al*., 2004). The unidirectional 10^0 specimens are 2.0 mm thick, 20 mm wide and 175 mm long. During the shear tests, the images of the specimen surface were recorded in the video mode then converted to image (Khoo *et al*., 2016). The image selected based on the time lap by referring the load on the UTM machine. The stresses in a ply with fibres oriented at an angle θ from the load direction as a function of the applied stress σ_{xx} are given by the following well known transformation equations which are easily derivable from force equilibrium considerations:

$$\boldsymbol{\sigma_{11} = \sigma_{yy}\cos^2\theta} \quad (3)$$

$$\boldsymbol{\sigma_{22} = \sigma_{yy}\sin^2\theta} \quad (4)$$

$$\boldsymbol{\sigma_{12} = \frac{1}{2}\sigma_{yy}\sin 2\theta} \quad (5)$$

For the 10° off-axis specimen, substituting 10° for θ in Equations 3 to 5 yields the following to three decimal figures:

$$\boldsymbol{\sigma_{11} = 0.970\sigma_{yy}} \quad (6)$$

$$\boldsymbol{\sigma_{22} = 0.030\sigma_{yy}} \quad (7)$$

$$\boldsymbol{\sigma_{12} = 0.171\sigma_{yy}} \quad (8)$$

$$\boldsymbol{\varepsilon_{l12} = (\varepsilon_{yy} - \varepsilon_{xx})(0.342) + \varepsilon_{xy}(0.9396)} \quad (9)$$

2.3 Iosipescu Shear Test

Another method to characterize shear properties of composite material, such as extracting the G_{12}, shear modulus and shear strength, is by performing Iosipescu shear test. Iosipescu test specimen is tested using the Iosipescu test fixture (Selmy *et al*., 2015). The specimen is instrumented in a test section between the notches at 45^o. The specimens are placed in Iosipescu test fixture in which the specimen is centered using the alignment pin and lightly clamped with the adjustable wedges as shown in Figure 10.

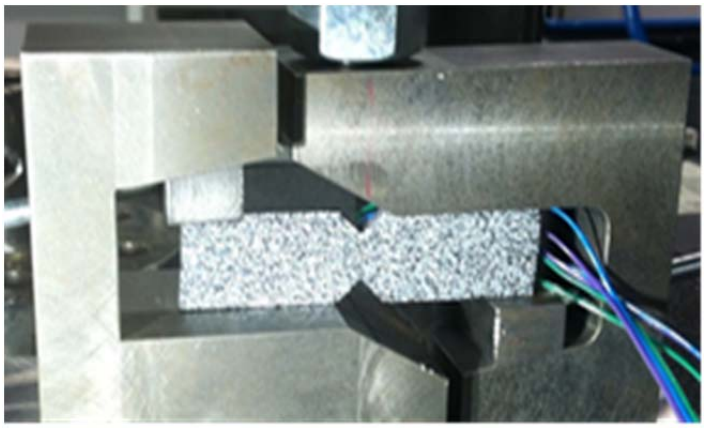


Figure 10: Iosipescu test rig and specimen V notch geometry.

The tests were performed on a servo-hydraulic with manual grips and a displacement rate of 0.5 mm/min. Load and strain data were taken up to a displacement of about 3.0 mm. Shear strain ε_{xy} contour of Iosipescu specimen region taken at central in between notches(upper and lower) as shown in Figure 11. In plane shear modulus is obtained by initial slope of shear stress-shear strain curve. Shear strain is evaluated from the normal strain obtained from DIC**.**

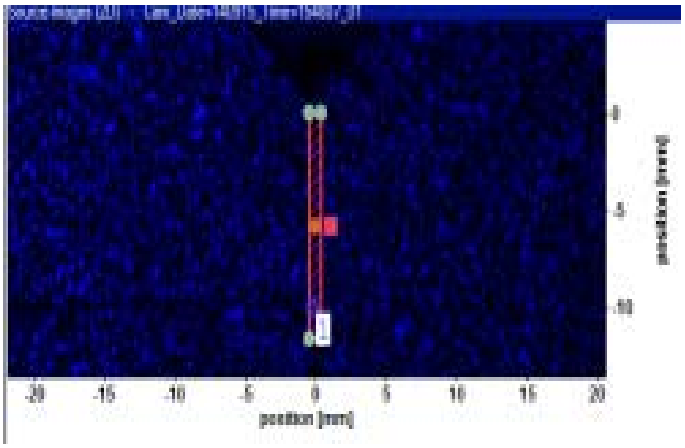


Figure 11: Section of shear strain field is computed.

The shear strain can be determined from the measured normal strain that is located at the centre of the notched section at 45° with the loading direction (Odegard & Kumosa, 2012). As the shear modulus or ultimate strain was to be calculated, it is required to determine the shear strain from the indicated normal strains at +45° and −45° at each required data point .The shear strain was calculated from the DIC reading by the relationship:

$$\boldsymbol{\gamma_i = |\varepsilon_{+45}| + |\varepsilon_{-45}|} \qquad \textbf{(10)}$$

The area of shear loading taking place was calculated as;

$$\mathbf{A = w\ x\ h}$$

Cross sectional area for the specimen, A, in units of mm^2 was recorded for each GFRP and CFRP specimens. A standard head displacement rate of 0.5 mm/min. Calculating the ultimate strength and determining the shear stress at each required data point can be performed using;

$$\boldsymbol{F^u = \frac{P^u}{A}} \qquad \textbf{(11)}$$

$$\boldsymbol{\tau_i = \frac{P_i}{A}} \qquad \textbf{(12)}$$

where:

F^u = ultimate strength, MPa

P^u = the lower of ultimate or force at 5 % engineering shear strain, N;

τ_i = shear stress at i^{th} data point, MPa;

P_i = force at i^{th} data point, N; and

A= cross-sectional area, mm^2

Shear strain ε_{xy} contour of Iosipescu specimen region taken at central in between notches (upper and lower). Thus the shear strain data can be generated and the corresponding modulus and strength can be found from the resulting stress strain curve. The below diagram, Figure 12 shows the stress strain curve of the notched specimen under static in-plane shear loading.

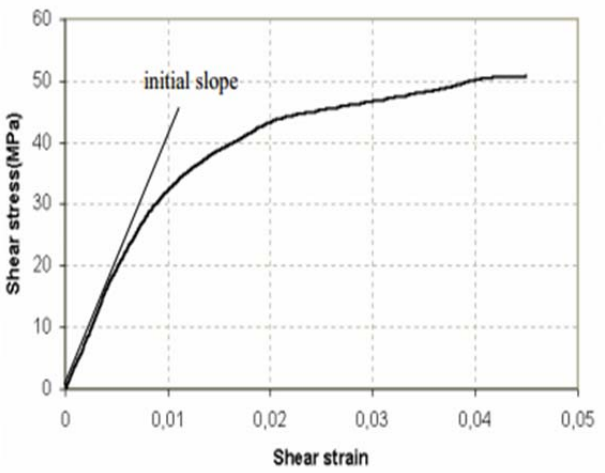


Figure 12: Shear stress against shear strain plot for composite from Iosipescu Shear Test (Nettles, A.T, 1994)

2.4 Bending Test

Another essential mechanical properties for composite that needed for better understanding and simulation input for finite element modelling and design optimisation is bending stiffness and flexural modulus. Three point bending test has been performed as accordance to ASTM D7264 (ASTM D7264, 2007). This test method designed to determine the flexural stiffness and strength properties of polymer matrix composites as shown in Figure 13 and Figure 14.

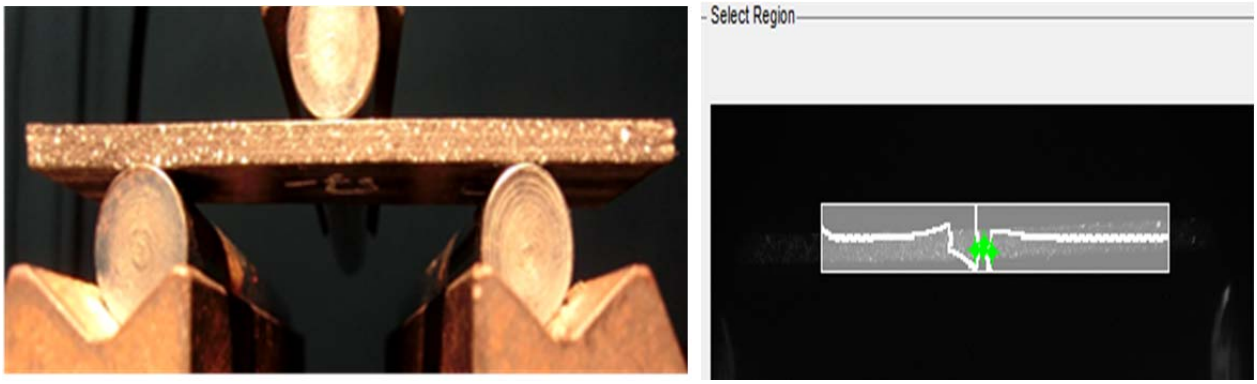


Figure 13: Three point bending at 0N load Hybrid Composite CFRP/GFRP and region of interest for bending deflection computation.

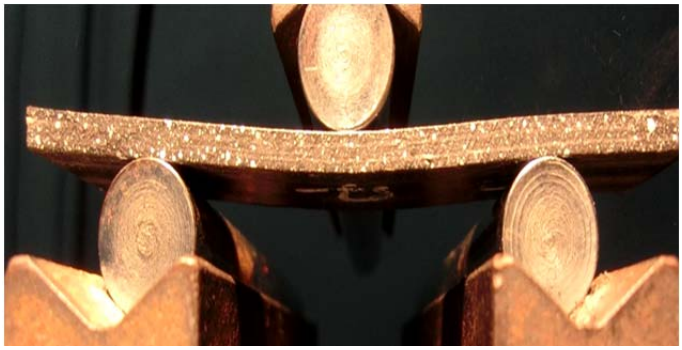


Figure 14: Three point bending at 1,200 N load on Hybrid Composite CFRP/GFRP

3. RESULTS AND DISCUSSION

The contour of displacement vector together with strain field in yy (longitudinal), xx (transverse), xy(shear) were assessed in order to compute the mechanical properties of the material. The in plane (2 Dimensional) displacement and strain field was also studied to ensure perfect tensile strain field, shear strain field and bending behaviour experienced by the sample in accordance to the type of loading imposed to the sample. The strain output processed by open source platform, *Ncorr* is depicted and discussed in this section.

3.1 Tensile Test

The value of strain in yy direction, ε_{yy} which correspond to strain in longitudinal direction and strain xx, ε_{xx} which correspond to strain in transverse direction can be obtained from DIC *Ncorr* processing strain contour as shown in Figure 15 and Figure 16.

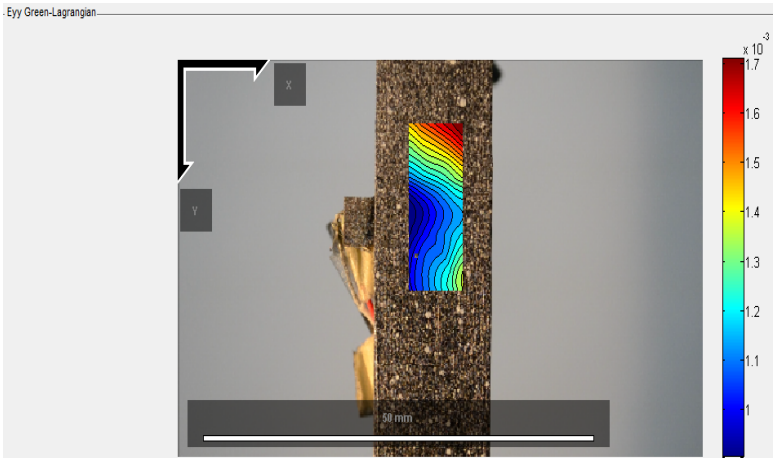


Figure 15: **Contour of strain yy (longitudinal), Ɛyy CFRP 0° at 5,000 N.**

Figure 15 depicts 0° CFRP specimen under tensile loading of 5000N where the contour of strain in yy direction, ε_{yy} shows the minimum value of 0.0009, the median around 0.0012 and the maximum at 0.0017 respectively. The value of strain, ε_{yy} increases with respect to increment of value of loading. From perspective of characterizing the material under study and computing the modulus of elasticity, E, the relation of Equation 1 and Equation 2 are used, where strain value is obtained from the average value computed seen in contour field. A minimum of eight points of average strain yy, ε_{yy} together with correspond value of stress yy, σ_{yy} are required in order to plot stress against strain graph for mechanical property, E, modulus of elasticity computation.

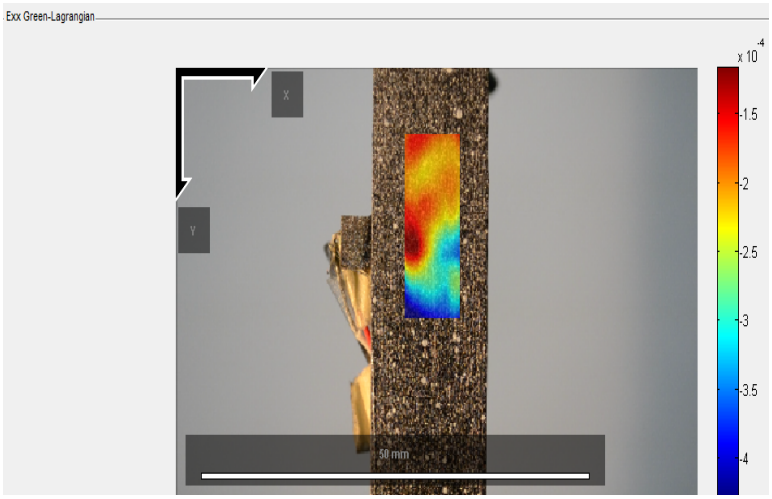


Figure 16: Contour of strain xx (transverse), Ɛxx CFRP 0° at 5,000 N.

Meanwhile, Figure 16 shows contour of strain xx (transverse direction) for 0° where it recorded the range of reduction in strain (negative value) of -0.0005: -0.0002 : -0.0001. It can be concluded that the effect of Poisson's taking place during tensile test on the composite specimen. The average value of transverse strain, ε_{xx} computed from strain xx contour in *Ncorr* platform to be used in Equation 1 and Equation 2 in determining value of Poisson's ratio.

3.2 Off Axis Shear Test

The shear strain, ε_{xy} field and contour of specimen under off axis shear loading was assessed to ensure perfect shear field experienced at 10^{o} plane with respect to principal direction. Figure 17 depicts the shear strain, ε_{xy} contour of 10^{o} off axis CFRP at corresponding load of 3,000 N. The computation of Shear Modulus, G_{12}, is performed by using relation described in Equations 3 to 9, where it is observed that all type of strain, ε_{yy}, ε_{xx} and ε_{xy} notation exists in the Equations 9 which requires the computation of average strain for each direction from *Ncorr* strain contour display.

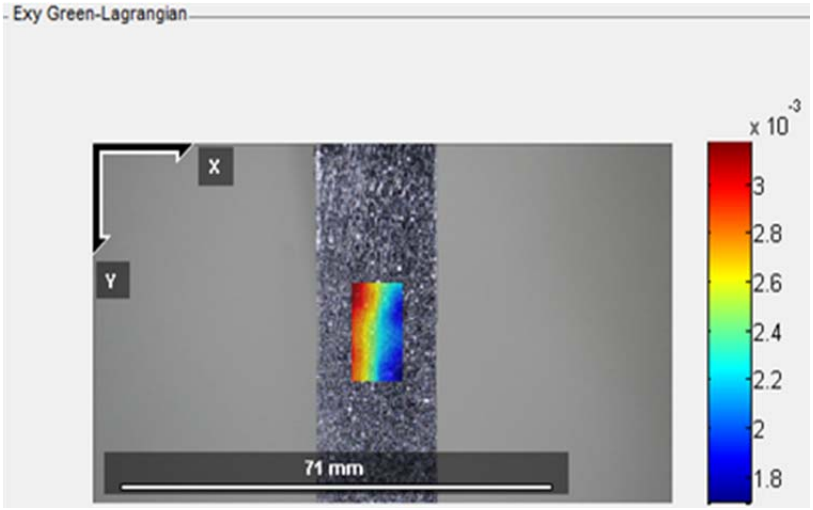


Figure 17: **Shear strain, ε_{xy} contour of 10^{o} off axis CFRP at corresponding load of 3,000 N.**

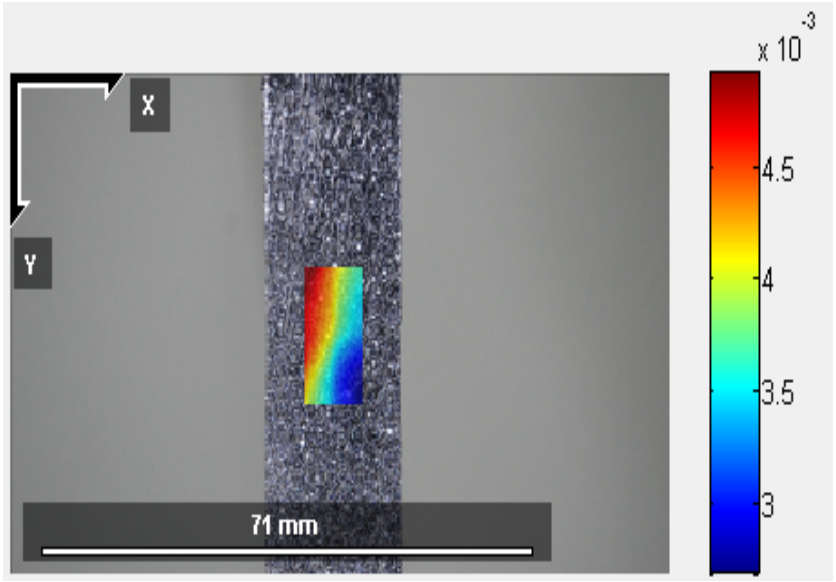


Figure 18: **Shear strain contour of 10^{o} off axis CFRP at corresponding load of 5,000 N.**

3.3 Bending Test

Figure 19 shows the post processing of deformation of hybrid composite CFRP/GFRP under bending using *Ncorr* platform. The deformation contour under interest in this research for the case of bending is deflection, which represented as notation V in *Ncorr* software. Figure 20 displays the deflection/displacement contour in loading/bending direction which is the variable under study.

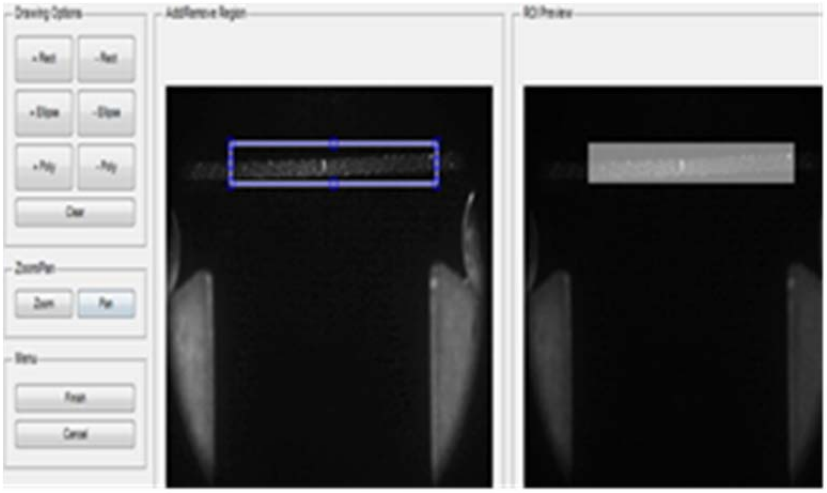


Figure 19: Region of interest selected for deformation field under study

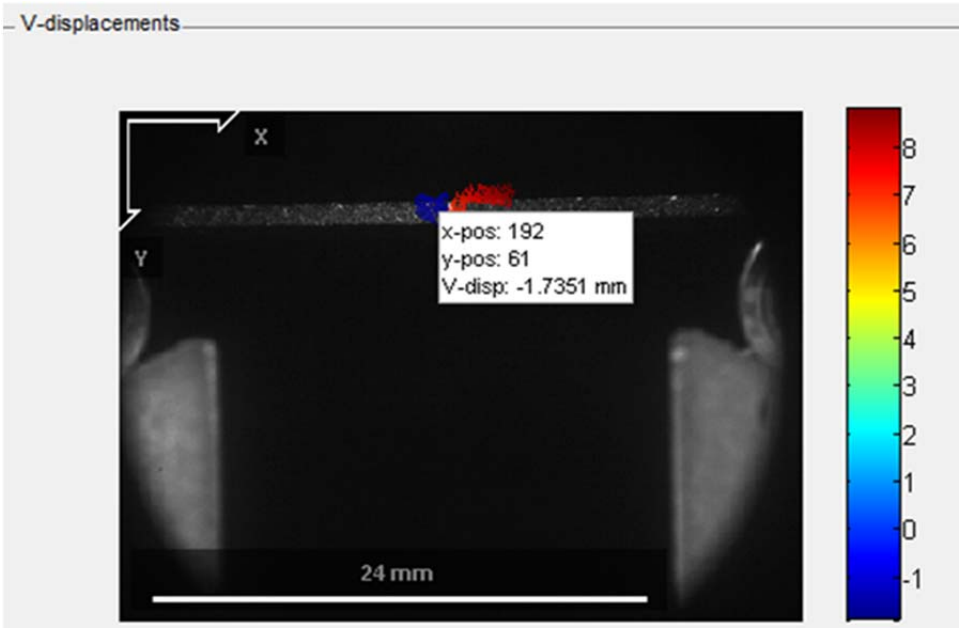


Figure 20: Displacement contour in V which parallel in Y axis direction.

Utilizing the equation for computing Flexural Modulus of composite, for hybrid composite CFRP/GFRP, Equation 13 is as follow:

$$\mathbf{E_f} = \frac{\mathbf{L^3\, m}}{\mathbf{4bd^3}} \qquad \mathbf{(13)}$$

E_f = flexural Modulus of elasticity,(MPa)
m = the gradient (i.e., slope) of the initial straight-line portion of the load deflection
L = support span, (mm)
b = width of test beam, (mm)
d = depth or thickness of tested beam, (mm)

4. CONCLUSION

This paper has successfully demonstrated the development of DIC technique as strain measurement method for composite material characterisation and NDE technique used in the aerospace and defence industries. The tool is also capable of measuring deflection/displacement of composite material under bending which could bring to significant studies of bending behaviour of composite material. Several experimental works had been carried out successfully in the aim to extract the mechanical properties for input into analytical and numerical modelling. The technique of DIC via *Ncorr* open source platform able to offer low cost NDE and NDT technique for aerospace and defence industries.

ACKNOWLEDGMENTS

Author would like to thank the Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka (UTeM), which provided composite characterisation technique testing using high speed camera, Olympus I-Speed 2. Gratitude also goes to the Faculty of Mechanical Engineering, Universiti Teknologi MARA (UiTM), Shah Alam for technical expertise on experimental mechanics and composite engineering.

REFERENCES

ASTM D3039-08. *Standard Test Tensile Properties of Polymer Matrix Composite Materials.* ASTM International, West Conshohocken, Pennsylvania.

ASTM D7264/D7264M –07. *Standard Test Method for Flexural Properties of Polymer Matrix Composite Materials*, 2007. ASTM International, West Conshohocken, Pennsylvania.

Blaber, J., Adair, B., & Antoniou, A. (2015). Ncorr: Open-source 2D digital image correlationMatlab software. *Exp. Mech..*, **55**: 1105–1122.

Gay, D. & Hoa, S.V. (2007). *Composite Materials, Design and Application, 2nd Ed.* CRC Press, Boca Raton, Florida.

Harris, B. (1987). Engineering composite materials. *Composites.* **18**: 261.

Khoo, S.W., Saravanan, K. & Ching,S.T. (2016). A review of surface deformation and strain measurement using two dimensional digital image correlation. *Metrol. Meas. Syst.*, **23**: 537–547.

Nettles, A.T (1994), *Basic Mechanics of Laminated Composite Plate*. NASA Reference Publication 1351, Washington D.C.

Odegard, G. & Kumosa, M. (2000). Determination of shear strength of unidirectional composite materials with the Iosipescu and 10° off-axis shear tests. *J. Composites Sci. Tech.*, **60**: 2917-2943.

Reu, P.L., Sweatt, W., Miller, T. & Fleming, D. (2015). Camera system resolution and its influence on digital image correlation. *Exp Mech.*, **55**: 9–25.

Selmy, A.I., Elsesi, A.R., Azab, N.A. & Abd El-Baky, M.A. (2012). Interlaminar shear behavior of unidirectional glass fiber (U)/random glass fiber (R)/epoxy hybrid and non-hybrid composite laminates. *Composites Part B: Eng.*, **43**:1714–1719.

Siddiqui, M. Z., Tariq, F., & Naz, N. (2011). Application of a two step digital image correlation algorithm in determining Poisson's ratio of metals and composites. *International Astronautical Congress*, Cape Town.

Tekieli, M., De Santis, S., de Felice, G., Kwiecień, A., & Roscini, F. (2016). Application of Digital Image Correlation to composite reinforcements testing. *Composite Structures.* **160**:670-688

Xavier, J.C., Garrido, N.M., Oliveira, M., Morais, J.L., Camanho, P.P., & Pierron, F. (2004). A comparison between the Iosipescu and off-axis shear test methods for the characterization of Pinus Pinaster Ait. *Composites Part A: Appl. Sci. Manuf.*, **35**: 827–840.

EFFECTS OF MAPPING ON THE PREDICTED CRASH RESPONSE OF CIRCULAR CUP-SHAPE PART

Rosmia Mohd Amman[1], Sivakumar Dhar Malingam[1*], Ismail Abu-Shah[2] & Mohd Faizal Halim[2]

[1]Faculty of Mechanical Engineering
[2]Faculty of Engineering Technology
Universiti Teknikal Malaysia Melaka (UTeM), Malaysia

[*]Email: sivakumard@utem.edu.my

ABSTRACT

Performing reliable prediction of crashworthiness is important in designing vehicles for the safety of passengers since the occupant safety is the ultimate goal in crashworthiness design. There are many factors that play a significant role in affecting the reliability of the crashworthiness models. One of the factors is the mapping of forming results into the crash models. Few studies have analysed the mapping effects on the crashworthiness of draw formed components. This paper presents an analysis of the crash response of draw formed circular cup from both experiments and finite element simulations. The predicted crash response for circular cup mapped with dissimilar geometry and mesh will be first shown. Predicted load-displacement and deformed shape will be compared to the measured ones. Thereafter, forming results were mapped on a secondary model, having similar geometry and mesh for crash simulations. For both analyses, the mapping process is performed using result mapper tools available in Hypercrash by including the preceding results in the form of sta file. Results revealed that it is important to include forming results in crash models by mapping preceding results on similar geometry and mesh instead on dissimilar geometry and nominal mesh for better crashworthiness predictions.

Keywords: *Crashworthiness; finite element simulation; mapping effects; preceding results.*

1. INTRODUCTION

The development of advanced high strength steel (AHSS) material, especially dual phase (DP) steel has become greater interest in the steel manufacturing industry for vehicle body panels and structures since they can meet requirements for improving vehicle safety, reducing weight and fuel consumption.

Stamping process in producing vehicle parts is also developing. According to Logue et al. (2007), stamped parts experienced greater strain which increases strength due to strain hardening. Crash safety is an important issue in the vehicle industry and therefore much attention is paid to the crash behaviour of vehicles and components. However, it is very costly to study the crash event experimentally since it requires a lot of materials and many sensors for recording huge data of impact loading. Therefore, since the 1980s, crash study by using numerical simulation has been intensively applied with an aim of reducing both time and money. The crashworthiness evaluation can be performed by a combination of experimental tests and finite element (FE) simulations. Indeed, through FE simulation the designer could study the response of the structural components when subjected to dynamic crash loads, predicts the global response to impact, estimate the probability of injury and evaluate numerous crash scenarios without conducting full crash testing (Obradovic *et al*., 2012). However, in order to reduce production cost and time, the effects of residual-forming properties on the crash performance of vehicle body structure are commonly neglected or rarely taken into account when performing numerical crash simulations.

Crashworthiness simulation can be a useful tool in vehicle design and there are many factors that play a significant role in affecting the reliability of the crashworthiness models. The reliability of FE analyses depends on the accuracy of input material parameters. Without taking into account preceding plastic deformation behaviour, crashworthiness study can lead to inaccurate predictions of load-carrying capacity and absorbed energy of structural component. Several studies do highlight the important of this issues as they studied manufacturing effects such as hydroforming (Dutton *et al*., 1999; Williams *et al*., 2010), tube bending (Oliveira *et al*., 2006) and deep drawing or draw forming (Kose & Rietman 2003; Doğan 2009; Mohd Amman *et al*., 2016; Amman *et al*., 2016) on subsequent crash process. All literature found demonstrated the significant effects of preceding manufacturing process on subsequent crash analysis.

Advance technologies have introduced a mapping tool which has successfully proven as an effective and time reducing method used to include forming results onto the subsequent process such as assembly or crash (Cowell *et al*., 2000; Nie *et al*., 2004; Sasek *et al*., 2010). Mapping method is one of the known methods that are commonly used to analyse the effect of forming to crash simulation either for industrial or research problem. Takashina *et al*. (2009) studied the influence of residual strain, work hardening and material thickness changes resulting from stamping process on the crashworthiness simulation at various impact load cases. They found that due to work hardening effect from stamping process, deformation is reduced to a similar level to actual experimental results in almost all impact load cases. Sasek *et al*. (2010) investigated the effects of the manufacturing process (stamping – welding – spring-back) on crash simulation of a simple box-beam. They used the technology of mapping to take the initial stress and strain from previous stamping simulation to be used in the next simulation. They concluded that pre-simulation can strongly affect the buckling resistance of the box-beam by changing the deformation mode and the internal energy absorbed by the structure. Dhanajkar *et al*. (2011) extracted pre-stresses from forming and then mapping it on the crash meshed model to carry out 35 mph frontal impact test. Their finding gives an evidence that the deformation modes changed due to the inclusion of pre-stresses which further improved the predictability of crash model. To examine the effects of the tube-bending process on subsequent crashworthiness, Oliveira *et al*. (2006) also employed mapping approach in order to transfer the deformation history, including strain, thickness changes and residual stresses obtained from tube bending models into the crash models. They found that by accounting work hardening and thickness changes in the material due to bending process during the modelling of the crash event, the predicted peak force and energy absorption was increased by 25-30% and 18%, respectively.

From literature studies, two different mapping approaches are commonly found to be performed by researchers to study forming effects on crash response. The first approach is to map the preceding results from the simulation of the forming process onto the FE nominal mesh (coarse mesh), based on ideal CAD geometry. Whereas the second approach map the preceding results on the FE deformed mesh (fine mesh) based on the geometry from forming process. For simplification and to save cost and time, almost all of the previous works found to use the first approach. However, the first approach neglects the geometrical effects and therefore does not include the overall forming effect of the crash analysis which could mislead the crashworthiness evaluation. This is because, the high gradient in strain distribution are not represented since the FE nominal mesh (coarse mesh) will smooth the strain distribution (Cajuhi *et al*., 2003).

This paper will firstly show the draw forming simulation of circular cup part. The strain rate effect is taken into account and Johnson-Cook material model is used to represent the hardening behaviour of the material. FE draw forming model developed are validated with experimental results before being used for further analysis. Secondly, two types of FE crash models; ideal CAD circular cup geometry with nominal mesh and draw formed circular cup geometry with deformed mesh, were developed based on two different mapping approaches. Forming results (i.e. non-uniform thickness distribution, residual stress and strain contour) were then transferred to the first and second crash models by mapping process. The predicted load-displacement response will be compared to the measured ones.

2. METHODOLOGY

The research material used in the present study is dual phase (DP600) steel, which is a family member in the Advanced High Strength Steel (AHSS). The material has a thickness of 1.2 mm and was manufactured by cold-rolling. This material is chosen since it has a high ratio of yield strength to tensile as well as excellent formability, which means they have good ability to distribute the strain experienced during working.

The draw forming and the crash process were performed at constant displacement rates of 8.333 mm/s to represent quasi-static loading condition. The Johnson-Cook strain-rate dependent material model is used to describe the flow stress hardening behaviour of the material in both simulations (draw forming and crash) with the equivalent Von Mises flow stress given by Equations 1 and 2. Where ε is the plastic strain, $\dot{\varepsilon}$ is the strain rate, $\dot{\varepsilon}_o$ is the reference strain rate and T^* is the homologous temperature as expressed in Equation 2. T is the material temperature, T_{melt} is the melting temperature and T_{room} is the room temperature. Parameter A is the quasi-static (reference strain rate) true yield stress of the material at 0.2% offset strain in room temperature, B is strain hardening constant, n is strain hardening coefficient, C is strain rate strengthening coefficient and m is the thermal softening coefficient. In most literature paper, the reference strain rate is defined as 1.0 s^{-1}. However, in this study, the reference strain rate is defined as 0.001 s^{-1} which is chosen as the strain rate of the quasi-static test. The room temperature is selected as the reference temperature. All the other four material parameters (A, B, n and C) are determined from the experimentally obtained true stress verses true strain curves. The parameter A, B, n and C are 417 MPa, 902 MPa, 0.49, 0.012 respectively. Since the test is conducted at room temperature which is equal to the reference temperature, no temperature effect is present, and therefore the temperature dependent term can be neglected.

$$\sigma = (A + B\varepsilon^n)(1 + Cln\frac{\dot{\varepsilon}}{\dot{\varepsilon}_o})(1 - T^{*m}) \tag{1}$$

$$T^{*^m} = \frac{T - T_{room}}{T_{melt} - T_{room}} \tag{2}$$

2.1 Draw Forming Process

Draw forming experiments were performed on deformable DP600 steel blank with diameter 85 mm and 1.2 mm thick using a draw forming test device developed by Abu-Shah et al. (2016). The draw forming test device was attached to the Instron servo static machine 5585 (UTM) as shown in Figure 1 (a). In this process, double action draw forming mechanism is used. The applied load mechanism of punch force and blank holder force (BHF) are driven by the machine's system and external hydraulic system respectively. The external hydraulic cylinder system which is attached to the UTM as shown in Figure 1 (b) was used to produce uniform BHF force on the blank. The sheet was draw formed into a circular cup-shaped by using a 50 mm diameter punch with 6 mm edge radius and a die cavity with shoulder radius 2 mm. The desired circular cup-shape is designed based on Erichsen cupping test geometrical features in order to demonstrate the vehicle draw forming part in a reduced size. Moreover, this shape which involves compression, bending and drawing process can also be used to represent the vehicle production process. The draw forming test is conducted at constant displacement rates of 8.333 mm/s with blank holder force of 100 kN. The tests were performed at room temperature.

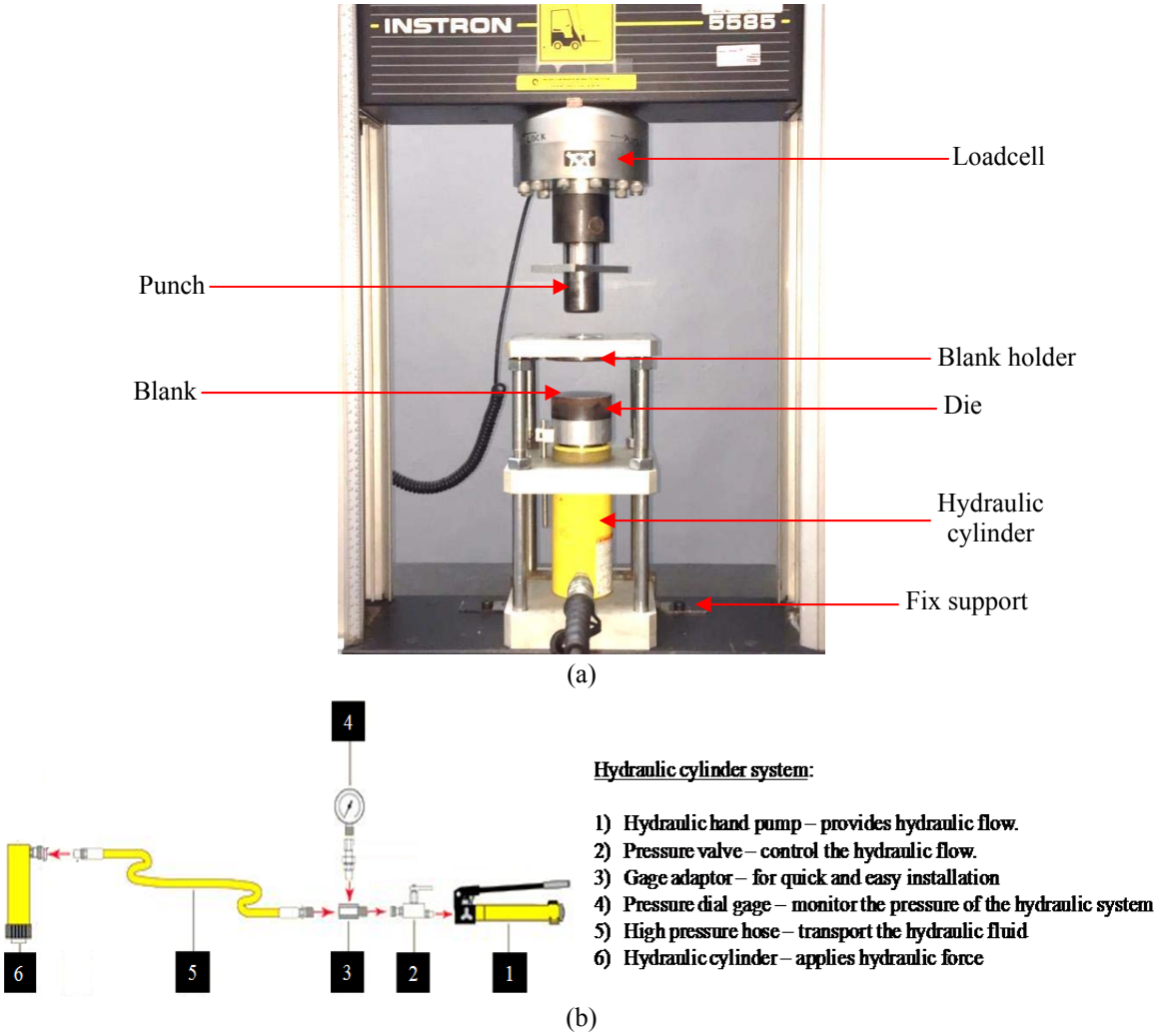


Figure 1: (a) Draw forming test tools setup, and (b) Hydraulic cylinder system (Abu-Shah *et al.*, 2016).

Finite element models of the circular cup draw forming experiment were created using HyperWorks Version 13. The draw forming consists of punch, die, blank holder and blank. The punch, die and the blank holder was modelled using rigid body shell elements, while the blank were modelled using deformable body shell elements. All parts were meshed using four-node quadrilateral 2D shell element type during the draw forming simulations. The boundary condition of the 3D FE draw forming model setup is shown in Figure 2.

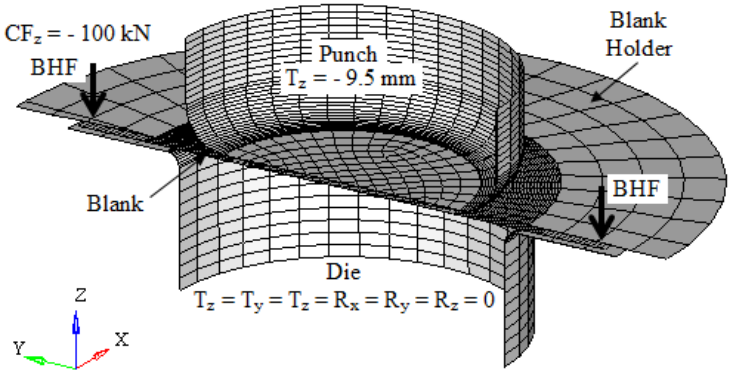


Figure 2: Sectional view of boundary conditions applied on draw forming process.

The FE draw forming simulation was performed at a constant displacement rate of 8.33 mm/s. The maximum drawing limit of 9.5 mm drawn depth was selected based from previous work by Amman *et al*. (2016) in order to conform to desired circular cup-shape without fracture. The draw forming FE model validation was conducted by comparing the FE and experiment global load-displacement curve as illustrated in Figure 3. The comparison showed similar trend between both results and show a close approximation to each other. This proved that the developed FE model is reliable for further analysis. The comparison of geometrical shape between CAD model and draw forming model is illustrated in Figure 4. The steel blank underwent large plastic deformation during the draw forming process which then leads to substantial geometrical changes especially at the sidewall area and spring-back effect on the FE draw form model. The non-uniform thickness distribution, residual stress and strain contour results obtained at the end of draw forming simulation is shown in Figure 5. These residual results are needed for mapping process in order to examine the mapping effects on the crash response. The preceding residual results from Figure 5 will be mapped to both model in Figure 4 using mapping process which will be discussed in the following section.

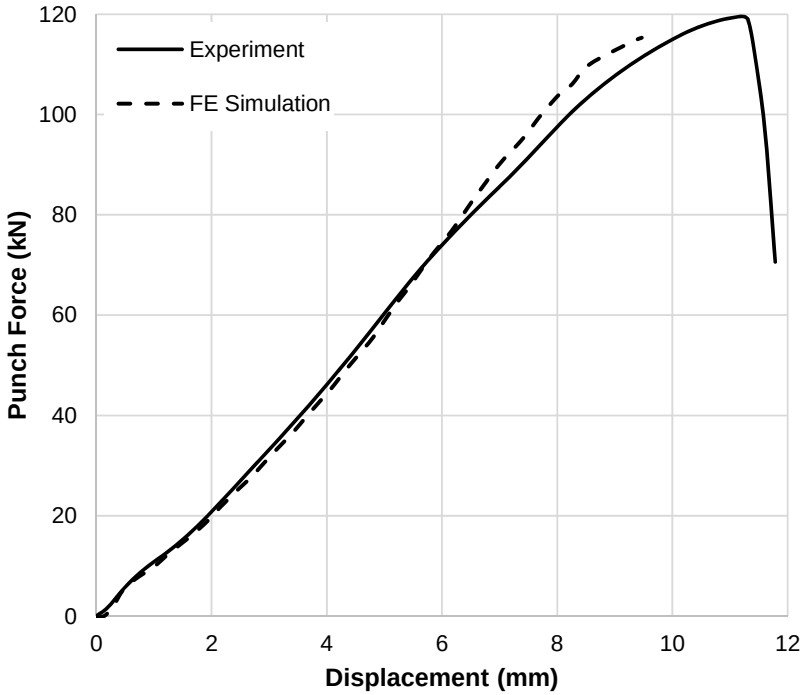


Figure 3: Comparison of global load-displacement response between FE simulation and experiment.

Figure 4: Comparison of geometrical shape between ideal CAD and FE draw formed developed model.

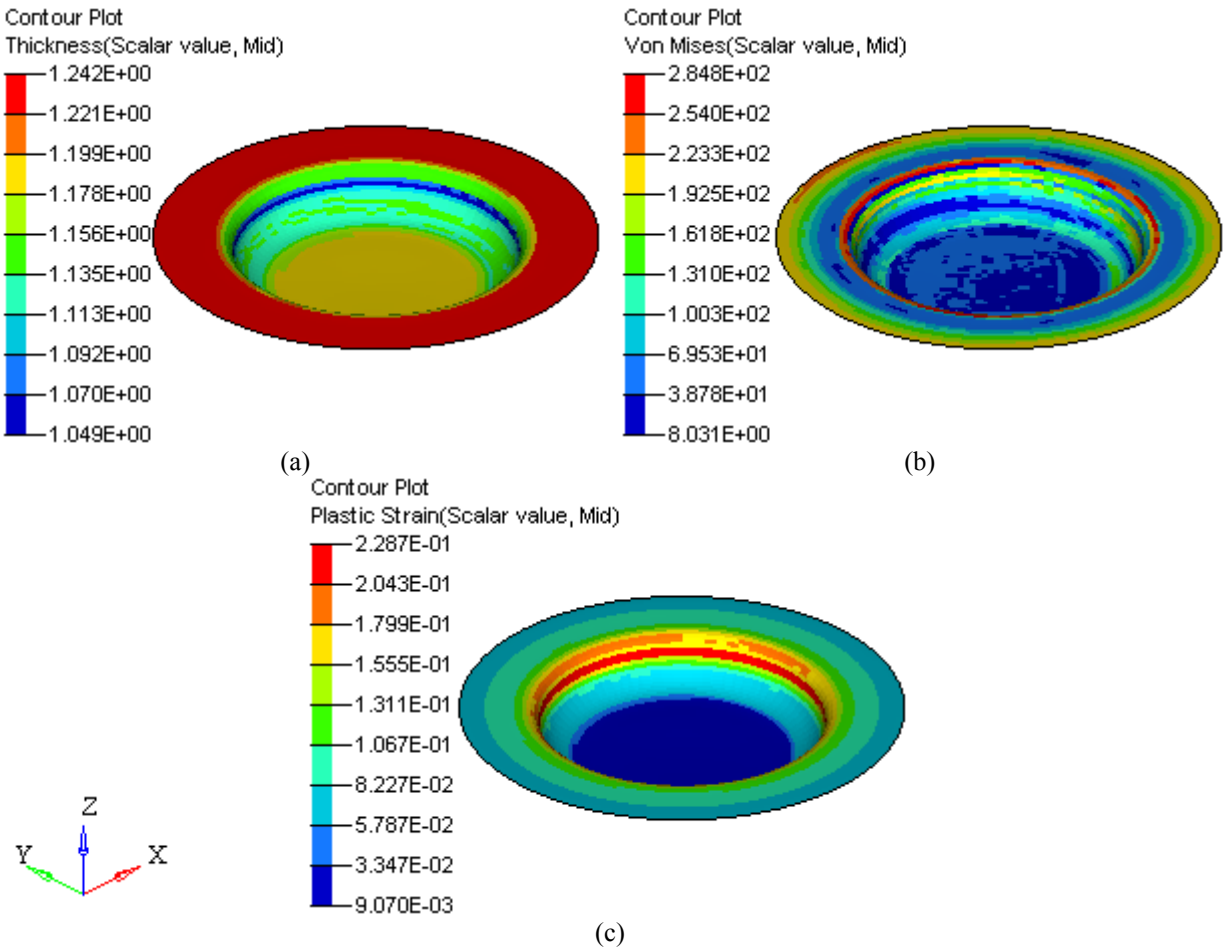


Figure 5: Residual contour results of (a) non-uniform thickness, (b) residual stress and (c) residual strain obtained at the end of draw forming process.

2.2 Crash Test

Crash test set-up as shown in Figure 6 was performed using UTM with load cell of 150 kN under displacement control. The circular cup which is formed from draw forming process was placed in between the impactor plate and base. In this test, a constant displacement rate of 8.333 mm/s is used to crash the circular cup-shaped part formed from plastic deformation draw forming process. This speed represents the quasi-static crash response which is used to investigate the mechanical behaviour in terms of energy absorption and deformation mode of the material by incorporating draw forming effects. The impactor was set to crash the draw-formed circular cup part up to 7 mm. This maximum displacement was selected based on the pre-test results until it reached the specified limiting load of UTM at 140 kN. No fixture was used to hold the specimens in place between the impactor and base. The specimens were crushed at room temperature. The crash test force-displacement was recorded for FE validation.

Figure 6: Crash test set up.

The FE circular cup crash model was developed to replicate the experimental design tools and setup. The crash tools consist of impactor, base and circular cup. The impactor was modelled using rigid shell elements, while the circular cup was modelled using deformable shell elements. All parts were meshed using four-node quadrilateral 2D shell element type during the crash simulation. The boundary condition setup for the crash simulation is illustrated in Figure 7. The impactor moved in the axial direction of the circular cup with a velocity of 8.333 mm/s. Two types of FE crash models were developed based on two different meshes (i.e. nominal mesh and deformed mesh). The FE developed crash models are mapped with the preceding draw forming results and the mapping approach employed in this study will be explain in the next sub-section. The crash analyses were carried out using radios explicit solver.

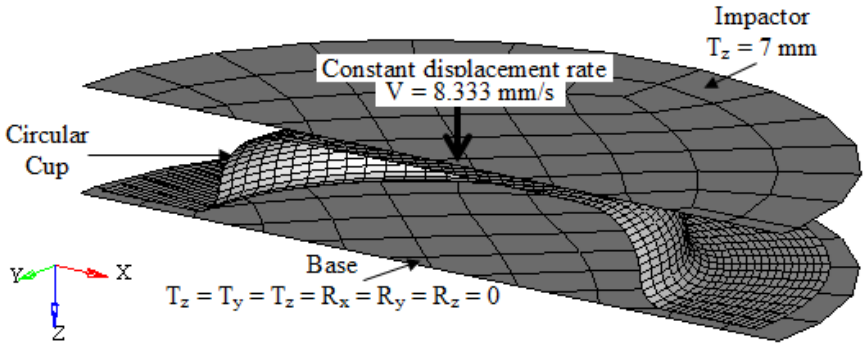


Figure 7: Sectional view of boundary condition setup for crash simulation.

2.2.1 Mapping Approach

The influences of preceding results on the subsequent crash analysis were investigated by performing a mapping process on the crash model. In FEA, "mapping" is defined as transferring the results of the previous simulation to the current simulation (Doğan 2009). In this study, the results of the draw forming simulation i.e. thickness changes, residual stress and strain are mapped to the crash model as initial conditions. The preceding results were mapped into the crash model by utilizing the sta file (geometrical result file) obtained from draw forming simulations by using a tool called result mapper that is available in Hypercrash software.

Two different mapping approaches were performed in this paper prior to crash analysis. The first crash model uses ideal CAD circular cup geometry with FE nominal mesh being mapped with preceding draw forming results. The second crash model used draw formed circular cup geometry with FE deformed mesh mapped with the same preceding draw forming results. The mapping strategies performed in this study are summarised in Table 1. The FE crash model with mapped preceding results is used for later crash analysis. The results obtained for crash analysis of both developed crash model are compared and the significant effect of mapping preceding residual results on the crash model were examined.

Table 1: Explanation of mapping approach performed in this study.

Mapping Strategy	Explanation
Method 1	• Mapped preceding results on ideal CAD geometry (Mapping on different geometrical shape). • Mapped from deformed mesh to nominal mesh (Mapping on different FE mesh).
Method 2	• Mapped preceding results on draw formed geometry (Mapping on similar geometrical shape). • Mapped from deformed mesh to deformed mesh (Mapping on similar FE mesh).

3. RESULTS AND DISCUSSION

The draw forming simulation results have been successfully mapped on the FE crash model of two different geometrical shapes of circular cup. Figure 8 shows the FE crash model is mapped with the contour of thickness, residual stress and strain resulted from the draw forming simulation. The FE results obtained after the crash analysis is compared and discussed. Initial condition after mapping on ideal CAD model (Method 1) displayed different contour when compared to mapping on draw formed model (Method 2). For the process of mapping on ideal CAD model (Method 1), state of variable of the integration point in the FE nominal mesh get the values of the nearest integration point from the preceding FE deformed mesh. The preceding results have been mapped only for the regions which overlap. This causes the distribution of thickness, residual stresses and strains not represented identically as the distribution are smoothed out. While mapping of the contour results from the forming model to the draw-formed crash model (Method 2) results on identical contour before and after mapping process. Since the same geometrical shape and FE mesh is used for both draw forming and crash. The results obtained for two types of mapping approach applied in this study will be compared and discussed by studying the load-displacement response, deformation mode and energy absorption.

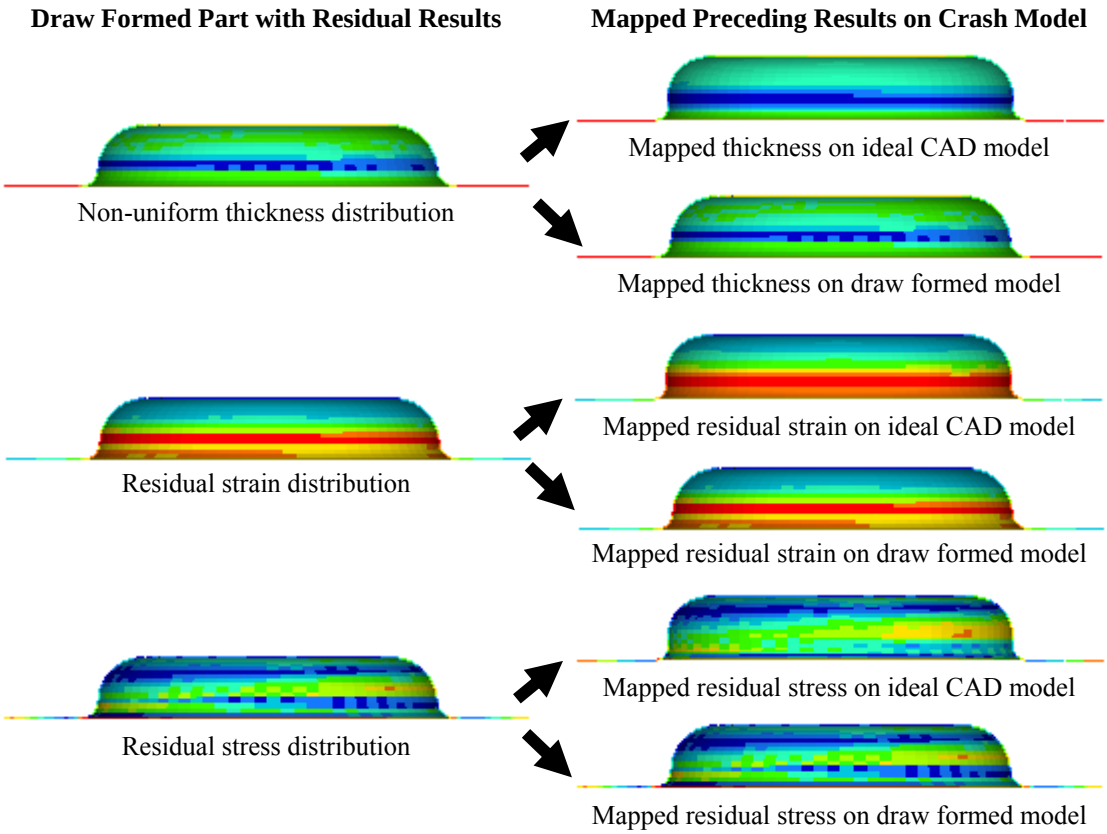


Figure 8: Mapping of preceding forming results on ideal CAD model and draw formed model as initial condition for crash analysis.

3.1 Effects of Different Mapping Approach on Predicted Crash Response

The effect of mapping on ideal CAD model (method 1) and mapping on draw formed cup model (method 2) is compared. It is highly anticipated that incorporating mapped contour on the FE crash model could change the mechanical properties and would give significant effect in later crash response. The FE global load-displacement responses for each mapping strategy applied are compared with experimental results as illustrated in Figure 9. From this result, it can be clearly seen that mapping method 1 gives very high crash force compared to mapping method 2. Due to the absence of draw forming effects (i.e., geometrical, thickness changes, work hardening and strain rate effect) in the initial condition of ideal CAD model, the crash results of mapping method 1 showed stronger and stiffer response than the results obtained from mapping method 2. From the load-displacement curve, it can be seen that by using mapping process, the shape of the load-displacement curve for the FE simulation results of mapping method 2 gives an almost similar pattern with the experimental results compared to mapping method 1.

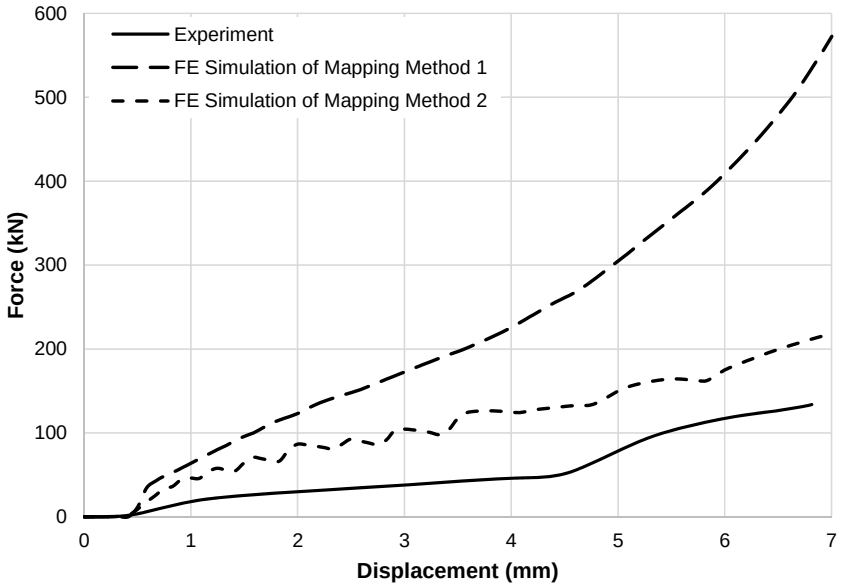


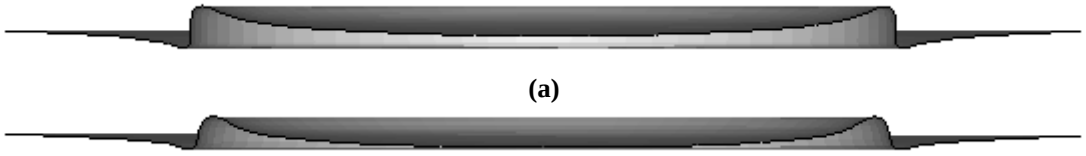
Figure 9: Load displacement of cup during crash test.

Method 1 and method 2 led to 327.94% and 63.11% higher load value, respectively, at the end of the crash process when compared to experimental value. Even though mapping on same geometrical shape with deformed mesh (method 2) leads to higher load value compared to the experiment, this method has been proven to be able to capture the deformation mode more accurately and gives an approximately similar pattern with the experimental curve compared to method 1 which is commonly practiced in industries. This work shows that forming data together with the true geometry of the crash member after the forming and spring back processes should be considered on the crash analysis in order for achieving reliable simulation results.

The final deformed shape of both mapping method 1 and method 2 are compared as illustrated in Figure 10. It appears that circular cup model that was mapped using method 2 deformed more than method 1 which indicated a weaker structure. Najafi & Rais-Rohani (2012) noted that the manufacturing responses are found to be most sensitive to the wall thickness and the corner radius as the thickness changes affect the stress distribution and consequently plastic deformation, while corner radius is more influential on the rupture response due to excessive plastic deformation that occurred in the corner regions. They found that the geometric attributes can significantly affect the energy absorption behaviour of a structural component. Similar results were found in this study.

Mapped preceding results which bring plastic deformation or work hardening effect combined with the ideal geometrical shape effect that existed in a circular cup of method 1 increased the stiffness of

the part, slows the deformation process and thus leads to higher energy absorption as depicted in Figure 11. At the end of the crash process, method 1 recorded 55.82% higher energy absorption than method 2. An analysis of the computed results obtained in this study showed a decreased of the energy absorption during crash for models which mapped preceding results due to draw forming process which carries work hardening and thinning effects.



(a)

(b)

Figure 10: Comparison of the cross-sectional side-view deformation at the end of crash process: (a) Method 1, (b) Method 2.

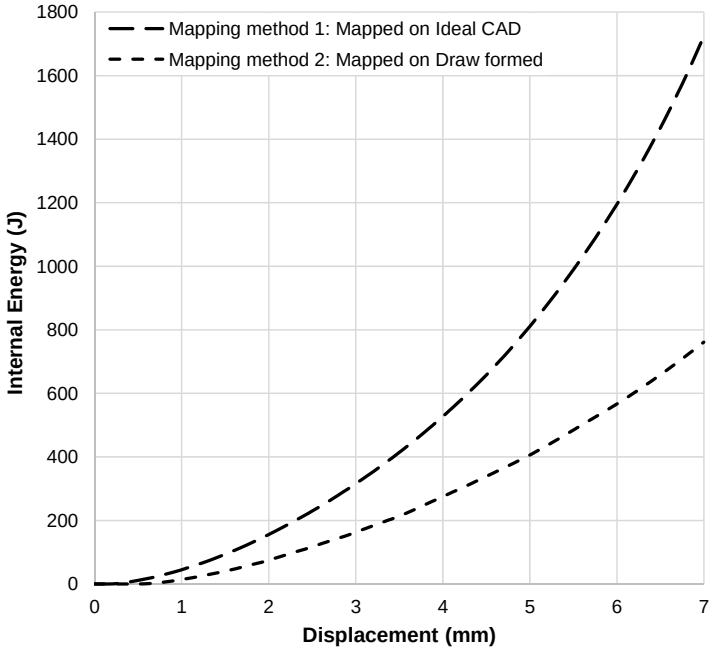


Figure 11: Comparison of energy absorption for mapping method 1 and method 2.

4. CONCLUSION

Two different mapping methods that can be used to include preceding residual results on subsequent crash analysis were analysed. The results showed that mapping method 2 is better to capture the deformation behaviour of the material during crash compared to method 1. In method 2, the geometrical effects, as well as residual stress-strain results would be considered in the subsequent process and therefore the crashworthiness prediction can be more accurate. It is concluded that mapping the preceding results on deformed geometrical shape with deformed mesh improves crashworthiness prediction compared to mapping on CAD geometrical shape with nominal mesh.

REFERENCES

Abu-Shah, I., Halim, M.F., Sulaimain, S.N. & Tamin, M.N. (2016). Design and development of auto-steel draw forming test device. *ARPN J. Eng. Appl. Sci.*, **11**: 10977–10981.

Amman, R.M., Halim, M.F., Sivakumar, D. & Abu-Shah, I. (2016). Influences of thickness and geometrical change from draw forming in quasi-static axial compression simulation. *Int. Rev. Modell. Sim.*, **9**: 442–449.

Cajuhi, A.J., Oliveira, A.A., Kato, G.M. & Romera, G.F.S. (2003). Forming analyses results influences on safety crash finite-element analyses. *SAE Brasil 2003 Congr. Exhib..,* 3621.

Cowell, B., Kellicut, A. & Fisher, A. (2000). The effects of forming and parameter mapping on further simulation. In *6th Annual LS-DYNA User's Conf.*, pp. 23-28.

Dhanajkar, N.S., Namboodiri, A., Muley, P. & Musale, G. (2011). Forming Automation using HyperStudy and Radioss to extract prestress for crash analysis. *Hyperworks Technol. Conf.*, pp. 1–7.

Doğan, U.Ç. (2009). *Effect of Strain History on Simulation of Crashworthiness*. MSc Thesis, Department of Mechanical Engineering, Middle East Technical University.

Dutton, T., Iregbu, S., Sturt, R., Kellicut, A., Cowell, B. & Kavikondala, K. (1999). The effect of forming on the crashworthiness of vehicles with hydroformed frame siderails. *SAE.*, **108**: 3354–3360.

Kose, K. & Rietman, B. (2003). Plasticity effects in subsequent simulations of car structures. *VII Int. Conf. Comput. Plasticity*, pp. 1–7.

Logue, B., Dingle, M. & Duncan, J.L. (2007). Side-wall thickness in draw die forming. *J. Mater. Process. Technol.*, **182**: 191–194.

Mohd Amman, R., Halim, M.F., Sivakumar, D., Abu-Shah, I., Sulaiman, S.N. & Samekto, H. (2016). Study of thinning effect from deep drawing process on crash analysis. *Proc. Mech. Eng. Res. Day 2016*. Ayer Keroh, pp. 37–38.

Najafi, A. & Rais-Rohani, M. (2012). Sequential coupled process-performance simulation and multi-objective optimization of thin-walled tubes. *Mater. Des.*, **41**: 89–98.

Nie, W. Z., Sa, C-Y. & Bohm, M. (2004). FormingFX: A tool for mapping sheet-forming analysis results to ABAQUS structural analysis. *2004 ABAQUS User's Conf.*, pp. 461-474.

Obradovic, J., Boria, S. & Belingardi, G. (2012). Lightweight design and crash analysis of composite frontal impact energy absorbing structures. *Compos. Struct.*, **94**: 2, 423–430.

Oliveira, D.A., Worswick, M.J., Grantab, R., Williams, B.W. & Mayer, R. (2006). Effect of forming process variables on the crashworthiness of aluminum alloy tubes. *Int. J. Impact Eng.*, **32**: 826–846.

Sasek, J., Pasek, M., Benes, K. & Glac, V. (2010). Effects of manufacturing process in crash simulations. *Appl. Comput. Mech.*, **4**: 113–120.

Takashina, K., Ueda, K. & Ohtsuka, T. (2009). Investigation of accuracy improvement on crashworthiness simulation with pre-simulation of metal forming. *7th European LS-DYNA Conf.*, 14-15 May 2009, Salzburg, Austria.

Williams, B.W., Worswick, M.J., D'Amours, G., Rahem, A. & Mayer, R. (2010). Influence of forming effects on the axial crush response of hydroformed aluminum alloy tubes. *Int. J. Impact Eng.*, **37**: 1008–1020.

EFFECT OF NICKLE FOIL WIDTH ON THE GENERATED WAVE MODE FROM A MAGNETOSTRICTIVE SENSOR

Nor Salim Muhammad*, Ayuob Sultan Saif Alnadhari, Roszaidi Ramlan, Reduan Mat Dan, Ruztamreen Jenal & Mohd Khairi Mohamed Nor

Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia

*Email: norsalim@utem.edu.my

ABSTRACT

Simulation study of guided wave excitation in pipes using nickel foils to form magnetostrictive sensors was conducted to observe the differences in the excited guided wave when using the same magnetostrictive material. It was performed through excitation of L(0,2) mode into models of aluminium pipes by using nickel foils at different widths. The simulation results of L(0,2) mode propagations show a significant difference in the wave propagation when the width of the nickel foil was too wide compared to the wavelength of the excited L(0,2) mode. The obtained results can provide useful information in developing high performance magnetostrictive sensors for defect inspections in pipe and tube structures.

Keywords: *Non-destructive testing; ultrasonic; guided wave.*

1. INTRODUCTION

The presence of defects in main facilities for petrochemical industries, power plants, and aircrafts may contribute to the failures on their operations (Rao & Nair, 1998; Koc & Altan, 2002; Purbolaksono *et al.*, 2010; Saleh *et al.*, 2014; Heimbs *et al.*, 2014; Jones *et al.*, 2015). In monitoring tasks, guided wave techniques have the advantage of inspecting large or long structures such as oil storage tanks, and pipes (Cawley *et al.*, 2003; Rose, 2004). In addition, there are also demands for inspections on civilian and military aircrafts for ice wing detection and defect inspections over the multilayered structures and their fluid lines (Kolkman *et al.*, 1996; Anghileri *et al.*, 2005; Gao & Rose, 2009; Mueller *et al.*, 2016). An aircraft has critical areas such as wings and stabilizers that must be free from contaminants to control the direction of movement during its operation as well as departing and landing (Gao & Rose, 2009). The formation of ice in these areas not only provides additional body load, but disturbs airflow, which reduces lifting forces. Delaminations are also common imperfections that occur in carbon fiber composites. The presence of the defects in the structure will result in structural weaknesses that may cause failure to the aircraft structures (Mueller *et al.*, 2016). At the same time, inspection of fuel leak and other fluid lines including oxygen supplies are also important to prevent any unexpected failures (Kolkman *et al.*, 1996; Anghileri *et al.*, 2005). Structure health monitoring using this guided wave technique has good service records for monitoring of insulated and underground pipelines (Cawley *et al.*, 2003). This is due to its ability to inspect the pipe structure through a fixed point in the pulse-echo mode for long distance underground pipe inspections. This advantage reduces the cost of accessing the surface of the pipes before the inspection and completes the task in shorter time than the previous techniques which require the removal of the insulation layer on pipes or surface of the ground covering the pipes (Dobson & Cawley, 2016). Among the commercial equipment available in the market are Teletest, Wavemaker, and MsS systems. These equipment use piezoelectric and magnetostrictive sensor (MsS) technologies to produce ultrasonic guided waves for screening defects in structures (Ostachowicz *et al.*, 2009). However this technique is still new in developing countries that are lacking of expertise and with equipment constraints.

The use of magnetostrictive sensors in the ultrasonic guided wave is able to generate torsional $T(0,1)$, and longitudinal $L(0,2)$ modes which are widely used in the guided wave pipe inspections (Kwun & Bartels, 1998; Kim *et al.*, 2011). In low frequency guided wave inspections, $L(0,2)$ is selected due to its fastest wave propagation in pipe which can easily separate the defect echoes from the complicated wave propagations (Shin & Rose, 1999; Mu & Rose, 2008). There are several studies on excitation of guided wave modes using magnetostrictive sensors in pipes which are made of ferromagnetic materials such as cobalt-iron alloy and nickel foils (Kwun & Teller, 1994; Kim *et al.*, 2013; Kim & Kwon, 2015). The previous researchers prepared the magnetostrictive sensor from the foils which are bonded along the complete circumferential direction on pipes and used ribbon cables to provide electricity for the wave excitations. The low cost magnetostrictive sensor has the advantage in reducing the cost of transducers compared to the piezoelectric transducers which are typically very costly in forming the ring transducers for excitation of axisymmetric wave modes in pipes.

This study aims to investigate the effect of width of nickel foils on the excitation of $L(0,2)$ from magnetostrictive sensors on pipe and tube structures. Simulations on the excitation of ultrasonic guided waves at frequencies of about 50 kHz were performed using several nickel foils at different widths which were placed along circumferential direction of the pipe models. Observation on amplitudes and number of wave mode cycles which were produced from different widths of nickel foil was performed to study the effect of the width of nickel foil on the generated wave mode.

2. METHODOLOGY

2.1 Frequency Selection

This study investigates relationship between the width of magnetostrictive sensor and wavelength of the excited waveform by modelling the wave excitation in an aluminium pipe. The study started with computing the dispersion curves for guided wave propagation in an aluminium pipe at thickness of 6 mm and outer diameter of 100 mm in CIVA simulation software. The computed results of group dispersion curves are plotted as in Figure 1. The figure shows the computed dispersion curves for longitudinal $L(0,1)$ and $L(0,2)$ modes in the defined pipe model. The computed dispersion curves show different wave behaviours depending on the selected frequency range whereas the excitation of $L(0,2)$ mode can have the largest and large dispersions as shown at the lower and upper of the computed frequency range. The figure also indicated $L(0,2)$ mode at highest wave speed for the frequency range of 30 to 250 kHz. However, the use of the fastest wave is generally proposed to obtain defect echo at shorter time arrival than the other modes (Shin & Rose, 1999; Cawley *et al.*, 2003; Mu & Rose, 2008). In addition, the wave mode also has small dispersion in group velocity as visible at frequency of 40 to 60 kHz of the dispersion curve. At the same time, the selection of frequency range with small dispersion curve also has an advantage to avoid the complicated time waveforms from the pipe inspections (Cawley *et al.*, 2003).

Therefore, the central frequencies of the tone burst signals in this study were selected in the range of 40 to 60 kHz for excitation of small dispersion $L(0,2)$ mode in the pipe which propagates at the fastest wave speed compared to the other propagating modes in the pipe.

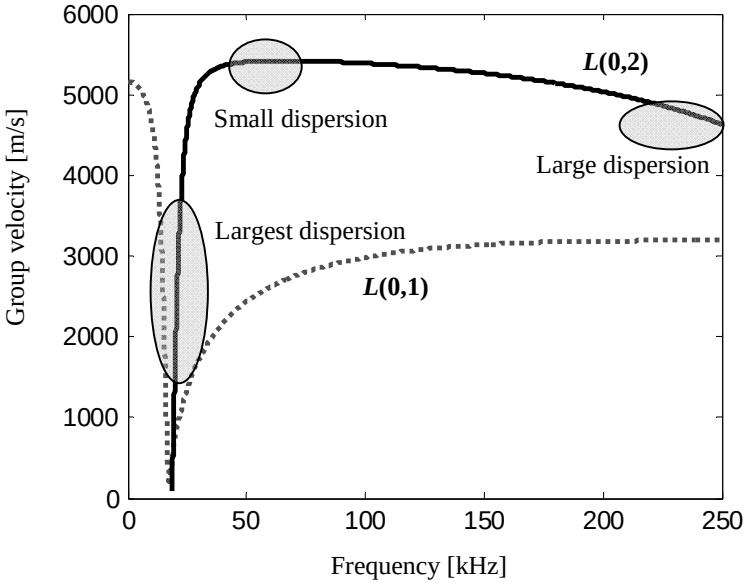


Figure 1: Dispersion curves of longitudinal modes for the aluminium pipe.

2.2 Model of Wave Excitation

We studied the effect of width of magnetostrictive sensor through simulation of wave propagation using ANSYS simulation software. The developed simulation model of wave excitation in the aluminium pipe is shown in Figure 2 which depicted a nickel foil bonded along the circumferential direction on left side of the pipe end. The pipe was modelled with circumferential defect at location about 5 m from the magnetostrictive sensor with defect width and depth at 10 and 4 mm, respectively. The reason for circumferential defect in the model of the pipe is to have less complicated wave reflection and transmission around defect in the pipe compared to the oval shape of defect which is typically used in defect studies. Models with different widths of nickel foils are used to study the effect of foil width on the excited $L(0,2)$ mode from measurement of the time waveform at point P and visualizing the propagating wave contours in the 6 m pipes. Measurement of waveform at point P was applied to identify the effect of the width of nickel foil on the defect detection in the pipe and visualization of the wave propagation was used to identify the excited wave structures from the modelled magnetostrictive sensors on pipes. Nickel foils at widths of $\lambda/8$ to λ were used in the simulation models for excitation of the wave mode from 40 to 60 kHz as detailed in Table 1. The widths of nickel foils used in the model started from 17.5 to 140.50 mm at 40 kHz, 14 to 112 mm at 50 kHz, and 11.63 to 93 mm at 60 kHz. Different widths are used in the simulations to observe changes in the amplitude of the reflected wave from defect with respect to the initial wave. At the same time, changes on the wave structure of the excited $L(0,2)$ mode in the pipe was monitored through changes on the wave contours visualized from the wave propagation into the circumferential defect. It aims to identify the appropriate widths for wave excitation of $L(0,2)$ mode using magnetostrictive sensor at controllable wave cycle and high amplitude excitation.

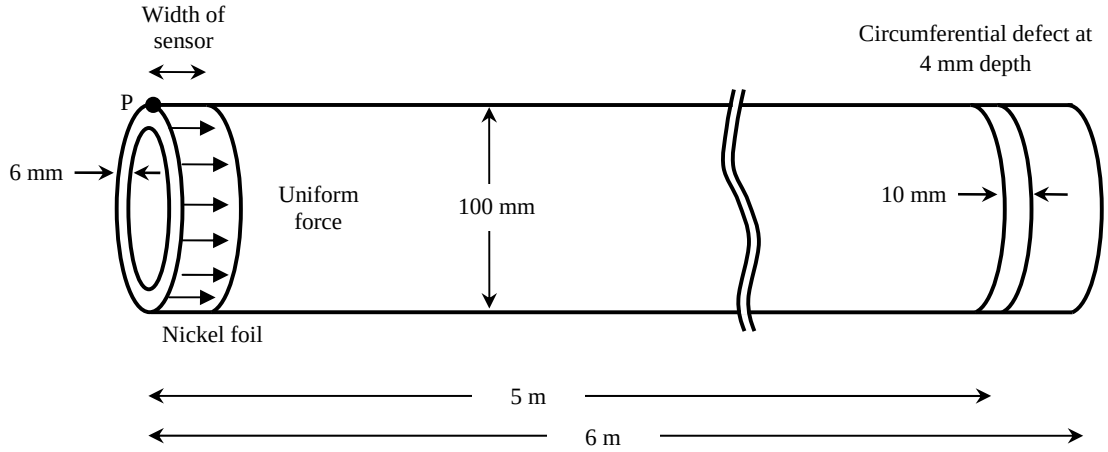
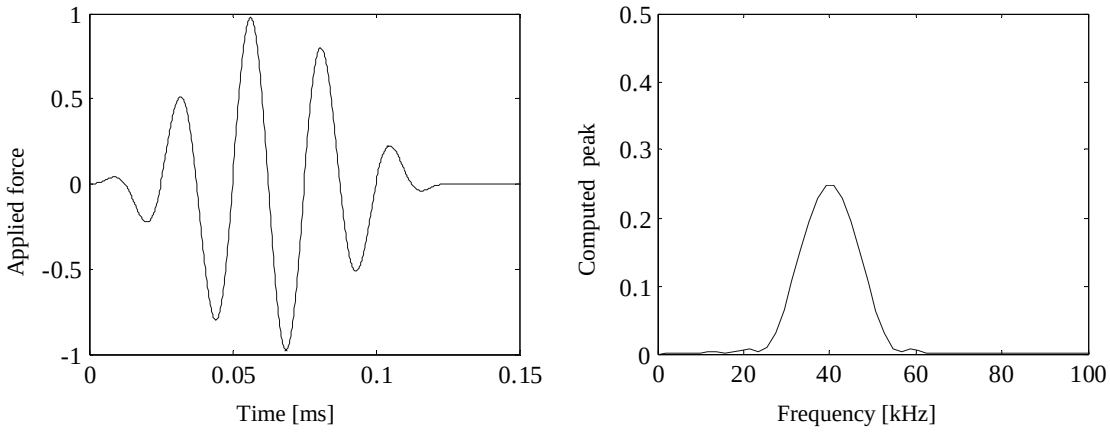


Figure 2: Model of wave excitation in ANSYS.

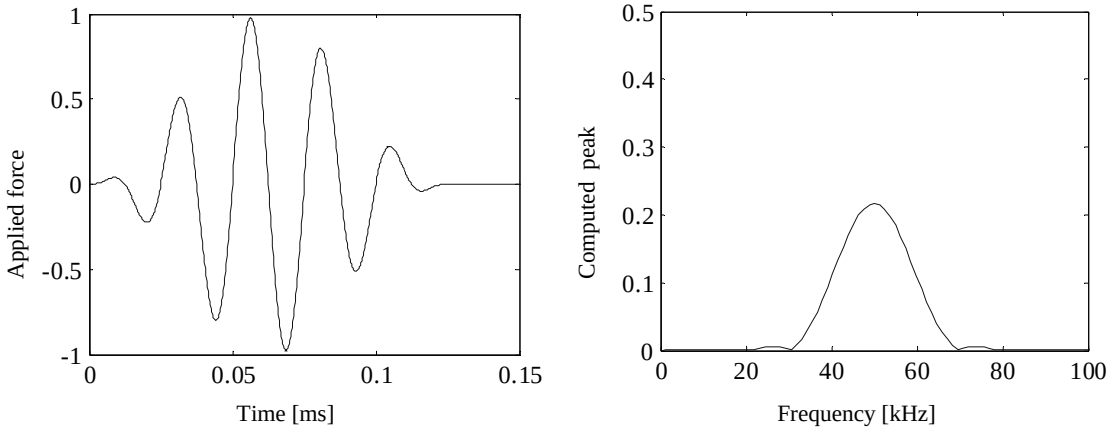
Table 1: Width of nickel foils and mesh number in FEM models.

Frequency (kHz)	Ratio of foil width to wavelength	Width (mm)	Mesh number
40	λ/8	17.56	216, 720
	λ/4	35.13	
	λ/2	70.25	
	λ	140.5	
50	λ/8	14	
	λ/4	28	
	λ/2	56	
	λ	112	
60	λ/8	11.63	
	λ/4	23.25	
	λ/2	46.5	
	λ	93	

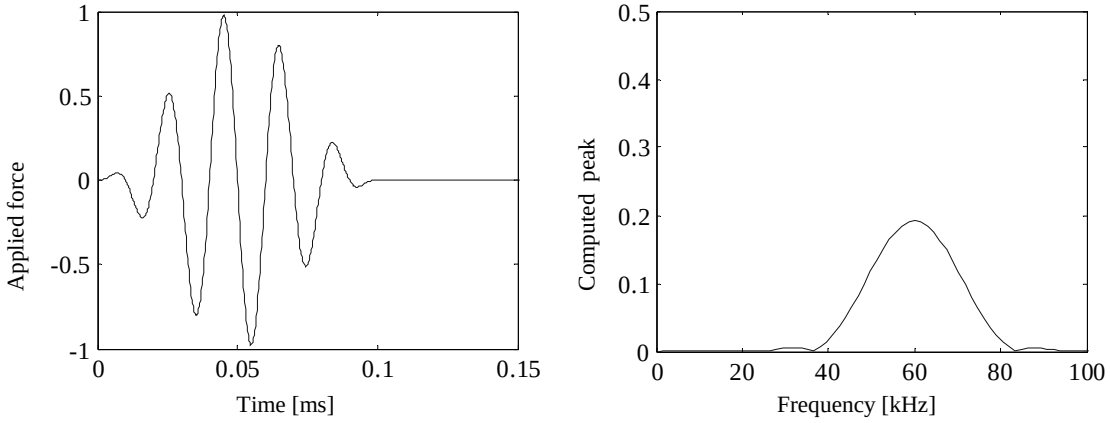
The final FEM models of the pipes were meshed at mesh size of 5 mm length of the hex dominant mesh which resulted in the number of mesh at 216, 720 elements. A uniform force in circumferential direction of the pipe was applied on the modelled nickel foil to investigate the behaviour of magnetostrictive transducers. The modes were simulated using 5 cycles of tone burst signals at frequencies 40, 50 and 60 kHz as shown in Figure 3.



a) Load signal at 40 kHz



b) Load signal at 50 kHz



c) Load signal at 60 kHz

Figure 3: Five cycles of tone burst signal used in simulation models.

3. RESULTS AND DISCUSSION

3.1 Computed Wave Propagation in Pulse-Echo Mode Guided Wave Inspection

Recorded waveforms of the computed wave excitations at different widths of nickel foils which used to form magnetostrctive sensors are shown in Figures 4 to 6 for excitation of $L(0,2)$ mode at central frequency from 40 to 60 kHz. The computed waveforms consist of initial wave, defect echo, and reflected wave from end of pipe. Validation of the defect location and pipe length can be made through calculation of travel distance from reflected $L(0,2)$ mode at defect and the pipe end. The distances can be evaluated using the velocity of $L(0,2)$ mode obtained in Figure 1 which propagates approximately at 5350 m/s and reflected from defect and end of pipe at 1.9 ms and 2.3 ms, respectively. Figures 4 to 6 also depicted the results of $L(0,2)$ mode propagation in pipes which excited from nickel foils at widths of $\lambda/8$, $\lambda/4$, $\lambda/2$, and λ as explained in Table 1.

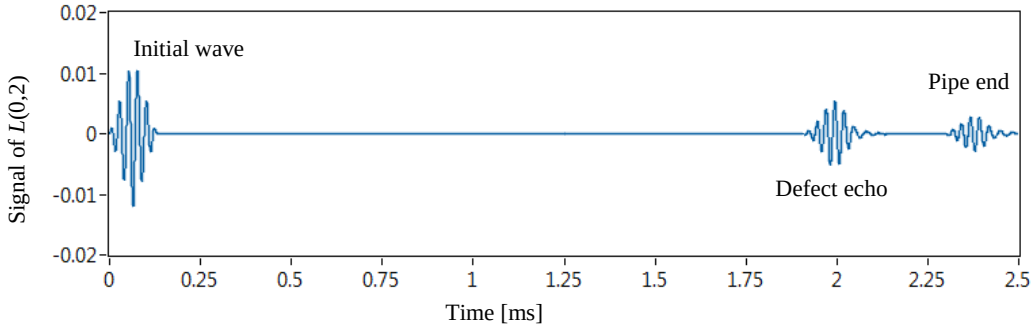
3.2 Reflected Wave From Defect

The reflected waves from defect are plotted for different widths of nickel foil at excitation frequencies of 40, 50, and 60 kHz as shown in Figures 4, 5 and 6 respectively. The computed signal level is not assigned to any physical quantity of the waveforms but is used to predict signal level of the excited ultrasonic wave in the pipe. Significant changes on the reflected wave from defect and end of pipe is observed for the widths of nickel foil which is less and greater than quarter of the wavelengths ($\lambda/4$). Good reflected waves from defect and end of pipe are observed in the simulations for widths of nickel at $\lambda/8$ and $\lambda/4$ as in Figures 4(a), 4(b), 5(a), 5(b), 6(a), and 6(b). At the same time, excited waves from nickel foil that is wider than the quarter wavelength ($\lambda/4$) of the selected frequencies computed very small reflected wave from defect and the end of the pipe as shown in Figures 4 (c), 4 (d), 5 (c), 5 (d), 6 (c), and 6 (d).

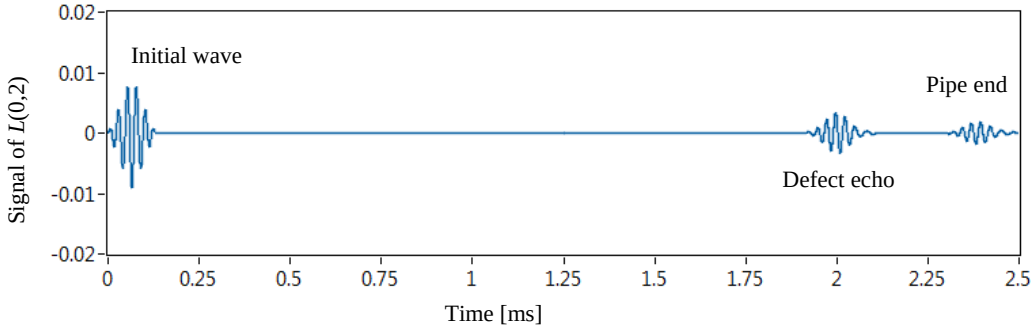
Further computation on reflected signals of $L(0,2)$ mode from defect and the end of the pipe were also computed at 40, 50, and 60 kHz in the pulse-echo mode using nickel foils at widths of $\lambda/10$, $\lambda/8$, $\lambda/6$, $\lambda/4$, $\lambda/2$, λ, and 2λ of the frequencies. The computed wave packets of the reflected waves are observed and plotted in Figures 7 and 8, respectively. The plotted results also indicated different signal levels for different widths of the nickel foil especially for the width which is not wider than quarter wavelength ($\lambda/4$). However, the widths of nickel foil which are wider than quarter wavelength ($\lambda/4$) of the selected frequencies computed almost undetectable reflected wave from defect and end of pipe as represented in the figures.

3.3 Excited Wave from Different Widths of Nickel Foils

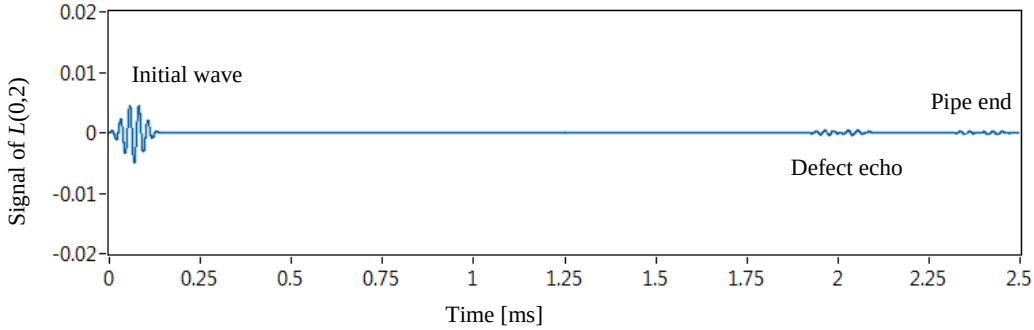
Simulation of excited wave from nickel foils at widths of $\lambda/8$, $\lambda/4$, $\lambda/2$, and λ are visualized at 50 kHz as shown in Figure 9. The wave propagation indicated similar wave structure for the simulation results at widths of $\lambda/8$ and $\lambda/4$ of the selected frequency as shown in Figures 9(a) and 9(b). On the other hand, simulation results at widths of $\lambda/2$ and λ depicted wave propagations which differ with the previous results as shown in Figure 9(c) and 9(d). The excited waves from the wider nickel foil showed propagating wave with non-uniform wave distribution and having shorter wavelengths which can be indentified from the narrow stripes of the wave contours. This explained that the wider nickel foil can excite false wave which is not the dominant $L(0,2)$ mode as indicated in Figures 9(c) and 9(d).



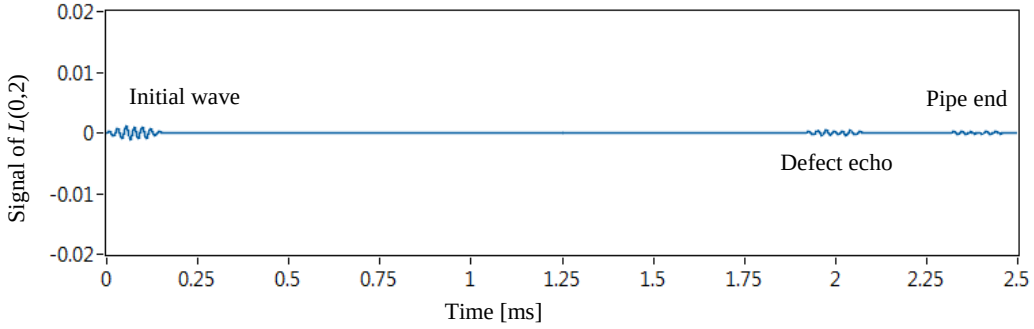
a) Width of nickel foil at λ/8



b) Width of nickel foil at λ/4

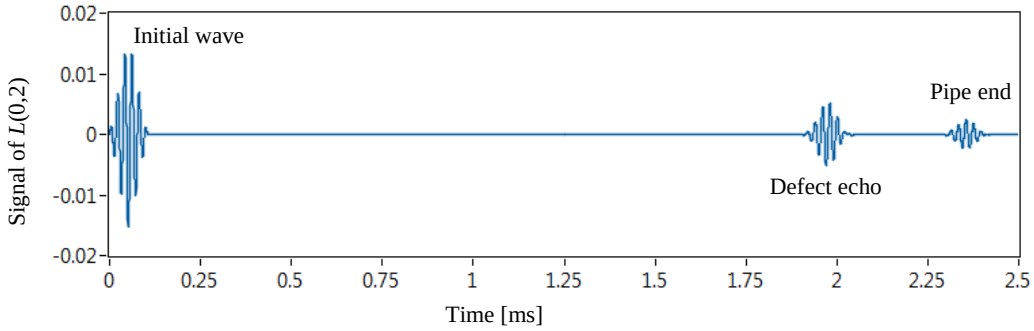


c) Width of nickel foil at λ/2

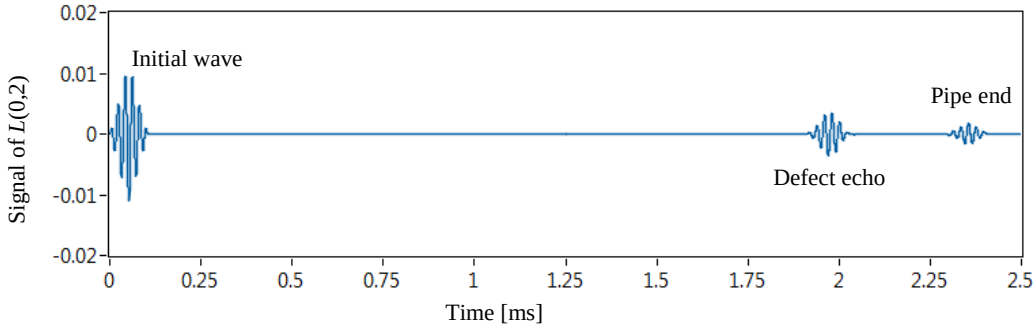


d) Width of nickel foil at λ

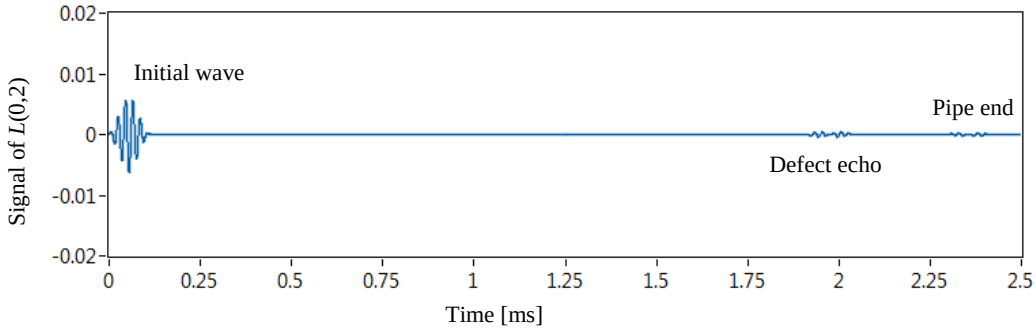
Figure 4: Excitation of *L*(0,2) at 40 kHz.



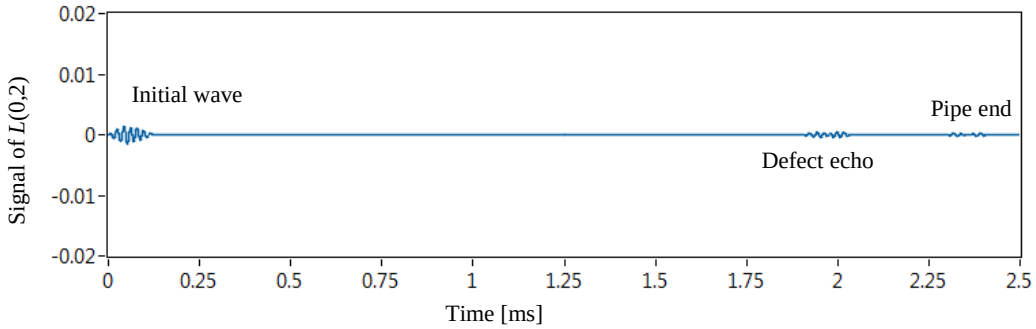
a) Width of nickel foil at λ/8



b) Width of nickel foil at λ/4

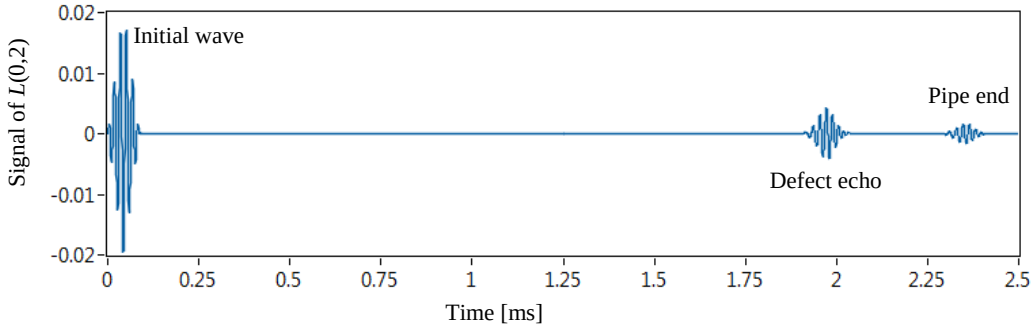


c) Width of nickel foil at λ/2

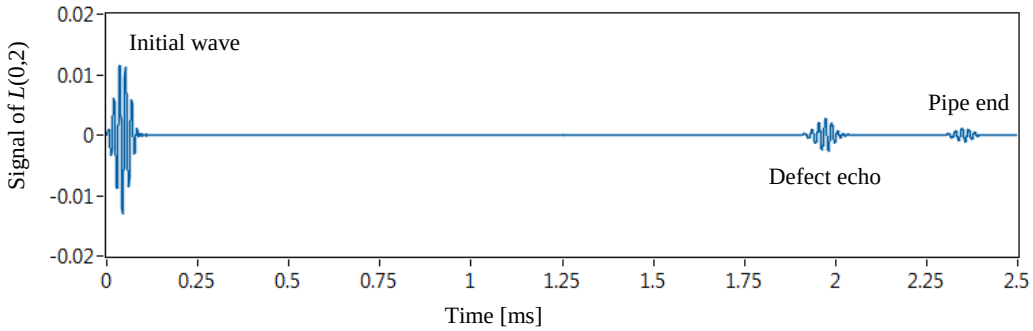


d) Width of nickel foil at λ

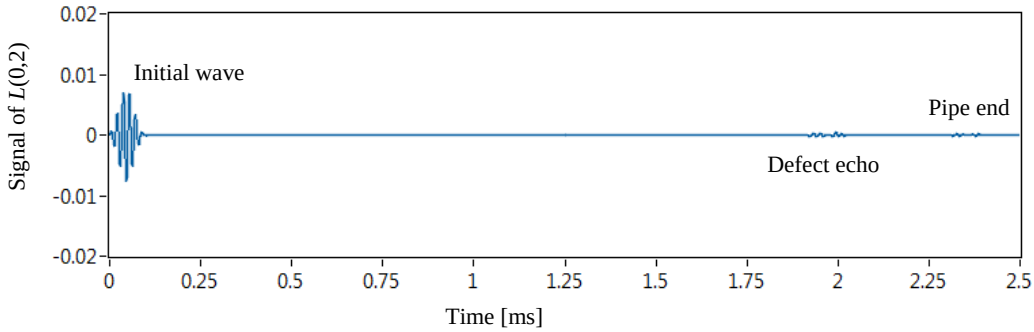
Figure 5: Excitation of *L*(0,2) at 50 kHz.



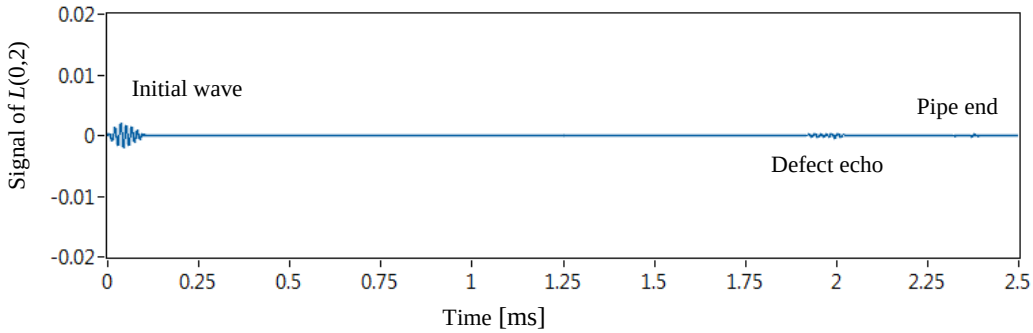
a) Width of nickel foil at λ/8



b) Width of nickel foil at λ/4



c) Width of nickel foil at λ/2



d) Width of nickel foil at λ

Figure 6: Excitation of *L*(0,2) at 60 kHz.

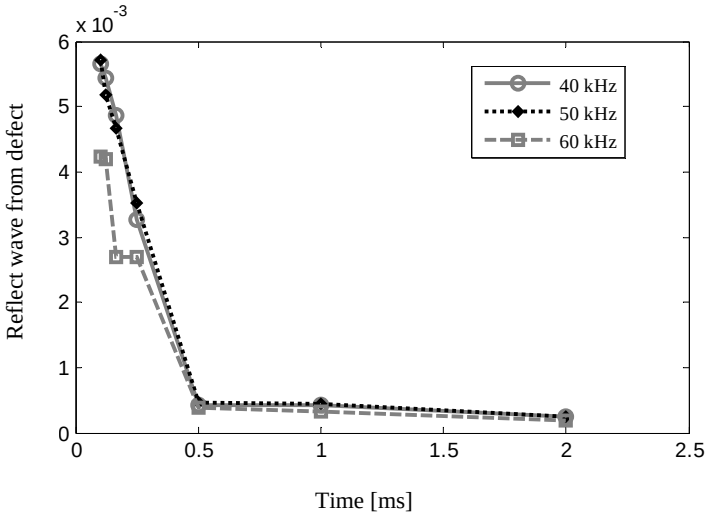


Figure 7: Reflected *L*(0,2) from defect in pulse-echo mode.

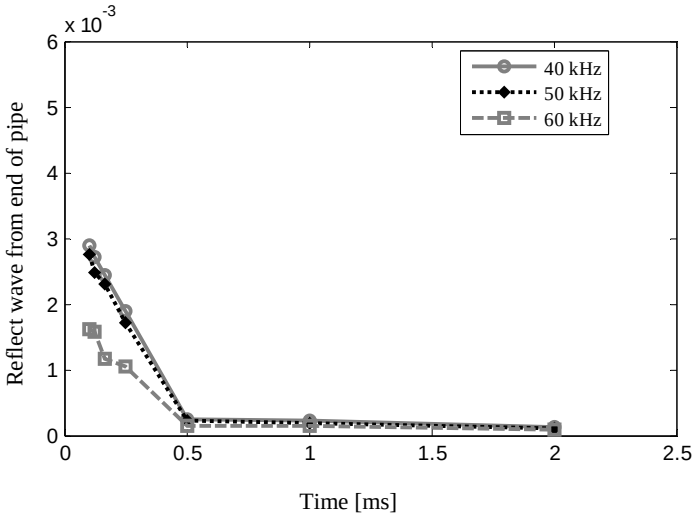


Figure 8: Reflected *L*(0,2) from end of pipe in pulse-echo mode.

Measurements of excited waves at point *P* are used to study the effect of width of nickel foil on the detection of reflected waves from defect and end of the pipe. The results indicated that the width of nickel foil that is not wider than quarter (λ/4) of the wavelength of the excited frequency is preferable for development of the magnetostrictive sensor which is consistent with previous work on magnetostrictive sensors which usually form the resonator at quarter of the wavelength (Kwun & Bartels, 1998; Vinogradov, 2009). The behaviour of the wave propagations from the nickel foil are explained from the visualization of the excited wave into defect in the pipe. Excited wave from nickel foils which are not wider than quarter (λ/4) of the wave length also demonstrate uniform wave propagation into defect compared to wave excited from the wider nickel foil on the pipe. The non-uniform wave propagation may consists of *L*(0,1) and flexural modes as well as *L*(0,2) mode in the pipe. The non-dominant of *L*(0,2) mode in the pipe resulted in very small reflected wave from defect and end of the pipe as depicted in the pulse-echo signal using nickel foils wider than the quarter (λ/4) of the wavelength. However, further studies with proper model are recommended to relate the effect of nickel width on the amplitude of the excited signals.

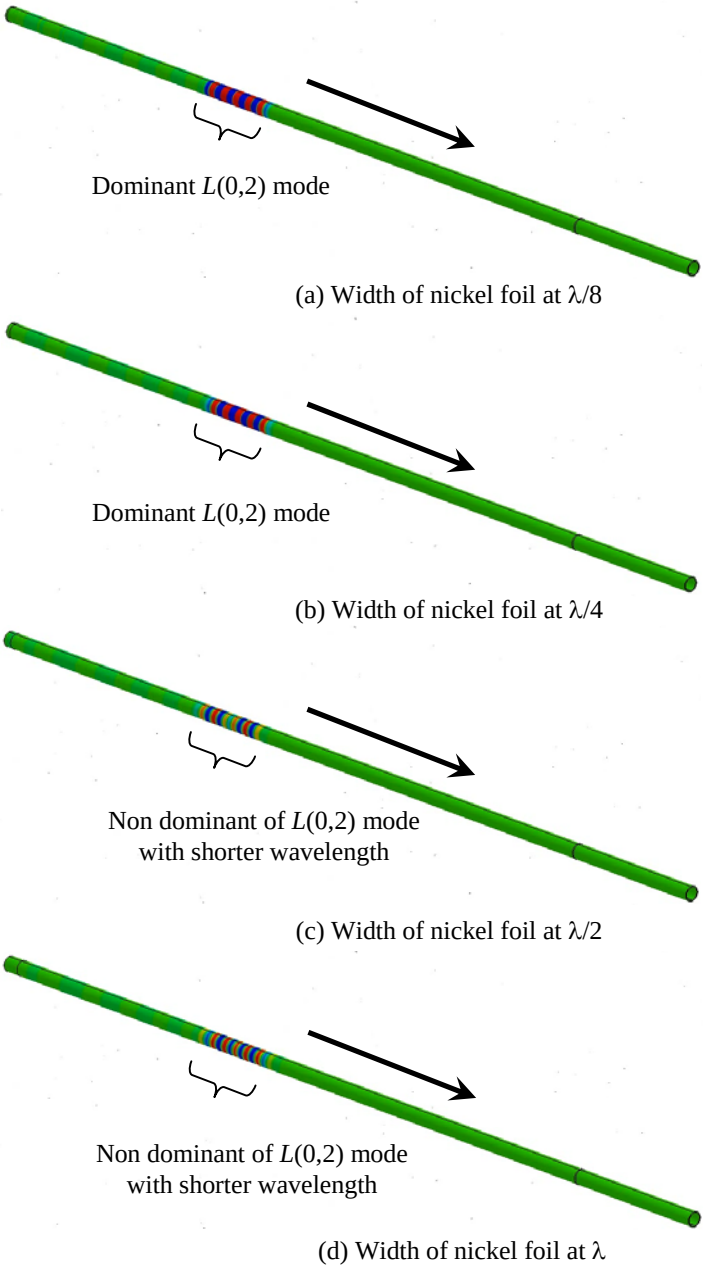


Figure 9: Snapshot of the excited guided wave at *t* = 0.5 ms for excitation at 50 kHz.

4. CONCLUSION

A study on magnetostrictive sensors was conducted through simulation of *L*(0,2) mode excitation in pipe from different widths of nickel foils at frequencies from 40 to 60 kHz. The results indicated fine wave excitations from the narrow width of nickel foils. The use of nickel foil wider than quarter wavelength (λ/4) of the intended wave has potential to excite another wave modes as well as the intended wave mode. Since, the intended mode were not excited as the dominant mode, very small reflected wave will be resulted from defect as well as the reflected wave from the end of the pipe.

ACKNOWLEDGMENTS

This work is supported by Universiti Teknikal Malaysia Melaka (UTeM) and Malaysia Ministry of Higher Education (MOHE) under research grant FRGS/1/2014/TK01/FKM/02/1/F0211, FRGS/2/2013/TK01/UTEM/02/F0171, and ERGS/1/2013/FKM/TK01/UTEM/02/01/E00014.

REFERENCES

Anghileri, M., Castelletti, L.M.L. & Tirelli, M. (2005). Fluid–structure interaction of water filled tanks during the impact with the ground. *Int. J. Impact Eng.*, **31**: 235-54.

Cawley, P., Lowe, M.J.S., Alleyne, D.N., Pavlakovic, B. & Wilcox, P. (2003). Practical long range guided wave inspection-applications to pipes and rail. *Mater. Eval.*, **61**: 66-74.

Dobson, J. & Cawley P. (2016). Independent component analysis for improved defect detection in guided wave monitoring. *Proc. IEEE*, **104**:1620-1631.

Gao, H. & Rose, J. L. (2009). Ice detection and classification on an aircraft wing with ultrasonic shear horizontal guided waves. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **56**: 334-344.

Heimbs, S., Nogueira, A. C., Hombergsmeier, E., May, M. & Wolfrum, J. (2014). Failure behaviour of composite T-joints with novel metallic arrow-pin reinforcement. *Compos. Struct.*, **110:** 16-28.

Jones, R., Peng, D., Huang, P., & Singh, R.R.K. (2015). Crack growth from naturally occurring material discontinuities in operational aircraft. *Procedia Eng.*, **101:** 227-234.

Kim, H. W., Lee, J.K. & Kim, Y.Y. (2013). Circumferential phased array of shear-horizontal wave magnetostrictive patch transducers for pipe inspection. *Ultrasonics*, **53**: 423-431.

Kim, Y., Moon, H., Park, K. & Lee, J. (2011). Generating and detecting torsional guided waves using magnetostrictive sensors of crossed coils. *NDT. E. Int.*, **44**: 145-151.

Kim, Y.Y & Kwon, Y.E. (2015). Review of magnetostrictive patch transducers and applications in ultrasonic nondestructive testing of waveguides. *Ultrasonics*, **62**: 3-19.

Koc, M. & Altan T. (2002). Prediction of forming limits and parameters in the tube hydroforming process. *Int. J. Mach. Tool Manu.*, **42**: 123 -138.

Kolkman, H.J., Kool, G.A. & Wanhill, R.J.H. (1996). Aircraft crash caused by stress corrosion cracking. *J. Eng. Gas Turb. Power*, **118**: 146-149.

Kwun, H. & Bartels, K.A. (1998). Magnetostrictive sensor technology and its applications, *Ultrasonics*, **36**: 171-178.

Kwun, H. & Teller, C.M. (1994). Magnetostrictive generation and detection of longitudinal, torsional, and flexural waves in a steel rod. *J. Acoust. Soc. Am.*, 96: 1202-1204.

Mu, J. & Rose, J.L. (2008). Guided wave propagation and mode differentiation in hollow cylinders with viscoelastic coatings. *J. Acoust. Soc. Am*, 124: 866-874.

Mueller, E. M., Starnes, S., Strickland, N., Kenny, P. & Williams C. (2016). The detection, inspection, and failure analysis of a composite wing skin defect on a tactical aircraft. *Compos. Struct.*, **145**: 186-193.

Ostachowicz, W., Kudela, P., Malinowski, P. & Wandowski, T. (2009). Damage localisation in plate-like structures based on PZT sensors. *Mech. Syst. Signal Pr.*, **23**: 1805-1829.

Purbolaksono, J., Ahmad, J., Beng, L.C., Rashid, A.Z., Khinani, A. & Ali, A.A. (2010). Failure analysis on a primary superheater tube of a power plant. *Eng. Fail. Anal.*, **17**: 158-167.

Rao, T.S. & Nair, K.V.K. (1998). Microbiologically influenced stress corrosion cracking failure of admiralty brass condenser tubes in a nuclear power plant cooled by freshwater. *Corros. Sci.*, **40**: 1821-1836.

Rose, J.L. (2004). Ultrasonic guided waves in structural health monitoring. *Key Eng. Mater.*, **270-273**: 14-21.

Saleh, J. H., Haga, R.A., Favaro, F.M. & Bakolas, E. (2014). Texas City refinery accident: Case study in breakdown of defense-in-depth and violation of the safety-diagnosability principle in design. *Eng. Fail. Anal.*, **36**: 121 - 133.

Shin, H.J. & Rose, J.L. (1999). Guided waves by axisymmetric and non-axisymmetric surface loading on hollow cylinders. *Ultrasonics*, **37**: 355-363.

Vinogradov, S. (2009). Magnetostricive transducer for torsional mode guided wave in pipes and plates. *Mater. Eval.*, **67**: 333-341.

RAPID DEFECT SCREENING ON PLATE STRUCTURES USING INFRARED THERMOGRAPHY

Nor Salim Muhammad[1*], Abd Rahman Dullah[1], Ahmad Fuad Ad Ghani[2], Roszaidi Ramlan[1] & Ruztamreen Jenal[1]

[1]Faculty of Mechanical Engineering
[2]Faculty of Engineering Technology
Universiti Teknikal Malaysia Melaka (UTeM), Malaysia

[*]Email: norsalim@utem.edu.my

ABSTRACT

This study focuses on an initial investigation in optical pulsed thermography for screening of thickness loss in plate structures. Optical infrared thermography on an aluminium plate with groove defect was conducted to simulate infrared thermography using halogen lights on the surface of material with high thermal conductivity. Thermal images were recorded to study the pattern of thermal diffusion in the plate with defect. The final results successfully visualized the groove defect from the intact surface of the plate. Temperature images indicated high temperature around the area of thickness loss during the heating process, which can be used to screen the corrosion defects in the actual structures.

Keywords: *Defect detection; optical infrared thermography; thermal anomaly.*

1. INTRODUCTION

Rapid non-destructive inspection on large structures has advantages to provide services on defect screening over large structures at low cost of human labours. Rapid developments on infrared thermal imaging sensing elements and the availability of low cost thermal imaging cameras allow many studies on thermal imaging for non-destructive inspections by many academic researchers. This kind of thermal imaging technique has advantage over other non-destructive inspection techniques as the failures are represented in the temperature images of the structures. At the same time, locations of defects can be identified from the recorded temperature distributions, which are expected to affect the non-uniformity of heat dispersion in the structures (Maldague & Marinetti, 1996; Cheng & Tian, 2011; Valiorgue *et al.*, 2013; Arora *et al*., 2015). This technique can also be used in online monitoring and off-line inspections to screen delaminated defects, voids, corrosions, cracks and weld conditions in structures. However, there are many restrictions on the real infrared thermography in order to visualise defects in structures. The effect of surface conditions, geometries of the screened structures, type of heat sources, and noises from environment are among the elements that can prevent us from having reliable thermal images (Maldague & Marinetti, 1996; Valiorgue *et al.*, 2013; Arora *et al*., 2015).

The use of infrared technology in developing countries is also relatively low. These technologies are used in border surveillance, antiaircraft facilities and infrared guided missiles (Bell & Glasgow, 1999; Mahulikar *et al.*, 2001; Smith *et al*., 2004) as well as in landmine zones (Khanafer *et al.*, 2003; López *et al.*, 2004; Deans *et al.*, 2006), but not widely used in the industries. At the same time, the use of infrared technology in machine condition monitoring and structural health monitoring can solve many industrial problems instantly and can be utilised for maintenance of military equipment and facilities as well. The array of infrared sensors that is equipped in infrared cameras can also be used to scan for temperature changes during the operation of machine

components, consisting of electrical circuits, motors, and bearings (Kim *et al.*, 2010; Bagavathiappan *et al.*, 2013). The sensors are also suitable for inspections on static structures, such as pressure vessels, pipes and aircraft components (Yang *et al.*, 2001; Rajic & Rowlands, 2010; Infante *et al.*, 2014; Endo & Kusaka, 2015). Infrared technology also has great potential in non-destructive testing, where it can be used to screen defects in structures faster than the conventional ultrasonic method. Thermal imaging cameras are also generally used to record thermal wave diffusion on structures. Access to temperature images from inspections are important to provide information of defect locations in the structures without using complicated signal processing as required in acoustic emission techniques for pipes and concrete structures (Mostafapour & Davoudi, 2013; Elfergani *et al*, 2015).

There are many studies on passive thermal imaging inspections for machine condition monitoring and buildings, including electrical circuit failures and leakages (Kim *et al.*, 2010; Balaras & Argiriou, 2013; Bagavathiappan *et al.*, 2013). Active infrared thermography techniques, which utilises thermal changes on the structures from external sources such as halogen lamps, have been studied for defect inspection in carbon fibre reinforced polymers and concrete structures (Brown & Hamilton, 2012; Arora *et al.*, 2015). They used halogen lights to heat the samples and obtained thermal images from the surface of the samples for the non-destructive defect inspection. However, reflections from surfaces of the samples might affect the results obtained from the infrared thermography especially when dealing with polished or shining surfaces. The problem will be more complicated when dealing with small temperature differences in high thermal conductivity materials.

The difference in thermal distribution for metal structures with defects is also expected to be lower than the non-metal structures, which indicates difficulties in active thermography. This study was carried out to use the thermal diffusion in an aluminium plate that is radiated using halogen lights for defect inspection using an infrared camera. Its objective is to visualise defects from the intact surface of the painted metal structures through the behaviour of non-uniform thermal diffusion across the thinned area for defect representation in real structures.

2. EXPERIMENTS

Two halogen lights at 150 W were used to heat an aluminium plate of 1,000 mm × 500 mm × 6 mm with a groove defect at depth and width of approximately 4 mm and 20 mm respectively. The specimen was arranged as in Figure 1, where the intact surface was placed to be on the top of the specimen. The aluminium plate with groove defect is shown in Figure 2 with residual thickness of 2 mm over the defect region. The intact surface above the surface with defect of the specimen was also painted with black paint to increase the emissivity of the infrared beam. The two halogen lights were tilted at 45° and placed at 500 mm on the left and right sides of the specimen as depicted in the figure, while the infrared camera was fixed at an angle of about 90° over the specimen. The infrared camera had vertical distance about 1 m from the painted surface of the specimen and placed in a direction that was not in the same path to light beams from the halogen lights. The autofocus and video streaming of the infrared camera were controlled from a host computer using a thermal imaging software from FLIR. The infrared camera was used to record temperature images of the defect region caused by the heat source to investigate the feasibility of defect screening using the thermal imaging technique. The thermal images were recorded in rainbow colour scheme to represent the temperature increase caused by the halogen lights for approximately up to 40 s after the heating process was started. The thermal images later than 40 s were not been recorded because of the temperature difference between the defect and intact regions that became smaller as the time increased longer than 40 s. Three points of temperature measurements indicated as Sp1, Sp2, and Sp3 were placed on the visualised area on the plate to plot temperature changes on the intact and defect surfaces when the plate exposed to the induced heat source from the lights. Points Sp1 and Sp2 were measurement points over the intact surfaces, whereas Sp3 was the measurement point over the defect region in the specimen.

Image processing also been applied in LabVIEW, as in Figure 3, in order to reduce noise in the recorded temperature images. Averages from five image frames were used to construct a new thermal image with lower level of noise for visualisation of the groove defect between 10 to 40 s of the heating process.

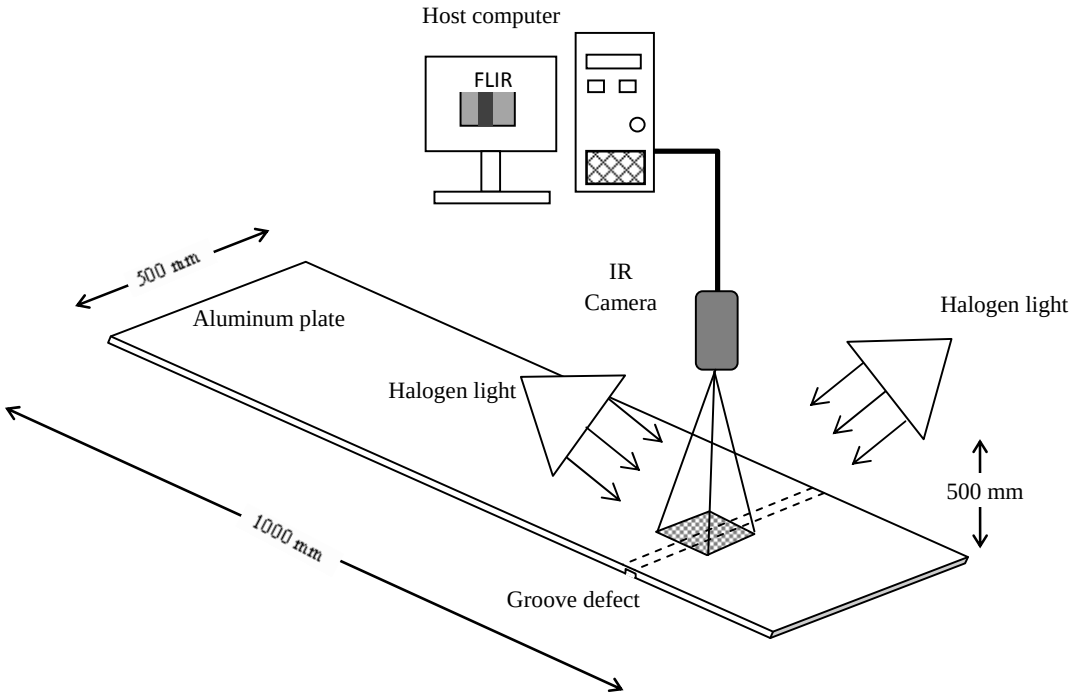


Figure 1: Optical infrared setup for defect visualisation.

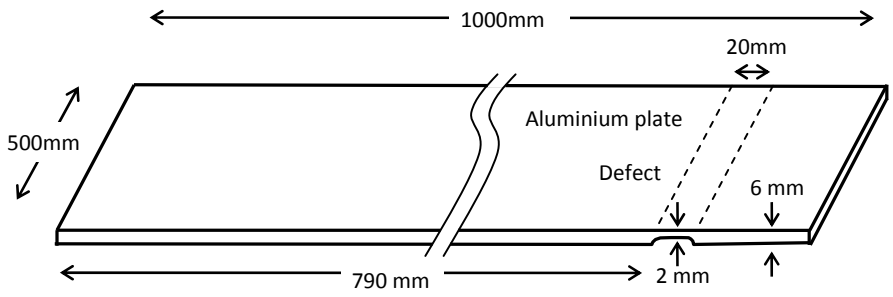


Figure 2: An aluminium plate with defect used in the study.

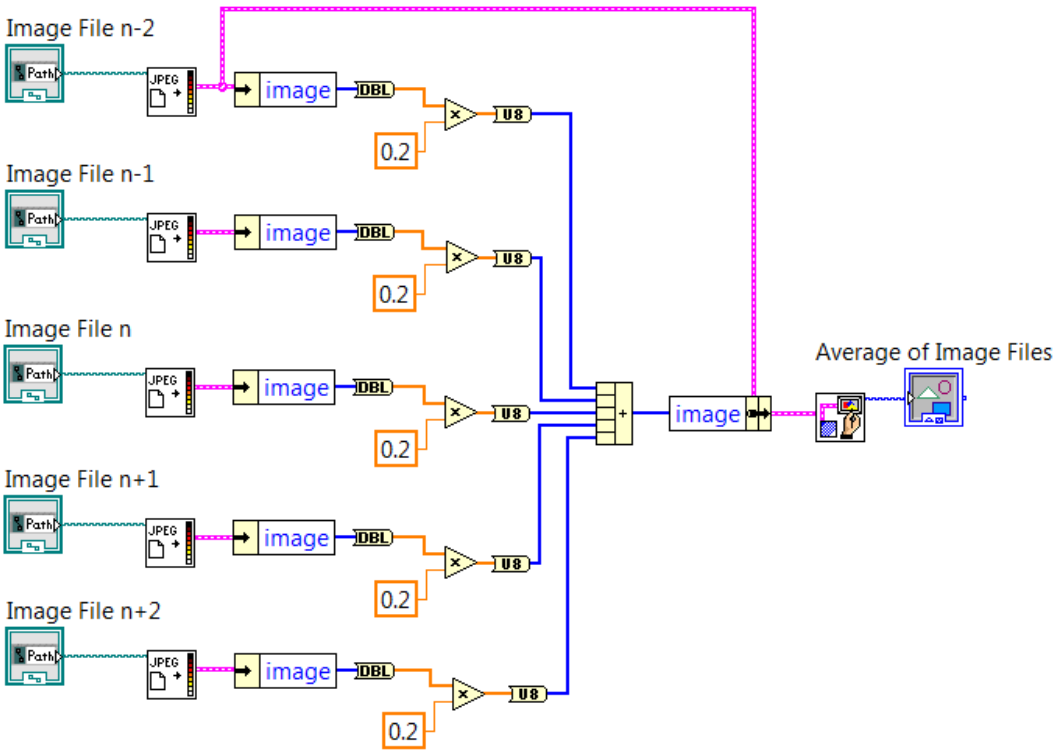


Figure 3: Calculation of average images in LabVIEW.

3. RESULTS AND DISCUSSION

3.1 Temperature Images and Temperature Increase in the Plate

Thermal images on the monitored area of the aluminium plate were recorded at every 5 s from 10 to 45 s after thermal radiation from the halogen lights as shown in Figure 4. The temperature images were obtained from a rainbow colour scheme with minimum and maximum threshold fixed at 23 and 25°C respectively. It is clearly observed that a relatively high temperature spot was visualised in the temperature images along the groove defect captured from the infrared camera. The presence of the defect on the back surface of the plate became more visible as the duration of the halogen lights increased. However, temperature image at 40 s begun to indicate high temperature on the intact surface as the plate received more thermal energy due to the increase in duration of the thermal radiation.

Temperature increase on the plate surface was monitored from three points at locations Sp1, Sp2, and Sp3 as in Figure 5, where the beam radiation from the halogen lights increased the plate temperature from about 23.3 to 24.5 °C within the recorded time of 40 s. The dark solid line in the figure represents the monitored temperature at point Sp2 in the region of the groove defect, while the gray dash line and light gray solid line show the temperature changes at points Sp3 and Sp2 in the intact region respectively. The increase of temperature in this optical pulsed thermography indicated higher temperature in the defect region as compared to the intact region as depicted in Figures 4 and 5. At the same time, the indicated temperature difference between the defect and intact regions (zones A and B) of the aluminium plate was only less than 0.2 °C. Therefore, the rainbow colour scheme was selected to be in between 23 and 25 °C. The 2 °C temperature range was used in this infrared thermography study to provide high sensitivity for the thermal wave visualisation across the monitored area.

However, the effects of surface conditions on the obtained temperature images were also observed on the right side of point Sp1 in Figure 4. This is probably due to the non-uniform condition on the painted surface, which affects the emissivity of the infrared beam. The non-uniform surface on the painted area can reflect the beam at slightly different behaviour than the uniform surface and indicated higher temperature (Valiorgue *et al.*, 2013).

At the same time, the temperature overheating prevented us from visualising the defect using infrared thermography due to the condition of temperature that reached the uniform high temperature in the whole structure, resulting in the temperature difference in defect areas becoming too small to be detected. Therefore, the halogen lights were turned off after 40 s of the heating process.

3.2 Enhancement of Temperature Images

The averaged temperature images generated using LabVIEW are shown in Figure 6. The results represent temperature distributions with lower noise as compared to the raw data in Figure 4 (c) to 4 (f). The averaged images were computed based on the effect of two frames of the previous and following frames respectively, as explained in Figure 3. The use of the image frames in Figure 4 allowed image reconstruction from 20 to 35 s as shown in Figure 6.

The averaged images in Figure 6 represent better temperature distributions on the groove and intact regions. Significantly high temperature distribution was observed along the groove region in the early 20 s of the averaged image. The groove region with thinner residual thickness recorded higher temperature as the temperature images were averaged at 35 s. This is agreement with the results from previous studies in non-metal structures (Brown & Hamilton, 2012; Arora *et al.*, 2015). The thermal wave propagation into the intact region also indicated higher temperature in the groove region as the duration of the thermal radiation increased.

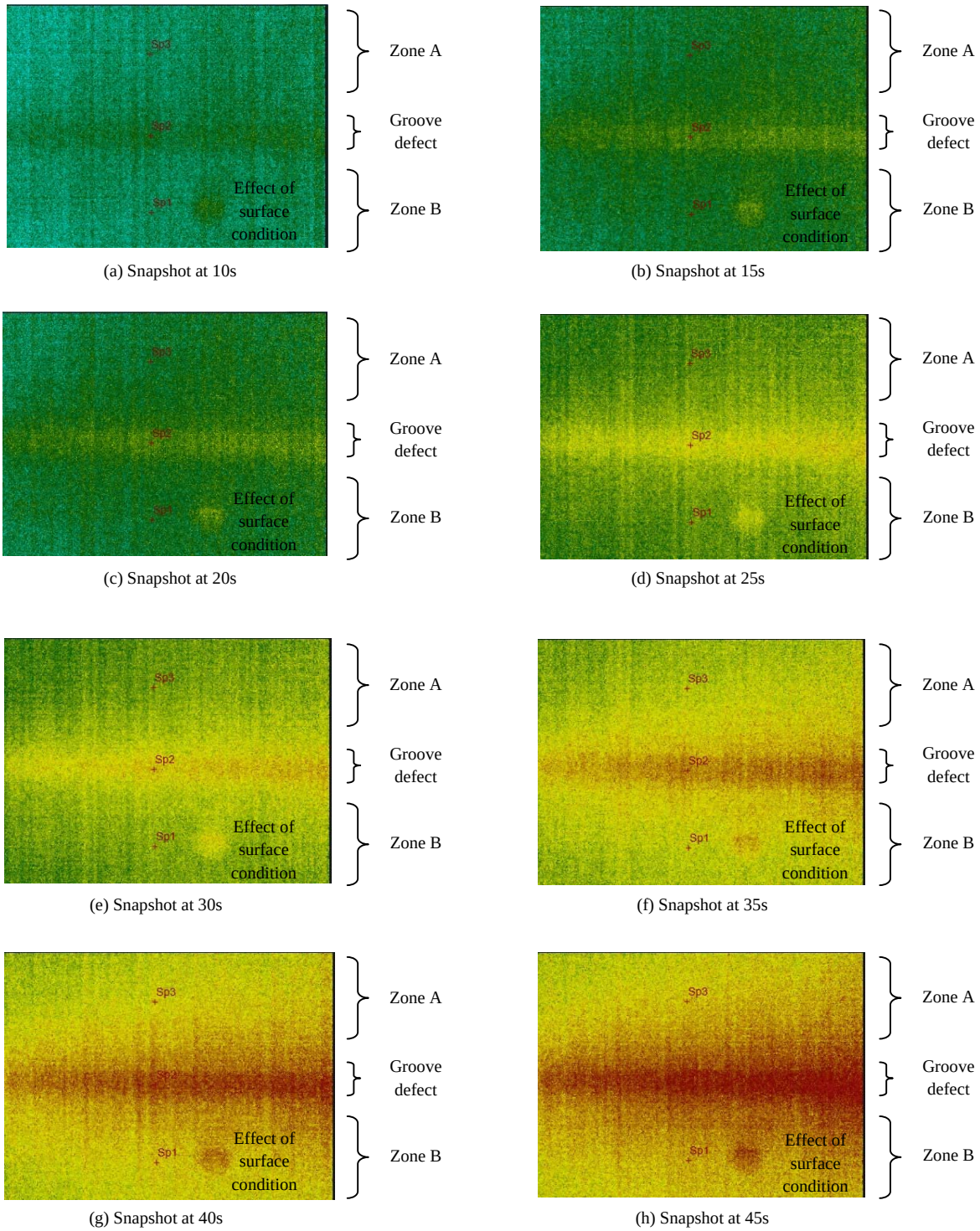
The non-uniform thermal wave propagation in structures with defect can be used to relate the thickness distributions or corrosion defects in structures. The technique shows feasibility for rapid defect screening on large structures like storage tanks, pressure vessels, and civil structures.

4. CONCLUSION

Optical infrared thermography was conducted on an aluminium plate with a groove defect. The increase of plate temperature due to heat radiation from halogen lights was used in visualising the defect location at the bottom surface of the plate. The results also demonstrated the feasibility of image processing for the recorded thermal images from the infrared thermography over large structures. The enhanced thermal images in plate indicate non-uniform thermal diffusions across the defect, whereas the thin area shows higher temperature as compared to the intact region of the plate. The presence of defects can be identified from the regions of high temperature contours that were recorded during the inspections.

ACKNOWLEDGMENTS

This work was supported by Universiti Teknikal Malaysia Melaka (UTeM) and the Malaysian Ministry of Higher Education (MOHE) under research grant FRGS/1/2014/TK01/FKM/02/1/F0211, FRGS/2/2013/TK01/UTEM/02/F0171, and FRGS/2/2013/TK01/UTEM/02/F00174.



(a) Snapshot at 10s

(b) Snapshot at 15s

(c) Snapshot at 20s

(d) Snapshot at 25s

(e) Snapshot at 30s

(f) Snapshot at 35s

(g) Snapshot at 40s

(h) Snapshot at 45s

Figure 4: Temperature images of the inspected area on plate.

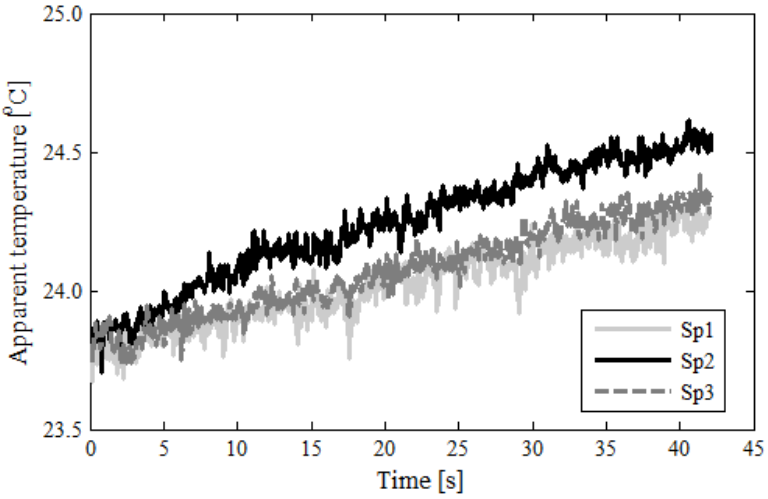


Figure 5: Monitored temperatures at points with different thickness condition.

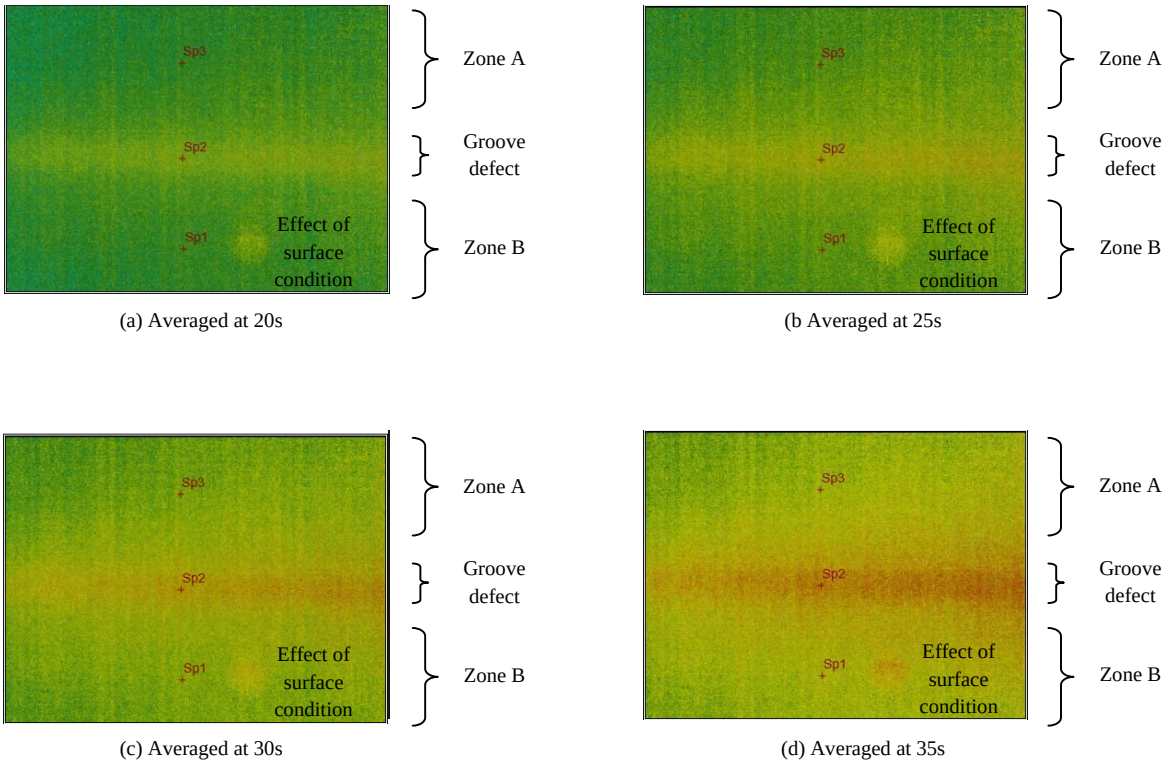


Figure 6: Averaged images computed from five frames of thermography results on the plate.

REFERENCES

Arora, V., Siddiqui, J.A., Mulaveesala, R. & Muniyappa, A. (2015). Pulse compression approach to nonstationary infrared thermal wave imaging for nondestructive testing of carbon fiber reinforced polymers. *IEEE Sens. J.*, **15**: 663-664.

Bagavathiappan, S., Lahiri, B. B., Saravanan, T. & Jayakumar, T. (2013). Infrared thermography for condition monitoring – A review. *Infrared Phys. Technol.*, **60**: 35-55.

Balaras, C.A. & Argiriou, A.A. (2002). Infrared thermography for building diagnostics. Energy and Buildings, Vol. **34** : 171-183.

Bell, W.A. & Glasgow, B.B. (1999). Impact of advances in imaging infrared detectors on anti-aircraft missile performance. *In* Holst, G.C. (Ed.), Infrared Imaging Systems: Design, Analysis, Modelling, and Testing, *Proc. SPIE*, **3701**: 244-253.

Brown, J. R. & Hamilton, H. R. (2012). Quantitative infrared thermography inspection for FRP applied to concrete using single pixel analysis. *Constr. Build. Mater.*, **38:** 1292-1302.

Cheng, L. & Tian, G. (2011). Surface crack detection for carbon fiber reinforced plastic (CFRP) materials using pulsed eddy current thermography. *IEEE Sens. J.*, **11**: 3261-3268.

Deans, J., Gerhard, J. & Carter, L. (2006). Analysis of a thermal imaging method for landmine detection, using infrared heating of the sand surface. *Infrared Phys. Tech.,* **48**:202-216.

Elfergani, H. A., Pullin, R. & Holford, K. M. (2013). Damage assessment of corrosion in prestressed concrete by acoustic emission. *Constr. Build. Mater.*, **40**: 925-933.

Endo, H. & Kusaka, T. (2015). Efficient inspection for gas pipes by infrared thermography. *KOBELCO Technol. Rev.*, **33**: 50-55.

Infante, V., Reis, L. & Freitas, M. (2014). Failure analysis of landing gears trunnions due to service. *Eng. Failure Anal.*, **41**: 118-123.

Khanafer, K., Vafai, K. & Baertlein, B. A. (2003). Effects of thin metal outer case and top air gap on thermal IR images of buried antitank and antipersonnel land mines. *IEEE Trans. Geosci. Electron.*, **41**: 123-135.

Kim, D, Yun, H., Yang, S., Kim, W. & Hong D. (2010). Fault diagnosis of ball bearings within rotational machines using the infrared thermography method. *KSNT*, **30**: 558-563.

Mahulikar, S. P., Sane, S. K., Gaitonde, U. N. & Marathe, A. G. (2001). Numerical studies of infrared signature levels of complete aircraft. *J. Aeronaut. Sci.*, **105**: 185–192.

Maldague, X. & Marinetti, S. (1996). Pulse phase infrared thermography. *J. Appl. Phys.*, **79**: 2694-2698.

Martinez P.L., Kempen, L.V., Sahli, H. & Ferrer, D.C. (2004). Improved thermal analysis of buried landmines. *IEEE Trans. Geosci. Electron.*, **42**: 1965–1975.

Mostafapour, A. & Davoudi, S. (2013). Analysis of leakage in high pressure pipe using acoustic emission method. *Appl. Acoust.*, **74** : 335-342.

Rajic, N. & Rowlands, D. (2010). *The Rapid Detection of Structural Flaws in the MRH-90 Helicopter Using Infrared Thermography – A Field Demonstration and Business Case*. Technical Report DSTO-TR-2467, Defence Science and Technology Organisation, Australia.

Smith, M.I., Heather, J.P., Ralph, J.F., Bernhardt, M., Griffith, E.J., Bradley, D.J. & Padda, H.S. (2004). Target tracking for missile warning applications, signal and data processing of small targets. *Proc. SPIE*, **5428**: 282-293.

Yang, B., Liaw, P.K., Wang, H., Jiang, L., Huang, J.Y., Kuo, R.C. & Huang, J.G. (2001). Thermographic investigation of the fatigue behavior of reactor pressure vessel steels. *Mater. Sci. Eng. A*, **314**: 131-139.

Valiorgue, F., Brosse, A., Naisson, P. & Rech, J. (2013). Emissivity calibration for temperatures measurement using thermography in the context of machining. *Appl. Therm. Eng.*, **58**: 321-326.

OPTIMISATION OF ELECTRODEPOSITION PARAMETERS ON THE MECHANICAL PROPERTIES OF NICKEL COBALT COATED MILD STEEL

Nik Hassanuddin Nik Yusoff[1,2*], Othman Mamat[2], Mahdi Che Isa[2] & Norlaili Amir[1]

[1]Mechanical Engineering Department, Universiti Teknologi PETRONAS (UTP), Malaysia
[2]Maritime Technology Division (BTM), Science & Technology Research Institute for Defence (STRIDE), Ministry of Defence, Malaysia

[*]Email: nikhassanuddin@yahoo.com

ABSTRACT

This paper discusses the effect of electrodeposition parameters of deposited nickel (Ni) cobalt (Co) from alkaline bath on commercial marine-grade mild steels. Prior to the deposition process, the Ni-Co samples were subjected to a two-level factorial design of experiment (DoE) to optimise the electrodeposition parameters, namely current density, electrolyte temperature and stirring speed. These electrodeposition parameters potentially affect the following responses of the coatings: microhardness, current efficiency and Co content. Current density contributes major outcome on the coating properties. Microhardness, current efficiency and Co content were significantly influenced by the current density. However, by changing the electrolyte temperature, a significant effect on current efficiency was noted. In conclusion, Ni-Co coated steels, with the highest current density of 50 mA/cm^2 and the highest electrolyte temperature of 50 $^\circ$C, were most preferred.

Keywords: *Electrodeposition; nickel-cobalt; alkaline bath; microhardness; current efficiency.*

1. INTRODUCTION

Electrodeposition or electroplating is a process of incorporating metals to the substract by using the electrochemical method. The ability of the technique to coat various types of materials, such as oxide, nitride, carbide and metallic, in a metal matrix allows it to provide coatings for multiple applications (Bakhit & Akbari, 2013). The advantages of this technique have encouraged researchers to further explore its potentials.

In recent years, there have been increasing interest focused on electrodeposited nickel-cobalt (Ni-Co) and Ni-Co based metal matrix composite (MMC) coatings due to their superior properties, such as higher hardness (Srivastava *et al*., 2006; Shi *et al*., 2006; Ranjith & Paruthimal Kalaignan, 2010; Srivastava *et al*., 2010), improved anti-wear, better corrosion resistance and oxidation resistance (Cai *et al*., 2015), as compared with pure Ni and Ni based composite coatings. The properties of Ni-Co or Ni-Co based composite coatings are mainly dependent on the incorporated particles and microstructures of the Ni-Co matrix (Kuo *et al*., 2004a; Low *et al*., 2006; Sui *et al*., 2007; Nguyen *et al*. 2013; Cai *et al*., 2015). Studies on Ni-Co coating suggested that the properties of this coating strongly depends on their cobalt content (Srivastava *et al*. 2006; Bakhit & Akbari, 2013). Controlling the cobalt content is necessary in order to obtain the best coating performance. The amount of cobalt can be controlled by experimental parameters, such as electrolyte composition, nickel cobalt ratio, temperature, agitation, current density and pH (Srivastava *et al*., 2006).

Ni-Co commonly deposited from acid bath which the solution pH is varied between 3.5 to 5 (Tian & Cheng, 2007; Zhang *et al*., 2014). The other applicable method to deposit metal coating are using alkaline bath as widely reported to Zn-Fe, Zn-Ni and Fe coating (Long *et al*., 2013; Feng *et al*., 2015; Lan *et al*., 2006). This alkaline bath can be used for the steel part of the complex shape due to uniform metal distribution in the deposit. However, it is difficult to develop an eco-friendly alkaline bath with high current efficiency (Feng *et al*., 2015). The complexing agents are the key factor to obtain a stable

alkaline bath. Different types of complexing agents have their own result on composition, structure and properties of coating.

Conversely, limited data is available in the literature concerning the effect of alkaline bath to the Ni-Co coating properties. In the present work, Ni-Co coating is deposited from an alkaline bath. The electrodeposition parameters are subjected to a two-level factorial design of experiment (DoE) due to the possible interaction between factors. The initial step in this DoE is to obtain the optimum processing parameters. The typical variables focused in this study are current density, electrolyte temperature and stirring speed. The aim of the current research is to carry out a factorial design study to determine the main effect, as well as the interaction effects, of several parameters on the mechanical properties of electrodeposited Ni-Co on commercial mild steel substrates. These properties are hardness, current efficiency and Co content.

2. EXPERIMENTAL PROCEDURE

Table 1 shows the bath composition and concentration that was used in this study. The substrate was prepared from 5 cm^2 of ASTM A36 mild steel plates, polished to a mirror finish surface by different grades of abrasive paper from 1,000 to 2,400, degreased with 20 wt.% sodium hydroxide (NaOH) solution in an ultrasonic bath for 5 min, then activated with 20% hydrochloric acid (HCl) for 60 s. These were used as the cathode, which were weighted before and after electrodepositing. A pure nickel plate was used as an anode. The required pH of the bath was adjusted to 9–10 by 10% hydrogen sulphate (H_2SO_4) solution. A hot plate that is equip with mechanical stirrer was used for bath agitation and temperature controlled. An Energy Dispersive X-ray Analyzer (EDX), equipped with Phenom Pro scanning electron microscope (SEM), was used for elemental analysis. In order to measure the microhardness of the coatings, a HMV Shimadzu microhardness tester was used with the indentation load of 0.25 N and indentation time of 10 s. The average value of ten different measurements was reported as the microhardness of the coating. The current efficiency (C.E) was estimated according to the Faraday law as ascribed by Sharifi *et al*. (2009). It was calculated by as follows:

$$C.E = M_e / M_t \quad (1)$$

$$M_t = AIt/nF \quad (2)$$

where M_e is experimental weight measured from deposit gain, A is sum of atomic weight ($g.mol^{-1}$) of metal, I is current supply in ampere, t is deposition duration in second, n is number of electrons transferred per atom of each metal, and F is the Faraday's constant (96,485 C·mol−1).

Table 1: The bath composition for the electrodeposition process.

Bath composition	Value
$NiSO_4.6H_2O$	0.25 M
$CoSO_4.7H_2O$	0.25 M
Triethanolamine (TEA)	1 M
K_2CO_3	0.7 M
$Na_4P_2O_7.10H_2O$	0.1 M
pH	9-10

The DoE was carried out using Design Expert (Version 6.0.10, Stat-Ease, Inc., Minneapolis). Current density (X_1), electrolyte temperature (X_2), and stirring speed (X_3) were chosen as independent parameters, whereas the studied responses were microhardness value (Y_1), current efficiency (Y_2), and Co content (Y_3). The independent parameters, experimental range and coded levels for plasma spray

coating formation are presented in Table 2. A 2^3 factorial design for the three independent variables was carried out. Adding one centre point to the design led to nine sets of experiments. Each response was modelled based on the three parameters using the following equation:

$$Y_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 + b_{123}X_1X_2X_3 \quad (3)$$

where Y_i is the theoretical response function (Yusoff *et al*., 2012).

Table 2: Experimental responses and levels of independent parameters for the electrodeposition process.

Deposition parameters	Range (coded level)		
	Low (-1)	Central point (0)	High (1)
Current density (X_1), mA/cm^2	10	30	50
Stirring speed (X_2), rpm	200	400	600
Temperature (X_3), ^{0}C	30	40	50

3. RESULTS AND DISCUSSION

3.1 Effect of Deposition Parameters to Ni-Co Coating Mechanical Properties

Table 3 presents the experimental responses measured for parameters at two levels factorial. Polynomial regression models were formulated using a factorial design to analyse the consequence of each parameter and their interactions with other parameters on each response. The important effects and their interaction were selected by considering the half-normal probability plots. Figure 1 shows how these variables were selected. Values positioned away from the straight line and at the right side of the plot were selected in the ANOVA calculations.

Table 3: Full-factorial design for electrodeposition of Ni-Co parameters.

Samples	Parameters			Response		
	X_1	X_2	X_3	Y_1	Y_2	Y_3
1	1	-1	1	541.5	9.99	51.7
2	-1	1	-1	480.2	14.22	67.1
3	1	1	-1	526.0	5.94	59.5
4	-1	-1	1	499.0	17.86	64.1
5	-1	-1	-1	325.6	14.58	69.1
6	1	1	1	621.6	10.06	64.3
7	1	-1	-1	516.3	2.04	56.6
8	0	0	0	522.0	10.09	62.5
9	-1	1	1	462.9	14.58	68.7

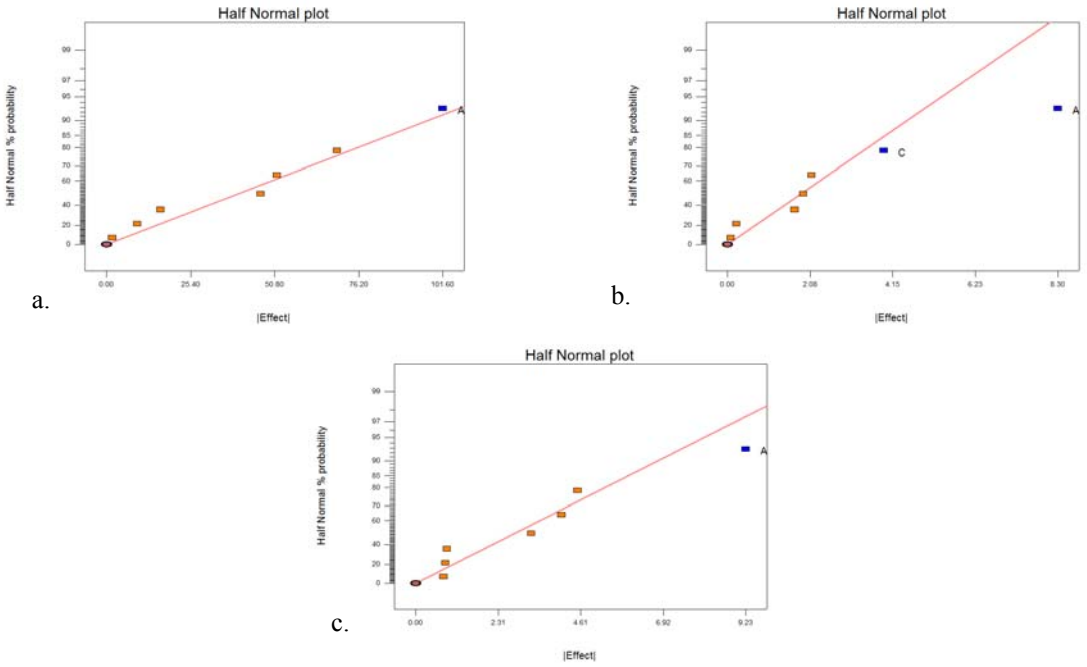


Figure 1: Half normal plots used for selecting the main effects (a = Y_1, b = Y_2 and c = Y_3).

The sum of square (*SS*), coefficients of the models, and probability (*P*) value for the two factorial designs are presented in Table 4.

Table 4: Results of ANOVA for different responses.

Term	*SS*	Coefficient	*P* value
(a) Y_1 (microhardness)			
b_0		424.35	
X_1	20645.12	2.54	0.0475
(b) Y_2 (current efficiency)			
b_0		9.53	
X_1	137.86	-0.21	0.0025
X_3	30.85	0.2	0.0453
(c) Y_3 (Co content)			
b_0		69.56	
X_1	170.2	-0.23	0.0183

The *P*-value represents the probability of error involved in accepting the observed value (Jarrah, 2009). A smaller value of *P* demonstrates higher significance of the model. *P*-value smaller than 0.05 indicates that the model is significant, with the confidence level of 95%. The current density exhibited significant result on the microhardness and Co content, while the other parameters had insignificant contribution to those responses. On the other hand, the current density and electrolyte temperature demonstrate significant influences to the current efficiency. Conversely, the interaction between them are not significant to the efficiency of the electrodeposition process. The final models in actual parameters, with the exception of the insignificant terms (with $P > 0.05$), are given in the following equations:

$$Y_1 = 424.35 + 2.54*X_1 \quad (4)$$

$$Y_2 = 9.53 - 0.21*X_1 + 0.2*X_3 \quad (5)$$

$$Y_3 = 69.56 - 0.23*X_1 \quad (6)$$

3.2 Effects on Microhardness Profile

Figure 2 shows the influence of varying current density from 10 to 50 mA/cm^2 on microhardness values of the coatings. Obviously, increasing the current density has a positive effect on microhardness, raising the coating microhardness from 460 to 551 HV. Other factors such as electrolyte temperature and stirring speed seem not offer a significant effect on the microhardness value. The influence of current density on microhardness value is demonstrated in Equation 4. A few factors have been discussed previously that influence the microhardness value. Bakhit (2015) and Marita *et al*. (2014) suggested that grain size gives a better effect to the microhardness value. According to the Hall-Petch relation, the nanocrystalline structure possibly will improve the coating microhardness (Pande & Cooper, 2009). However, the rise of microhardness value at a higher electrodeposition current in this study is more related to the percentage of the Co content in the coating. Figure 3 shows that the distribution of Co wt.% as a function of current density. Low current density applied during deposition process would increase the Co wt.% more than 60 wt.%, thus reducing their microhardness value. Higher Co concentration would change the coating crystal structure from face centred cubic (*fcc*) to hexagonal close-packed (*hcp*) as reported by Srivastava *et al*. (2006). Consequently, predominant *hcp* phase reduced the microhardness value of the coating (Cai *et al*., 2015).

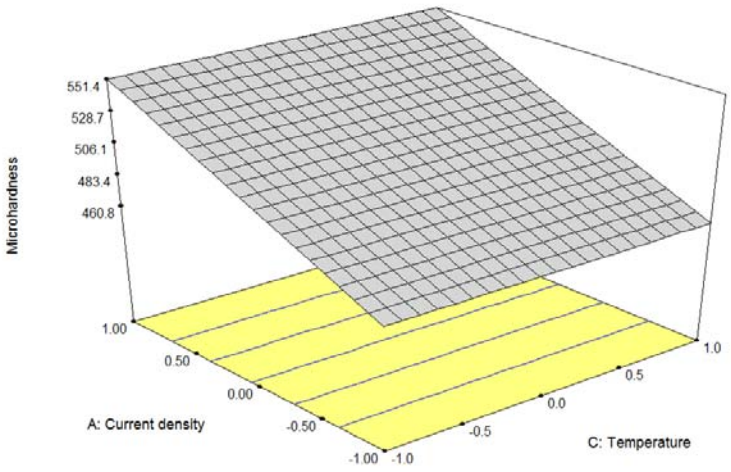


Figure 2: Effect of current density and electrolyte temperature on microhardness properties.

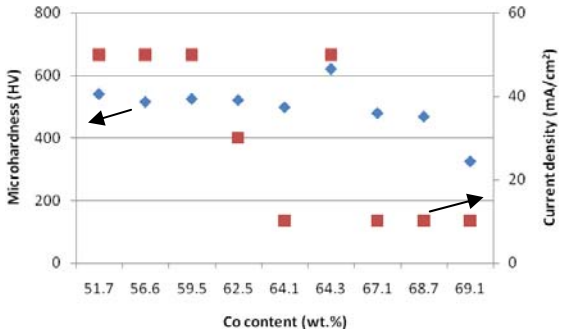


Figure 3: The distribution of Co content as a function of current density and their microhardness values.

3.3 Effects on Current Efficiency

Equation 5 demonstrates the significant factors affecting the current efficiency of the deposition process. Current density and electrolyte temperature are the main effects that influence the current efficiency positively. This finding indicates that the current efficiency decreases when the current density is increased. While lessening electrolyte temperature also reduced the current efficiency as showed in Figure 4.

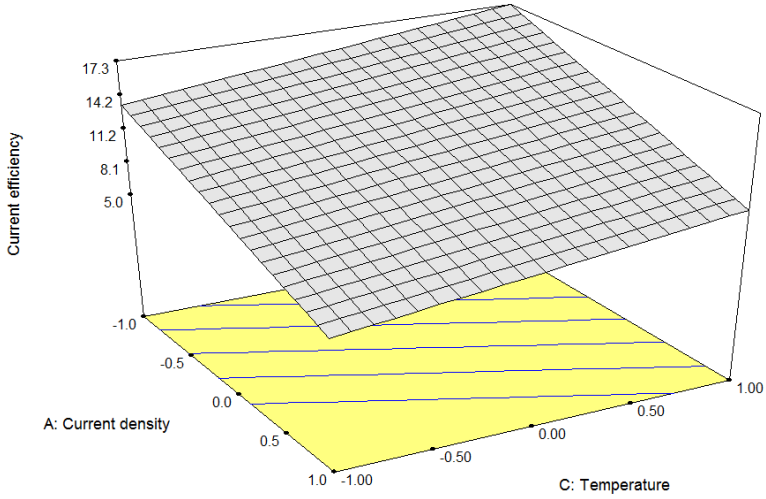


Figure 4: Effect of current density and electrolyte temperature on deposition current efficiency.

The highest current efficiency on this deposition process could be obtained with low current density together with high electrolyte temperature. Low current density used produced low overpotential deposition, which does not occupy the reduction of water. Hydrogen evolution that occurs on the cathode electrode prohibited the active surface for electrochemical reaction. As a result, reduction of Ni and Co were diminished. This finding is in good agreement with other researchers, found that the decline of Ni-Co deposition at high current density was due to intensive providing of hydrogen evolution (Kuo *et al*., 2004b; Jović *et al*., 2007; Rafailović & Minić, 2009). On the other side, high temperature applied to the electrolyte would increase the kinetic energy of the ions, which enhance their movement (Tian & Cheng, 2007). Therefore, increasing temperature results in better deposition rate (Low, 2006).

3.4 Effects on Co Content

Figure 5 reveals the key factors affecting the Co concentration of deposited coatings. Again, the current density is the major effect that influences the Co concentration. Increasing the current density from 10 to 50 mA/cm^2 would decrease the Co content within the coating from 67 to 58 % wt. The reduction of the Co content can be described using Equation 6. This finding indicates that electrodeposition of Ni-Co coating conducted at low current density exhibits anomalous behaviour in that the ratio of Co/[Ni+Co] within the coating exceeds the ratio of Co/[Ni+Co] in the electrolyte. A commonly accepted model for anomalous deposition is based on the kinetic factor that made the $Ni(OH)^+$ deposition rate lower than $Co(OH)^+$. The preferential deposition of $Co(OH)^+$ over $Ni(OH)^+$ not only slows the process but also blocks the access of $Ni(OH)^+$ to the substrate (Vazquez-Arenas, 2012b). Increase of the current density would reduce the anomalous behaviour of the coating. In order to explain the deposition mechanism, the cyclic voltammogram of Ni-Co in the alkaline bath (Figure 6) shall be referred to.

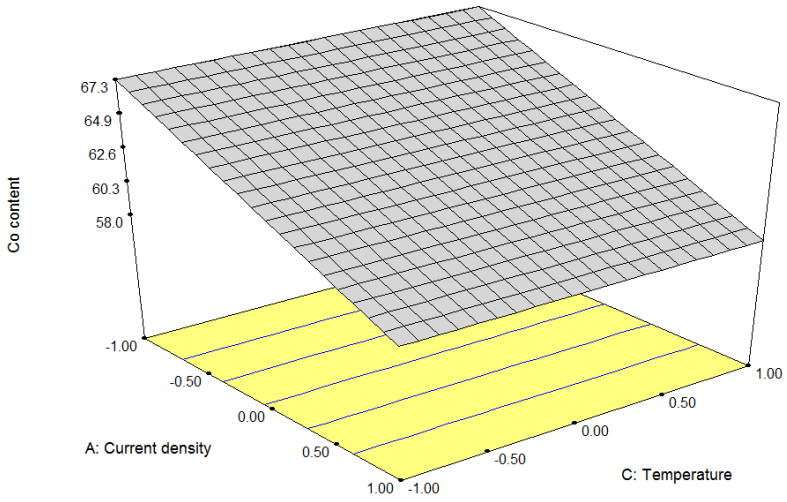


Figure 5: Effect of current density and electrolyte temperature on Co content within the coating.

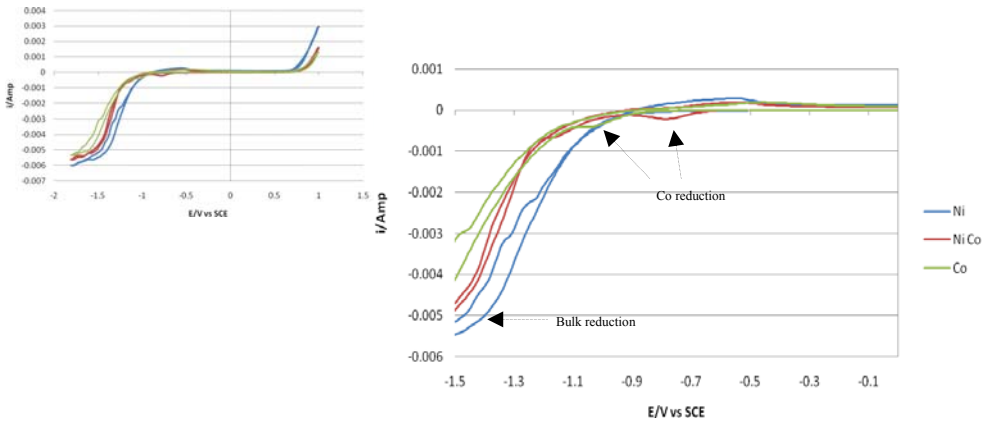


Figure 6: Cyclic voltammetry test on Ni, Co and mix Ni-Co solution with complexing and buffering agents.

The reduction potential of Co in Ni-Co deposition is lower than Ni. Consequently, reduction of Co ions is preferred in the deposition process due to more positive potential. Enhancing the current density of the deposition process will definitely increase the deposition overpotential. As shown in Figure 6, current density at 57 mA/cm^2 is the point of bulk reduction for Ni-Co and water to occur. High overpotential would increase mass transport of Ni ions. Therefore, the preferential deposition of Co was reduced as the anomalous behaviour of the coating also reduced. This finding is supported by Vazquez-Arenas (2012b) and Vazquez-Arenas *et al*. (2012), who found that anomalous behaviour of Ni-Co coating from acidic bath is strongest at low overpotential and tends to diminish as overpotential increases.

4. CONCLUSION

Electrodeposition of Ni-Co coating was successfully deposited on mild steel substrates prepared in alkaline bath. The effects of electrodeposition parameters on the properties of the coatings were studied with respect to the current density, electrolyte temperature and stirring speed using a two-level factorial DoE. When the current density increases, there is a significant increase in the coating microhardness value. Changing the electrolyte temperature and stirring speed has an insignificant effect on microhardness. However, electrolyte temperature has significant influence on the current efficiency. A significant effect on the current efficiency was also noted due to the current density and electrolyte temperature. Higher rate of hydrogen evolution at high overpotential deposition leads to low current efficiency. However, improvement of ion movement at higher electrolyte temperature resulted in better current efficiency. The optimum properties of the coatings (i.e., microhardness, current efficiency and Co content, in particular) can be achieved using the highest current density of 50 mA/cm^2 and the highest electrolyte temperature of 50 °C.

REFERENCES

Bakhit, B. (2015). The influence of electrolyte composition on the properties of ni–co alloy coatings reinforced by SiC nano-particles. *Surf. Coat. Tech.,* **275**: 324–331.

Bakhit, B. & Alireza A. (2013). Synthesis and characterization of Ni–Co/SiC nanocomposite coatings using sediment co-deposition technique. *J. Alloy Compd.,* **560**: 92–104.

Cai, F., Chuanhai, J., Peng, F. & Vincent, J. (2015). Effects of Co contents on the microstructures and properties of electrodeposited NiCo–Al composite coatings. *Appl. Surf. Sci.*, **324**: 482–489.

Feng, Z., Qingyang, L., Jinqiu, Z., Peixia, Y., Hailin, S. & Maozhong, A. (2015). Electrodeposition of nanocrystalline Zn–Ni coatings with single gamma phase from an alkaline bath. *Surf. Coat. Tech.*, **270**: 47–56.

Jarrah, N. A. (2009). Studying the influence of process parameters on the catalytic carbon nanofibers formation using factorial design. *Chem. Eng. J.*, **151**: 67–71.

Jović, V. D., Jović, B. M., Maksimović, V. & Pavlović. M. G. (2007). Electrodeposition and morphology of Ni, Co and Ni–Co alloy powders: Part II. Ammonium chloride supporting electrolyte. *Electrochim. Acta*, **52**: 4254–63.

Kuo, S. L., Yann, C. C., Ming, D. G. & Wen, H. H. (2004). Nano-particles dispersion effect on Ni/Al_2O_3 composite coatings. *Mater. Chem. Phys.*, **86**: 5–10.

Lan, C. J., Liu, W. Y., Ke, S. T. & Chin, T. S. (2006). Potassium salt based alkaline bath for deposition of Zn–Fe alloys. *Surf. Coat. Tech.*, **201**: 3103–3108.

Long, J.M., Xiu, Z. & He, Z.P. (2013). Effect of triethanolamine addition in alkaline bath on the electroplating behavior, composition and corrosion resistance of Zn-Ni alloy coatings. *Adv. Mat. Res.*, **738**: 87–91.

Low, C. T. J., Wills, R. G. A. & Walsh, F. C. (2006). Electrodeposition of composite coatings containing nanoparticles in a metal deposit. *Surf. Coat. Tech.*, **201**: 371–383.

Marita, Y., Ridwan, S. N. & Nurdin. (2014). Preparation of Ni-Co alloy by electrodeposition. *Appl. Mech. Mater*., **525**: 58-61.

Nguyen, V.C., Lee, C.Y., Chen, F. J., Lin, C.S. & Chang, L. (2013). An electroplating technique using the post supercritical carbon dioxide mixed watts electrolyte. *Surf. Coat. Tech.,* **232**: 234–239.

Pande, C.S. & Cooper, K.P. (2009). Nanomechanics of Hall–Petch relationship in nanocrystalline materials. *Prog. Mater. Sci.*, **54:** 689–706.

Rafailović, L.D. & Minić, D.M. (2009). Deposition and characterisation of nanostructured Nickel-Cobalt alloys. *Hem. Ind.*, **63**: 57–69.

Ranjith, B. & Paruthimal Kalaignan, G. (2010). Ni–Co–TiO2 nanocomposite coating prepared by pulse and pulse reversal methods using acetate bath. *Appl. Surf. Sci.*, **257**: 42–47.

Sharifi, B., Mojtahedi, M., Goodarzi, M. & Vahdati Khaki, J. (2009). Effect of alkaline electrolysis conditions on current efficiency and morphology of zinc powder. *Hydrometallurgy,* **99**: 72–76.

Shi, L., Sun, C. F., Gao, P., Zhou, F. & Liu, W. M. (2006). Electrodeposition and characterization of Ni–Co–carbon nanotubes composite coatings. *Surf. Coat. Tech.*, **200**: 4870–4875.

Srivastava, M., Ezhil Selvi, V., William Grips, V.K. & Rajam, K.S.. (2006). Corrosion resistance and microstructure of electrodeposited Nickel–Cobalt alloy coatings. *Surf. Coat. Tech.*, **201**: 3051–3060.

Srivastava, M., William Grips, V.K. & Rajam, K.S. (2010). Electrodeposition of Ni–Co composites containing nano-CeO_2 and their structure, properties. *Appl. Surf. Sci.*, **257**: 717–722.

Sui, J. H., Gao, Z. Y., Cai, W. & Zhang, Z. G. (2007). Corrosion behavior of NiTi alloys coated with diamond-like carbon (DLC) fabricated by plasma immersion ion implantation and deposition. *Mater. Sci. Eng.*, **452–453**: 518–23.

Tian, B. R. & Cheng, Y. F. (2007). Electrolytic deposition of Ni–Co–Al_2O_3 composite coating on pipe steel for corrosion/erosion resistance in oil sand slurry. *Electrochim. Acta.,* **53**: 511–517.

Vazquez-Arenas, J., Altamirano-Garcia, L., Treeratanaphitak, T., Pritzker, M., Luna-Sánchez, R. & Cabrera-Sierra, R. (2012a). Co-Ni alloy electrodeposition under different conditions of pH, current and composition. *Electrochim. Acta.*, **65**: 234–43.

Vazquez-Arenas, J., Treeratanaphitak, T. & Pritzker, M. (2012b). Formation of Co–Ni alloy coatings under direct current, pulse current and pulse-reverse plating conditions. *Electrochim. Acta,* **62**: 63–72.

Yusoff, N.H.N., Ghazali, M. J., Isa, M. C., Daud, A. R., Muchtar, A. & Forghani, S. M. (2012). Optimization of plasma spray parameters on the mechanical properties of agglomerated Al_2O_3–13%TiO_2 coated mild steel. *Mater. Design.*, **39**: 504–508.

Zhang, Z., Jiang, C. & Ma, N. (2014). Microstructure and corrosion behavior of electrodeposited Ni-Co-ZrC coatings. *J. Mater. Eng.*, **23**: 4065–4071.

PREPARATION AND CHARACTERIZATION OF PBXN-109EB AS A NEW HIGH PERFORMANCE PLASTIC BONDED EXPLOSIVE

Mahdi Ashrafi[1*], Hossein Fakhraian[1], Ahmad Mollaei[1] & Seyed Amanollah Mousavi Nodoushan[2]

[1]Department of Chemistry, Imam Hossein University Tehran, Iran
[2]Department of Chemistry and Chemical Engineering, Malek-Ashtar University of Technology, Iran

[*]Email: mhdashrafi@yahoo.com

ABSTRACT

A new formulation of plastic bonded explosive (PBX) using energetic polymers as binder and plasticizer is herein reported. In the new formulation (named PBXN-109EB), inert HTPB (Hydroxyl-terminated polybutadiene) binder and inert dioctyl adiphate (DOA) plasticizer in PBXN-109 formulation have been substituted respectively by nitro hydroxyl terminated polybutadiene (NHTPB) as an energetic binder and nitro polybutadiene (NPB) as an energetic plasticizer. The kinetic and thermodynamic parameters for the decomposition of this new formulation were determined via non-isothermal Differential Scanning Calorimetry (DSC) curves from room temperature to 400°C. Density, hardness, vacuum stability, mechanical property, viscosity, glass transition temperatures (T_g) and performance properties of PBXN-109EB were compared to conventional PBXN-109. Scanning Electron Microscopy of PBXN-109EB shows a uniform dispersion of the compounds.

Keywords: Energetic binder; plastic bonded explosive (PBX); thermodynamic parameters; performance properties; blast pressure.

1. INTRODUCTION

The use of energetic binders (comprising energetic polymer and plasticizer), is driven by the need for more energetic materials in modern insensitive munitions (IM) (Provatas, 2003; Ang & Pisharath, 2012). The use of inert binder systems (such as HTPB: hydroxyl terminated polybutadiene) and inert plasticizers (such as DOA: dioctyl adiphate) decreases the final energy output of the PBX such as PBXN-109 (comprising ~15% HTPB/DOA binder system). It has been presumed that the final energy output can be improved if energetic binder and plasticizer are used. Some energetic binders such as poly-GLYN plasticized by K10 or GLYN oligomer were previously used to increase the performance characteristics of the corresponding PBX (Ang & Pisharath, 2012). The physico-chemical compatibility of the binder and plasticizer are important and affect the final properties of the formulated PBX. Thermal analysis of PBXs is important not only for understanding the kinetics of their thermal decomposition but also for determining their exothermic decomposition potential during storage, processing and handling. Recently, DSC is one of the main techniques to evaluate the thermal behavior of energetic materials (Budrugeac & Segal, 2007; Roger & Kissinger, 2012; Shekhar *et al*., 2013; Abusaidi *et al*., 2016).

We have previously reported about the preparation and properties of NHTPB/NPB energetic binder system (Ashrafi *et al*., 2017). Herein, the new formulation of a PBX (named PBXN-109EB) based on NHTPB/NPB (nitro-HTPB/nitro-PB) energetic binder system is investigated and the physicochemical properties and performance parameters (such as pressure, impulse and velocity of detonation) of PBXN-109EB and PBXN-109 are compared.

2. EXPERIMENTAL SECTION

2.1 Materials and Instruments

HTPB (M_n = 5252 g mol^{-1} by GPC) and bimodal RDX (distribution of 95/5 from pre-mixed grades Type II class 1/class 5) were obtained from energy materials research laboratory in Iran. NHTPB (as an energetic binder; M_n = 5614 g mol^{-1} by GPC) and NPB (as an energetic plasticizer; M_n = 2807 g mol^{-1} by GPC) were prepared using the previously reported method (Ashrafi *et al*., 2017). Optimum nitration of the HTPB binder system for mix viscosity, T_g and mechanical properties was between 7-8%. The other chemicals were of analytical grade, supplied by Merck company and were used without any further purification. The density measurements were carried out with a Model AG 285 METTLER TOLEDO balance according to the ASTM D792-91 method.

Tensile stress and elongation at break were measured using a Hiwa instrument according to the ASTM D 638 method and conducted at a cross-head rate of 500 mm min^{-1} on stamped samples. Hardness was measured using Hiwa instrument according to the ASTM-D2240 method to assess Shore "A". SEM VEGA (Tescan Brno, Czech Republic) was used for morphology assessment. Secondary electron imaging in the scanning electron microscope (SEM Tescan Brno, Czech Republic) was used for morphology assessment.

The viscosity determinations were performed using a LDVD-II Brookfield viscometer with rotational speed of 2 rpm at 60 ºC. The glass transition temperatures (T_g) of PBXN-109EB and PBXN-109 were determined by a NETZSCH differential scanning calorimeter DSC 200 F3 maia, under nitrogen atmosphere with a flow rate of 50 ml min^{-1}. The tests were carried out with a programmed temperature ramp from -150 to 150 °C at 10 °C min^{-1}.

The DSC analyses were performed using a Mettler Toledo DSC1 instrument under nitrogen atmosphere with temperature programmed rates of 5, 10, 15 and 20 °C min^{-1} from room temperature to 350 °C. The impact sensitivities were determined using a fall hammer apparatus (BAM) with the Bruceton staircase method. The friction sensitivities were determined using a standard Julius Peter's apparatus operating up to 36 kg load. Vacuum thermal stabilities were determined according to the MIL-STD-1751A (48 h at 100 ºC). Blast pressure and impulse of detonation were determined by piezoelectric pressure transducers. Four piezoelectric sensors (fabricated by the Swiss Kistler Company (A 603B)) were used to record the pressures. All sensors were calibrated before the experiments by comparing with the reference standard sensor (7061 BK model). Three types of charge amplifiers were used to amplify electrostatic charges of piezoelectric sensors: Swiss Kistler 5011-B-12, Austrian AVL 3057-A-07, and Iranian ACA-81 amplifiers. The sensors and amplifiers were connected together by low-noise E-178 cables with 2 mm diameter and 10 m length. A four-channel analog-to-digital card was used to record the data obtained from the amplifiers. Velocity of detonation (VOD) was also calculated according to the Iranian Defense Standard (IDS-196) using a optical fiber sensor, which is capable of detecting and transmitting a light signal accompanying the detonation wave front.

2.2 Preparation of Isocyanate Cured PBXs

The PBX formulations (PBXN-109 and PBXN-109EB) (ingredient percentage of which were on the basis of military specification MIL-E-82886(OS) (MIL-E-82886, 1995)) were prepared using a double planetary mixer. Firstly, HTPB-DOA or NHTPB-NPB mixtures (14.7% with 50:50 ratio) were degassed in a vacuum oven at 70 ºC for 3 h. Then, RDX (64%, pre-mixed classes 1 and 5) and aluminum powder (20%, spherical and micronized) were added and mixed with the binder system. Then, IPDI (0.96%, at a ratio of 1:1 isocyanate/binder, as curing agents), dibutyl tin dilaurate (0.02%, as catalyst), 2,2'-methylene-bis[6-tertbutyl-4-methylphenol] (0.1%, as anti-oxidant), and N,N'-di(2-hydroxyethyl) dimethyl hydantoin (0.26%, DHE: as binding agent), were added and the uncured PBX was mixed for 30 min. Finally, the mixture was casted in to the pre-heated polyethylene cylindrical

molds (internal diameter 94 mm, height 95 mm and wall thickness 2 mm), and also in to the steel mold (220 mm × 24 mm) and were cured in an oven at 70 ºC for 7 days.

The PBXN-109 and PBXN-109EB were casted into the polyethylene cylindrical molds for the determination of blast pressure, impulse and fireball characterization. PBXN-109EB was also casted into the steel molds for the study on the physico-mechanical properties, SEM, thermal decomposition, sensitiveness tests and velocity of detonation. The other properties that awere studied for PBXN-109 are from literatures. All detonation experiments were performed on charges boosted with C4 cylinders (length / diameter = 1.5) and initiated with M8 detonators. The density of the charges are 1.65 ± 0.04 g cm^3.

3. RESULTS AND DISCUSSION

3.1 Physico-Mechanical Properties

Stress-elongation, density and hardness properties of formulated PBXN-109EB were compared with PBXN-109 (Figure 1 and Table 1) showing that, based on MIL-E-82886(OS) standard, PBXN-109EB meets all of PBXN-109 requirements (MIL-E-82886, 1995).

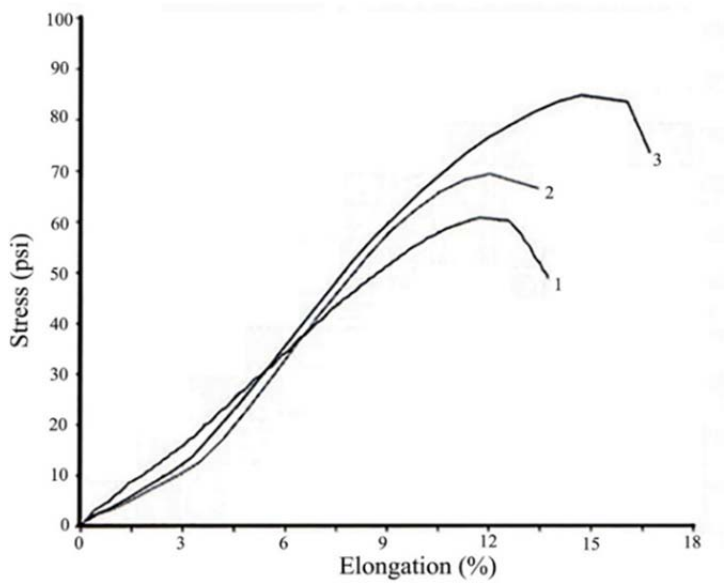


Figure 1: Stress-elongation curves for three samples of PBXN-109EB.

Table 1: Comparison of physico-mechanical properties comparison of PBXN-109 and PBXN-109EB.

Property	PBXN-109EB	PBXN-109 (MIL-E-82886, 1995)
Density (g/cm^{-3} , at 25°C)	1.69	1.60-1.70
Stress (max, psi, at 25°C)	71.7	> 60
Elongation (maximum stress, %, at 25°C)	12.8	> 12
Hardness, shore A (30 sec at 25°C)	78	> 30

3.2 Scanning Electron Microscopy (SEM)

Scanning electron micrographs of PBXN-109EB and PBXN-109 (Figure 2) show that bimodal explosive particles (coarse / fine) are completely embedded into the polymer matrixes and uniform dispersion of ingredients in PBXs is performed. Due to the near polarity of explosive particles to

energetic polymer matrix in PBXN-109EB, adhesion of this particle is better than PBXN-109 with inert polymer matrix. This problem has also been approved by hardness experiment.

3.3 Viscosity Measurements Of Uncured PBXs

Viscosities of uncured PBXN-109EB and uncured PBXN-109 vs. time at 60ºC with rotational speed of 2 rpm are shown in Table 2 and Figure 3. In the initial period of curing, the polymer molecules are small in size and viscosity is low. As the curing proceeds, the molecular size increases, diminishing the mobility of the molecules and increasing the viscosity with respect of time. Increase of the molecular weight in NHTPB/NPB binder system (compared to HTPB/DOA) due to NO_2 grafting caused an increase of initial viscosity of PBXN-109EB with respect to PBXN-109, but has not posed problem for cast-curing procedure of PBXN-109EB.

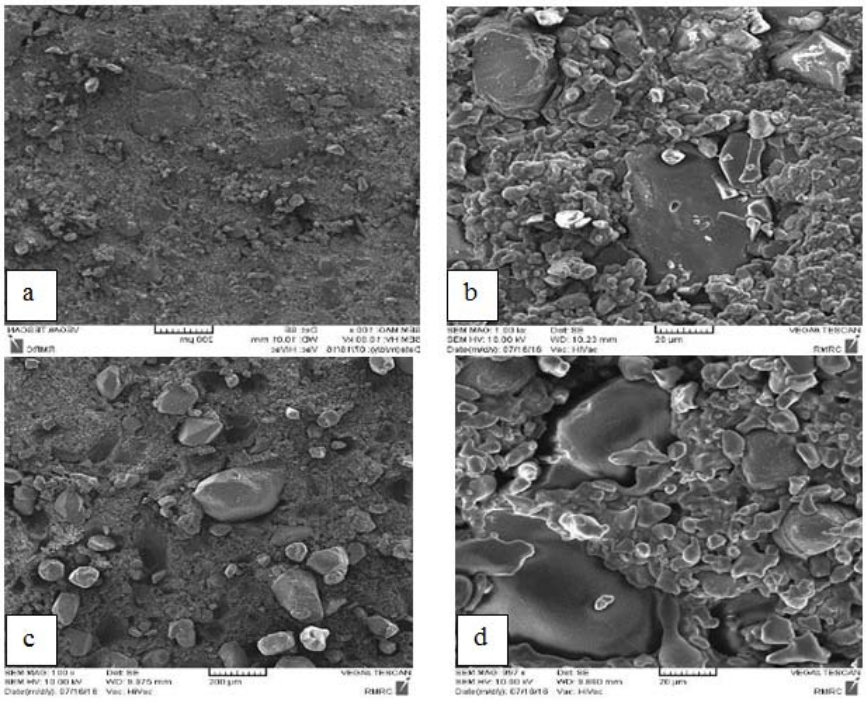


Figure 2: The SEM images of PBXN-109EB (a and b) and PBXN-109 (c and d) with 100 and 1,000 magnification.

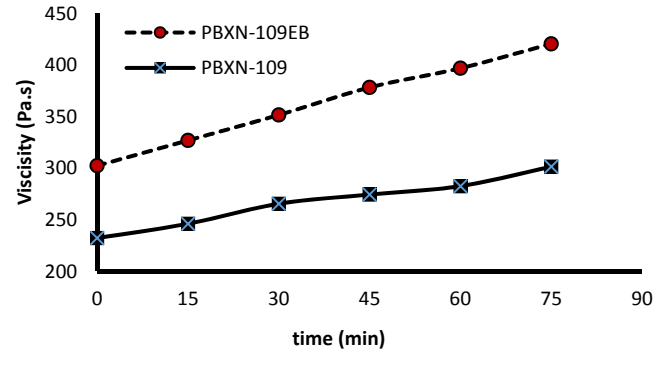


Figure 3: Plots of viscosity vs. time at 60ºC with 2 rpm for uncured PBXN-109EB and uncured PBXN-109.

Table 2: Viscosities of uncured PBXs.

Time (min)	Viscosity of uncured PBXN-109 (Pa.s)	Viscosity of uncured PBXN-109EB (Pa.s)
0	232.3	302.5
15	246.2	327.0
30	265.6	351.7
45	274.5	378.4
60	282.7	397.0
70	301.5	420.7

3.4 Glass Transition Temperatures (T_g) Measurements Of Cured PBXs

Glass transition points of PBXN-109EB and PBXN-109 were determined via DSC (Figure 4), as -88.4 and -93.2 °C respectively. The existence of polar nitro groups in the energetic NHTPB/NPB binder system increased its T_g as compared to PBXN-109. However, this increase is not so important to affect processing properties.

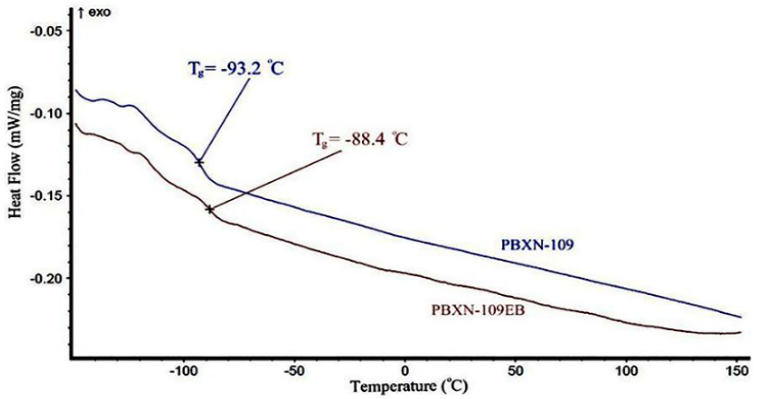


Figure 4: Glass transition temperatures for PBXN-109EB and PBXN-109 with DSC.

3.5 Thermal Decomposition Kinetic And Thermodynamic Properties

The Arrhenius parameters (E: activation energy and A: pre-exponential factor) for the thermal decomposition of PBXN-109EB sample were determined by DSC thermograms (Figure 5), using the Kissinger's method (Budrugeac & Segal, 2007; Ang *et al*., 2009; Roger & Kissinger, 2012; Abusaidi *et al*., 2016; Ashrafi *et al*., 2017).

This method allows to obtain the values of E and A from a plot of $\ln(\beta/T^2)$ against $1{,}000/T$ for a series of experiments at different heating rates (β), where T is the peak temperature of the exothermic decomposition in DSC thermograms. The equation is as follows:

$$\ln\left(\frac{\beta}{T^2}\right) = \ln\left(\frac{AR}{E}\right) - \frac{E}{RT} \quad (1)$$

where R is the universal gas constant (8.314 J mol^{-1} K^{-1}). A straight line (R^2= 0.9992) is obtained from the plot of $\ln(\beta/T^2)$ against $1{,}000/T$ (Figure 6), indicating that the mechanism of thermal decomposition of PBXN-109EB does not vary during the decomposition under various heating rates. The activation energy and pre-exponential factor were derived from the slope and intercept of plotting regression line respectively.

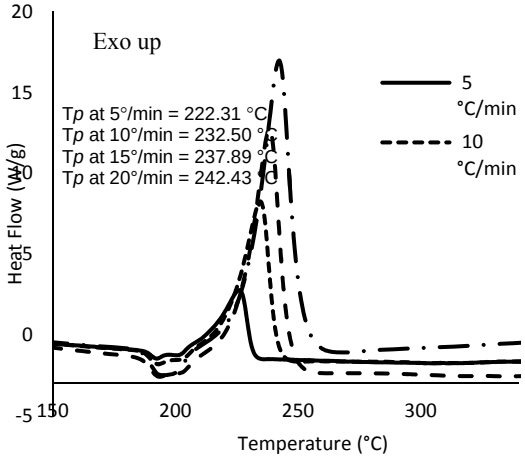


Figure 5: DSC thermograms of PBXN-109EB with different heating rates.

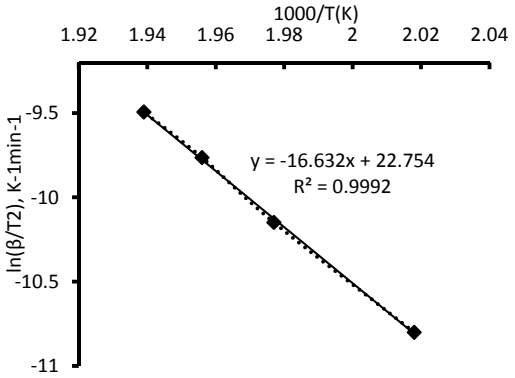


Figure 6: The plot of ln (β/T^2) vs. 1,000/T for PBXN-109EB.

In order to calculate the pre-exponential factor, it was assumed that the decomposition of PBXN-109EB followed first-order kinetics. Knowing the value of A and E, for a given temperature, the rate constant, k for the decomposition reaction of PBXN-109EB can be calculated using the following equation (Shekhar *et al*., 2013):

$$k = Ae^{-E/RT} \tag{2}$$

The calculated values of activation energies, frequency factors and k for PBXN-109EB at 150–300°C are listed in Table 3, following which the thermodynamic parameters of activation can be calculated from the following equations (Olszak-Humienik & Mozejko, 2000):

$$\mathrm{A}\exp\left(-\frac{\mathrm{E}}{\mathrm{RT}}\right) = (k_B\mathrm{T}/\mathrm{h})\exp(-\frac{\Delta\mathrm{G}}{\mathrm{RT}}) \tag{3}$$

$$\Delta\mathrm{H} = \mathrm{E} - \mathrm{RT} \tag{4}$$

$$\Delta\mathrm{G} = \Delta\mathrm{H} - \mathrm{T}\Delta\mathrm{S} \tag{5}$$

where ΔG, ΔH and ΔS are Gibbs energy, enthalpy and entropy of activation respectively (Table 3). These parameters are important in estimation of aging, thermal stability and mechanism of decomposition of energetic materials (Shekhar *et al*., 2013; Ashrafi *et al*., 2017).

Table 3: Kinetic and thermodynamic parameters for PBXN-109EB.

Activation energy, E (kJ mol^{-1})	Pre-exponential factor, ln A (min^{-1})	Square of the correlation coefficient, R^2	Rate constant for 50°C, k (min^{-1})	ΔG (kJ mol^{-1})	ΔH (kJ mol^{-1})	ΔS (J mol^{-1} K^{-1})
138.278	32.47	0.9992	5.61×10^{-9}	165.9	134.074	-63.011

3.6 Sensitiveness Testing

To assess safe handling and processing in comparison to other explosive formulations, PBXN-109EB and PBXN-109 were subjected to a series of sensitiveness tests, including Rotter Impact (50 % probability of explosion of the samples were determined with a 5 kg drop weight at a height of 60 cm), BAM Friction and Vacuum Thermal Stability (VTS) (Table 4).

Table 4: Sensitiveness test of PBXN-109 and PBXN-109EB.

Test	PBXN-109EB	PBXN-109
Rotter Impact (N.m, 5 Kg from 60 cm)	29.6	>30
BAM Friction (kgf)	>36*	> 36 (Hamshere *et al*., 2003)
VTS (mL/g, at 100°C for 48 h)	0.91	< 0.5 (MIL-E-82886, 1995)

*No reactions were observed up to 36 kg load.

Despite high gas volume of PBXN-109EB, compatibility between the NHTPB binder and NPB plasticizer is acceptable. It is well known that nitrate evolve significant gas volumes upon vacuum stability testing as a result of the inherent instability of the nitrate moiety (Cliff, 1999). However, VTS of formulated PBXN-109EB is approximately the same as PolyGLYN-K10 (1.2 mL g^{-1}) and PolyGLYN-GLYN oligomer (0.97 mL g^{-1}) (Provatas, 2003).

3.7 Velocity of Detonation

The velocities of detonation for PBXN-109EB and PBXN-109 confined in heavy walled seamless steel tube (diameter 25 mm) were determined using coaxial probes. The mean velocity of detonation was determined for three firing (with M8 detonator) by time-of-arrival optical fiber pins spaced at 100 mm intervals along the length of the charge. The results (Table 5) show that Velocity of Detonation (VOD) in modified-PBX is 4-6 % (about 400 m s^{-1}) higher than PBXN-109.

Table 5: Velocity of detonation of PBXN-109 and PBXN-109EB.

Sample	Density (g cm^{-3})	Diameter (mm)	VOD (m s^{-1})	Technique	Ref
PBXN-109	1.68	25	7480	Optical fiber pins	This work
PBXN-109EB	1.69	25	7865	Optical fiber pins	This work
PBXN-109	1.70	25	7335	CHEETAH (cal.)	(Lu, 2001)
PBXN-109	1.65	20	7391	Digital streak imaging	(Hamshere *et al*., 2003)
PBXN-109	1.65	35	7577	Digital streak imaging	(Hamshere *et al*., 2003)

This difference of velocity is only due to replacement of binder system (15% of the formulation) by NHTPB/NPB energetic binder system (obtained from nitration of HTPB and PB by about 7-8 %). For further study, velocities of detonation were compared to other literatures. According to Lu (2001), for PNXN-109, the velocity of detonation calculated with CHEETAH was compared to PBXN-109EB. According to the Hamshere *et al*. (2003), it is known that due to problems with the calibration software for the digital camera, the velocity of detonation measured for PNXN-109 with digital streak imaging technique are not very accurate and the reported VOD of PBXN-109 (7678 m s^{-1} for 50 mm diameter) using this technique was considered unrealistically high.

3.8 Blast Pressure And Impulse

In every test (four assemblies), one charge (PBXN-109 and PBXN-109EB, two samples each) was placed on a steel plate placed 2 m above the ground. Piezoresistive pressure gauges were placed at 2.5, 3.5, 4.5 and 5.5 m away from the charge explosive in the side-on configuration (Figure 7). The data was processed to yield peak pressure and impulse. The pressure impulse was calculated from the area under the *p-t* curve (Maranda *et al*., 2011; Stewart, 2013). The plots of blast pressure and impulse of PBXN-109EB and PBXN-109 (Figures 8 and 9, and Table 6) showed that blast pressure and impulse of PBXN-109EB is 8-9% higher than PBXN-109.

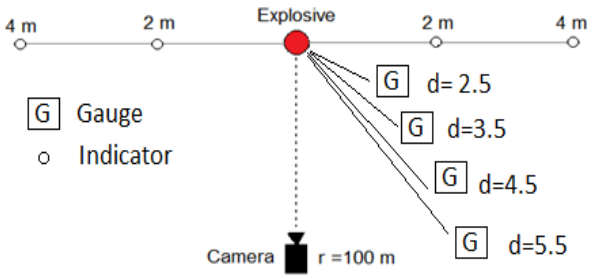


Figure 7: Schematic of the charge configuration for air-blast tests.

Table 6: Overpressure and impulse of PBXN-109 and PBXN-109EB in different distances.

Sample	Distance				Total Average
	2.5m	3.5m	4.5m	5.5m	
	Overpressure, bar				
PBXN-109 (1)	1.84	0.60	0.38	0.26	0.770
PBXN-109 (2)	1.72	0.70	0.36	0.27	0.762
Average	1.78	0.65	0.37	0.265	**0.766**
PBXN-109EB (1)	2.07	0.73	0.43	0.30	0.882
PBXN-109EB (2)	1.86	0.60	0.40	0.29	0.787
Average	1.965	0.665	0.415	0.295	**0.834**
Impulse, bar-ms					
PBXN-109 (1)	0.66	0.45	0.34	0.32	0.442
PBXN-109 (2)	0.64	0.50	0.34	0.30	0.445
Average	0.65	0.475	0.34	0.31	**0.443**
PBXN-109EB (1)	0.73	0.51	0.37	0.32	0.482
PBXN-109EB (2)	0.68	0.52	0.37	0.33	0.475
Average	0.705	0.515	0.37	0.325	**0.478**

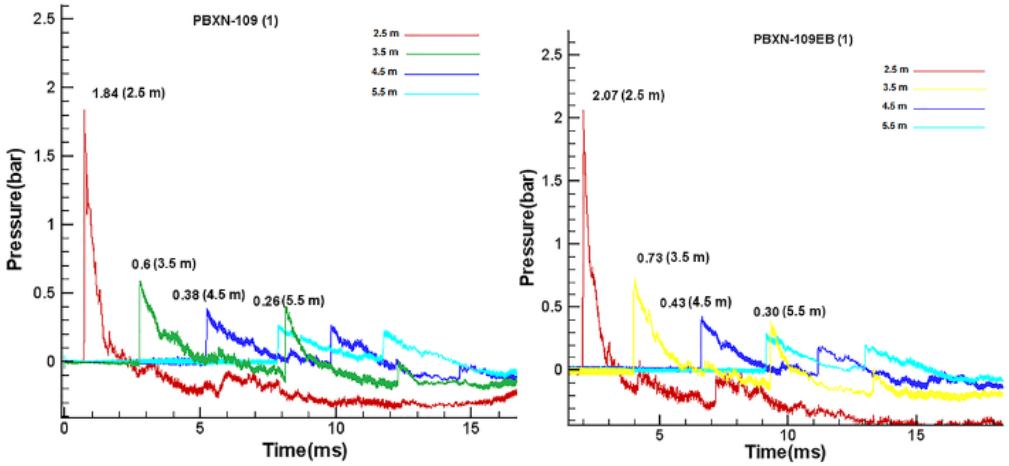


Figure 8: Blast pressure at various distances for PBXN-109 and PBXN-109EB.

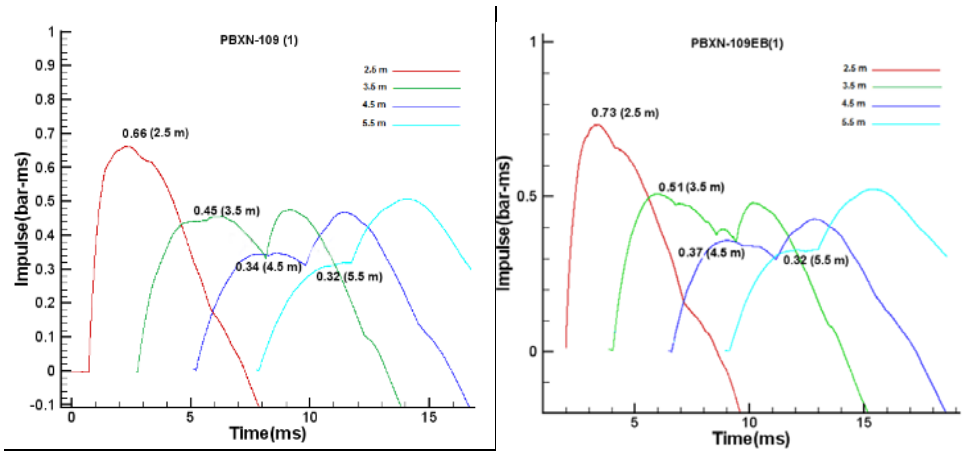


Figure 9: Impulse graphs at various distances for PBXN-109 and PBXN-109EB.

3.9 Fireball Characterization

The diameters of the fireballs (captured at the moment of maximum fireball extent) were scaled graphically using a high-speed camera (300 fps). The known distance between the four indicators located around the explosive charge was used as the scale for fireball measurements. Figure 10 shows the fireball of the PBXs at different times after initiation. The results summarized in Table 7 show that the fireball diameter, height of explosion and duration of fireball increased significantly for PBXN-109EB with energetic binder system compared to PBXN-109 with inert binder system.

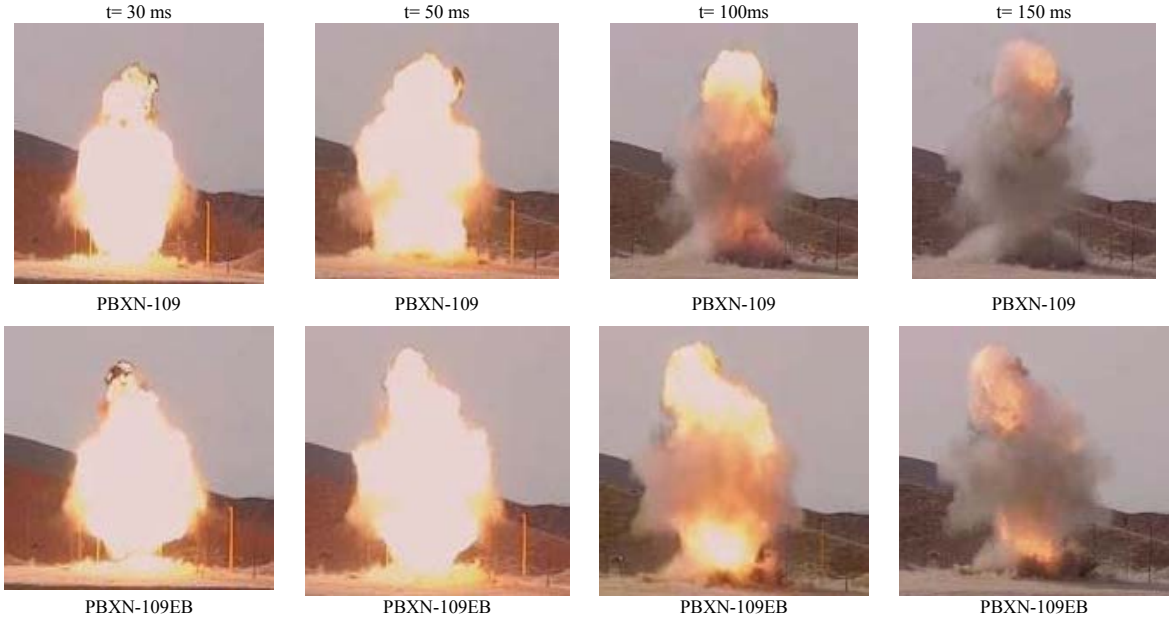


Figure 10: Fireball of examined PBXs at different times after initiation.

Table 7: Fireball characterization of PBXN-109 and PBXN-109EB.

Sample	Fireball diameter (m)	High (m)	Duration of fireball (ms)
PBXN-109	2.2	3.3	180
PBXN-109EB	2.6	3.7	200

4. CONCLUSION

A new high performance PBX named PBXN-109EB, containing energetic binder, was proposed, meeting all of PBXN-109 requirements. Compared to PBXN-109, the VOD of PBXN-109EB increased by 4-6%, while the blast pressure and impulse increased by 8-9% due to the energetic binder system. The fireball diameter, height of explosion and duration of fireball increased significantly in PBXN-109EB. Due to the high average molecular weights of HTPB and PB that are used (5252 and 2670 g mol^{-1} respectively), a nitration of higher than 7-8% will cause deterioration of the binder and PBX properties (such as viscosity, T_g and mechanical properties). Thus, if HTPB and PB with lower average molecular weights are accessible, higher percentage of nitration and therefore, higher VOD and more acceptable operational properties of the final produced PBX are possible. Sensitivity testing also showed that PBXN-109EB retains similar IM characteristics found in PBXN-109. The kinetic parameters such as activation energy, frequency factor and rate constants as well as the thermodynamic parameters for the thermal decomposition of PBXN-109EB were obtained from the DSC data by Kissinger's method.

REFERENCES

Abusaidi, H. Ghaieni, H.R. Pourmortazavi, S.M. & Motamed-Shariati, S.H. (2016). Effect of nitro content on thermal stability and decomposition kinetics of nitro-HTPB. *J. Therm. Anal. Calorim.,* **124:** 935–941.

Ang, H. & Pisharath, G.S. (2012). *Energetic Polymers: Binders and Plasticizers for Enhancing Performance.* Wiley-VCH, Germany.

Ashrafi, M. Fakhraian, H. & Dehnavi, M.A. (2017). Synthesis, characterization and properties of nitropolybutadiene as energetic plasticizer for NHTPB binder, *Propellants, Explos. Pyrotech.,* **42:** 269–275.

Budrugeac, P. & Segal, E. (2007). Applicability of the Kissinger equation in thermal analysis. *J. Therm. Anal. Calorim.,* **88:** 703–707.

Cliff, M. (1999). *PolyGLYN Binder Studies and PBX Formulation.* Technical Achievements from a LTA to DERA Fort Halstead, DSTO-TR-0884, Australia.

Hamshere, B.L. Lochert, I.J. & Dexter, R.M. (2003). *Evaluation of PBXN-109: The explosive Fill for the Penguin Anti-Ship Missile Warhead,* DSTO-TR-1471, Australia.

Lu, J.P. (2001). Evaluation of the Thermochemical Code-CHEETAH 2.0 for Modelling Explosives Performance, DSTO-TR-1199, Australia.

Maranda, A. Paszula, J. Zawadzka-Malota, I. Kuczynska, B. Witkowski, W. Nikolczuk, K. & Wilk, Z. (2011). Aluminum powder influence on ANFO detonation parameters. *Central European, J. Energ Mater.,* **8:** 279-292.

MIL-E-82886(OS), (1995). *Military Specification, Explosive, Plastic Bonded, Cast PBXN-109.*Department of Defence, US.

Olszak-Humienik, M. & Mozejko, J. (2000). Thermodynamic functions of activated complexes created in thermal decomposition processes of sulphates. *Thermochim. Acta.,* **344:** 73–79.

Provatas, A. (2003). *Characterization and Binder Studies of the Energetic Plasticizer-GLYN Oligomer.* DSTO-TR-1422, Australia.

Provatas, A. (2003). *Formulation and Performance Studies of Polymer Bonded Explosives (PBX) Containing Energetic Binder Systems.* DSTO-TR-1397 Part I. Australia.

Roger, L.B. & Kissinger, H.E. (2012). Homer Kissinger and the Kissinger equation, *Thermochim. Acta.,* **540:** 1–6.

Shekhar Pant, C. Mada S. S. Santosh, N. M. Banerjee, Sh. & Khanna, P. K. (2013). Single step synthesis of nitro-functionalized hydroxyl-terminated polybutadiene, *Propellants, Explos. Pyrotech.,* **38:** 748–753.

Stewart, J. B. (2013). *Air Blast Calculations*. Weapons and Materials Research Directorate, ARL-TN-549.

Wang, Q. Wang, L. Zhang, X. & Mi, Z. (2009). Thermal stability and kinetic of decomposition of nitrated HTPB. *J. Hazard. Mater.,* **172:** 1659–1664.

NONLINEAR ROV MODELLING AND CONTROL SYSTEM DESIGN USING ADAPTIVE U-MODEL, FLC AND PID CONTROL APPROACHES

Nur Afande Ali Hussain[1,3*], Syed Saad Azhar Ali[1], Mohamad Naufal Mohamad Saad[1] & Mark Ovinis[2]

[1]Centre for Intelligent Signal and Imaging Research (CISIR), Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS (UTP), Malaysia
[2]Department of Mechanical Engineering, Universiti Teknologi PETRONAS (UTP), Malaysia
[3]Maritime Technology Division, Science & Technology Research Institute for Defence (STRIDE) Ministry of Defence, Malaysia

[*]Email: nur.afande_g03481@utp.edu.my

ABSTRACT

This paper presents the development of remotely operated vehicle (ROV) control modelling and control synthesis using nonlinear adaptive U-model and compares it with the proportional-integral-derivative (PID) control and fuzzy logic control (FLC) approaches. A nonlinear ROV model based on dynamic equations using the Newtonian method, and derivation towards kinematics equations and rigid-body mass matrixes are explained. This nonlinear ROV model represents the underwater thruster dynamics, ROV dynamics and kinematics related to the earth-fixed frame. Multivariable nonlinear adaptive control synthesis using the U-model approach incorporated with radial basis function (RBF) neural networks along with the PID and FLC approaches are implemented using MATLAB™ Simulink and integrated with the nonlinear ROV model. Simulations are carried in six degree of freedom (DoF) manoeuvring position in x, y, z coordinates from (0,0,0) to (5,5,1), with the final reference position at (10,10,2). All three controllers are compared and analysed in terms of control synthesis and model tracking capabilities without external disturbances intervention. The simulations are then done with external disturbances intervention for the nonlinear ROV model and the control performances are analysed. The results show good control signal convergence and tracking performance using the U-model control approach.

Keywords: *Multivariable systems; nonlinear modelling; adaptive control; remotely operated vehicle (ROV).*

1. INTRODUCTION

Remotely operated vehicle (ROV) has become an important tool for underwater robotic applications, which can be used to perform complex tasks such as underwater structure visual inspection, mechanical works using manipulator, sediment and gas sampling, underwater mapping, and underwater security & surveillance. A ROV is controlled directly from the umbilical cord for power system and data transfer from the control station at the surface. It can be equipped with many different sensors and transducers depending on its applications due to unlimited power demand (Yuh, 2000). However, the dynamic behaviour of the system may become unstable, in addition with the external environment condition, thus making it difficult to control. Similar to other underwater robotic platforms, ROV dynamics are highly nonlinear, highly coupled in motion and vulnerable to external disturbances. A ROV platform is an underactuated system (number of degree of freedoms (DoF) exceeds the number of actuators) due to design optimisation. In addition, ROV motion is susceptible not only to external disturbances but also from the added mass components. Therefore, ROVs must overcome the environment conditions that can have changing hydrodynamics during operation (Fossen, 2002; Antonelli *et al.,* 2008).

The impact of nonlinearities is not extreme if the working conditions are consistent and constrained, or mild. In any case, when the conditions are not clear because of the disturbances or obscure impact, the linear approximation model approach may not perform well. Most of the underwater vehicle control schemes are based on the underactuated system due to the restriction of shapes, cost, complexity, power consumption, etc. (Borhaug *et al.,* 2006; Arslan *et al.*, 2007; Li *et al.*, 2015; Juan *et al*., 2015; Chen *et al.*, 2016). A ROV considered as a nonlinear system and in order to overcome the uncertainty of the unstructured environment, adaptive control and nonlinear method are desirable to achieve the best tracking control ability. Neural networks represent as a nonlinear system with the ability to adapt itself according to a performance index based on the training algorithm used. It implements multilayer networks containing neurons and the complexity of the problems depending on the network size (Afande *et al.,* 2016)

The performance of the controller is validated through convergence speed, tracking error and stability of the system by tuning the learning rates (Mutaz & Ahmad, 2015; Chu *et al.,* 2016; ul Amin *et al.,* 2016). Another adaptive technique is using a back-stepping controller (Liu *et al.,* 2016). In this method, a thrust control distribution strategy is developed using the pseudo-inverse technique, and movement or motion controller is built based on adaptive control and back-stepping. In Zhu & Gu (2011), adaptive back-stepping combined with sliding control was introduced. In Shen *et al.* (2017), nonlinear model predictive control (NMPC) was presented for direction following control, but some stability issues need to solved for future results. The results showed that the controllers performed with good dynamic response and tracking precision, but involved complex mathematical modelling. Another work demonstrates the use of proportional-integral-derivative (PID) with Ziegler-Nicholes tuning method for hold altitude in ROV platforms (Ali *et al.,* 2016). The control method improved the control responses by reducing overshoot and settling time. The PID controller was further improved by adding adaptive capabilities in Qiao *et al.* (2016). The adaptive algorithm will estimate and compensate the external forces caused by the passive arm, umbilical cable and uncertainty in buoyancy. Fuzzy logic control (FLC) is another adaptive method that could be implemented for nonlinearity of dynamic behaviour approximation for control decision making (Johnson *et al.,* 2016), and can be integrated with another control scheme for better control performance, stability and nonlinear robustness (Lakhekar *et al.,* 2015).

This paper implements nonlinear multivariable (MIMO) adaptive control synthesis using U-model for a ROV platform. Due to the nonlinearity of the ROV platform, the U-model control approach is further improved by implementing a neural networks algorithm for capturing nonlinearity in the plant. With the specific goal of upgrading the convergence speed, the controller synthesis used only single-layer neural networks with radial basis function (RBF) activation function (Ali *et al.,* 2014; Afande *et al.,* 2016; Chu *et al.,* 2016; Abbasi *et al.,* 2017; Afande *et al.*, 2017). The U-model controller synthesis will be compared with the PID and FLC schemes to study and analyse the control tracking performance.

2. METHODOLOGY

2.1 Nonlinear ROV Model

The notations and variables used for marine vehicles, as defined by Society of Naval Architects and Marine Engineer (SNAME), are shown in Table 1. Utilising the Newtonian approach, the movement of an inflexible body using the body-settled reference at the cause in (Figure 1) is given by the accompanying arrangement of conditions (Chin, 2013):

$$M_{mass}[\dot{v}_1 + v_2 \times v_1 + \dot{v}_2 \times r_G + v_2 \times (v_2 \times r_G] = \tau_1 \quad (1)$$

$$I\,\dot{v}_2 + v_2 \times (\mathrm{I}v_2) + M_{mass} \times r_G \times (v_1 + \dot{v}_2 \times v_1) = \tau_2 \quad (2)$$

where $r_G = [x_G \ y_G \ z_G]^T$ is the area of the focal point of gravity; $\tau_1 \in \check{R}^3$ and $\tau_2 \in \check{R}^3$ are the outer force and moment vector; $v_1 = [u\ v\ w]^T \in \check{R}^3$ is linear velocity vector; and $v_1 = [p\ q\ r]^T$ is the angular velocity vector. $M_{mass} \in \check{R}^{3x3}$ is the ROV mass matrix:

$$M_{mass} = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} = mI_{3x3} \tag{3}$$

where I_{3x3} is the identity matrix together with $I \in \check{R}^{3x3}$. The platform body equation comprising of Coriolis and centrifugal forces, and inertia forces can be represented as follows:

$$M_{RB}\,\dot{v} + C_{RB}\,(v) = \tau \tag{4}$$

where $M_{RB} \in \check{R}^{6x6}$ is the mass-inertia matrix, $C_{RB}\ (v)$ is the Coriolis and centrifugal matrix, $\tau = [\tau_1\ \tau_2]^T \in \check{R}^{6x6}$ is a vector of moments and external forces, while $v = [v_1\ \ v_2]^T \in \check{R}^{6x6}$ is the direct and precise speed vector. The open loop ROV nonlinear dynamic equation can be represented as:

$$\tau = \tau_A + \tau_H = M\dot{v} + C(v)v + D(v)v + G_f(\eta_2) \tag{5}$$

where $M_V = M_{RB} + M_A$, $C_V = C_{RB} + C_A$, $D(v)$ is the damping matrix due to surrounding fluid, and $G_f(\eta_2)$ is the gravitational and buoyancy matrix.

Table 1: Notations and variables used for marine vehicles (Fossen, 2002; Chin, 2013).

Motion Descriptions	Forces & moments	Linear & Angular Velocities
Motions in the *x*-direction (surge)	x	u
Motions in the *y*-direction (sway)	*y*	*v*
Motions in the *z*-direction (heave)	*z*	*w*
Rotations about *x*-axis (roll)	φ	*p*
Rotations about *y*-axis (pitch)	θ	*q*
Rotations about *z*-axis (yaw)	ψ	*r*

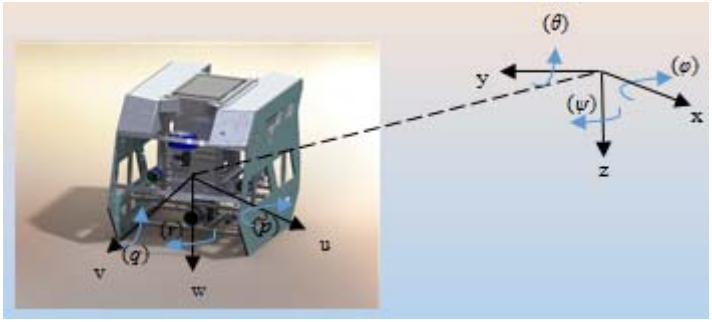


Figure 1: ROV motion dynamics.

The transformation or change utilising Euler angles gives a vital transformation between the dynamics expressed by Equation 5. The kinematic conditions that represent Euler's transformation can be utilised as:

$$\dot{\eta} = J(\eta_2)v \tag{6}$$

On the other hand, Euler's change between the body and earth-fixed frames *(Π,v)* to *(Π,* $\dot{\eta}$) is expressed as:

$$\begin{bmatrix} \eta \\ \dot{\eta} \end{bmatrix} = \begin{bmatrix} I_{6x6} & 0_{6x6} \\ 0_{6x6} & J(\eta_2) \end{bmatrix} \begin{bmatrix} \eta \\ v \end{bmatrix} \tag{7}$$

where the Euler change framework matrix, $J(\eta_2)$ is inferred by rotation of the Euler angles $\Pi_2=[\varphi\ \theta\ \psi]^T$ about the *x*, *y* and *z*-axes to give:

$$J(\eta_2) = \begin{bmatrix} J_1(\eta_2) & 0 \\ 0 & J_2(\eta_2) \end{bmatrix} \tag{8}$$

where:

$$J_1(\eta_2) = \begin{pmatrix} cos\psi cos\theta & -sine\psi cos\varphi + cos\psi sine\theta sine\varphi & sine\psi sine\varphi + cos\psi cos\varphi sine\theta \\ sine\psi cos\theta & -cos\psi cos\varphi + sine\varphi sine\theta sine\psi & -cos\psi sine\varphi + sine\theta sine\psi cos\theta \\ -sine\theta & cos\theta sine\varphi & cos\theta cos\varphi \end{pmatrix} \tag{9}$$

$$J_2(\eta_2) = \begin{pmatrix} 1 & sine\varphi tan\theta & cos\varphi tan\theta \\ 0 & sine\varphi & -sine\varphi \\ 0 & sine\varphi/cos\theta & cos\varphi cos\theta \end{pmatrix} \tag{10}$$

The body platform mass, and Coriolis and centrifugal force parameters can be obtained using a hydrodynamics software, such as computation fluid dynamic (CFD). The gravitational and buoyancy matrix can be expressed as:

$$G_f(\Pi) = \begin{bmatrix} (W-B)sine\theta \\ -(W-B)cos\theta sine\varphi \\ -(W-B)cos\theta cos\varphi \\ -(y_G W - y_B B)cos\theta cos\varphi + (z_G W - z_B B)cos\theta sine\varphi \\ (z_G W - z_B B)sine\theta + (x_G W - x_B B)cos\theta cos\varphi \\ -(x_G W - x_B B)cos\theta sine\varphi - (y_G W - y_B B)sine\theta \end{bmatrix} \tag{11}$$

In underactuated control of ROV, just four thrusters are adequate to completely move the stage in the six DoF. Two unactuated DoF of roll and pitch speed are asymptotically steady and bounded due to mechanical design. Most of the ROV is designed with four thrusters configuration. The underactuated thrusters configuration model can be expressed as:

$$\tau_A = Tu \tag{12}$$

where $u = F_T\bar{u}\epsilon\check{R}^4$ with $F_T = f_T I_{4x4}$, $\bar{u}\check{R}^4$ is the reference voltage as input for the thrusters and I_{4x4} is an identity matrix representing the four thrusters system. The representation of the nonlinear ROV model can be modelled as in Figure 2.

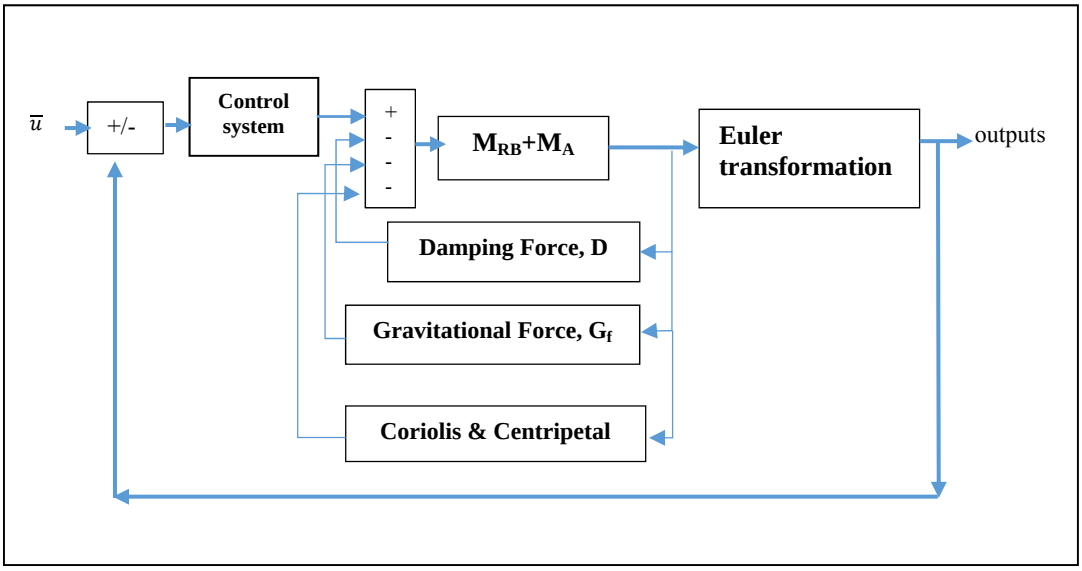


Figure 2: Nonlinear ROV model.

2.2 Nonlinear Adaptive U-Model Control Synthesis Approach

Recently developed control oriented model for multivariable system call U-model can be implemented to acquire experiment data and to perform model identification of the underwater dynamics with the disturbances (Ali *et al,*. 2010, 2014). U-model is a control oriented model and it is based on the control input signal in Equation 13. U-model expands the nonlinear NARMAX (Equation 14) and any other nonlinear model such as the Hammerstein, bilinear, Lur'e and nonlinear autoregressive models with Exogenous Input (NARX), Nonlinear Finite Impulse Response (NFIR) and Output Affine Model (Ali, 2007). U-model simplifies the control parameters term into a polynomial form (Shafiq & Butt, 2005; Abbasi *et al.,* 2015):

$$y(t) = \sum_{j=0}^{M} \alpha_j(t) u^j(t-1) + d(t) \tag{13}$$

$$y(t) = f(y(t-1) \dots y(t-m) u(t-1) \dots (t-m) + d(t-1) d(t-m)) \tag{14}$$

in which M is the system order of model input *u(t-1)*. Parameter $\alpha_j(t)$ is the function of past data sources or inputs and outputs *u(t-2),…,u(t-n), y(t-1),…,y(t-n)*, and errors *d(t-1),…, d(t-n)*. $\alpha_j = [A_0, A_1, \dots, A_M]$. α_j are matrices and updated online using the gradient descent formula. In this paper, the normalised least mean square (nLMS) approach is used for parameters update. The SISO U-model can be extended for a multivariable system or MIMO using the following equation:

$$Y_m(t) = \sum_{j=0}^{M} A_j U^j(t-1) = F(U(t-1)) \tag{15}$$

Y_m (t) is a vector p x 1 and *U(t-1)* is the current control signal with m x 1 control input vector. M is the level of multivariable polynomial, while U^j is the vector with j^{th} energy of control inputs $u_i(t-1)$:

$$U^j(t-1) = [u_1^j(t-1) u_2^j(t-1) \dots u_m^j(t-1)]^T \tag{16}$$

Nonlinear approximation using RBF is incorporated with U-model modelling in order to acquire the actual dynamics of the unknown nonlinear function in the ROV model due to its nonlinear mapping ability (Lam & Wunsch, 2017; ul Amin *et al.,* 2016; Mutaz & Ahmad, 2015). In order to reduce computational time, single layer feedforward networks are selected as compared with multilayer networks (Ali *et al.,* 2014). Integration of U-model with RBF nonlinear approximator will enhance the nonlinear modelling process, as demonstrated in our previous work (Afande *et al.,* 2016). Only matrix A_0 is updated using neural networks function as follows:

$$A_0 = Y(t) = W(t).\phi(t) \tag{17}$$

RBF is chosen as activation function due to its better learning speed (Mutaz & Ahmad, 2015; Ruano *et al.,* 2016; ul Amin *et al.,* 2016). The activation function can be express as follows:

$$\phi(t) = e(-\frac{u(t-1)-c_i{}^2}{\beta^2}) \quad \text{for } i = 1,2\ldots n \tag{18}$$

where n is the number of shrouded layer neuron, c_i is the centre of the i^{th} $(c_1, c_2, c_3..c_n)$ shrouded layer node and β is the width of the activation function. As a result, the weights A_0 of the RBF approximator and the rest of parameters A_j are updated using nLMS as follows:

$$W(t+1) = W(t) + \mu(t).error(t).\phi(t)^T \tag{19}$$

$$A_j(t+1) = A_j(t) + \mu(t).error(t).U^j(t-1)^T \tag{20}$$

where $\mu(t)$ is nLMS learning rate $(0 < \mu(t) < 1)$.

From the U-model modelling approach, the plant output is a polynomial equation structured by the parameter *u(t-1)*. Hence, the control law can be integrated by utilising converse model control. The controller output of the system can be obtained using the Newton-Raphson algorithm for the root solving method. The selection of the previous control signal as the initial value for the next time instant using Newton-Raphson is given by:

$$u_{i+1}(t-1) = u_i(t-1) - \frac{y_m - x(t)}{y'_m(t)} \tag{21}$$

$$u_{i+1}(t-1) = u_i(t-1) - \mathrm{m}\frac{\sum_{j=0}^{M}\alpha_j(t)u^j(t-1) - \mathrm{x(t)}}{d\sum_{j=0}^{M}\alpha_j(t)u^j(t-1)/\mathrm{d}u_i(t-1)} \tag{22}$$

where i is the iteration index, $x(t)$ is the input of the controller and $u(t-1)$ is the output of the controller. We introduced a learning rate parameter, m $(0<m\leq 1)$ in Equation 22 to decrease the convergence rate and preserve the system stability. In this work, model-based control is implemented using internal model control (IMC) as a nonlinear modelling control framework. Figure 3 shows the online adaptive IMC using U-model.

2.3 PID Control Synthesis

PID is a widely used controller in many applications such as in industry and robotics due to its simplicity. The PID method implies that the controller gains matrices adjustment K_p, K_d and K_i for better control system performances. The controller law for single DoF PID for the nonlinear ROV control can be express as:

$$u = T^+(K_p e(t) + K_d e(t) + K_i \int_0^t e(\tau)d\tau) \tag{23}$$

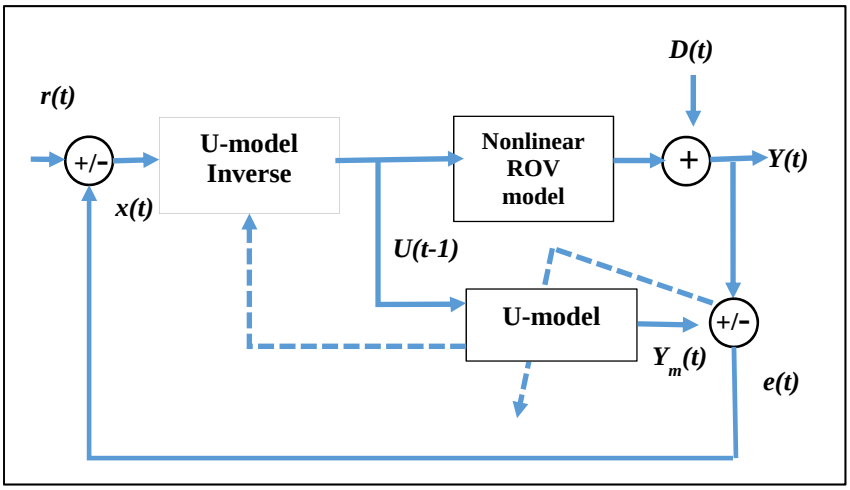


Figure 3: Online adaptive IMC U-model control synthesis.

The compensator formula transfer function can be express as:

$$P + I\frac{1}{s} + D\frac{N}{1+N\frac{1}{s}} \tag{24}$$

Based on Figure 4, the PID control parameter can be tuned manually to improve control response. The PID control block is a feedforward path of the feedback loop where u is the input and y is the output of the controller.

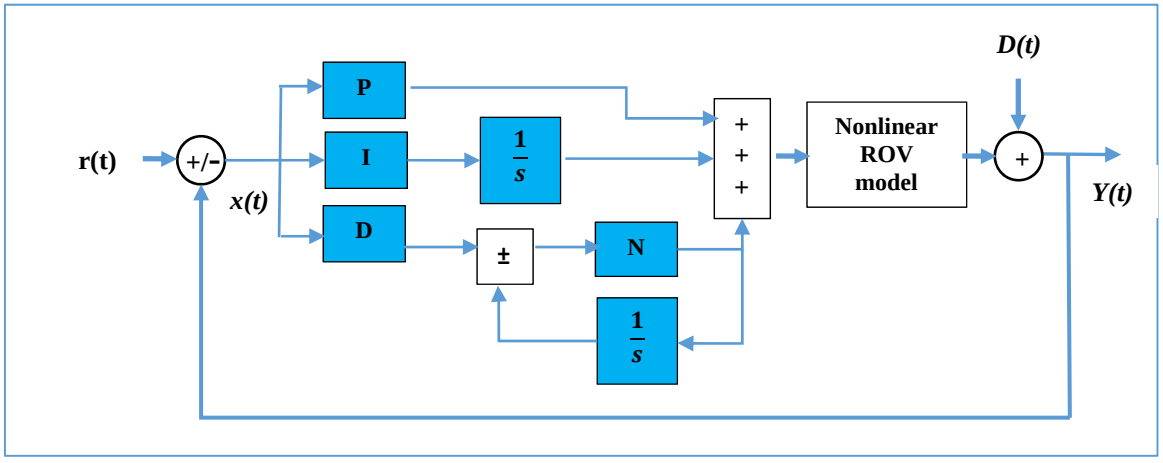


Figure 4: PID control synthesis.

2.4 FLC Control Synthesis

The FLC design has three main components for the controller design, which are fuzzification, inference rules and defuzzification operations. Fuzzification involves defining the controlled inputs, normal system error, e and change of error, $\dot{e}$ ranging from 0 to 1. Next, the inference rules will define the membership functions involving activation functions, such as triangular and trapezoidal with linguistic variables defined as: Negative Big (NB), Negative Medium (NM), Negative Small (NS), Zero (Z), Positive Small (PS), Positive Medium (PM) and Positive Big (PB). Then, these linguistic values are transformed to the numerical value of the control variable as the output of the FLC control synthesis in Figure 5.

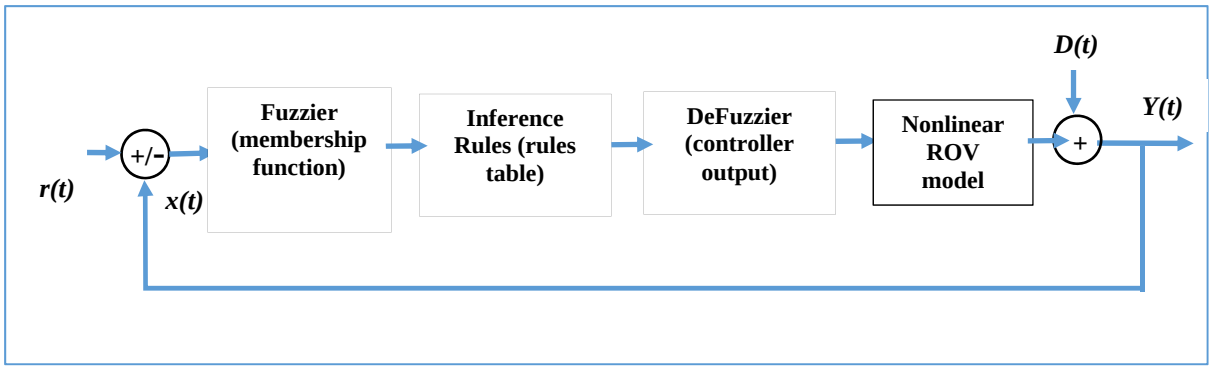


Figure 5: FLC control synthesis.

3. SIMULATION RESULTS

In order to investigate the effectiveness of the control approach (using U-Model, PID and FLC), the simulations are carried using MATLAB™ Simulink based on Figure 2 with external disturbances equal to 0. The ROV was commanded to move diagonally in the *x*-*y* plane while submerged in the water surface, that is $x_d = y_d = 5$ m and $z_d = 1$ m, then towards another set of desired values; $x_d = y_d =$ 10 m and $z_d = 2$. The simulation considers six DoF ROV dynamics and the controller input (thrusters) consist of τ_x, $\tau_y$and τ_z. Thrusters 1 and 2 will control the *x*- and *y*-axes (surge & sway) dynamics, while thrusters 3 and 4 are for the *z*-axis (heave). The three unactuated DoF of roll, pitch and yaw are equal to 0 as they are asymptotically stable and bounded as in Equation 12. All the nonlinear ROV dynamics parameter can be acquired via simulation by Computational Fluids Dynamics (CFD) analysis or by real experimentation using towing tank facilities to estimate and acquire the hydrodynamic forces. Due to cost and time limitation, the parameters of the dynamic model are based on the ROV platform developed by Chin *et al*. (2006).

The proposed U-model based IMC (Figure 3) is implemented to the six DoF ROV nonlinear model using the third order system ($M = 3$) in Equation 15 - $[A_1, A_2, A_3]$. RBF neural networks with five neurons $[n_1, n_2, n_3, n_4, n_5]$ for the A_0. The centre for each neuron, c_i is chosen between 0 to 5 based on the reference input range and $\beta = 2$. The total number of parameters that need to be updated is eight. The learning rate $\mu(t)$ for Equations 19 and 20 is 0.01. All the initial conditions for parameters A_j and W are equal to 0.1 or [0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1]. The learning rate parameter for the controller synthesis algorithm in Equation 22 is $m = 0.00008$. The reference input $r(t)$ is equal to [5 5 1 0 0 0] for the first 1,500 s then towards to [10 10 2 0 0 0] until 3,000 s.

The PID control implementation is done by estimating and tuning to calibrate and ensure a better controller design of the ROV platform. PID gains will be tuned manually for better settling time, oscillation and steady state values responses. Comparison of two PIDS (PID 1 and PID 2) is shown in Table 2. The PID gains are divided into three different directions, which are surge (x), sway (y) and heave (z).

Table 2: Control parameters for PID controllers.

Control parameters	PID 1	PID 2
$K_{p\text{-}x,\ y\ \&\ z}$	0.4271	0.0451
$K_{d\text{-}x,\ y\ \&\ z}$	4.6219	0.0089
$K_{i\text{-}x,\ y\ \&\ z}$	0.0075	0.0089
Filter coefficient (N)	2.7391	100

The FLC inference operation is implemented using 49 rules (or If-Then) statements in Figure 6. The implication-aggregation compositional rule of inference and weighted average method were used in the deffuzzier process based on the Sugeno inference method. The rule table has the same output

membership in a diagonal direction or Toeplitz structure, as shown in Table 3 and Figure 7. The output variable is the voltage applied to the thruster's system.

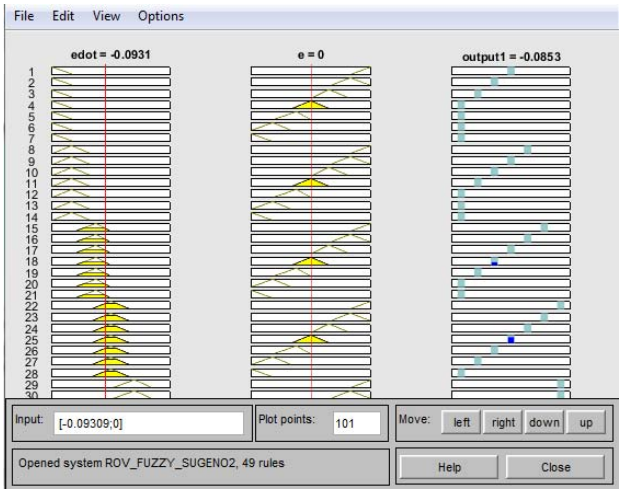


Figure 6: Rule viewer for the FLC control synthesis in MATLAB Simulink.

Table 3: Controlled inputs *e* & *ė* for the FLC control synthesis.

Rule table with Toeplitz structure for Sugeno								
e	***ė***	**PL**	**PM**	**PS**	**Z**	**NS**	**NM**	**NL**
NL		0	-0.25	-0.5	-0.75	-0.75	-0.75	-0.75
NM		0.25	0	-0.25	-0.5	-0.75	-0.75	-0.75
NS		0.5	0.25	0	-0.25	-0.5	-0.75	-0.75
Z		0.25	0.5	0.25	0	-0.25	-0.5	-0.75
PS		0.75	0.25	0.5	0.25	0	-0.25	-0.5
PM		0.75	0.75	0.25	0.5	0.25	0	-0.25
PL		0.75	0.75	0.75	0.25	0.5	0.25	0

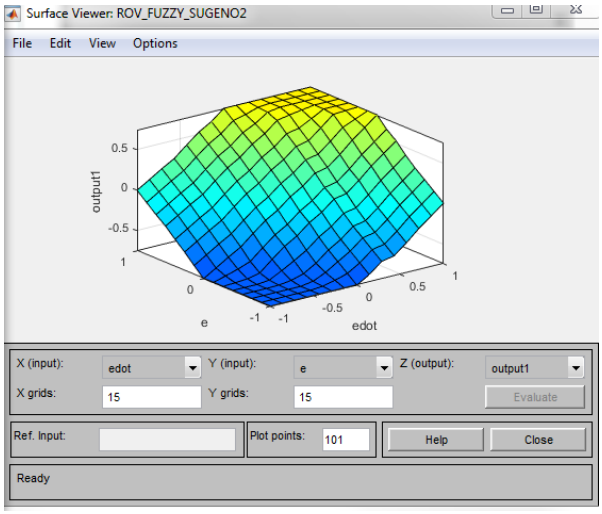
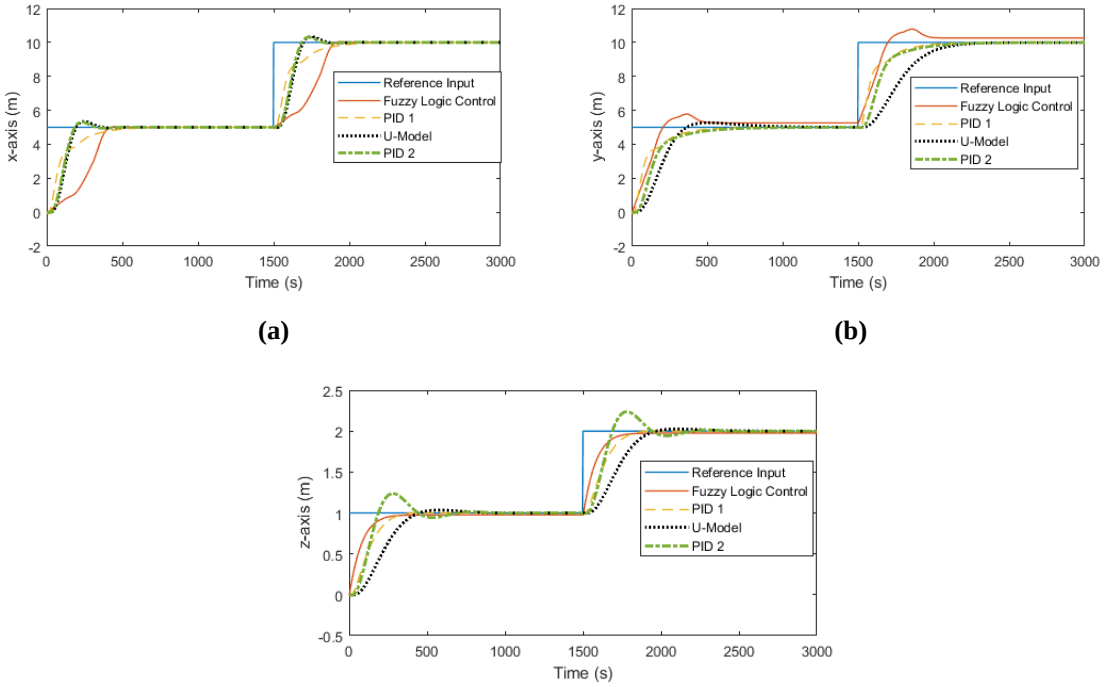


Figure 7: FLC defuzzier surface plot output in MATLAB Simulink.

The results of the comparison of the nonlinear ROV control approaches are presented in Figure 8. The simulation was carried out using MATLAB Simulink to analyse the nonlinear ROV dynamics and control system. Based on the results, all three-MIMO control approaches converge to the desired

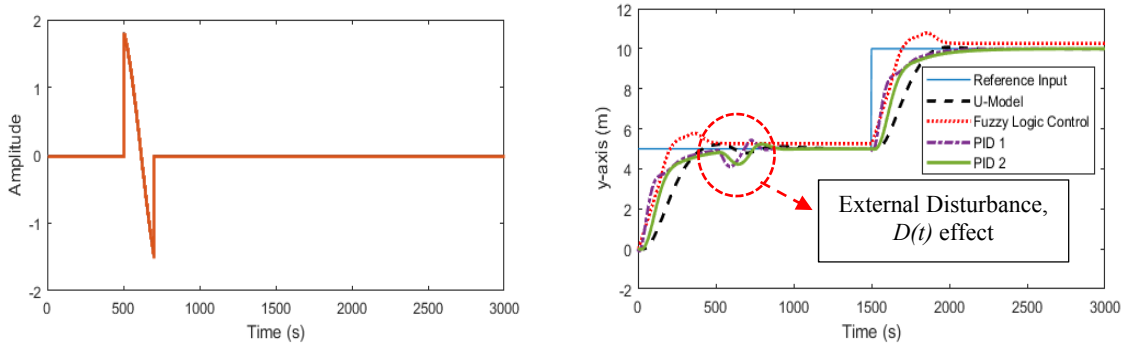
value of the reference input. In this simulation, external disturbances are set to zero and all control synthesis approaches are working properly without any interventions. The U-model and FLC approaches are implemented using adaptive capability based on the tracking error between nonlinear ROV model response and the controller, while the PID control approach only relies on the initial gain value. The results are further analysed by including the external disturbance signal replicating the ocean environment disturbances. The outputs, *Y(t)* for each controller are observed as in Figure 3-5 with disturbance intervention *D(t)*. The external disturbances will usually reduce controller performances unless the control framework approach is adapting the changes in the parameter uncertainty. The external disturbance signal, *D(t)* will appear from 500 to 700 s at the output of the nonlinear ROV model for each control approach as in Figure 9(a). The results of the comparison of controller performances against the external disturbances (position response of the *y*-axis) is presented in Figure 9(b) and the zoomed-in feature during external disturbance event is shown in Figure 10. In this paper, only the results of position response of the *y*-axis is presented for detailed discussion.



(a) **(b)**

(c)

Figure 8: Position response: (a) *x*-axis (b) *y*-axis (c) *z*-axis.



(a) **(b)**

Figure 9: Position response of the *y*-axis with external disturbance: (a) External disturbance signal, *D(t)* (b) Controller response with the disturbance.

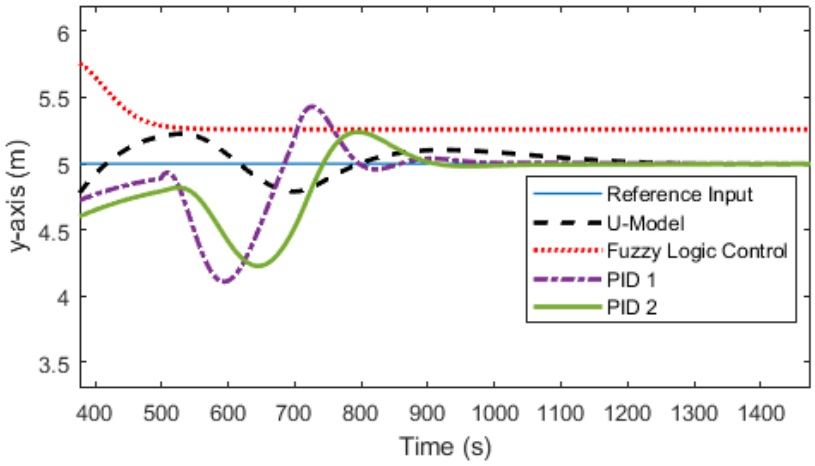


Figure 10: Zoomed-in position response of the *y*-axis with external disturbance interference.

From the comparison of the controller approaches, all the controllers converge to the desired reference input value, except for PID 1 and PID 2 that are affected directly by the external disturbances due to changing of gain values. The control structure is not adapted to the changing parameter due to external perturbation. The controller gains for the PID approach need to be tuned manually until a good control response is achieved. On the other hand, the U-model and FLC approaches are adapted to the changing dynamics due to external perturbation, demonstrating the robustness of the control approaches and adapting capabilities. However, the U-Model IMC approach only converged to a value that is very close to the reference signal value with minimum error. The FLC approach utilised a complex decision-making process (49 rules statements) and computation power from fuzzification, rule-based storage, inference instrument and defuzzification operations, thus giving a higher DoF in tuning its control parameters. For the U-Model, approach only defines the initial values of parameters A_j, W_i and learning rate between (0 – 1) to ensure a good model identification and convergence. The controller then calculates the $U(t\text{-}1)$ online using the Newton-Rhapson algorithm. Nonlinear approximation using the RBF approach is incorporated with the U-model to enhance nonlinear tracking capacity for a system with uncertainties and disturbances.

4. CONCLUSION

The findings of this study demonstrate that adaptive multivariable nonlinear control strategies can be implemented for ROVs with parameter uncertainties. PID control synthesis can be tuned to improve system performance, but the control structure can be degraded due to external disturbances. The U-model and FLC approaches can be implemented in a nonlinear ROV model for controller design. The FLC approach requires a good understanding of the decision-making process and appropriate output rule structure, while the U-model approach only requires suitable initial condition range from 0 to 1 for all the parameters, weights and learning rate values. U-model IMC incorporated with the RBF approach includes adaptive and learning capabilities to reduce the error between the model and multivariable nonlinear ROV system for better control synthesis to sustain the stability of the system.

ACKNOWLEDGEMENT

The research was supported by Fundamental Research Grant Scheme (FRGS/1/2014/TK03/UTP/02/9) and Prototype Development Research Grant Scheme (PRGS/1/2017/TK04/UTP/02/7).

REFERENCES

Abbasi, I., Ali, S.S.A., Ovinis, M. & Naeem, W. (2017). U-model based controller design for an unmanned free swimming submersible (UFSS) vehicle under hydrodynamic disturbances. *Indian J. Geo Mar. Sci.*, **46**: 742–48.

Abbasi, I., Ali, S.S.A., Ovinis, M. & Naeem, W. (2015). Adaptive identification of underwater glider using u-model for depth & pitch control under hydrodynamic disturbances. *Jurnal Teknologi*, **74**: 113–18.

Afande, N.A.A, Ali, S.S.A. & Naufal, M.M.S. (2016). Underactuated nonlinear adaptive control approach using u-model for multivariable underwater glider control parameters. *IEEE 6th Int. Conf. on Underwater Syst. Technol.: Theory and Applications (USYS 16)*, 13-14 December 2016, Pulau Pinang.

Afande, N.A.A, Ali, S.S.A., Naufal, M.M.S., Ovinis, M., Nordin, N. & Adil, S.H. (2017). Underactuated nonlinear adaptive control approach using u-model incorporated with rbfnn for multivariable underwater glider control parameters. *Indian J. Geo Mar. Sci.*, **46**: 2482–92.

Ali, S.S.A., (2007). *U-Model Based Multivariable Nonlinear Adaptive Control*. PhD Thesis, Electrical Engineering. King Fahd University of Petroleum & Minerals, Saudi Arabia.

Ali, S.S.A., Fouad M.A., Shafiq, M. & Jamil M.B. (2010). U-model based learning feedforward control of mimo nonlinear systems. *Electr. Eng.*, **91**: 405–15.

Ali, S.S.A., Moinuddin, M., Raza, K. & Adil, S.H. (2014). An adaptive learning rate for RBFNN using time-domain feedback analysis. *Sci. World J.*, **2014:** ID 850189.

Ali, Z.A., Ronnny, M. & Djoko, P. (2016). Implementation of PID controller for hold altitude control in underwater remotely operated vehicle. *Int. Semin. Intell. Tech. Appl.*, pp. 665–670.

Antonelli, G., Fossen, T. and Yoerger, D. (2008). *Underwater Robotics*. Springer Handbook of Robotics.

Arslan, M.S., Naoto F. & Ichiro, H. (2007). Optimal control of an underwater vehicle with single actuator. *Symp. Underwater Technol. Workshop Scientific Use Submarine Cables Related Tech.*, pp. 581–87.

Borhaug, E., Kristin Y.P. & Alexey, P. (2006). An optimal guidance scheme for cross-track control of underactuated underwater vehicles. *IEEE 14th Mediterranean Conf. Cont. Automation*, pp. 1–5.

Chen, Y., Rongmin Z., Xingyu Z. & Jun, G. (2016). Adaptive fuzzy inverse trajectory tracking control of underactuated underwater vehicle with uncertainties. *Ocean Eng.*, **121**: 123–33.

Chin, C.S. (2013). *Computer-Aided Control System Design: Practical Application Using MATLAB and Simulink*. CRC Press, Boca Raton, Florida.

Chin, C.S., Lau, M.W.S., Eicher, L. & Gerald, S. (2006). A robust controller design method and stability analysis of an underactuated underwater vehicle. *Int. J. Appl. Math. Comput. Sci.* **16**: 345–56.

Chu, Z., Zhu, D. & Yang, S.X. (2016). Observer-based adaptive neural network trajectory tracking control for remotely operated vehicle. *IEEE Trans. Neural Netw. Learn. Syst.*, **28**: 1633–45.

Fossen, T.I. (2002). *Marine Control Systems Guidance, Navigation and Control of Ships, Rigs and Underwater Vehicles, 2nd Ed.* Marine Cybernetics, Trondheim, Norway.

Johnson, J., Madhumitha, G., Niyas, Y. & Shyam, H. (2016). Design, development and fuzzy logic based control of a remotely operated underwater vehicle. *Int. Conf. Robotics Automation Humanitarian Applications (RAHA),* pp. 1-6.

Juan, L., Qingyan, Z., Xinghua, C. & Naeim, F.M. (2015). Path following backstepping control of underactuated unmanned underwater vehicle. *IEEE Int. Conf. Mechatron. Automation (ICMA)*, pp. 2267–72.

Lakhekar, G.V., Waghmare, L.M. & Londhe, P.S. (2015). Enhanced dynamic fuzzy sliding mode controller for autonomous underwater vehicles. *IEEE Underwater Technol. (UT)*, pp. 1-7.

Lam, D. & Donald, W. (2017). Unsupervised feature learning classification with radial basis function extreme learning machine using graphic processors. *IEEE Trans Cybern.,* **47**: 224–31.

Li, B., Yuanxin, X., Chenzhan, L., Shuangshuang, F. & Wen, X. (2015). Terminal navigation and control for docking an underactuated autonomous underwater vehicle. *IEEE Int. Conf. Cyber Technol. Automation, Cont. Intell. Syst. (CYBER)*, pp. 25–30.

Liu, H., Yanhui, W., Xinghe, Z. & Guangchun, L. (2016). Operated ROV thrust distribution control

system based on adaptive back-stepping controller. *35th Chinese Cont. Conf. (CCC)*, pp. 4633–39.

Mutaz, T. & Aziz, A. (2015). Solar radiation prediction using radial basis function models. *IEEE Int. Conf. on Develop. of E-Systems Eng. (DeSE)*, pp. 77–82.

Qiao, L., Lixing, Z. & Weidong, Z. (2016). Robust adaptive PID control for positioning of remotely operated vehicle working in close proximity of an underwater structure. *Chinese Cont. Conf., (CCC)* pp. 5780–85.

Ruano, M.G., Hajimani, E. & Ruano, A.E. (2016). A radial basis function classifier for the automatic diagnosis of cerebral vascular accidents. *Glob. Med. Eng. Phys. Exchanges / Pan American Health Care Exchanges (GMEPE/PAHCE)*, pp. 1–4.

Shafiq, M. & Butt, N.R. (2005). U-model based adaptive IMC for nonlinear dynamic plants. *10th IEEE Conf. Emerging Technol. Factory Automation (ETFA),* pp. 955-959.

Shen, C., Brad, B. & Yang, S. (2017). Modified C/GMRES algorithm for fast nonlinear model predictive tracking control of AUVs. *IEEE Trans. Contr. Syst. Tech.*, **25**: 1896–1904.

ul Amin, R., Li, A., Lu, H. & Li, J. (2016). An adaptive sliding mode control based on radial basis function network for attitude tracking control of four rotor hover system. *IEEE Chinese Guidance, Navigation Cont. Conf. (CGNCC)*, pp. 580–85.

Yuh, J. (2000). Design and control of autonomous underwater robots: A survey. *Autonomous Robots,* **8**:7-24.

Zhu, K. & Gu, L. (2011). A MIMO nonlinear robust controller for work-class ROVs positioning and trajectory tracking control. *Chinese Cont. Decision Conf. (CCDC),* pp. 2565-2570.

FLIGHT SIMULATOR INFORMATION SUPPORT

Vladimir R. Roganov[1*], Elvira V. Roganova[2], Michail J. Micheev[1], Tatyana V. Zhashkova[1], Olga A. Kuvshinova[3] & Svetlan M. Gushchin[3]

[1]Penza State Technological University, Russian Federation
[2]LLC "Video3", Russian Federation
[3]Penza State University of Architecture and Construction, Russian Federation

[*]Email: vladimir_roganov@mail.ru

ABSTRACT

The relevance of creating an information support system for a flight simulator is caused by the need in expanding the list of the training tasks in order to minimize pilot training flights in real flying vehicle. The possibility of solving the set task is retained by insufficient level of development of theoretical and research-methodological aspects for forming the visual model of a significant area of terrain, flying over, which allows training pilots to solve the navigation tasks while correcting the flying vehicle's exposure by the leveling points. The objective of this paper is the development of a structural-functional model of the initial data preparation process of the leveling points, visually observed models that will be used by the pilots for solving navigation tasks during their training flights on a flight simulator, considering testing of the model at development of real flight simulators with expanded list of training tasks. The primary method of the research is simulation of a sufficient number of leveling points when flying in a flight simulator over a distinctive area, which allows for teaching pilots to solve the tasks of the model flying vehicle location in the model space, while at the same time developing the professional skills of solving the navigation tasks that will be used in their real flights. The paper represents a structural-functional model of the initial data preparation process of the leveling points, the use of which will allow customers to develop the requirements towards the visually observed model of a large area of a real terrain, over the model of which the flights in the flight simulator will be performed in order to develop the pilots' professional skills of orientation by the visually observed objects of the environment. The model is aimed at expanding the list of the training tasks that were used to be solved on the flight simulator, by adding to them teaching to flying vehicle piloting in normal and pre-emergency situations, and also teaching to solve the tasks of navigation with the orientation by the observed leveling points.

Keywords: *Structural-functional mod; cognitive visual environment; data mining; reorientation; big data.*

1. INTRODUCTION

The latest achievement of science and technology, having been manifested in development of the cutting-edge hardware systems, high quality two-dimensional display systems perceived by a person as three-dimensional, and recognizable models of real objects, have significantly expanded the list of training tasks for professional training of pilots in operating the flight simulators (Zaytsev, 2005). The use of flight simulators has already been reducing the number of pilot training flights in real flying vehicles to the point that the teaching of an airman trainee to operate a new flying vehicle may be carried out without real flights in that vehicle (Roganov, 2017). Adding a recognizable, visually observed model of a significant real terrain area from 400×400 km to 1,500×1,500 km in size allows for the expansion of the list of training tasks with the flying vehicle position fixing by means of visually observed leveling points (Roganov, 1995). This was impossible before due to the drawbacks of the hardware systems simulating the exterior of leveling points models seen through the cockpit windows as well as on the displays of the relevant emulators (Roganov, 2014), and also due to insufficient level of development of theoretical and scientific-methodological aspects of forming the visually observed model of a significant real terrain area (Krasovskiy *et al*., 2008). This would allow

for teaching of pilots to solve navigation tasks (Mamayev *et al*., 2002), with correction of the flying vehicle location by means of the leveling points, for example, at cross-country flights.

Earlier pilots studied to solve navigation tasks by means of specialized and pilotage flight simulators in the "cloud flying" mode (i.e., with no ground visible) according to the information they obtained from the emulators of radar facilities: aviation compass locators and local observation radio stations (Roganov *et al*., 2017). The tasks of navigation using visually recognizable observed models were solved only at teaching to landing approach to the model runaway, when a pilot operated a flying vehicle model, visually estimating the distance to the model runaway and correcting the model flying vehicle position towards the model runaway location on the ground of the visualization wall (Yurkov *et al*., 2000). In other cases of pilot training flights, it was impossible to solve the task of navigating by means of visually observed models of leveling points (Roganov, 2002).

Adding new flight simulator education tasks to the list of already existing ones has become possible thanks to both emergence of high quality 2D-indicators (Roganov *et al*., 2017), reflecting the model leveling points on the displays of emulators of radio locator and thermal imager, and emergence of 3D indicators of the emulator of visual environment, making the trainee believe that he sees the 3D-objects (in the quality, which allows teaching the trainee to professionally visually determine the distance to the highlighted model leveling point and to train the eye sight measurement for real flights) (Roganov *et al*., 2015), and development of structural-functional model of the initial data preparation process of the leveling points, with visually observed models that will be used by the pilots at solving the navigation tasks during simulator flights (Roganov, 2015).

The results of the research conducted by self-appointed group members at Video3, OOO R&D (Public Contract No. 8009p / 8265 dated 04/30/2010 with Federal Government Budgetary Institution "Foundation for Assistance to Small Innovative Enterprises in Science and Technology"), formed the basis for this paper, by having selected the list of leveling points and located them in the model region of the flights, called the visualization scene, allowing for the expansion of the number of the training tasks solved in the flight simulator, through the addition of teaching to solve the navigation tasks at en-route flight with visual definition of the flying vehicle location in the model space by means of the visible model leveling points.

2. METHODOLOGICAL FRAMEWORK

Information support of flight simulator (FS) should create the conditions for teaching pilots not only to operate a flying vehicle (FV) (Zaytsev, 2005), but also to solve air pilotage navigation tasks (Roganov *et al*., 2017). Analysis of the set task allowed for the revealing of the components of FS information support:

- Map of the virtual space, being a simplified marker of the map of a real land area selected for the FS with the denoted models of the ground-based radio radars used as beacons (Roganov et al., 2017);

- Visuals' (V) databases with primitives, allowing real-time simulation with the use of hardware system called Computer Image Generator at the display of intermediate image formation of 3D display devices of 2D projection of 3D visualization scene, caught by the viewing frustum and its subsets: the database of radio locator (RL) emulator and the database of thermal imager (TI) emulator with the information necessary and sufficient for the real-mode construction of the visualization scene observed models (Shevrov, 2011), with the coloring of the model objects in correspondence with the accepted monochromic color of the relevant emulator's display (Vyatkin *et al*., 1997);

- Information-structural models formalizing the reasonability of including some or other model real objects in to the V, RL and TI databases considering their spatial-temporal, semantic

structures of aprior and experimental data for ensuring the opportunity to teach pilots to FS, FV pilotage and flight navigation (Roganov, 2002);

- Verbal-logical models describing the information obtained by means of expert evaluation at analysis of the model objects, which are supposed to be included into the V, RL and TI databases for creating the conditions of training the professional skills of the pilots, operating FS and acquiring professional skills of FV pilotage and flight navigation (Yurkov et al., 2000);

- Constantly updating experimental information obtained from the emulators of the cockpit devices for solving the navigation task using the "dead reckoning" method (Yurkov *et al*., 2000);

- Technology of obtaining monitoring information, which allows generating information for pilot sufficient for determining the FV location in the virtual space through data mining, using the models of the leveling points caught by all the registers: visual, radio and IR (Gerasimenko *et al*., 2017);

- Technology of data mining of the monitoring information, which allows simulating the FV location in the virtual space by means of dead reckoning due to constant information read-out from the emulators of cockpit devices and drawing the distance covered on the virtual space map with confirmation of the FV location in the virtual space according to the observed model leveling points (Roganov, 2014);

- Algorithm for ranging the objects of the environment, chosen as a prototype at the virtual space simulation in order to take the decision about necessity and sufficiency of their including into the virtual space to ensure an opportunity of solving on FS not only the pilotage tasks, but also the tasks of flight navigation (Prokhorov & Kharin, 1995).

The objective of the paper is to describe the complex system of FS information support, ensuring including flight navigation into the list of the training tasks. The method of creating the complex system of information support development incudes the synthesis of information support (Roganov, 2015), which allows including the tasks of teaching to flight navigation, is one of the big data methods (Ivanov & Vampilov, 2014). Compared to the currently applied FV location definition technology in the model space according to the data from cockpit devices (Yurkov *et al*., 1999), which allow orienting by "cloud flying" without visible ground, a more precise method of the pilots' teaching to detecting the FV location by the visible model leveling points is suggested (Mamayev *et al*., 2002).

Implementation of such a method requires the creation of a complex system of information support development, which allows the pilot to solve navigation tasks at the model FV flight in the virtual space, which implies the opportunity to orient by:

- Visual images of the 3D model leveling points seen through the cockpit windows (for this purpose, it is necessary not only to synthesize the 3D model leveling points in real-time mode and optically transform 2D images into 3D ones, but also to solve the issue of developing the list of such points necessary and sufficient for teaching the pilots to solve the navigation tasks and development of the necessary models choice algorithm, caught by the viewing frustum of the visual surround simulator, from the general visualization scene);

- Images of the model leveling points at the display of the radio locator emulator (considering simple detailization of the model leveling points at the display of the radio locator emulator; as the research showed, one model leveling point is enough, but at the same time it is suggested using the same choice algorithm from the general visualization scene of the model points caught by the viewing frustum of the radio locator emulator, space angle of which may change);

- Visual images of the model leveling points on the display of the thermal imager emulator (considering insignificant distance of IT-image observation, as the research showed, one

model leveling point is enough, but at the same time it is suggested to use similar algorithm of choice from the general visualization scene of the model points caught by the viewing frustum of the thermal imager emulator);

- Information from the cockpit devices emulators with the application of "dead reckoning" method (for teaching of the pilots to en-route flights with the correction of the model FV location by the visually observed model leveling points);

- Information from the radio equipment emulators (for teaching pilots to define the model FV location with correction by the visually observed model leveling points);

- Solution of these tasks implies application of cognitive graphics ensuring solution of the flight navigation tasks. It implies performing the works on development of the relevant model real terrain for V, RL and TI ensuring the opportunity to teach to the flight navigation tasks including definition of the model FV location and en-route.

Such an approach allows the following:

- Forming semantic structure of the information environment, which allows teaching to the flight navigation tasks on the FS;

- Defining in this structure the place for visual environment simulator and the requirements towards it (in particular, setting of the minimum distance for observation of the nearest model visualization scene);

- Defining the place of radio locator emulator, heat imager emulator and model visual space map in this structure;

- Focusing on providing the trainee with an opportunity to use the hierarchical method when choosing the necessary information at the pilot's solving the task of the space reorientation considering his experience, personal preferences and sufficiency of the information obtained to solve the set tasks, which is ensured by the topological structure of the initial data stored in the V, RL and TI databases, defined by the methods of its collection, processing, storage and retrieval from the data bases (Vyatkin *et al*., 1999).

Figure 1 presents the scheme of the data mining, supplied to the pilot training on FS at special reorientation. The information obtained by the pilot allows him to define the model FV location in the virtual space at any time point, using the same methods and algorithms applied by him at definition of the FV location in the real space (Figure 2).

3. RESULTS

Solving the tasks of flight navigation on FS requires synthesizing virtual space, the size of which will allow transferring the model FV and teaching to solving al the known navigation tasks (particularly, the size of the visualization scene for flight simulators developed by a Canadian company, CAE is from 10,000×10,000 km to 15,000×15,000 km, which allows for solving the set tasks) (Vyatkin & Dolgovesov, 2002). Except for the large size, the visualization scene should have sufficient saturation of the simulated virtual space with the model leveling points. Synthesis of such a space is a complex task, gradually solved by FS developers.

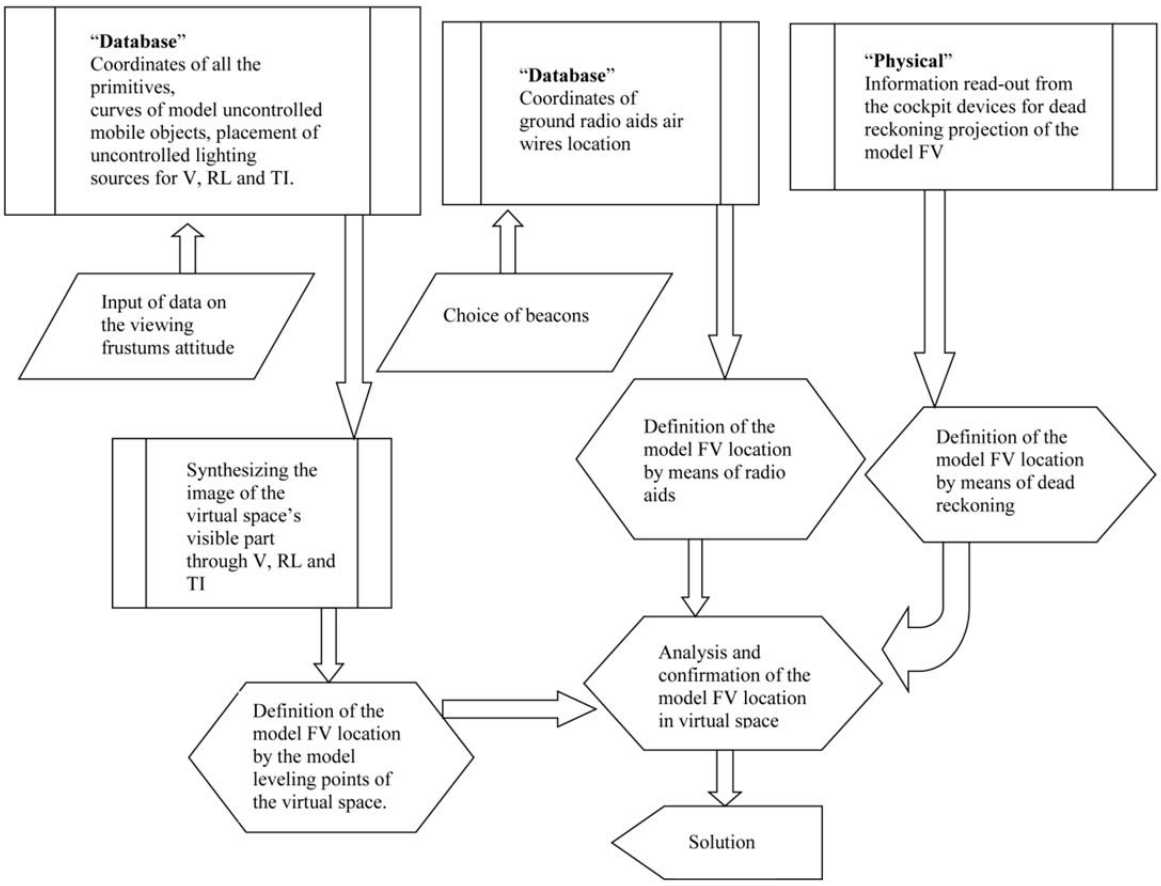


Figure 1: *AD* data mining.

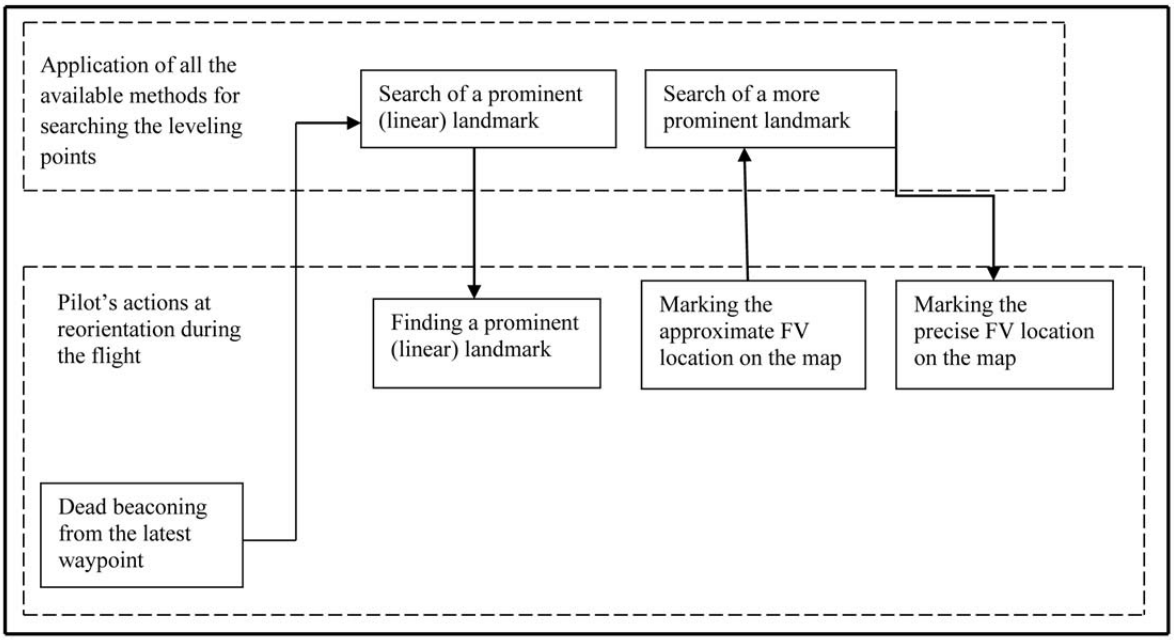


Figure 2: Actions of the pilot at defining the model FV location in the virtual space.

Considering the fact that it is impossible to synthesize an accurate copy of a real terrain, the task of virtual space synthesizing is solved on the basis of the requirements specification on the FS (Mikheev

et al., 2017). This task starts from a preliminary stage, where necessary information is being collected: from the pilots' surveying – what object they need for performance of real tasks on FV to the computations – what models of the selected objects may be synthesized using Computer Image Generator, during a specific real-time period (Mikheev *et al*., 2016). For applicably to the task of synthesizing the visually observed models of the 3D object, we will obtain the following:

- For V, the database for the Computer Image Generator of visuals and the map of the virtual space;
- For RL emulator, the database for Computer Image Generator of radio locator emulator DBr ;
- For thermal imager emulator, the database for Computer Image Generator of thermal imager emulator DBq .

All the information about the selected objects of simulation should correspond to the points marked at the map of virtual space. The analysis showed that some of the Mv model objects stored in the DBv, may be included in the DBr database with the corresponding Mr coloring, and also colored as Mq - into the DBq database. Thus, the majority of the visualization objects seen at the displays of radio locator and thermal imager are sub-majority of the visualization objects seen through the FS cockpit window.

$Mr \subset Mv$, this requirement is met only for the area of virtual space where flights on FS are possible, and does not cover the part of virtual space bordering upon this area, but where the flights are possible, but this area can be seen at the display of radio locator emulator on an appropriate scale.

$Mq \subset Mv$, this requirement is always met.

Due to some peculiarities of radio locator image formation a part of Mr_i , visualization objects seen on the radio locator display cannot be seen on the thermal imager display and vice versa. Figure 3 shows the stages of preparing the information for visuals, emulators of radio locator and thermal imager.

4. DISCUSSION

The list of the tasks solved in the flight simulator that is to be included into the requirements specification for the simulator is prepared by the customer, usually by experienced pilots of large companies. Based on the analysis of the pilots' actions in situations that result in crashes or prerequisites to a flight accident, they developed requirements to teaching pilots for professional skills of flight navigation in normal and critical flight situations, so that during the flight in the FS, the pilot could automatize the setting of the command of the control gear to prevent negative development of a specific situation. The set tasks are successfully performed on the pilotage flight simulators and today, there is no pilot who would not have trained in a flight simulator. According to the existing practice, all the tasks solved in the FS are discussed by the customer with the representatives of the flight simulator manufacturing companies in order to constantly expand the list of training situations on the basis of flight accidents analysis.

Recently, not much attention has been paid to simulators for navigating officers. Specialized simulators for training navigating officers could imitate case studies at "cloud flying". Teaching to solve navigation tasks with definition of flying vehicle location by visually observed leveling points at an area of the visualization scene exceeding the airport terrain with the runaway had been impossible. The reasons of insufficient attention to solving navigation tasks is insufficient number of simulators to model a large region of a real terrain (the size of which exceeds 400×400 km) with sufficient number of recognizable model leveling points, which could be mapped to the displays of radio locator and thermal imager emulators or could be observed as 3D through the cockpit window. Such a situation was explained by both technical characteristics of the applied displays of radio locators and thermal

imager emulators and the necessity to solve scientific and research tasks of packing, storage and timely retrieval of the corresponding data (today these tasks may be solved by means of recently developed methods and algorithms of big data).

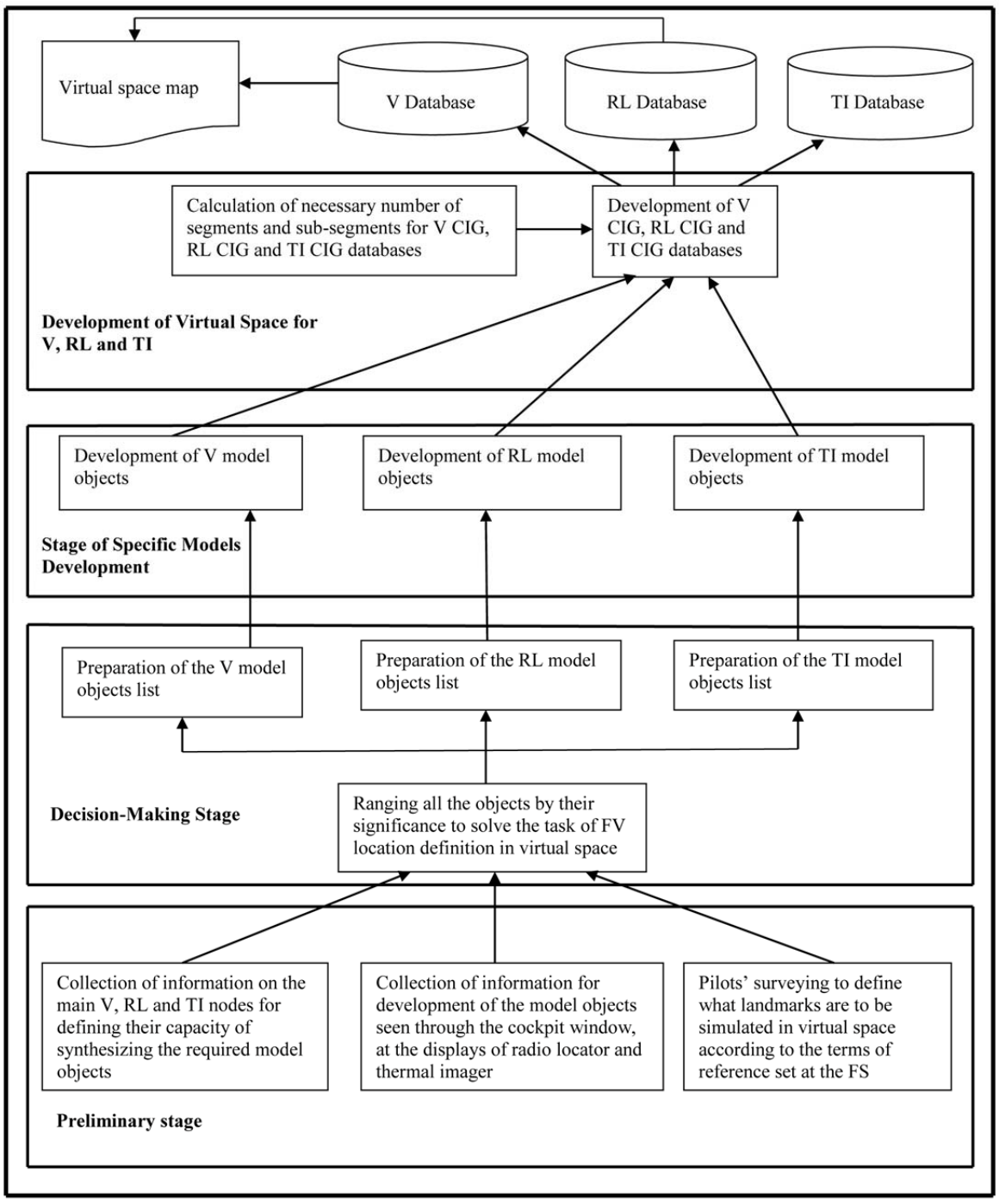


Figure 3: Stages of preparing the information for visuals, emulators of radio locator and thermal imager.

5. CONCLUSION

Thus, it is currently possible to expand the list of the tasks, solved by the trainees in flight simulators. What it involves is a solution of navigation tasks using the methods and algorithms for definition of the flying vehicle location in the space by the visually observed model leveling points earlier used by pilots at orientation by the visually observed leveling points in real flights.

The possibility of solving new tasks on flight simulators is connected with new technical characteristics of the radio locator and thermal imager emulators' displays with new capacities of the devices for 3D indication of the visual environment emulator, and capacities of Computer Image Generator hardware systems with the databases with all the necessary information stored about sufficient number of model leveling points that allow solving navigation tasks at flights over the entire visualization scene with the size not exceeding 400×400 km. In order to solve the tasks of choice, pack and retrieval of information from the databases it is necessary to use the algorithms of the model objects interconnection, based on consideration of the requirements specifications and on flight simulator with expanded list of training tasks, which requires synchronization of the information included into the databases for computer image generator (CIG) for visuals, radio locator emulator and thermal imager emulator. At the same time, each database of the relevant emulator should contain sufficient model leveling points that is necessary for solving the most simple navigation task – teaching pilots to en-route flight.

The materials of the paper represents value for developers of professional simulators for pilots and drivers; for developers of computer games, and simulators of aircrafts and cars; and for scientists working in the area of ground profile simulation.

REFERENCES

Gerasimenko, A., Ivanov, K., Kislyukov, V. & Roganov, V. (2017). A new faults blocking method for out-of-step protection. *Proc. IEEE East-West Design Test Symp. (EWDTS'2017)*, 29 September - 2 October 2017, Novi Sad, Serbia.

Ivanov, P.D. & Vampilov, V.Z. (2014). Big data technologies and their application at the modern industrial enterprise. *Eng. J. Sci. Innov.*, **8**: 52-56.

Krasovskiy, A.A., Lapshin, E.V. & Yurkov, N.K. (2008). *Mathematic Simulation of an Aircraft Flight Dynamics*. PGU, Penza.

Mamayev, V.Y., Sinyakov, A.N., Petrov, K.K. & Gorbunov, D.A. (2002). *Air Navigation and Elements of Flight Navigation*. SPbGUAP, Saint Petersburg.

Mikheev, M.Y., Roganov, V.R., Andreev, P.G., Goryachev, N.V. & Trusov, V.A. (2017). *Developing the structure of the quality control system of power supply units in mobile robots. Int. Siberian Conf. Control Commun.* (SIBCON), 29-30 June 2017, Astana, Kazakhstan.

Mikheev, M.Yu., Zhashkova, T.V., Shcherban, A.B., Grishko, A.K. & Rybakov, I.M. (2016). *Generalized structural models of complex distributed objects*. *Proc. IEEE East-West Design Test Symp. (EWDTS'2017)*, 14-17 October 2017, Yerevan.

Prokhorov, A.V. & Kharin, N.P. (1995). *Ranging of Documents in Descending Order by their Semantic Match to the Request on the Basis of Automotive Accounting of the Associative Relations Built*. Central Russian House of Knowledge, Moscow.

Roganov, V., Miheev, M., Roganova, E., Nurgozhin, B.I. & Fillipenko, V. (2017). *Main Provisions for Formation of Cognitive Model of Visually Observable Environment Synthesized for Aircraft Simulator*. Atlantis Press, Paris.

Roganov, V.R.. (2014). Analysis of the operators training simulator indicative devices. *Defense Complex Res. Sci. Prog. Russ.*, **4**: 80-87.

Roganov, V.R.. (2002). *Methods of Forming the Virtual Reality*. Penza State University, Penza.

Roganov, V.R.. (1995). *Organization of Visual Databases and Management of Computer Generators of the Images of the Visual Environment Simulators*. PhD Thesis, State Technical University, Penza.

Roganov, V.R., Asmolova, E.A., Seredkin, A.N., Chetvergova, M.V., Andreeva, N.B. & Filippenko, V.O. (2014). Problem of virtual space modelling in aviation simulators. *Life Sci. J.*, **11**: 371-373.

Roganov, V.R., Sagyndyk, A.B., Akhtarieva, R.F., Beisenbayeva, A.K. & Sannikova S.I. (2017). Integrated organization of the system for forming the information support of aeronautical simulator. *Int. J. Appl. Eng. Res.,* **12**: 5207-5213.

Roganov, V.R., Semochkina, I.Yu. & Tyurin, M.V. (2015). On the necessity of considering the complex approach at creation and research of information models of the simulation systems' virtual space. *Reliab. Qual. Complicated Systi*, **4**: 38-45.

Roganov, V.R. (2015). Analysis of theoretic aspects of development of a cognitive model of orientation in the visually observed environment and their application for improvement of flight simulators. *21st Century Res. Past Challe. Present*, **4**: 88-93.

Shevrov, R. (2011). Use by the Ministry of Defense of Germany of the commercial satellites to create three-dimensional model of the earth landscape. *Foreign Mil Rev* **6**: 65-66.

Vyatkin, S.I. & Dolgovesov, B.S. (2002). Convolution surfaces synthesis with recursive division of the object space. *Autometry*, **4**: 58-65.

Vyatkin, S.I., Dolgovesov, B.S. & Ovechkin, V.V. (1997). Chizhik, S.E. *Photorealistic Imaging of Digital Terrains, Freeforms and Thematic Textures in Real-Time Visualization System Voxel-Volumes*. SPIE, Bellingham, Washington.

Vyatkin, S.I., Dolgovesov, B.S., Yesin, A.V., Schervakov, R.A. & Chizhik, S.E. (1999). Voxel volumes volume-oriented visualization system. *Int. Conf. Shape Modeling Appl.*, pp. 371-378.

Yurkov, N.K., Andreev, A. N., Blinov, A. V. & Yakimov, A. N. (1999). Conceptual approach to introduction of information technology into the field of simulation. *Meas. Tech.*, **42**: 421-426.

Yurkov, N.K., Andreev, A.N., Danilov, A.M, Klyuev, B.V., Lapshin, É.V. & Blinov, A.V. (2000). Information models for designing conceptual broad-profile flight simulators. *Meas. Tech.*, **43**: 667-672.

Yurkov, N.K., Blinov, A.V. & Maksud, D.S. (2000). Diagnosis of restorable components of special-purpose on-board data-acquisition systems. *Meas. Tech.*, **43**: 578-580.

Zaytsev, V. (2005). Application of simulators at training of the US navy pilots. *Foreign Mil. Rev.*, **2**: 59-64.

IMPLEMENTATION OF PARAMETER MAGNITUDE-BASED INFORMATION CRITERION IN IDENTIFICATION OF A REAL SYSTEM

Md Fahmi Abd Samad* & Abdul Rahman Mohd Nasir

Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia

*Email: mdfahmi@utem.edu.my

ABSTRACT

Model structure selection is one among crucial steps in system identification and in order to carry out this, an information criterion is needed. It plays an important role in determining an optimum model structure with the aim of selecting an adequate model to represent a real system. A good information criterion should not only evaluate predictive accuracy but also the parsimony of the model. In the past, there had not been, or scarcely have been, any information criterion that evaluates the parsimony of model structures (bias contribution) based on the magnitude of parameter or coefficient. However, recently, some efforts had been made that took into account such strategy and proved the criterion to be effective for simulated datasets. This paper presents the comparison between two information criteria that are based on parameter magnitude information in selecting a good model to represent a real system based on gas furnace. Genetic algorithm (GA) was used to optimise the implementation. The selected models were then tested using correlation tests for model validation. It is shown that parameter magnitude based information criterion 2 (PMIC2) is able to select a more parsimonious model than PMIC but with similar validation results.

Keywords: *Parameter magnitude; information criterion; system identification; linear regressive model; genetic algorithm (GA).*

1. INTRODUCTION

In many fields, nonlinear dynamic modelling is applied in approximating a wide range of systems. Basically, system identification is used for estimating a model to represent their systems. System identification can be considered a regression problem, where the relationship between input and output variables of a dynamical system has to be estimated. This task is typically accomplished by minimising a certain information criterion, which measures how well the estimated relationship approximates the one that truly links the available input-output data pairs (Ljung, 1999). Its basic idea is to compare the time dependent responses of the actual system and identified model based on a performance function, hereby referred to as information criterion, giving a measure of how well the model response fits the system response (Alfi & Fateh, 2010).

An identification procedure typically consists of estimating the parameters of different models, and next, selecting the optimal model complexity within that set. Increasing the model complexity will decrease the systematic errors, but, at the same time, the model variability increases (Riddef *et al.*, 2004). Model accuracy and parsimony, known as variance and bias: $f(J) = Var(J) + Bias(J)$, are important considerations in selecting a model structure (Ljung, 1999). Hence, selecting a model with the smallest variance is not a good idea because when the number of parameters increase, the variance will continue to decrease but will present a complex model. At a certain complexity, the additional parameters no longer reduce the systematic errors but are used to follow the actual noise realisation on the data. Often, in order to deal with the bias-variance trade-off, the information criterion is augmented with a penalty term intended to guide the search for the "optimal" relationship penalising undesired regressors, where regressors refer to possible terms and variables. Regularised estimation has been widely applied in the context of system identification (Prando *et al.*, 2015). Model validation

is the final step of system identification. It is used to check the goodness of the estimated model (Zhang *et al.*, 2005).

In Samad & Nasir (2017a, b, c, 2018a), parameter magnitude-based information criterion 2 (PMIC2) was tested on simulated datasets generated from autoregressive with exogenous input (ARX) and nonlinear autoregressive with exogenous input (NARX) models. Comparisons were made between PMIC2, Akaike information criterion (AIC) (Akaike, 1974), Bayesian information criterion (BIC) (Schwartz, 1978) and corrected Akaike information criterion (AICc) (Anderson *et al.,* 1994). In the simulation results, PMIC2 proved that it performed better than AIC, BIC and AICc for all the simulated models (ARX and NARX). In this paper, the effectiveness of PMIC2 will be studied by testing on real system data, in this case, gas furnace data. The models were generated as NARX models. Since there are many possible models to test, genetic algorithm (GA) was used to optimise the search. The models selected by PMIC2 were tested by correlation tests for model validation. The results of PMIC2 were then compared to another parameter-magnitude based information criterion (Samad *et al.*, 2013), hereby denoted PMIC, in order to evaluate the performance between both information criteria.

The next sections are as follows: Section 2 introduces information criteria; Section 3 explains about GA, which is used as part of the search for model structure; Section 4 explains the model validation tests; Section 5 provides simulation setup; Section 6 presents result and discussions; and lastly, Section 7 concludes the findings of this paper.

2. INFORMATION CRITERIA

According to Samad *et al.* (2013), parameter magnitude-based information criterion (PMIC) evaluates the bias contribution by the sum of squared residuals while the variance contribution is calculated by a penalty function (PF). This is written as follows:

$$PMIC = (\sum_{t=k}^{N}(y(t) - \hat{y}(t))^2) + PF \tag{1}$$

where $PF = \ln n$, n is the number of terms satisfying $(|a_j| < penalty) + 1$ while $|a_j|$ represents the absolute value of the parameter for term j and $penalty$ is a fixed value termed penalty function parameter. The penalty function penalises terms with the absolute values of the estimated parameter less than the penalty. The penalty parameter value will be set equal to or slightly lower than the parameter value in cases where the smallest tolerable absolute parameter value is known or can be roughly estimated.

By referring to Samad & Nasir (2018b), the effectiveness of PMIC was compared to AIC, AICc and BIC, where, overall, PMIC was found to perform better in model structure selection, where it can select the true model in the form of ARX and NARX models for given datasets.

The PMIC2 was developed from the approach of using PMIC. In order to overcome the hassle of determining the suitable penalty function parameter of PMIC, it modifies the bias term or known as penalty function, and begins as follows:

$$PMIC2 = RSS + PF \tag{2}$$

where *RSS* is the residual sum of square or defined as the maximised value of the likelihood function for the estimated model. It can be written as follows:

$$RSS = \sum_{t=k}^{N} \varepsilon^2(t) = \sum_{t=k}^{N}(y(t) - \hat{y}(t))^2 \tag{3}$$

The basic theory of the penalty function for PMIC and PMIC2 revolves around the consideration that the magnitude of parameter could have a big role in choosing whether a term is significant enough to be included in a model. This assists one's judgment in choosing or discarding a term / variable. In PMIC2:

Considering, $\theta_1 > \theta_2 > \cdots > \theta_j$

One should give: $pen_1 < pen_2 < \cdots < pen_j$

where θ is the value of parameter in a model and pen is the value of penalty applied for having the variable / term associated with the parameter. Applying a big penalty for having a variable / term that has small parameter magnitude may cause the model to be unfavourable in comparison to other models. In other words, the variable / term that was penalised is said to be insignificant to the model's accuracy in comparison to a variable / term that has big parameter magnitude. One way to realise this is using the following equation:

$$pen = \frac{1}{\theta} \tag{4}$$

Hence, the equation of PMIC2 is:

$$PMIC2 = \sum_{t=k}^{N}(y(t) - \hat{y}(t))^2 + \sum \frac{1}{\theta_j} \tag{5}$$

where $\hat{y}(t)$ and $y(t)$ are the *k*-step-ahead predicted output and actual output value at time t respectively, N is the number of data, θ_j is the magnitude of parameter in the model, and *j* is the number of parameters.

3. GENETIC ALGORITHM (GA)

Among all methods in evolutionary computation, GA is probably the most widely known. Its application is recorded in various fields, including image processing, pattern recognition, operational research, biology and computer sciences (see e.g., Haupt and Haupt, 2004, Kumar *et al.,* 2010). There are three main characteristics of GA (Holland, 1992, Eshelman, 2000):

i. The algorithm uses binary bit string representation.
ii. The selection method used is fitness-proportional selection.
iii. It uses crossover as its main genetic search operator.

Other genetic operators are mutation, which is used as a 'background operator' to prevent loss of important gene information or allele, and inversion, which is particularly significant for permutation-based coding. The outline of the algorithm was initially known as reproductive plan using genetic operators (Holland, 1973).

The number of strings is known as the population size, typically denoted *popsize*. Each chromosome consists of genes separated by different positions also defined as *locus.* The genes carry information of the chromosome. In the simulation, the chromosomes represent different model structures and are evaluated based on a specified information criterion. Based on the evaluation, the selection of highly fit individuals and genetic manipulation stage, that includes the crossover and mutation then takes place. The selection and genetic manipulation of these chromosomes are usually performed to a fixed number of generations, denoted *maxgen*. Other related terminologies can be referred in Haupt & Haupt (2004) and Bäck & Fogel (2000). The flow chart of a simple GA is given in Figure 1.

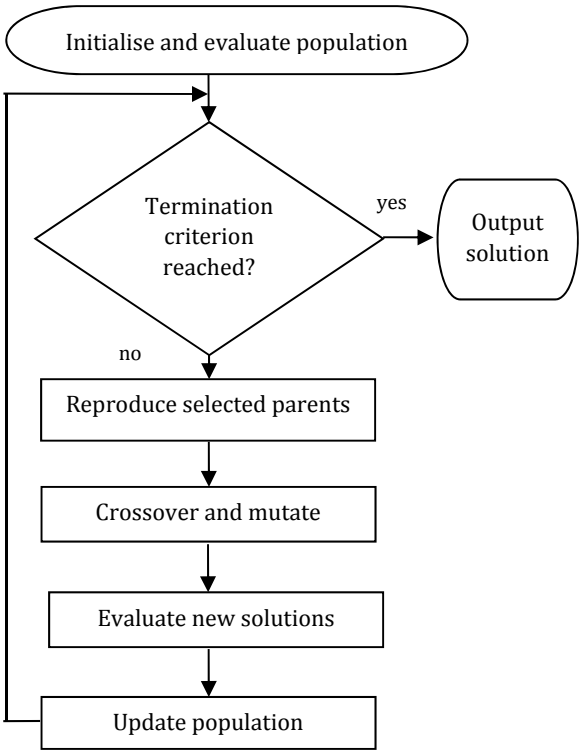


Figure 1: Flow chart of a simple GA.

4. MODEL VALIDATION TESTS

Once the final model is identified, the final step of system identification, which is model validation, is carried out. A model can only be accepted as valid once it is proven that the selected terms and / or variables do not contribute bias to its accuracy. The correlation tests were used to ensure that no other significant terms and / or variables were omitted from the model. Nonlinear models require more tests than linear models since nonlinear models contain polynomials of variables. The tests are as follows (Billings & Voon, 1986):

i. Autocorrelation of residuals:

$$\phi_{\varepsilon\varepsilon}(\tau) = \frac{E[\varepsilon(t)\varepsilon(t-\tau)]}{E[\varepsilon^2(t)]} = \delta(\tau) \tag{6}$$

where $\delta(\tau)$ is the Kronecker delta such that $\delta(0) = 1$ and $\delta(\tau) = 0$, for $\tau \neq 0$.

ii. Cross correlation of residual with input:

$$\phi_{u\varepsilon}(\tau) = \frac{E[u(t)\varepsilon(t+\tau)]}{\sqrt{E[u^2(t)\varepsilon^2(t)]}} = 0, \qquad \forall\tau \tag{7}$$

iii. $$\phi_{\varepsilon(\varepsilon u)}(\tau) = \frac{E[\varepsilon(t)\varepsilon(t-1-\tau)u(t-1-\tau)]}{\sqrt{E[\varepsilon^2(t)]E[\varepsilon^2(t)u^2(t)]}} = 0, \tau \geq 0 \tag{8}$$

iv. $$\phi_{u^{2\prime}\varepsilon}(\tau) = \frac{E\left[\left(u^2(t)-\overline{u^2}\right)\varepsilon(t+\tau)\right]}{\sqrt{E\left[\left(u^2(t)-\overline{u^2}\right)^2\right]E[\varepsilon^2(t)]}} = 0, \qquad \forall\tau \tag{9}$$

v. $$\phi_{u^{2\prime}\varepsilon^2}(\tau) = \frac{E\left[\left(u^2(t)-\overline{u^2}\right)\varepsilon^2(t+\tau)\right]}{\sqrt{E\left[\left(u^2(t)-\overline{u^2}\right)^2\right]E[\varepsilon^4(t)]}} = 0, \qquad \forall\tau \tag{10}$$

where $u(t)$ is the input, τ is the lag order and $E[\cdot]$ is the expectation operator that can be calculated based on the formula

$$\hat{\phi}_{x'y'}(k) = \frac{\frac{1}{N}\sum_{t=1}^{N-k}(x(t)-\bar{x})(y(t+k)-\bar{y})}{\sqrt{(\phi_{x'x'}(0)\phi_{y'y'}(0))}} \tag{11}$$

and $\hat{\phi}_{x'y}(\tau) = \hat{\phi}_{x'y'}(\tau)$ with $\bar{y}$ set to zero. The residual, $\varepsilon(t)$, is calculated by the following:

$$\varepsilon(t) = y(t) - \hat{y}(t) \tag{12}$$

with $y(t)$ as the actual output and $\hat{y}(t)$ as the predicted output. The overbar denotes the time average so that $\overline{u^2}$ is given by:

$$\overline{u^2} = \frac{1}{N}\sum_{t=1}^{N} u^2(t) \tag{13}$$

In all the tests above, the accepted bandwidth for a model's fit to a system is approximately $\pm 1.96/\sqrt{N}$ when allowed 95% confidence interval with N as the number of data points.

5. SIMULATION SETUP

In this section, PMIC2 was tested with real data, in this case gas furnace data, through simulation using MATLAB software. This input-output data was described as an actual process plant data available in Jenkins & Watts (1968) and Box *et al.* (1994), consisting of a discrete stochastic input series of gas feed rate in ft^3/min and output series of carbon dioxide concentration in outlet gas. There were 296 pairs of input-output data sampled at an interval of 9 s. In Jenkins & Watts (1968) and Box *et al.* (1994), the process was found to be adequately represented by a second-order input and output lags, but not tested using correlation tests. In Jamaluddin *et al.* (2007), PMIC was used for this data set to choose a mathematical model to represent the gas furnace system and the model was also tested using correlation tests.

To allow NARX modelling, the following specification, which was recommended in Ahmad *et al.* (2004), was used in this study: output lag, n_y = 2; input lag, n_u = 2; and nonlinearity, l = 2. With this specification, the number of regressors amounts to 15 and the search space has 32,767 solution points. The least squares method was used as parameter estimation method.

The specification of the GA is fixed where the population size, with *popsize* set to 500; the maximum generation is 100; the mutation probability, p_m = 0.01; and the crossover probability, p_c = 0.6. The algorithm implements roulette-wheel selection, single-point crossover and binary bit mutation (Bäck *et al.,* 2000).

For model validation test, the model selected by PMIC2 was tested using correlation tests. All five tests listed in Section 4 were made.

6. RESULTS AND DISCUSSION

From the simulation test using GA, PMIC2 selected the model, represented in binary string as [111 010 000 000 000] as the best chromosome. Each bit in the string represents a specific variable or term. By knowing the sequence of variable/term in the simulation program, the selected regressor may be easily be traced back when bit 1 is found. Other than the calculated direct current level (which is equivalent to a constant), the variables selected are $y(t-1)$, $y(t-2)$ and $u(t-2)$. From Jamaluddin *et al.* (2007) and Samad (2017), PMIC selected the model [111 010 010 000 000] as the

best chromosome. The selected regressors and its parameter values selected by both PMIC2 and PMIC are provided in Table 1. Comparisons could be made directly, where it shows that PMIC2 selected a more parsimonious model than PMIC. PMIC2 selected a model with 4 out of 15 regressors, while PMIC selected more regressors, which is 5 out of 15 regressors. As can be seen from the table, a nonlinear term, $y(t-1)u(t-1)$, is also selected by PMIC. With this additional term, considering that a control system is to be developed for the gas furnace system, additional control plan needs to be developed, which in turn, will require higher cost and thus undesirable. The model selected by PMIC2 promises simpler control plan as there are less variables to be controlled.

Table 1: Variables, terms and parameter values of selected model by PMIC2 and PMIC for gas furnace data.

Variable / Terms	Parameter Value using PMIC2	Parameter Value using PMIC
Constant or d. c. level	6.4103	7.1795
$y(t-1)$	1.5073	1.4449
$y(t-2)$	-0.6274	-0.5793
$u(t-2)$	-0.3790	-0.6600
$y(t-1)u(t-1)$	-	0.0047

Figure 2 shows the result of correlation tests on the model selected by PMIC2. It seems to be acceptable where only a few points were out of the interval. Compared to the correlation test for a model selected by PMIC in Figure 3 (Samad, 2017), it looks quite similar for all tests. According to Jamaluddin *et al.* (2004), these validation deficiencies may be inherent from wrong selection of lag orders or nonlinearity of the model. It is also plausible that based on the similarities of the situation to the heat exchanger problem in Billings & Fadzil (1985), the deficiencies may be caused by lack of noise terms.

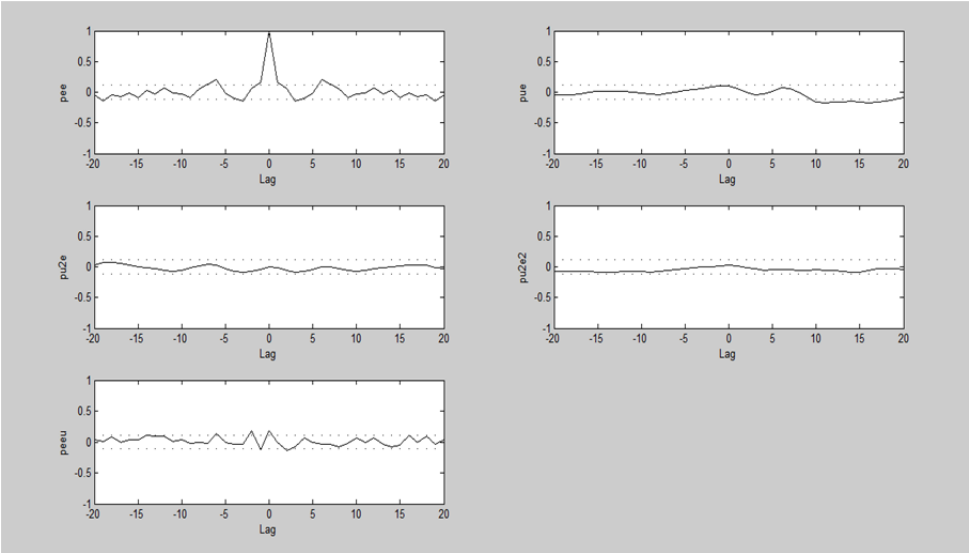


Figure 2: Correlation tests of selected model by PMIC2 using GA for gas furnace data.

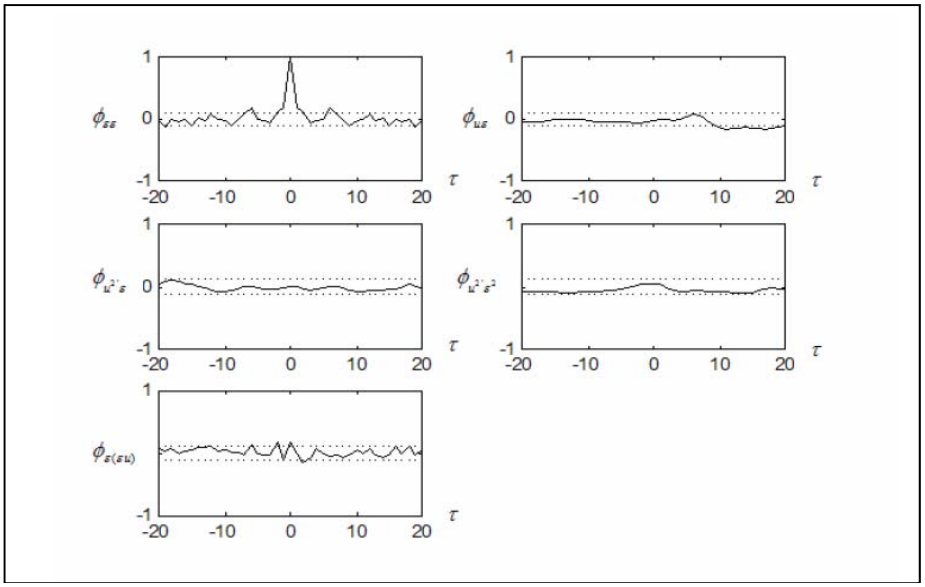


Figure 3: Correlation tests of selected model by PMIC using GA for gas furnace data

7. CONCLUSION

From the results, PMIC2 selected a more parsimonious model for gas furnace data than PMIC. This was also achieved without having to try all possible models, but by utilising GA. The results of correlation tests for the model seem good. It must be emphasised that PMIC2 overcomes the hassle of selecting the suitable value of penalty function parameter before use in PMIC. Together with other studies done using PMIC2 on simulated models, PMIC2 can be considered as a reliable information criterion for model structure selection of discrete-time model. Future work shall be concentrated on PMIC2 use for other real data, datasets with specifications of bigger search space and different forms of discrete-time modelling such as linear, and nonlinear autoregressive moving average with exogenous variable.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support from Universiti Teknikal Malaysia Melaka and Ministry of Higher Education Malaysia for research grant FRGS/1/2015/TK03/FKM/02/F000271.

REFERENCES

Ahmad, R., Jamaluddin, H. & Hussain, M.A. (2004). Selection of a Model Structure in System Identification Using Memetic Algorithm. *Proc. 2nd Int. Conf. Artificial Intelligence and Engineering Technology*. Aug 3-5. Universiti Malaysia Sabah. 714-720.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.*, **19**: 667-674.

Alfi, A. & Fateh, M.M. (2010). Parameter identification based on a modified PSO applied to suspension system. *J. Softw. Eng. Appl.*, **3**: 221-229.

Anderson, D.R., Burnham K.P. & White, G.C. (1994) AIC model selection in overdispersed capture–recapture data. *Ecology*, **75**: 1780-1793.

Bäck, T, Fogel, D.B. & Michalewicz, Z. (Eds.) (2000). *Evolutionary Computation 1: Basic Algorithms and Operators*. Institute of Physics Publishing.

Bäck, T. & Fogel, D.B. (2000). Glossary. *In* Bäck, T, Fogel, D. B. and Michalewicz, Z. (Eds.). *Evolutionary Computation 1: Basic Algorithms and Operators*. Institute of Physics Publishing, Bristol, pp. xxi-xxxvii.

Billings, S.A. & Fadzil, M.B. (1985). The Practical Identification of Systems with Nonlinearities. *Proc. 7th IFAC Symp. Identification System and Parameter Estimation*. York, United Kingdom, 155-160.

Billings, S.A. & Voon, W.S.F. (1986). Correlation Based Model Validity Tests for Non-Linear Models. *Int. J. Contr.*, **44**: 235-244.

Box, G.E.P., Jenkins, G.M. & Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control, 3rd Ed.* Prentice-Hall Englewood Cliffs, New Jersey.

Eshelman, L.J. (2000). Genetic Algorithms. *In* Bäck, T, Fogel, D. B. and Michalewicz, Z. (Eds.), *Evolutionary Computation 1: Basic Algorithms and Operators*, Institute of Physics Publishing, Bristol, pp. 64-80.

Haupt, R.L. & Haupt, S.E. (2004). *Practical Genetic Algorithms, 2nd Ed.*, John Wiley and Sons, Hoboken, New Jersey.

Holland, J.H. (1973). Genetic Algorithms and the Optimal Allocation of Trials. *SIAM J. Comput.,* **2**: 88-105.

Holland, J.H. (1992). *Adaptation in Natural and Artificial Systems*. MIT Press, Massachusetts.

Jamaluddin, H., Samad, M.F.A., Ahmad, R. & Yaacob, M.S. (2007). Optimum Grouping in a modified genetic algorithm for discrete-time, non-linear system identification. *Proc. IMechE Part I: J. Syst. Contr. Eng.*: **221**: 975-989.

Jenkins, G.M. & Watts, D.G. (1968). *Spectral Analysis and Its Applications*. Holden-Day, San Francisco.

Kumar, M., Husian, M., Upreti N. & Gupta D. (2010). Genetic Algorithm: Review and Application. *Int. J. Inf. Tech. Knowl. Manage.,* **2:** 451-454.

Ljung, L. (1999). *System Identification: Theory for the User,* 2nd Ed., Prentice Hall, Upper Saddle River, New Jersey.

Prando, G., Pillonetto G. & Chiuso, A. (2015). The role of rank penalties in linear system identification. *IFAC-Papers Online*, **48**: 1293-1300.

Riddef', F.D., Pintelon, R., Schoukens, J. & Gillikinb, D.P. (2004). Modified AIC and MDL model selection criteria for short data records, *IEEE Trans. Instrum. Meas.*, **54**: 144-150.

Samad, M.F.A (2017). *Applications of Evolutionary Computation in System Identification*. UTeM Press, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Samad, M.F.A. & Nasir, A.R.M. (2017a). Parameter-magnitude based information criterion for identification of linear discrete-time model. *Proc. Innovative Res. Industrial Dialogue'16.* 129-130.

Samad, M.F.A. & Nasir, A.R.M. (2017b). Parameter magnitude-based information criterion in identification of discrete-time dynamic system. *J. Mech. Eng.,* **4**: 119-128.

Samad, M.F.A. & Nasir, A.R.M. (2017c). Discrete-time system identification based on novel information criterion using genetic algorithm. *J. Fundam. Appl. Sci.*, **9**: 584-599.

Samad, M.F.A. & Nasir, A.R.M. (2018a). Performance of parameter-magnitude based information criterion in identification of linear discrete-time model. *Commun. Math. Appl.*, to be published.

Samad, M.F.A. & Nasir, A.R.M. (2018b). Comparison of information criterion on identification of discrete-time dynamic system. *J. Eng. Appl. Sci.* **12**: 5660-5665.

Samad, M.F.A., Jamaluddin, H., Ahmad, R., Yaacob M.S. & Azad, A.K.M. (2013). Effect of penalty function parameter in objective function of system identification. *Int. J. Automot. Mech. Eng.*, **7**: 940-954.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals Stat.*, **6**: 461-464.

Zhang, L.F., Zhu Q. M. & Longden A. (2005). Nonlinear model validation using novel correlation tests. *2005 IEEE Int. C. Syst. Man Cybern.*, **3:** 2879-2884.

DESIGN METHOD FOR DISTRIBUTED ADAPTIVE SYSTEMS PROVIDING DATA SECURITY FOR AUTOMATED PROCESS CONTROL SYSTEMS

Aleksei A. Sychugov* & Dmitrii O. Rudnev

Department of Information Security, Tula State University, Russian Federation

*Email: xru_2003@list.ru

ABSTRACT

Adaptive control allows for the synthesis of control systems that vary depending on certain parameters and external influences. The main purpose of this work is to study the construction method for destroyed adaptive systems that are necessary for data security systems of automated process control systems (APCS). This paper compares two existing design methods for data security systems, distributed and centralized, and suggests using a hybrid approach to overcome the shortcomings of the existing methods. It then describes an experiment analyzing the efficiency of the considered methods. The experiment shows that the hybrid and centralized approaches have similar levels of accuracy, which is higher than that in case of the distributed approach. It should be noted that during the experiment, the centralized approach required much more computational resources than the hybrid and distributed ones. This is due to the fact that the learning sample for the centralized approach contained a complete set of properties (128), while in the hybrid approach only 10 features corresponding to the selected number of basic objects were given as input for the classifier. Thus, it can be concluded that the use of the hybrid approach allows for high accuracy of determination the state due to its adaptability, and also allows the use of the optimal number of computational resources.

Keywords: *Distributed information system; state analysis; featureless pattern recognition; automated control system.*

1. INTRODUCTION

Over the last years, plenty of attention has been paid to ensuring the data security of automated process control systems (APCS) at industrial enterprises and its crucial facilities, in order to provide safe and reliable operation of the enterprise. In order to solve this problem, APCS must provide complete and up-to-date information about the current state of the industrial enterprise for the analysis. This condition is automatically met in the case of normal operation of the equipment and APCS itself, and in the absence of external destructive factors. However, in practice, the equipment as well as APCS components can fail, or be exposed to intentional or unintentional, technogenic or anthropogenic destructive impact, which can lead to severe consequences depending on the level of complexity and danger of the industrial enterprise (Hoornaert *et al*., 2017). Furthermore, APCS-managed enterprises are complex dynamic systems and as a result, the description of their operation requires a large number of controlled parameters. Therefore, considerable computing resources are needed for their processing, which can also lead to a failure in obtaining reliable information about the current state of the enterprise. Thus, developing methods to estimate APCS state in order to ensure its information security is a task of primary importance (Stouffer *et al*., 2013).

One of the possible solutions to this problem is an intelligent distributed adaptive system that analyzes the current state of the APCS and the industrial object, which is able to detect anomalies in their work and predict the possible outcome of an abnormal situation at an early stage (before the operator gets the information) (Gil, 2014). The intelligence of such systems lies in the ability to analyze measuring information, to recognize an abnormal situation and to take appropriate measures, while ensuring a high level of its own performance by distributing intelligence to separate parts of the system. Adaptability in this case means the ability to dynamically change the settings of the information security system in order to detect and prevent an abnormal situation.

2. MATERIALS & METHODS

Currently, there are two main approaches to analyzing the state of distributed information systems (DIS): centralized and distributed ones (Alghuried, 2017).

The centralized approach consists of collecting information on the operation of all APCS elements and analyzing the entire data in one central element. The advantage of this approach is high adaptability, which is the ability to respond quickly to the changes in system's properties (Fu et al, 2011). The drawbacks of this approach is high requirements for computing resources of the APCS central element, which analyses the need to take additional measures to ensure the security of the transmitted information. Ensuring the security of the transmitted information in small-scale systems is a cheap part of interaction protocol (Alghuried, 2017). However, as the amount of information circulating in the protected system increases, the overhead grows and affects the performance of the system as a whole. The operation scheme of APCS state analysis according to the centralized approach is as shown in Figure 1.

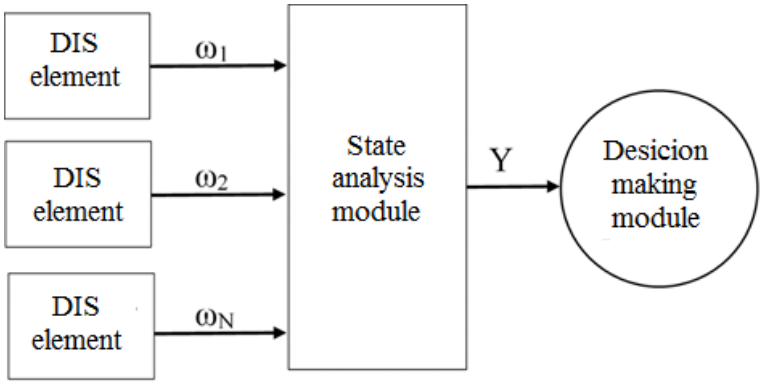


Figure 1: APCS state analysis scheme according to the centralized approach.

According to the distributed approach for the analysis of DIS state, each element analyzes its state on the basis of some previously specified decision rule. The decision rule is calculated only once, based on the state of some subset of the DIS elements. The advantages of this approach include high speed and relatively low requirements for computing resources. The main feature of the distributed approach is that the analysis involves the computational power of all the DIS elements (Verstraete *et al*., 2015; Jagasics & Vajda, 2016). The main disadvantage of this approach is its low adaptability, that is when the properties of the system change, it is necessary to repeat the process of obtaining the decision rule. The question of confirming the reliability of the transmitted information remains unsolved. Since each element analyzes its own state, other elements receive the final information that cannot be verified. The operation scheme of APCS state analysis, according to the distributed approach is shown in Figure 2.

3. METHODOLOGY

In order to overcome the above mentioned shortcomings, a hybrid approach to analyzing DIS state is proposed, which on the one hand, is characterized by adaptability, and on the other hand, it can ensure the confidentiality of the information circulating in the system, describing the separate elements of APCS. Instead of analyzing a set of real number vectors describing the original objects, it is suggested to analyze the values of similarity measures of the considered objects and a set of predefined objects (Figure 3). This approach is based on the compactness hypothesis by Arkadyev & Braverman (1964). It assumes that the objects of the real world with close values of the hidden target property are similar in their observable properties. This approach, which uses the projection space samples based on projection attributes, represented by similarity to a predetermined (basic) object, instead of objects' properties, is called featureless pattern recognition. Such an approach allows using the already available algorithms of machine learning, giving them secondary characteristics as input (Rudnev & Sychugov, 2016). The objects can be of any possible nature. To determine the objective characteristic, it is enough to determine the similarity measure between the objects.

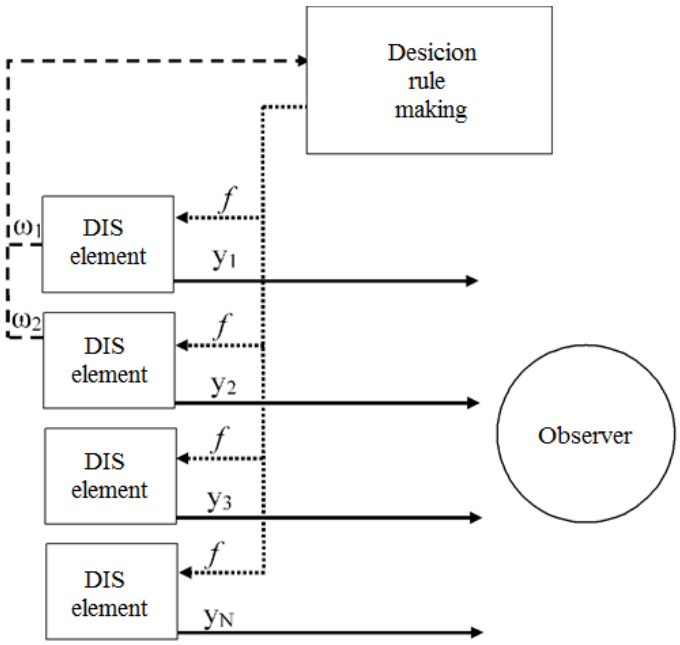


Figure 2: APCS state analysis scheme according to the distributed approach.

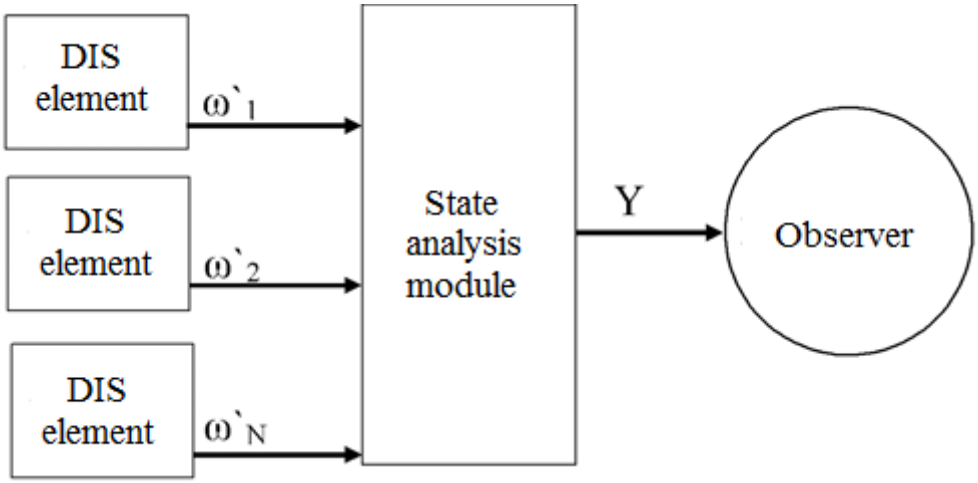


Figure 3: APCS state analysis scheme according to the hybrid approach.

Let us assume that Ω is a set of DIS elements:

$$\Omega = \{\omega_0, \omega_1, \ldots, \omega_N\} \quad (1)$$

where N – the number of elements in the system. *Y* is a hidden property, defining the state of the DIS element:

$$Y = \{y_0, y_1, \ldots, y_N\} \quad (2)$$

there is a function y^*: $\Omega \to Y$, that maps input instances $\omega_i \in \Omega$ to output labels $y_i \in Y$. Then the problem of designing the DIS state analysis system can be formulated as follows: it is necessary to get an algorithm a : $\Omega \to Y$, optimal from the point of view of computational resources and efficiently operating under the condition that the properties of the elements and the components of the set Ω vary with time.

Let us introduce the similarity function ρ for all objects of Ω:

$$\forall \omega_i, \omega_j \in \Omega : \rho_{i,j} = \rho(\omega_i, \omega_j), \rho_{i,j} \in \Re . \quad (3)$$

Let $B \subseteq \Omega$ be a set of basic objects found earlier:

$$B = \{b_1, b_2, \ldots, b_M\} . \quad (4)$$

Let us define the transformation $\Omega \to \Omega'$ as follows:

$$\forall \omega_i \in \Omega : \omega_i' = \{\rho(\omega_i, b_1), \ldots, \rho(\omega_i, b_M)\}, \omega_i' \in \Omega' . \quad (5)$$

Then, in accordance with the secondary properties it is possible to transform: $\Omega' \to Y$.

In order to ensure the confidentiality of the APCS elements' property values, the transformation $\Omega \to \Omega'$ needs to be such that $\Omega' \to Y$ and that it is impossible or computationally difficult to do the inverse transformation: $\Omega' \not\to \Omega$.

Let us consider the conditions under which the transformation $\Omega' \not\to \Omega$ is impossible. In general, in order to restore the properties of the original object according to secondary properties, it is necessary to solve the system of equations:

$$\begin{cases} \rho(\omega_i, b_1) - \omega_1' = 0 \\ \quad \ldots \\ \rho(\omega_i, b_M) - \omega_M' = 0 \end{cases} \quad (6)$$

The solution of such systems depends on the nature of the objects ω and properties of the chosen similarity function. There are two choices: in the first case ω is an object of a linear space, and ρ has the properties of a metric one, and in the second case ω can be of any nature. In practice, both variants can occur, often real-world objects are described using a real number vector – this approach is typical for pattern recognition problems (Redkovskii, 1990). However, there are situations when the real object is difficult or impossible to describe in the form of a number vector. In this case, methods of featureless pattern recognition are used (Mottl, 2002).

Power subsystems of enterprises can serve as an example of industrial facilities, for which distributed APCS is used. Such systems are a crucial part of their infrastructure and any anomaly in their operation can significantly affect the operation of the entire enterprise. Therefore, searching for anomalies in system operation and timely redistribution of system resources is an important task of energy facilities' APCS.

4. RESULTS & DISCUSSION

A numerical experiment was performed to compare the effectiveness of the described approaches. For the experiment we used data on the operation of the APCS of the electric power system (Pan *et al.*, 2015). The data contains voltage, current and relay states at various nodes of the power object. The data was collected during a study on the information security of the electric power facility at the University of Mississippi together with the Oak Ridge National Laboratory (USA). During the research, various scenarios of intrusion into the infrastructure of the energy facility were analyzed and the system was monitored both at the time of the attack and at the time of the system's normal operation.

The data represents 73,101 records, each containing 128 numerical characteristics. To emulate the decentralized operation of the system, the entire initial data set was divided into 15 parts. Each part corresponds to a separate stage of the research, which contains both regular behavior and the attack according to one of the previously specified scenarios. In this experiment, adaptability is understood as the ability to perform correct classification of states related to the scenarios that were unknown earlier.

During the experiment, the centralized, distributed and hybrid approaches to the analysis of the DIS state were compared. The task for all three approaches was to find the anomalies in the energy facility states. In order to minimize the impact of data on the result, the information from various DIS elements was taken separately as a learning sample, thus for each of the approaches, the learning and recognition cycles were performed 15 times, according to the number of different attack scenarios.

For the distributed approach, at the initial moment of time, only the marked data were available, on the basis of which the decision rule was obtained, and then the whole data set was classified. For centralized and hybrid approaches also only the marked data were available at the initial moment, but the entire set of data was available for classification, which allowed the use of partial learning algorithms. As a method of pattern recognition, we used random decision trees. This method was chosen due to the fact that the authors of (Beaver *et al.,* 2014) studied the effectiveness of pattern recognition methods on the same initial data, and the random forest method (Rokach & Maimon, 2008) showed the best results. Another advantage of random forest is the ability to work effectively with a large number of features.

In the experiment with the hybrid method, unlike the experiment with the centralized one, a preliminary stage was taken to obtain secondary properties of the object. According to the described algorithm, at the first stage of the experiment with the hybrid approach, we obtained a basis by means of *k*-means clustering, with cluster centers taken as the basic elements. The number of basic elements was chosen empirically and reached 10 elements. Then, we emulated the collection of system elements' states in order to identify elements that at the moment were in anomaly state, that is, the elements under attack. Euclidean distance was chosen as the metric of state comparison. Previously, on the marked sample, the vector of properties' significance was obtained, and all the features were scaled within the range [-1, 1].

For each pattern recognition algorithm, the selection of the optimal coefficients was carried out using the method of grid search. During the search for optimal parameters of pattern recognition algorithms, cross-validation by *k*-blocks was applied. The *F*-measure metric was used to estimate the accuracy. It is a widely used score function to analyze the results of the binary classification algorithms. It considers both the precision *p* and recall *r* of the test to compute the score. The results of the experiment are shown in Figure 4, where the box plot charts represent the distribution of *F*-measure metrics for the approaches.

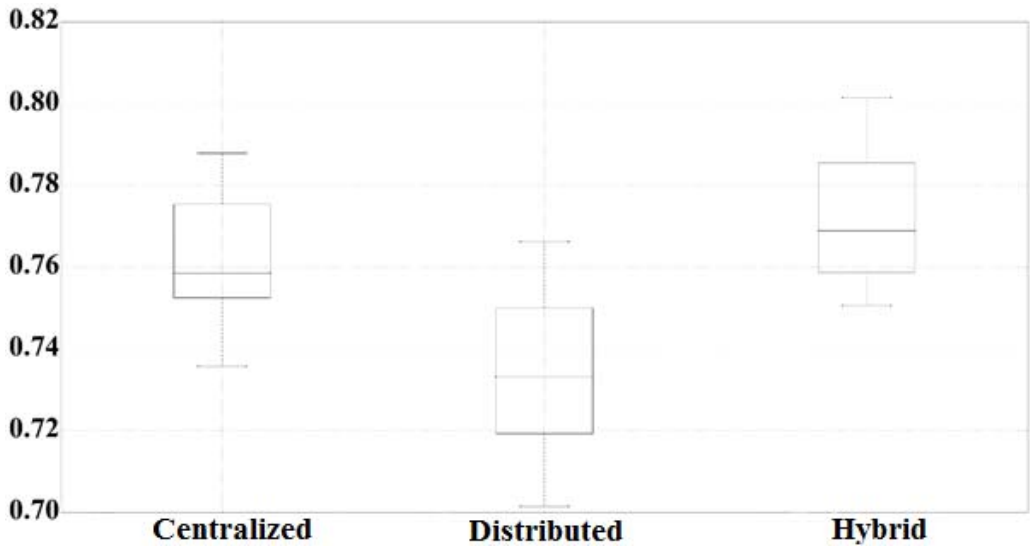


Figure 4: *F*-measure of the resulting accuracy of the energy facility's state analysis using different approaches.

The experiment showed that the hybrid and centralized approaches have similar levels of accuracy, which is higher than that in case of the distributed approach. During the experiment, the centralized approach required much more computational resources than the hybrid and distributed ones. This is due to the fact that the learning sample for the centralized approach contained a complete set of properties (128), while in the hybrid approach, only 10 features corresponding to the selected number of basic objects were given as input for the classifier. Thus, it can be concluded that the use of the hybrid approach allows for high accuracy of determination the state due to its adaptability, and also allows the use of optimal number of computational resources.

5. CONCLUSION

In this paper, the experiment to compare the accuracy and performance of three approaches for analysis of the state of DIS (hybrid, centralized and distributed) is discussed. It was found that the hybrid and centralized approaches have similar levels of accuracy, which is higher than that in case of the distributed approach. The hybrid approach required much less computational resources than the centralised and distributed ones.

ACKNOWLEDGEMENT

The results of this research project are published with the financial support of Tula State University within the framework of the scientific project № 2017-58PUBL.

REFERENCES

Alghuried, A.A. (2017) *Model for Anomalies Detection in Internet of Things (IoT) Using Inverse Weight Clustering and Decision Tree*. PhD Thesis, Dublin Institute of Technology, Dublin.

Arkadyev, A.G. & Braverman, E.M. (1964). *Teaching Computers to Recognize Patterns*. Science, Moscow.

Beaver, J., Borges, R., Buckner, M., Morris, T., Adhikari, U. & Pan, S. (2014). Machine learning for power system disturbance and cyber-attack discrimination. *Proc. 7th Int. Symp. Resilient Control Syst.*, 19-21 August 2014, Denver, Colorado.

Fu, R., Zheng, K., Zhang, D., & Yang, Y. (2011). An Intrusion Detection Scheme Based on Anomaly Mining in Internet of Things. In IEEE International Conference on Wireless, Mobile & Multimedia Networks (ICWMMN 2011), Beijing, 27 – 30 Nov. 2011, pp. 315-320. DOI: 10.1049/cp.2011.1014.

Gil, D.L.I. (2014). *A Formal Approach for Designing Distributed Self-Adaptive Systems*. Linnaeus University Press, Växjö.

Hoornaert, S., Ballings, M., Malthouse, E.C. & Van den Poel, D. (2017). Identifying new product ideas: waiting for the wisdom of the crowd or screening ideas in real time. *J. Prod. Innovat. Manag.*, **34(5)**: 580-597.

Mottl, V., Seredin, O., Dvoenko, S., Kulikowski, C. & Muchnik I. (2002). Featureless pattern recognition in an imaginary Hilbert space. *16th Int. Conf. Pattern Recog.*, pp. 205- 208.

Stouffer, K., Falco, J. & Scarfone, K. (2013). *Guide to Industrial Control Systems (ICS) Security*. NIST Special Publication 800-82. Revision 1, National Institute of Standards and Technology (NIST), US.

Jagasics, S. & Vajda, I. (2016). Cogging torque reduction by magnet pole pairing technique. *Acta Polytech. Hung.*, **13**: 107-120.

Pan, S., Morris T., & Adhikari, U. (2015). Developing a hybrid intrusion detection system using data mining for power systems. *IEEE T. Smart Grid.* **6**: 3104-3113.

Rokach, L. & Maimon, O. (2008). *Data mining with decision trees: theory and applications.* World Scientific, Hackensack, New Jersey.

Redkovskii, N.N (1990). A numerical method for solving systems of nonlinear equations. *Cybernetics Syst. Anal.*, 26: 384-393.

Rudnev, D.O. & Sychugov, A.A. (2016). Analysis of the security of distributed information systems based on featureless pattern recognition. *Proc. Tula State University Tech Sci.,* **11:** 152-159.

Verstraete, H.R., Wahls, S., Kalkman, J. & Verhaegen, M. (2015). Numerical evaluation of advanced optimization algorithms for wavefront aberration correction in OCT. *Adaptive Optics: Anal, Method. Syst. 2015,* 7-11 June 2015, Arlington, Virginia.

LOW CONTRAST IMAGE ENHANCEMENT USING RENYI ENTROPY

Vijayalakshmi Dhurairajan[1*], Teku Sandhya Kumari[2] & Chekka Anitha Bhavani[3]

Department of Electronics & Communication Engineering, Vignan's Institute of Engineering for Women, India

[*]Email: vijaya1183@gmail.com

ABSTRACT

In this paper, an algorithm is proposed to enhance low contrast images using Renyi entropy. Renyi entropy concentrates on prominent amplitudes within a distribution, which allows the proposed enhancement algorithm to work in the edges efficiently without any artefacts. It is calculated from the 2D histogram which provides the information regarding the occurrence of each intensity value in the local grids. A mapping function is obtained from Renyi entropy that maps the input intensity to the corresponding output intensity in order to enhance the image. Further enhancement is achieved using discrete cosine transform (DCT) with the proposed algorithm. The experimental results show that the proposed algorithm provides better performance in terms of subjective and objective measures along with brightness preservation in the enhanced image. The method proposed for contrast enhancement can be effectively used in defence applications for the detection and identification of targets.

Keywords: *2D histogram; Renyi entropy; contrast enhancement; brightness preservation; discrete cosine transform (DCT).*

1. INTRODUCTION

Contrast enhancement is one of the most important issues of image processing and analysis. It is a fundamental step in image processing applications. Image enhancement is used to transform an image on the basis of the characteristics of human visual system. The techniques for image enhancement are designed to improve the quality of the images perceived by human. The objective of image enhancement is to improve the interpretability of the information present in the images for human viewers. Enhancement results in better quality image, which can be done by either suppressing noise or increasing the contrast of the image (Gonzalez, 2001). If the contrast of an image is highly concentrated in a specific range, i.e., an image is very dark, the information may be lost in those areas, which are excessively and uniformly concentrated. The solution to such images is to optimize the contrast of an image in order to represent all the information in the input image (Kim, 1998; Taric, 2009). Image enhancement techniques are used in applications such as real-time imagery in operational environments in unmanned aerial vehicles (UAVs), naval ships and land based units (Piet, 2013). Contrast enhancement prior to fusion of low-light visible and infrared images has been proposed for applications in navigation and surveillance in defence (Sandhya, 2016).

Contrast enhancement changes the image value distribution to cover a wide range. If the image values are concentrated near a narrow range, then it is referred to as low contrast. If the contrast of the image increases, lighter areas becomes lighter and dark areas becomes darker. If the contrast of the image decreases, all the pixels will be in mid-shade of grey, which makes the image to fade. Adjusting the pixels of an image in order to improve the quality of an image for natural look is to perform image contrast enhancement. The contrast of an image can be revealed by its histogram (Jobson, 1997).

Digital image enhancement techniques may be broadly divided into two principal categories; transform domain and image domain methods. Approaches based on transform-domain methods consist of computing a 2D transform (e.g., Fourier or Hadamard transforms) of the image to be enhanced, altering the transform, and computing the inverse to yield an image which has been

enhanced in some manner. Image-domain methods, on the other hand, operate directly on the image in question by means of algorithms that are usually based on grey-level content (Gonzalez, 2001).

Most of the algorithms are employed in spatial domain instead of frequency domain owing to its complexity. One of the popular spatial domain algorithm is global histogram equalization (HE). It is widely used due to ease of implementation and better performance. A cumulative distribution function (CDF) is derived from the input image. The CDF of the input image is mapped to the uniform distribution in order to occupy the complete range of grey levels. A key advantage of the method is that it is a straightforward technique and an invertible operator. However, it causes over-enhancement of the low contrast images, which results in a noisy appearance in the enhanced images. Histogram equalization increases the contrast of background noise and it significantly alters the brightness of an image. Thus, it is not applicable for all kinds of images (Kim, 1997).

Local histogram equalization (LHE) is also one of the techniques used for image enhancement. Based on LHE, many algorithms have been developed to overcome the drawbacks of HE. LHE uses a window that slides the entire image. Each pixel is locally histogram equalized within the window. However, HE-based algorithms produce checker board effects. On account of the complexity and selection of window size in LHE, more number of algorithms has been developed to improve the performance of HE (Arici, 2009). Some of the algorithms based on HE is bi-histogram equalization (BHE) and minimum mean brightness error bi-histogram equalization (MMBEBHE) (Chen, 2003).

Contrast enhancement also uses optimization algorithms. Convex optimization is used in flattest histogram specification with accurate brightness preservation (FHSABP). It is used to produce flattest histogram subject to mean brightness constraint. FHSABP may result in low contrast as it is designed to preserve the average brightness (Wang, 2003).

Gamma correction techniques are being developed to compensate for the input-output characteristic of cathode ray tube displays. The adaptive gamma correction with weighting distribution (AGCWD) method was proposed for the enhancement of dimmed images that result in an image with less brightness due to low exposure (Huang, 2016). Recently, a 2D histogram algorithm was developed to enhance low contrast images that use the contextual information of pixels to enhance the image by incorporating the grey levels of each pixel and its neighbouring pixels. Enhancement is achieved by equal distribution of the grey level differences. In general, 2D histogram based methods resulting visually pleasing outputs when compared to 1D histogram based methods (Celik, 2012).

A hybrid contrast enhancement algorithm known as spatial entropy based contrast enhancement (SEDCT) was developed by Celik (2014) in order to achieve unified perception. In this algorithm, global enhancement is achieved by calculating the weights from spatial entropy of grey levels. Spatial entropy is normalized before calculating the weights. Furthermore, local contrast enhancement is performed in transform-domain with the help of discrete cosine transform (DCT). The drawback of this method is that it does not preserve the average brightness of the image. The residual SEDCT algorithm proposed by Celik (2016) performs brightness preservation but fails to utilise the dynamic range of the image resulting in contrast loss. The spatial mutual information and page rank based contrast enhancement (SMIRANK) algorithm proposed makes use of the entire dynamic range of the image to improve the quality of the image in terms of contrast. However, algorithm runs slowly for large sized input images (Celik, 2016).

In order to overcome the above-mentioned problems, a new algorithm is proposed in this paper, where the low contrast of an image is enhanced using Renyi entropy. Unlike the above-mentioned algorithms, the proposed method not only preserves the average brightness, but also increases the information content of the image by increasing the contrast. Renyi entropy is mainly preferred because it is easy to implement with less computational complexity.

2. PROPOSED METHOD

2.1 Renyi Entropy

Renyi entropy is a parameter generalization of the Shannon entropy. Renyi entropy is used in many fields. It mainly focuses on the prominent amplitude with in a distribution without requiring any decomposition. For a given set of finite discrete probabilities $P= \{p_1, p_2.....,p_N\}$ and a real number $\alpha \geq 0, \alpha \neq 1$,Renyi entropy H_α is defined as:

$$H_\alpha = \frac{1}{1-\alpha} log_2 \left(\sum_{n=1}^{N} p_n^\alpha \right) \tag{1}$$

Case 1:
If α=0, H_α converges to Shannon entropy.

Case2:
H_α is a non-increasing function of α, i.e. $\alpha_1 < \alpha_2 \Rightarrow H_{\alpha 1} \geq H_{\alpha 2}$

Case 3:
For every value of α, H_α is maximum when P is uniformly distributed, and it is minimum and equal to zero when P has a single non-zero value.

The main advantage of Renyi entropy is its dependence on the order α, which provides a different concentration for each value of α. This entropy measure is different from Shannon entropy by the modification applied to the histogram before addition, where the power is raised to the factor α (Obin & Liuni, 2012).

2.2 Enhancement Algorithm

The algorithm for the proposed method is given in the following steps:

Step 1: Consider an image I with the size MxN.

Step 2: Find the unique intensities in the input image I and sort them in ascending order. Suppose the image has l unique intensities, then $I_1<I_2<I_3<\ldots\ldots\ldots I_l$.

Step 3: Calculate the aspect ratio r:

$$r = \frac{M}{N} \tag{2}$$

M and N represent the number of rows and columns respectively.

Step 4: Calculate the number of sub-images that can be formed from the input image:

$$x = roundoff\left[(lr)^{\frac{1}{2}} \right] Y = roundoff\left[\left(\frac{l}{r} \right)^{\frac{1}{2}} \right] \tag{3}$$

The locations of the sub images are given by $\left[(x-1)\frac{M}{X}, x\frac{M}{X} \right] \times \left[(y-1)\frac{N}{Y}, y\frac{N}{Y} \right]$

Step 5: Find the 2D histogram for the intensity values, in which the 2D histogram of the i^{th} intensity value is defined as $h_i=h_i(x, y)$, where x varies from 1 to X and y varies from 1 to Y. The total number of sub-images in the images I is XY.

Step 6: Calculate the Renyi entropy from the 2D histogram:

$$H_i = \frac{1}{1-\alpha}\sum_{x=1}^{X}\sum_{y=1}^{Y} log_2\left(\frac{1}{h_i(x,y)}\right)^{\alpha} \tag{4}$$

where the value of α is considered as 2 in our proposed work. This is as experimentally, after several simulations with different values of α, the results showed that the mean brightness value is better for α=2. Hence, experimentally, the value α= 2 is considered in this proposed work.

Step 7: Calculate discrete density function f_i:

$$f_i = \frac{H_i}{\sum_{i=1}^{l} H_i \; for \; i \neq l} \tag{5}$$

Step 8: Normalize the density function to obtainthenormalised density function f_{ri}:

$$f_{ri} = \frac{f_i}{\sum_{i=1}^{l} f_i} \tag{6}$$

Step 9: Find CDF F_i by summing up the normalized density function:

$$F_i = \sum_{i_1=1}^{i} f_{r_{i1}} \tag{7}$$

Step 10: Map the intensities J_i with the mapping function:

$$J_i=F_i \text{x} 255 \tag{8}$$

Step 11: In order to enhance the image further, apply DCT to the image J, where J is the enhanced image obtained from the mapping function; k varies from 1 to M and l varies from 1 to N:

$$C(k, l) = \text{DCT}(J) \tag{9}$$

Step 12: Alter the coefficients by multiplying the weighting function $w(k,l)$:

$$C_m(k,l) = C(k,l)w(k,l) \tag{10}$$

$$\text{where } w(k,l) = \left(1+\frac{\beta-1}{M-1}k\right)\left(1+\frac{\beta-1}{N-1}l\right) \tag{11}$$

where $\beta \geq 1$. In order to get higher enhancement, the value of β should be high. For the automatic selection of β, it is estimated from the entropy.

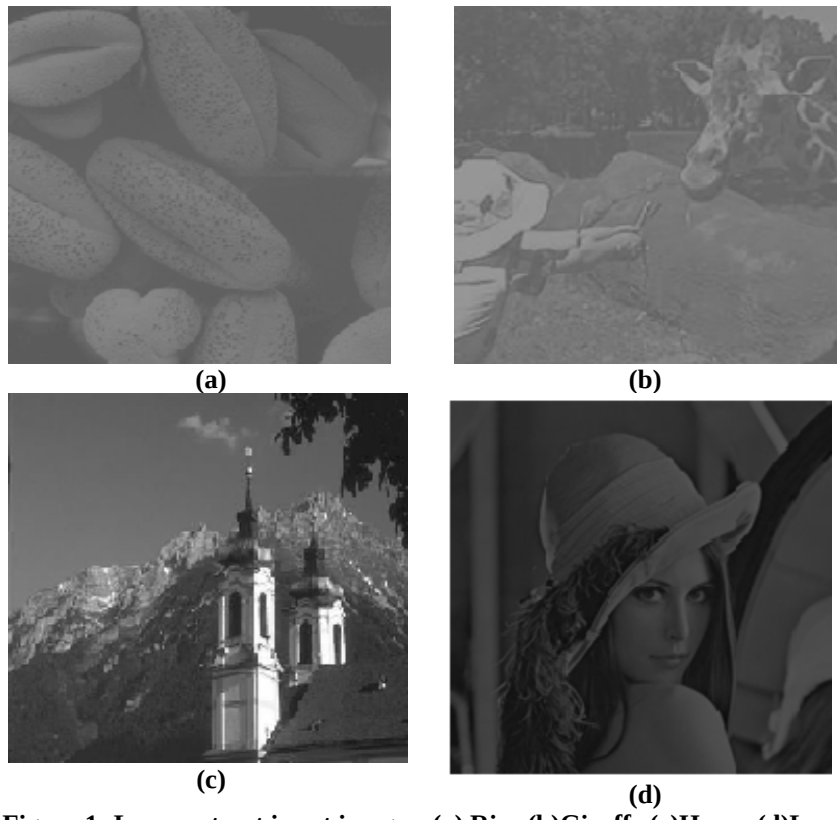
$$\beta = \left(\sum_{i=1}^{l} f_i \log_2(f_i) \right)^{0.5} \tag{12}$$

Step 13: Apply inverse DCT (IDCT) on the modified image, which will give the overall enhanced image:

$$J = IDCT\,[C_m(k,l)] \tag{13}$$

3. RESULTS AND DISCUSSION

The experimental results of the proposed algorithm along with the existing methods are enumerated. Figure1 shows the standard low contrast input images of Food, Giraffe, House and Lena considered for the simulation.



(a) **(b)** **(c)** **(d)**

Figure 1: Low contrast input images: (a) Rice (b)Giraffe (c)House (d)Lena.

Some of the existing techniques such as gamma correction, HE and SEDCT are considered for comparison to the proposed technique. The simulated results of test images considered are shown in Figures 2 - 5. From the simulated results, it is observed that gamma correction has limited precision and hence, multiple input values may map to the same output values. This leads to poor utilization of the entire dynamic range of the input image. HE utilizes the available greyscale of the image efficiently but it tends to over-enhance the image contrast, which results in darkening effect in the enhanced image. When compared to gamma correction and HE, SEDCT provides better results in terms of utilization of the greyscale and reduction of the darkening effect as compared to the general techniques. However, it is observed that SEDCT produces some artefacts in the enhanced image. The simulation results of the proposed method overcome these issues and it is observed that it effectively utilizes the entire greyscale without any over-enhancement and reduces the artefacts.

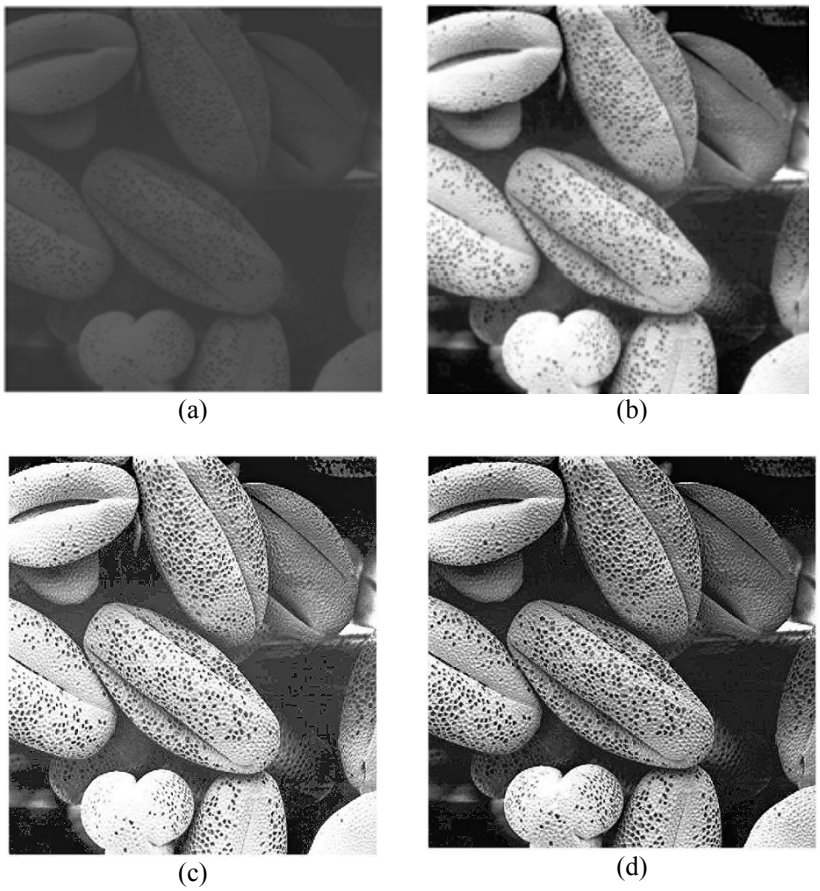


Figure 2: Enhanced by :(a) Gamma correction(b) HE(c) SEDCT(d) Proposed method.

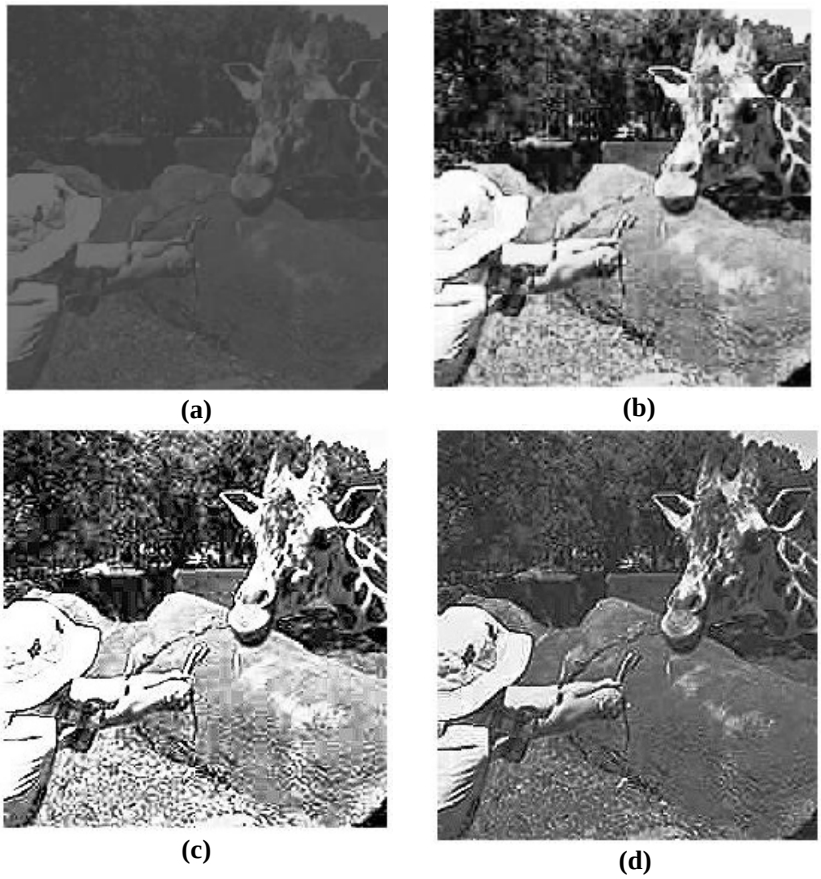
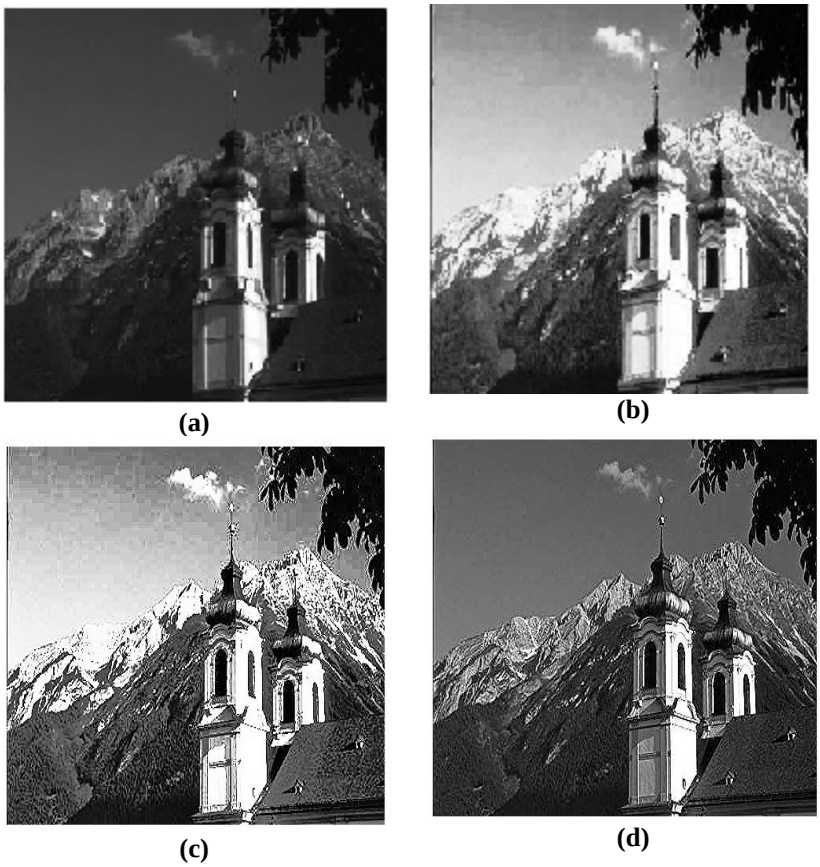


Figure 3: Enhanced by: (a) Gamma correction (b) HE(c) SEDCT(d) Proposed method.



(a) **(b)** **(c)** **(d)**

Figure 4: Enhanced by: (a) Gamma correction (b) HE(c) SEDCT(d) Proposed method.

(a) **(b)** **(c)** **(d)**

Figure 5: Enhanced by: (a) Gamma correction (b) HE(c) SEDCT(d) Proposed method.

It is evident from the simulated images that the proposed method gives better visually pleasing images. However, it is difficult to analyse the image quality with subjective measures, so the image quality needs to be analysed with objective measures. Some of the measures considered for objective analysis are described as:

Entropy: Image entropy specifies the uncertainty in the image values. It measures the average amount of information required to encode the image values. Entropy is a measure of histogram dispersion. A higher value of E indicates that the image has richer details. Entropy is measured by using a basic expression, which is given by:

$$E(I) = -\sum_{l=1}^{P} p(i_l) log_2\, p(i_l) \tag{14}$$

where $p(i_1)$ is the probability of the intensity i_l.

Normalized entropy: Using the entropy values of the original and enhanced images, the normalized entropy is defined as:

$$E_N = \frac{1}{1 + \dfrac{log_2(256) - E(J)}{log_2(256) - E(I)}} \tag{15}$$

$E(I)$ and $E(J)$ represent the original I and enhanced J image entropies respectively. The higher the value of normalized entropy, the better the enhancement in terms of utilization of dynamic range, which provides better image details.

Mean: Mean value gives the contribution of individual pixel intensity for the entire image. The mean value indicates the average brightness of an image.

$$M(I) = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} I(m,n) \tag{16}$$

Absolute Mean Brightness Error (AMBE): It is a measure of preservation of the original image brightness. It is defined as:

$$AMBE = \left| M(I) - M(J) \right| \tag{17}$$

Where $M(I)$ and $M(J)$ represent the mean values of the original I and enhanced J images respectively. The preservation of the original image is linked to the lower value of AMBE.

The described measures are computed for the test images considered in the work and are tabulated in Table 1. It is evident from the quantitative measures that the proposed method produces high entropy value, which indicates improvement of information content of the image. Low value of AMBE of proposed method indicates preservation of the mean brightness of the image without increasing the complexity of the algorithm when compared to the existing methods. The proposed method shows high normalized entropy, which indicates better enhancement results with good image details. The mean value obtained for the proposed technique shows that the image is not over-brightened as compared to the other existing techniques. On the whole, the experimental results show that the proposed method provides better results in terms of both quantitative and qualitative measurements.

Table 1: Quantitative measures for the enhanced images.

Image	Metric/ Method	Gamma correction	Histogram Equalization	SEDCT	Proposed Method
Rice	Entropy	5.0628	5.033	7.328	**7.5904**
	DEN	0.4962	0.4936	0.8099	**0.8760**
	AMBE	37.55	18.35	31.1039	**2.04**
	Mean	71.55	127.42	140.172	**107.03**
Giraffe	Entropy	5.29	5.13	7.0692	**7.3**
	DEN	0.49	0.47	0.7379	**0.7892**
	AMBE	37.57	16.75	23.6938	**13.361**
	Mean	79.19	127.52	140.47	**103.415**
House	Entropy	6.6014	5.883	7.064	**7.19**
	DEN	0.44	0.342	0.5394	**0.55**
	AMBE	32.73	35.90	38.013	**8.255**
	Mean	58.72	127.36	129.47	**83.20**
Lena	Entropy	5.13	5.66	7.353	**7.449**
	DEN	0.428	0.4784	0.7676	**0.7953**
	AMBE	69.32	79.28	83	**59.162**
	Mean	61.68	127.29	131.5124	**107.210**

4. CONCLUSION

In this paper, an enhancement algorithm using Renyi entropy, which employs 2D histogram of an input image to enhance low contrast images is proposed. The 2D histogram presents distribution of intensity values along with the spatial locations that helps to improve the performance of the enhancement algorithm. Renyi entropy is used for its better concentration towards prominent amplitudes such as the edges of an image. Performance comparisons with the current enhancement algorithms show that proposed method achieves satisfactory image enhancement. It also produces a better quality visual image with less complexity when compared to the other enhancement algorithms. The proposed method for improving the image quality in terms of image enhancement can be used for real time target identification in surveillance for defence applications.

REFERENCES

Arici,T., Dikbas,S. & Altunbasak,Y. (2009). A histogram modification framework and its application for image contrast enhancement. *IEEE T. Image Proc.*, **18**:1921-1935.

Celik, T. (2012). Two dimensional histogram equalization and contrast enhancement. *IEEE T. Pattern Recog.*, **45**:3810-3824.

Celik, T. (2014). Spatial entropy based global and local image contrast ehancement. *IEEE T. Image Proc.*, **23**:5298-5308.

Celik, T. & Li, H.C. (2016). Residual spatial entropy based contrast enhancement and gradient based contrast measures. *J. Modern Optics*, **63**:1600-1617.

Celik,T. (2016). Spatial mutual information and Page Rank-based contrast enhancement and quality-aware relative contrast measure. *IEEE T. Image Proc.*, **25**:4719-4728.

Soong-Der Chen & Ramli, A. (2003). Minimum mean brightness error bi-histogram equalization in contrast enhancement. *IEEE T. Consumer Electr.*, **49**: 1310-1319.

Gonzalez, R. & Woods, R. (2001). *Digital Image Processing*. Prentice-Hall, Inc., New Jersey.

Huang, L., Cao, G. & Yu, L. (2016). Efficient contrast enhancement with truncated adaptive gamma correction. *IEEE I. Cong.Image Signal Process., Biomedical Eng. Informatics*, pp. 189-194.

Jobson, D., Rahman, Z. & Woodell, G. (1997). A multiscale retinax for bridging the gap between colour observations of scenes. *IEEE T. Image Proc.*, **6**:965-976.

Kim, T.K., Palik, J.K. & Kang, B.S. (1998). Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering. *IEEE Consumer Electr.*, **44**: 82-87.

Kim, Y.T. (1997). Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE T. Consumer Electr.*, **43**: 1-8.

Obin, N. & Liuni, M. (2012). On the generalization of Shannon Entropy for speech recognition. *IEEE Spoken Lang. Tech. Workshop*, Miami.

Piet, B.W.S., Kemp, R.A.W. & Schuttle, K. (2013). Image enhancement technology research for army applications. *SPIE Defense, Secur. Sens. 2013*, Baltimore, Maryland.

Sandhya, K.T., Koteswara Rao, S. & Santhi Prabha, I. (2016). Contrast enhanced low-light visible and infrared image fusion. *Defence Sci. J.*, **66**:266-271.

Wang, C., Peng, J. & Ye, Z. (2008). Flattest histogram specification with accurate brightness preservation. *IET Image Proc.*, **2**: 249-262.

DETERMINATION OF ARTIFICIAL RECHARGE LOCATIONS USING FUZZY ANALYTIC HIERARCHY PROCESS (AHP)

Marzieh Mokarram[1*] & Dinesh Sathyamoorthy[2]

[1]Department of Range and Watershed Management, College of Agriculture and Natural Resources of Darab, Shiraz University, Iran
[2]Science & Technology Research Institute for Defence (STRIDE), Ministry of Defence, Malaysia

[*]Email: m.mokarram@shirazu.ac.ir

ABSTRACT

In arid and semi-arid regions, distributions of irregular spatial and temporal rainfall cause vegetation to be damaged by droughts, which leads to political problems between different countries. In these areas, one of the ways to prevent land degradation is the use of flooding for artificial recharge of groundwater aquifers. Determination of suitable areas for artificial recharge is very important, which needs to be done with sufficient accuracy. In this study, suitable artificial recharge locations were investigated for north and northwest of Fars province, Iran. For this aim, different map layers, including precipitation, lithology, distance of fault, land use, slope, elevation and drainage density, were prepared using ArcGIS. For preparing fuzzy maps, the importance of each factor in artificial recharge was defined using membership functions. The artificial recharge map was prepared using analytic hierarchy process (AHP). The results show that about 56% of the study area is suitable for artificial recharge.

Keywords: Artificial recharge; analytic hierarchy process (AHP); fuzzy quantifiers; geographical information systems (GIS).

1. INTRODUCTION

Water is one of the most important materials in the world. Water's essential nature makes it a strategic natural resource globally, and in its absence, an important element of political conflicts in many areas. Arid and semi-arid areas have problems of water scarcity, with water withdrawal from ground water aquifers dramatically leading to drop in landing level. One of the methods for controlling and preventing the lowering of groundwater level is artificial recharge, which is the planned human activity of augmenting the amount of groundwater available through works designed to increase the natural replenishment or percolation of surface waters into the groundwater aquifers, resulting in a corresponding increase in the amount of groundwater available for abstraction. Artificial recharge also has applications in wastewater disposal, waste treatment, secondary oil recovery, prevention of land subsidence, storage of freshwater within saline aquifers, crop development, and stream flow augmentation (Asano, 1985; Oaksford, 1985; Contributors *et al.,* 2013).

Several studies have been carried out for the determination of areas that are most suitable for artificial recharge. Krishnamurthy & Srinivas, 1995 investigated the role of geological and geomorphological factors in groundwater exploration. Krishnamurthy *et al*., 1996 and Saraf & Choudhury, 1998 employed geographical information systems (GIS) and remote sensing to determine suitable locations for artificial recharge. Ghayoumian *et al*., 2002 demonstrated some examples of artificial recharge of aquifers by flood spreading in Iran. Other methods proposed include fuzzy methods (Zehtabian *et al*., 2001; Nouri, 2003; Tiwari *et al*., 2017) and artificial neural network (Verma *et al*., 2016).

One of the methods for determining suitable locations for artificial recharge is through a combination analytic hierarchy process (AHP) and fuzzy logic methods. There have been several studies conducted on mapping artificial recharge through the use of fuzzy-AHP (Krishan *et al*., 2014; Verma *et al*., 2016; Moghaddam *et al*., 2017; Tiwari *et al*., 2017). In all these studies, the results show that fuzzy-AHP method is an appropriate method for determining artificial recharge locations. This is as it allows for the overlaying of layers in GIS within the least possible time in order to evaluate for the best model among site localizing models in a given region (Moghaddam *et al*., 2017).

The aim of this study is the prediction of artificial recharge locations using fuzzy AHP in the north and northwest of Fars province, Iran. In Section 2, the fuzzy modeling and AHP are explained. In Section 3, the case study and input data properties are explained, while Section 4 describes the fuzzy-AHP results. Finally, Section 5 provides concluding remarks on the study's effectiveness and potential applications.

2. METHODOLOGY

2.1 Study Area

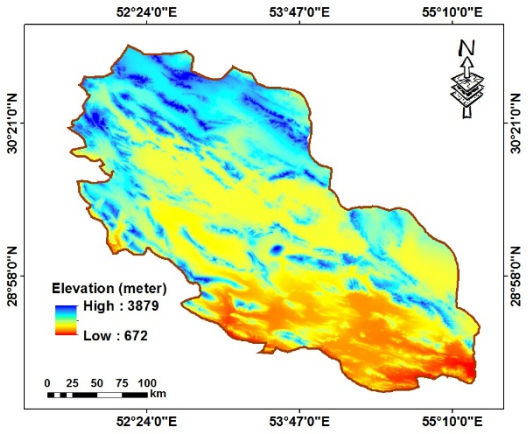
The study was carried out in the north and northeast of Fars Province, Iran. It is an area of 67,262.36 km^2, and is located at latitudes of 27° 54′- 31° 00′ N and longitude of 51° 42′ -55° 24′ E (Figure 1). The altitude of the study area ranges from 672 m to 3,879 m a.s.l.

Figure 1: Location of the study area.

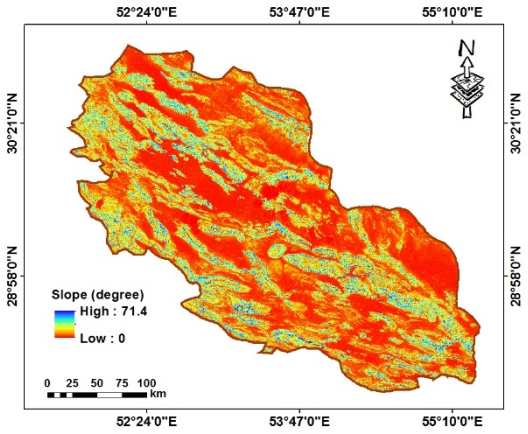
2.2 Data Preparation

Different map layers were used in this study, including precipitation, lithology, distance of fault, land use, slope, elevation and drainage density, which were prepared using ArcGIS. Lithology and fault maps were derived from geological maps with scale of 1:100,000. Slope layers were extracted from Shuttle Radar Topography Mission (SRTM) digital elevation models (DEMs) (Figure 2 (a)). Based on Figure 2 (a), many parts of south and southeast of the study area have elevation more than 1,000 m,

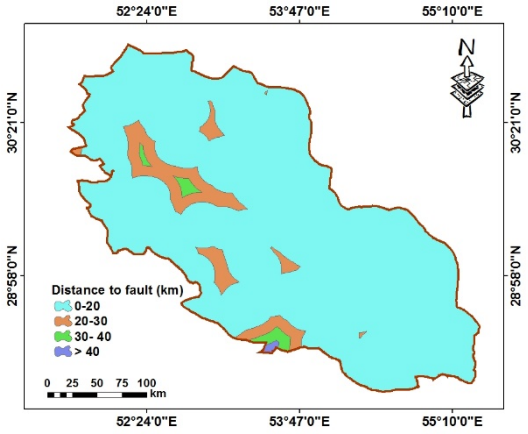
while the slope value is between 0 - 71.4° (Figure 2 (b)). For preparing the raster map for distance of faults, ArcGIS' buffer tools were employed (Figure 2 (c)). Using Landsat 7 ETM+ satellite images and information of the Organization of Agriculture Jahad Fars, a land use map was prepared, consisting of agriculture, garden, bare land, forest, dry farming, salt land, range land, rock and salt land (Figure 2 (d)). For the determination of sensitivity of geological formations, a geological formations map with five classes, produced by the Iranian Geological Organization, was used (Figure 2 (e)). Drainage density and rainfall maps were also prepared (Figures 2 (f) and 2 (g) respectively).



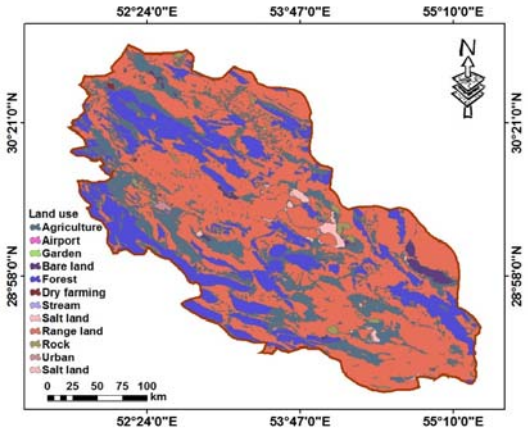
(a)



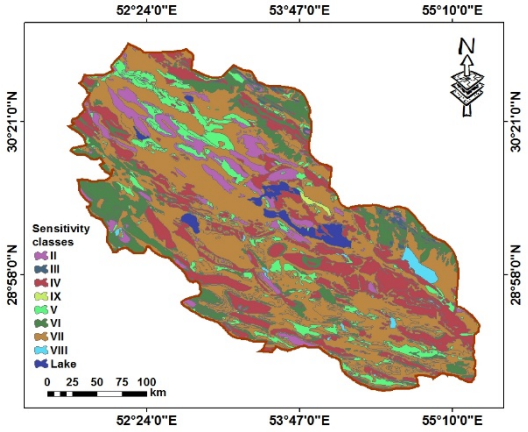
(b)



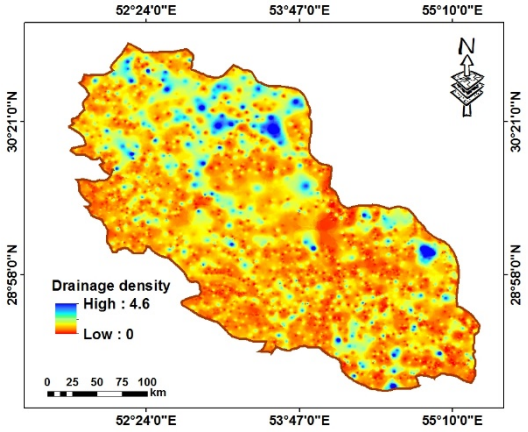
(c)



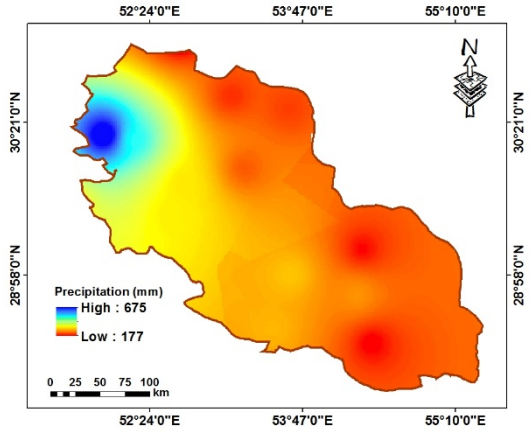
(d)



(e)



(f)



(g)

Figure 2: The parameter maps used for determination of artificial recharge in the study area: (a) Elevation (b) Slope (c) Distance to fault (d) Land use (e) Geology (f) Drainage density (g) Precipitation.

2.3 Methods

2.3.1 Fuzzy Inference

Zadeh (1965) defined a fuzzy set by trapezoidal membership functions from properties of objects. Mathematically, a fuzzy set can be defined as follows (McBratney & Odeh, 1997):

$$A = \{x, \mu_A(x)\} \quad for \quad each \quad x \varepsilon X \tag{1}$$

where μ_A is the trapezoidal membership function (MF) that defines the grade of membership of x in fuzzy set A.

The following MF was used for precipitation, lithology, land use, elevation and drainage density:

$$\mu_A(x) = f(x) = \begin{Bmatrix} 0 & x \leq a \\ x-a/b-a & a \prec x \prec b \\ 1 & x \geq b \end{Bmatrix} \tag{2}$$

where x is the input data and a, b are the limit values.

For slope and distance of fault, the following MF was used:

$$\mu_A(x) = f(x) = \begin{Bmatrix} 1 & x \leq a \\ b-x/b-a & a \prec x \prec b \\ 0 & x \geq b \end{Bmatrix} \tag{3}$$

where x is the input data and a, b are the limit values.

2.3.2 Analytical Hierarchy Process (AHP)

AHP (Saaty, 1980) provides a hierarchical structure by reducing multiple criteria into a pairwise comparison method for individual or group decision-making, and allows the use of quantitative (objective) and qualitative (subjective) information (Malczewski, 1999). AHP allows some small inconsistency in judgment because humans are not always consistent. The ratio scales are derived from the principal Eigen vectors and the consistency index is derived from the principal Eigen value (Table 1). In general, the different processes of the fuzzy-AHP method is as shown in Figure 3.

Table 1: Scales for pairwise comparisons (Saaty & Vargas, 1998).

Intensity of importance	1	3	5	7	9	2, 4, 6, and 8
Definition	Equal importance	Moderate importance of one over another	Essential importance	Demonstrated importance	Absolute importance	Intermediate values between the two adjacent judgments

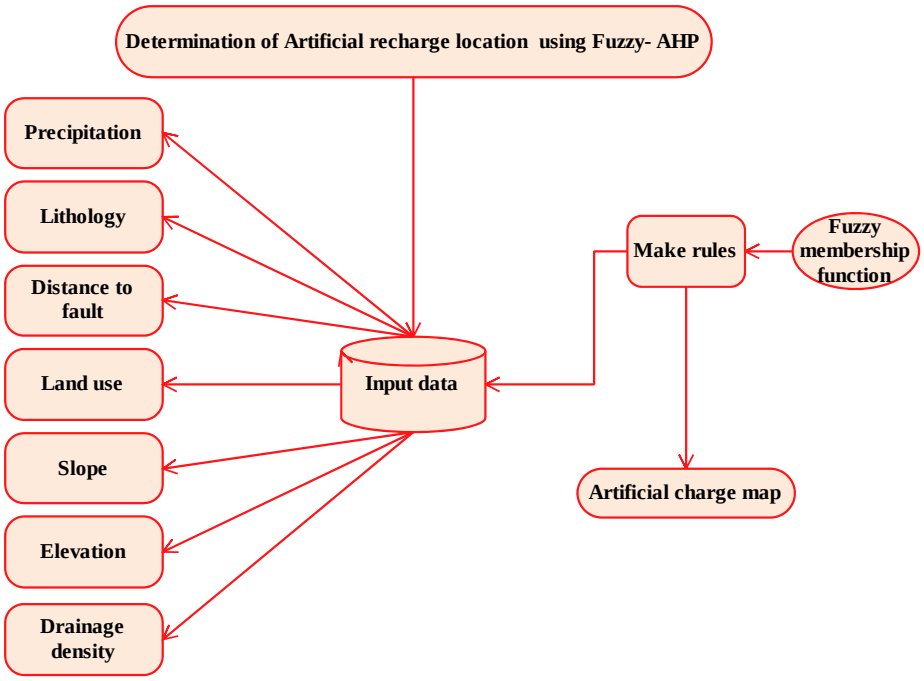
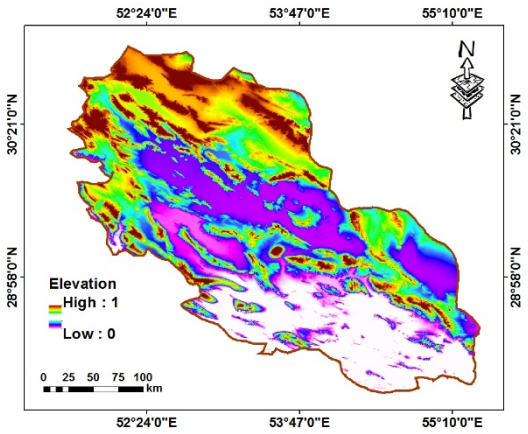


Figure 3: Flowchart for the methodology used in this study to determine the suitable location of artificial recharge map using fuzzy-AHP.

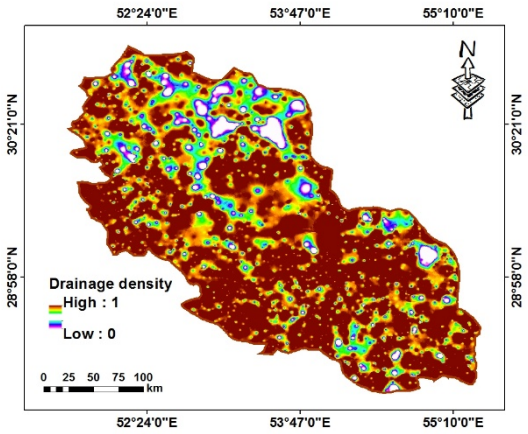
3. RESULTS & DISCUSSION

3.1 Fuzzy Method

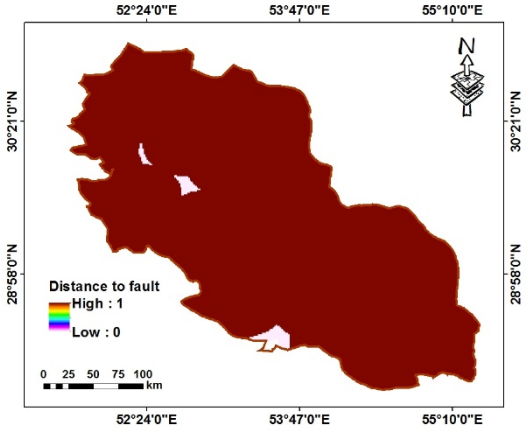
The fuzzy maps for each parameter prepared to determine suitable locations of artificial recharge are shown in Figure 4, where MF is closer to 0 with decreasing suitable location, while MF is closer to 1 with increasing suitable location. According to Figure 4 (a), the north of the study area has high elevation. Drainage density in north of the study area has low value (Figure 4 (b)) while almost all of the study area has value of 1 for distance of fault (Figure 4 (c)). The results of the fuzzy method for land use showed that almost 80% study area is suitable for artificial recharge (Figure 4 (d)). In addition, according to Figure 4 (e), almost 60% of the study area has values more than 0.7 that is suitable for artificial recharge. The results of the precipitation fuzzy map shows that the west and northwest of the study area are more suitable than the other parts (Figure 4 (f)). According to the slope map in Figure 4 (g), the center of the study area is suitable for artificial recharge.


52°24'0"E
53°47'0"E
55°10'0"E
30°21'0"N
28°58'0"N
N
Elevation
High : 1
Low : 0
0 25 50 75 100
km

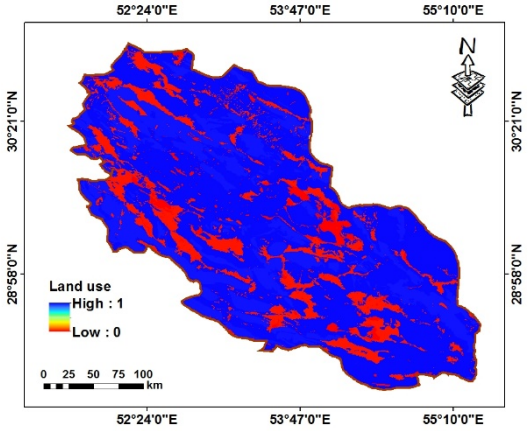
(a)


52°24'0"E
53°47'0"E
55°10'0"E
30°21'0"N
28°58'0"N
N
Drainage density
High : 1
Low : 0
0 25 50 75 100
km

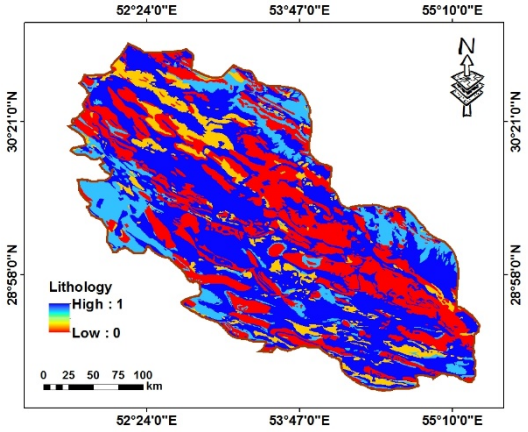
(b)


52°24'0"E
53°47'0"E
55°10'0"E
30°21'0"N
28°58'0"N
N
Distance to fault
High : 1
Low : 0
0 25 50 75 100
km

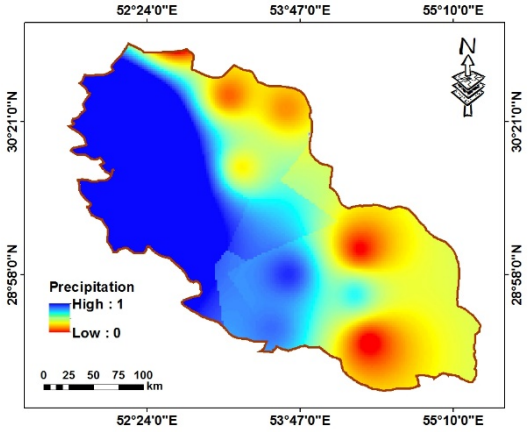
(c)


52°24'0"E
53°47'0"E
55°10'0"E
30°21'0"N
28°58'0"N
N
Land use
High : 1
Low : 0
0 25 50 75 100
km

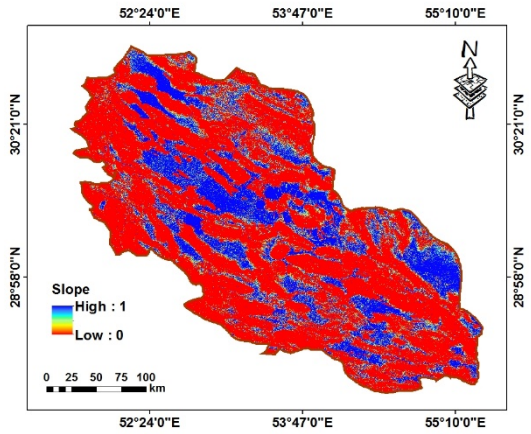
(d)


52°24'0"E
53°47'0"E
55°10'0"E
30°21'0"N
28°58'0"N
N
Lithology
High : 1
Low : 0
0 25 50 75 100
km

(e)


52°24'0"E
53°47'0"E
55°10'0"E
30°21'0"N
28°58'0"N
N
Precipitation
High : 1
Low : 0
0 25 50 75 100
km

(f)



(g)

Figure 4: Fuzzy maps of study area for each parameter: (a) Elevation (b) Drainage density (c) Distance to fault (d) Land use (e) Geology (f) Precipitation (g) Slope.

3.2 AHP

In this study, AHP was utilized to incorporate different types of input data and provide pairwise comparison for two criteria. According to Figure 5, slope and elevation have the highest and lowest weight respectively. The fuzzy-AHP map for artificial recharge was prepared, as shown in Figure 6. According to Figure 7, about 56 % of the study area is suitable for artificial recharge. Thus, the parts in the centre of the study area can be applied for wastewater disposal, storage of freshwater within saline aquifers, crop development, and stream flow augmentation.

Figure 5: Pairwise comparison factor weights for the data layers used.

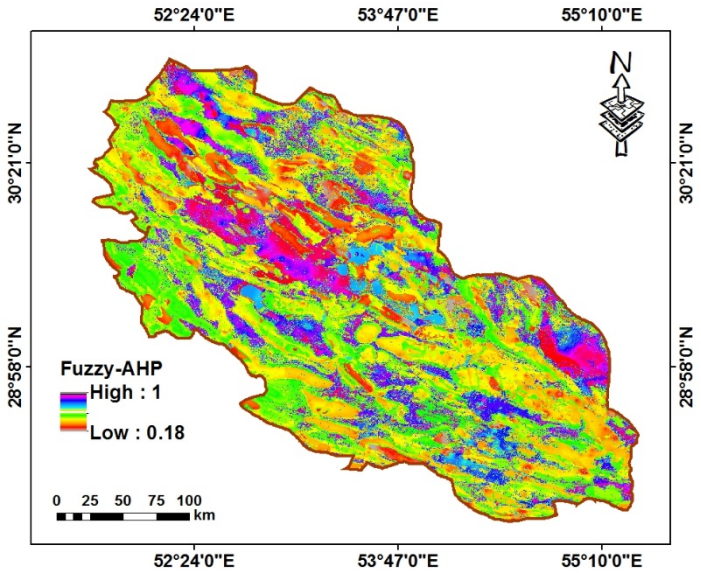


Figure 6: Fuzzy-AHP map for artificial recharge.

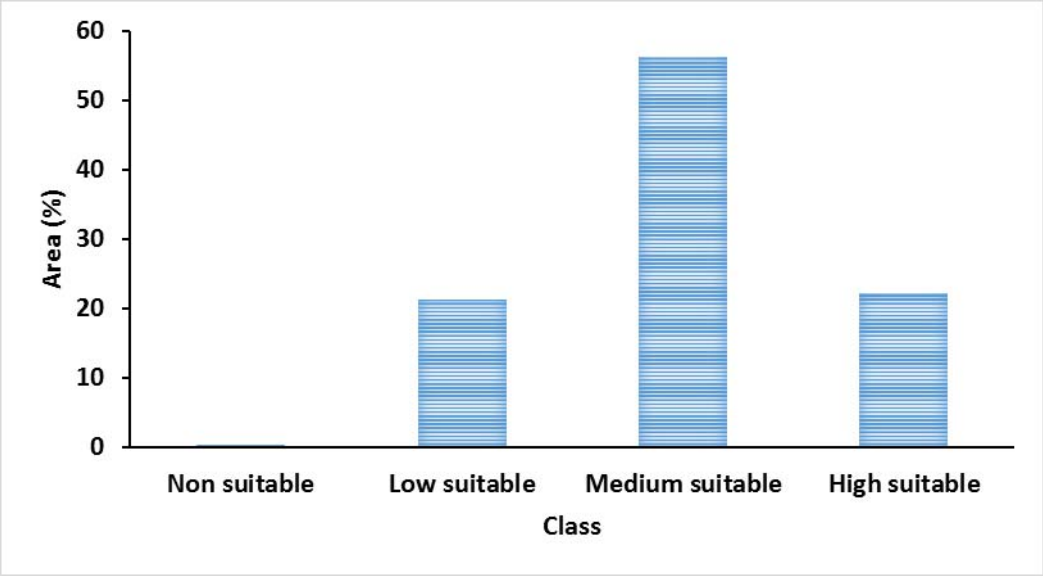


Figure 7: Classes of artificial recharge for the fuzzy-AHP map.

4. CONCLUSION

This study evaluated the spatial distribution of suitable locations of artificial recharge using the fuzzy-AHP approach. In order to identify artificial recharge sites in a semi-arid aquifer in the north and northeast of Fars Province of Iran, elevation, slope, geology, distance to fault, land use, drainage density and precipitation maps were prepared. Integrated assessment of these maps using fuzzy and AHP based on GIS techniques proved to be a suitable method for identifying preferred artificial recharge sites. The results showed that about 56 % of the study area is suitable for artificial recharge.

REFERENCES

Asano, T. (1985). *Artificial Recharge of Groundwater*. Butterworth Publishers, Boston.

Contributors, N., Kulongoski, J. T., Massmann, G., & Newman, B. (2013). *Using Isotopes for Design and Monitoring of Artificial Recharge Systems*. IAEA TECDOC No. 1723, International Atomic Energy Agency (IAEA), Vienna.

Ghayoumian, J., Shoaei, Z., Karimnejad, H.R., Ghermezcheshmeh, B. & Abdi, P. (2002). Some examples of artificial recharge of aquifers by flood spreading in Iran. *In* Van Rooy, J.L. & Jermy, C.A. (Eds.), *Proc. 9th Cong. Int. Assoc. Eng. Geol. Environ.*, Balkema, Rotterdam, pp. 1529–1537.

Gorsevski, P.V. & Jankowski, P. (2010). An optimized solution of multi-criteria evaluation analysis of earthquake events using fuzzy sets and Kalman filter. *Comput. Geosci.,* **36**:1005–1020.

Krishan, G., Rao, M.S., Loyal, R.S., Lohani, A.K., Tuli, N.K., Takshi, K.S., Kumar, C.P., Semwal, P., & Kumar, S. (2014). Groundwater level analyses of Punjab, India: A quantitative approach. *Octa J. Environ. Res.*, **2**:221–226.

Krishnamurthy, J. & Srinivas, G. (1995). Role of geological and geomorphological factors in groundwater exploration: a study using IRS LISS data. *Int. J. Remote Sens.*, **16**: 2595–2618.

Krishnamurthy, J., Venkatesa Kumar, N., Jayaraman, V., & Manivel, M. (1996). An approach to demarcate groundwater potential zones through remote sensing and geographical information system. *Int. J. Remote Sens.*, **17**: 1867–1884.

Malczewski, J. (1999). *GIS and Multicriteria Decision Analysis.* John Wiley & Sons Inc., New York.

Malczewski, J. (2006). Ordered weighted averaging with fuzzy quantifiers: GIS-based multicriteria evaluation for land-use suitability analysis. *Int. J. Appl. Earth Obs. Geoin.*, **8**: 270–277.

Malczewski, J., Chapman, T., Flegel, C., Walters, D., Shrubsole, D., & Healy, M.A. (2003). GIS-multicriteria evaluation with ordered weighted averaging (OWA): Case study of developing watershed management strategies. *Environ. Plann. A.*, **35**: 1769–1784.

Malczewski, J. & Rinner, C. (2005). Exploring multicriteria decision strategies in GIS with linguistic quantifiers: a case study of residential quality evaluation. *J. Geogr. Syst*. **7**: 249–268.

McBratney, A.B. & Odeh, I.O.A. (1997). Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma, 77, 85e113.

Moghaddam, H.K. Dehghani, M., Rahimzadeh kivi, Z., Moghaddam, H.K. & Hashem, S.R. (2017). Efficiency assessment of AHP and fuzzy logic methods in suitability mapping for artificial recharging (Case study: Sarbisheh basin, Southern Khorasan, Iran). *J. Water Harvesting Res.*, **2**: 57-67.

Nouri, B. (2003). *Identification of Suitable Sites for Groundwater Artificial Recharge Using Remote Sensing and GIS in Gavbandi Watershed*. M.Sc. thesis, Faculty of Natural Resources, Tehran University, Iran.

Oaksford, E.T. (1985). *Artificial Recharge: Methods, Hydraulics, and Monitoring, In: Artificial Recharge of Groundwater*. *In* Asamo, T. (Ed.), Artificial Recharge of Groundwater, Butterworth Publishers, Boston, pp. 69-127.

Saaty, T.L. (1980). *The Analytic Hierarchy Process*. MacGraw-Hill, New York.

Saaty, T. L., & Vargas, L. G. (1998). Diagnosis with dependent symptoms: Bayes theorem and the analytic hierarchy process. *Oper. Res.,* **46**: 491-502.

Saraf, A.K. & Choudhury, P.R. (1998). Integrated remote sensing and GIS for groundwater exploration and identification of artificial recharge sites. *Int. J. Remote Sens.*, **19**, 2595–2616.

Tiwari, A.K., Lavy, M., Amanzio, G., De Maio, M., Singh, P.K. & Mahato, M.K. (2017). Identification of artificial groundwater recharging zone using a GIS-based fuzzy logic approach: a case study in a coal mine area of the Damodar Valley*, Appl. Water. Sci.,***7**: 1-12.

Verma, D. K., Bhunia, G. S., Shit, P. K., Kumar, S., Mandal, J. & Padbhushan, R. (2017). Spatial variability of groundwater quality of Sabour block, Bhagalpur district (Bihar, India). *Appl. Water. Sci.,* **7**: 1997-2008.

Zadeh, L.H. (1965). Fuzzy sets. *Inf. Comput*, **8**: 338–353.

Zehtabian, G.R., Alavipanah, S.K. & Hamedpanah, R. (2001). Determination of an appropriate area for flood water spreading by remote sensing data and GIS. *Proc. Int. Conf. New Tech. New Century*, Seoul, Korea,1–6.

AVAILABILITY-ORIENTED CONTRACT MANAGEMENT APPROACH: A SIMPLIFIED VIEW TO A COMPLEX NAVAL ISSUE

Al-Shafiq Abdul Wahid[*1], Mohd Zamani Ahmad[2], Khairol Amali Ahmad[3] & Aisha Abdullah[4]

[1]Faculty of Mechanical Engineering, University Technology Malaysia (UTM), Malaysia
[2]Institut Sultan Iskandar, University Technology Malaysia (UTM), Malaysia
[3]Universiti Pertahanan Nasional Malaysia (UPNM), Malaysia
[4]Enigma Technical Solutions Sdn. Bhd., Malaysia

[*]Email: al_shafiq@hotmail.com

ABSTRACT

Navies around the world aspire to improve their fleet operational availability. Many navies struggle to achieve their targeted high operational availability even though they are certain that they have continuously implemented improved maintenance concepts and philosophies, allocated the necessary budget and implemented advanced human capital development plans. Nevertheless, the efforts may be futile when they could not be allocated precisely in tackling the issues concerning "human and equipment" related Downtime Influence Factors (DIFs) impacting ship operational availability. The extended exploratory research encapsulated all the efforts in discovering a simplified methodology in tackling this complex naval issue. The 13 objectives achieved in this paper covers the initial work in identifying the DIFs until the development of a Contract Management Control and Monitoring System (ConCaMS) that is able to assist policymakers and all stakeholders including Contract Managers in managing the ship maintenance contract efficiently and effectively. The ConCaMS is able to guide so that actions could be taken earlier enough for recovery to be possible, as opposed to traditional methods. Additionally, the ConCaMS could also be used by policymakers and Top Management of the private sectors as well as the governments as a proven method in comparing contract performance between various contracts, by using availability as the performance benchmark.

Keywords: *Naval ship availability; downtime influence factors (DIFs); availability-oriented framework; contract management control and monitoring system (ConCaMS); recovery availability (Ao) for in-service support.*

1. BACKGROUND ON MAINTENANCE

For many decades, maintenance was regarded as an unavoidable part of the production function and difficult to manage. Hence, maintenance was initially considered as 'necessary rework' and was only given minimal focus. In most organisations, maintenance remains to be considered a burden, and sometimes even considered as a needless cost, that was given the least priority in time, resources and budget. This phenomenon is rampant worldwide, across various industries and through the cultural divide. This negative connotation only changed gradually where maintenance became a separate, fully recognized and essential business function (Xia-Feng *et. al,* 2008). It was only after World War Two that more attention was attributed to it in aviation and in addition in other industrial sectors like defence, nuclear, chemical and petrochemical. Ship maintenance was not well structured or organized in comparison to the other industrial entities which observed that huge savings may be made when carrying out proper maintenance tasks (Lazakis *et al.*, 2010).

Similar to other industries, ship maintenance was considered part of operational tasks needed to be performed on a daily basis, a mere necessity to move the ship to perform its mission from one place to another. For the merchant marine, the shipping industry has made great progress based on studies

and recommendations by academicians as well as consultation by international maritime organisations, governing bodies and classification societies. The improvements also include the areas of safety and environmental protection, with the objective in general of increasing the quality, reliability and availability of the ships. This has consequently increased the positive image of the ship operators, ship owners and supporting organisations from the private sector.

Naval vessels or navy ships on the other hand, are a completely different breed altogether. They have various designs to complete their different missions, with a vastly different range of equipment and systems onboard especially those related to battle and combat management such as Anti Surface Warfare (ASuW), Anti-Submarine Warfare (ASW), Anti-Air Warfare (AAW), Electronic Warfare (EW), Search and Rescue (SAR), humanitarian and many other navy related functions. Examples of complex cross-functional capability frameworks as depicted by (Olivier *et al.*, 2014) in Figure 1.

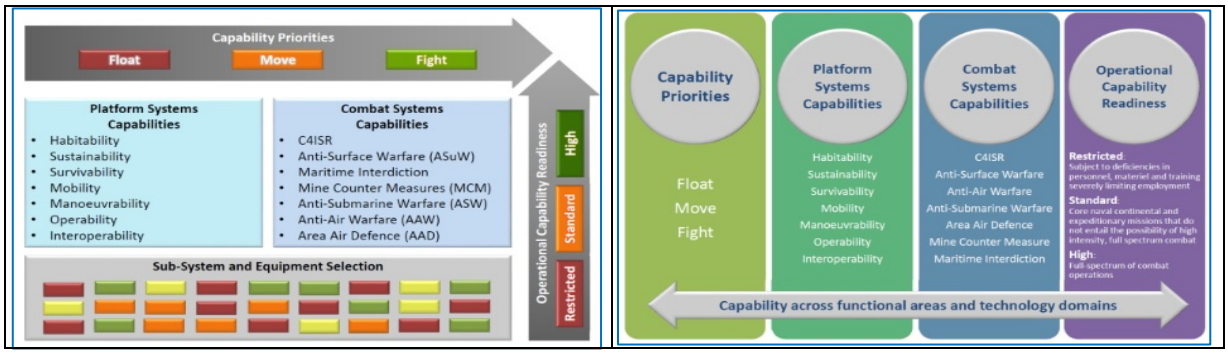


Figure 1: Examples of cross-functional capability frameworks (Olivier *et al.*, 2014).

An example of naval vessels is reflected in Figure 2 illustrating Royal Malaysian Navy (RMN) Patrol Vessels (PV) and Figure 3 depicting the PV undergoing maintenance.

Figure 2: Example of naval vessels (RMN Patrol Vessels KD KEDAH and KD PAHANG).

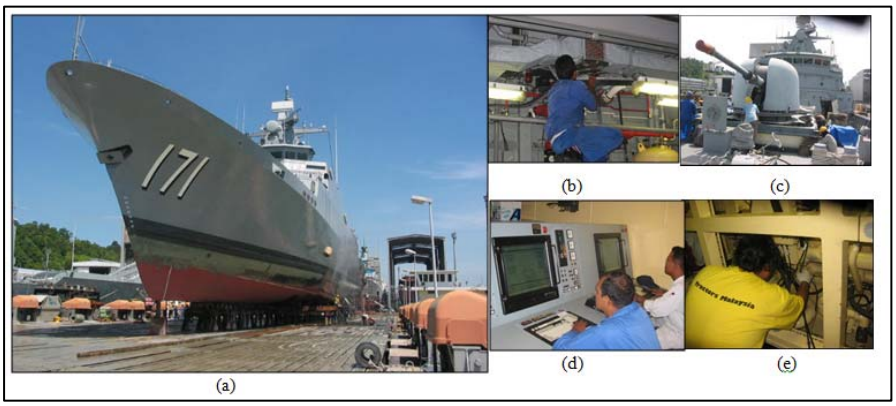


Figure 3: (a) Upslip of RMN PV KD KEDAH (b) HVAC maintenance (c) Gun maintenance (d) Ship control and monitoring system (SCAMS) maintenance (e) Main engine maintenance

Compared to merchant vessels, the naval vessels differ in their concept of operations, their range of equipment and their concept of equipment redundancies, and vary in their policy and priority of

maintenance. In general, naval vessels are not expected to comply with merchant ship's requirements of meeting the environmental standards, and most importantly are not strictly bounded to achieve a targeted profit as compared to commercial establishments.

Nevertheless, these warships have been facing similar issues encountered by their sisters in the merchant marine sector, due to the fact that the shipbuilding contract appears to have no direct bearing on the maintenance contract. Even though it is quite normal for naval ships to have their life-cycle cost (LCC) calculated prior to delivery, the most visible part is the acquisition cost which is normally more evident. The 'not-so-evident' part which includes the operational cost, maintenance cost, spare parts supply costs, engineering documentation, most of which are part of the Integrated Logistics Support (ILS) costs, are not attended to as strictly as the acquisition cost. Frequently the cost of sustaining equipment is 2 to 20 times the acquisition cost (Barringer, 1997). This continues to happen even though the 'not-so-evident' costs over the lifetime of the vessel are significantly higher than the acquisition cost, most likely due to the length of time involved from 'cradle to grave' averaging between 25-30 years and also due to the unfamiliarity of organisations towards this area. This has resulted in limited technical and financial data being collected especially in developing countries including Malaysia based on authors' experience in the RMN and supporting industries, to study and compare projected versus actual maintenance activities and its associated costs.

1.1 Complexity of Naval Ships

A modern naval vessel or warship/submarine would consist of in excess of 100 integrated systems that are linked structurally, mechanically, electrically, hydraulically, pneumatically and electronically (SIA, 2018). The systems need power and cooling, and required to communicate with each other to achieve full operational capability (Henry & Bill, 2015). Consequently, the naval ship operational availability turns into a complex problem (Dell'Isola & Vendittelli, 2015). The ship design of major surface combatants capable of effectively responding to all possible missions within the spectrum of modern conflicts and military operations other than war (MOOTW) is increasingly difficult due to the complex nature of the rapidly evolving and unpredictable global threat environment. Naval ship design can also be understood to be a networked System-of-Systems (SoS) multidisciplinary process whereby a decision on one aspect of the design may have simultaneous, multiple effects on other aspects of the design (Ford *et al.*, 2013; Olivier *et al*., 2014).

Traditional ship design methodologies have evolved from the sequential nature of the design to more advanced computational methods enabling the simultaneous manipulation of several degrees of freedom to better understand the interdependencies between factors (Olivier *et al.*, 2012), consequently this design complexity has been identified by Pascual *et al*. (2006) as causes of greater risk for asset downtime. The naval vessel is also designed to effectively respond to all possible missions and all kind of complex military operations according to its roles. The evolution of the Roles of the Navy (Canadian Navy, 2012) has developed into "Trinity of Roles" and the evolution from Booth Model to Leadmark Model can be described through Figure 4.

Navies worldwide face similar challenges in achieving high asset availability, where the situation is aggravated due to the complex nature of warships including the variety of military roles (Directorate of Maritime Strategy Canada, 2001). To improve any assets operational availability undoubtedly further complicates the problem due to a long list of interconnected contributing factors (GAO,1982), whereby interdependencies and uncertainties involving human and equipment related factors appear with unclear significance and unknown weightage.

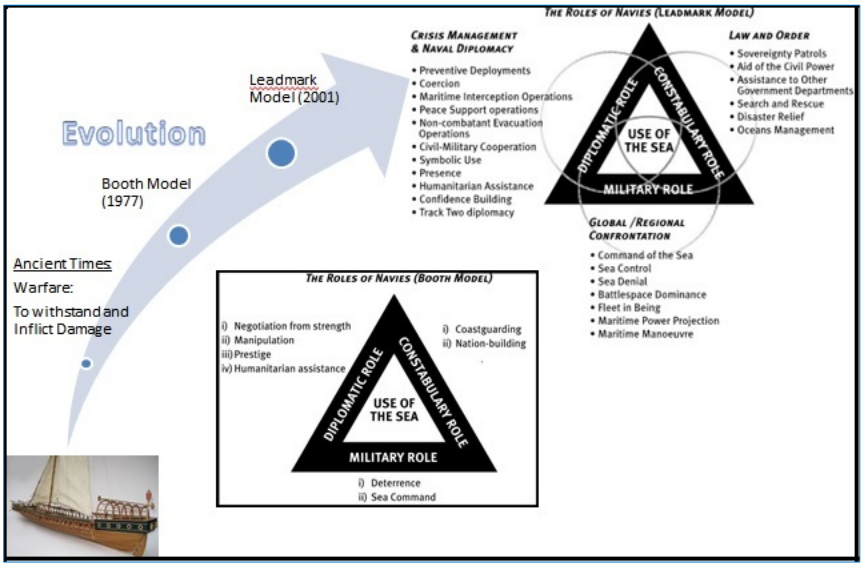


Figure 4: Evolution of the roles of the Navy - "Trinity of Roles".

1.2 Contract Management Concepts

Following delivery of each vessel on completion of construction and trials activities, and especially after the end of the warranty period, the vessel shall be completely under the responsibility of the Navy for the operations and maintenance activities. Unless a special maintenance contract is awarded by the Navy to the shipbuilder or any authorized party, the coordination of maintenance activities would then become less efficient and troublesome due to the limited number of Navy support team personnel allocated to maintain the ship as well as the inexistence of a significantly large budget allocation for the maintenance of the vessel. Many navies consider this as the In-Service phase and sign an In-Service Support (ISS) contract for the maintenance of the vessels. Several navies including United Kingdom (Datta & Roy, 2010; Tomkins, 2012) and Australia (Henry & Bil, 2015) on the other hand implement the most recently popular but costly "Performance or Availability Based Contract" whilst others remain with the traditional "execution upon receipt of order only" philosophy or commonly known as per-order basis. There exist other types of contracts with other sorts of forms and contents but mostly are modifications from the two major types above.

The United States Navy (USN) generally continues to apply a traditional service procurement practice, as opposed to the shift in concept in the UK, since 2000 to apply Availability Contracting, an approach that began replacing the traditional procurement service practice (Datta & Roy, 2009). Therefore it is an accepted fact that the complexity of the naval vessel itself as an asset, with complex roles and missions is further aggravated by the intricacies of the various types of maintenance contracts they belong to.

1.3 Achieving High Operational Availability of Naval Ships – a complex problem

All navies in the world aspire to improve the operational availability of their fleet. Most navies such as the USN (Marais *et al.*, 2013), Korean Navy (Paik, 2014) and RMN (RMN, 2011) have specific operational availability targets, but still remains a problem to be achieved. Astoundingly despite the sophistication and considerably higher maintenance budgets by modern navies such as USN, it remains a question as to why availability is still less than expected.

Naval vessel or warship in itself as an asset is inherently complex, and the operational availability of warship is also a complex problem (Dell'Isola & Vendittelli, 2015). Therefore, improving Ship Availability or Operational Availability of naval vessel further magnifies the complexity of the problem making it "complexly complicated". Ship Availability is defined by Inozu (1996) and Blanchard & Fabrycky (1998) as the probability that the ship is available and capable of performing the intended function at any random point of time. Hou Na *et al.* (2012) described availability as "uptime" which can be formulates as one minus downtime or known as unavailability, with the resulting mathematical implication that the more unavailability or "downtime", the lesser the availability achieved. This can be easily described as stated in Equation 1 and Figure 5:

$$\text{Availability (Uptime)} = 1 - \text{Unavailability (Downtime)} \quad (1)$$

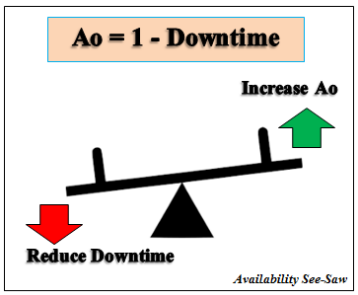


Figure 5: Availability see-saw.

Ship operational availability is also described as the number of days the warships are available for operational tasking in a year (GAO, 2015). Therefore, the objective in achieving high operational availability can only be achieved by reducing the impact of all factors that create downtime or unavailability. Nevertheless, as we know, this remain a challenging task for all navies in the world, similarly faced by any other assets from any engineering fields globally.

1.4 Simplifying a Complex Problem

The complex naval vessel systems and machineries, dynamic and ever-changing roles and mission, location of naval bases, sailing alone or part of a battlegroup, meeting various operational tempo, when compounded with the various types of complicated maintenance contracts they belong to with specific targeted operational availability, makes the situation multiple times more complex than meets the eye. It is further aggravated due to a long list of contributing factors to Ship Availability that are intertwined, with so many ambiguities and uncertainties on the relationship between each factor involving human and equipment, the unclear significance and weightage of each factor, the unknown direct and indirect impact of each factor onto each other, and onto the resulting Ship Availability.

The question now comes to whether it would be possible for the 'complexly complicated' situation to be simplified for the benefit of better understanding of the various levels of stakeholders. Would it be possible to holistically study the human and equipment related Downtime Influence Factors (DIFs) affecting Ship Availability? Would this better understanding of stakeholders benefit various organisations in their ultimate effort for improving the Ship Availability? Would the research findings assist Project Managers and Contract Managers in managing their contracts better, even with some commonly known constraints? Would this research benefit other industries in similar manner?

A few researchers have attempted to consolidate some factors to find interdependencies and also try to implement best practices in Project Management (PM), but none have been able to holistically consolidate as many factors as necessary for a thorough study. The race to maximize operational availability or uptime is hampered by the simple fact that there exists a long list of possible

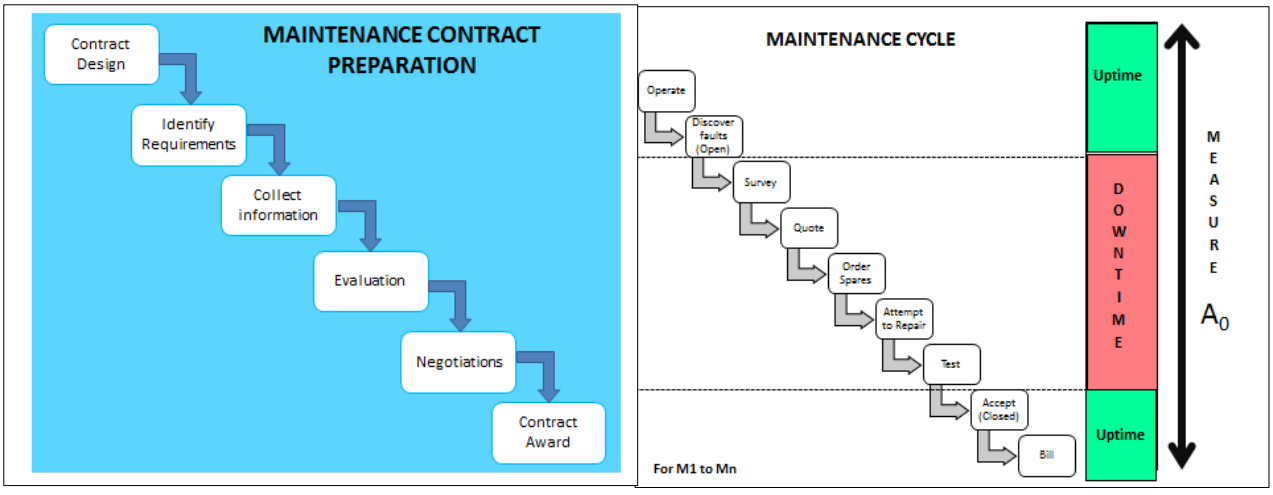
contributing factors affecting downtime. There was no literature published previously that attempted to consolidate human and equipment related factors on naval ships, in fact for any engineering field in general until studied recently by Al-Shafiq *et.al* (2017a). The most current and closest research to the naval ship availability is for the Italian Navy which concluded that Ship Operational Availability of warships require a more innovative and comprehensive approach for the design and support by Dell'Isola & Vendittelli (2015). Donkelaar (2017) studies the operational availability for the Royal Netherlands Navy pointing out that the availability requirements are not met and insufficient at present.

The snowballing effect as a result of ineffective Contract Formulation impacts the Contract Manager three-fold, a weak contract to be implemented resulting in the brewing and subsequent surfacing of a magnitude of issues that could have been avoided, inability for the assets to be managed with high availability, and the non-existence of a model or mechanism to assist the Contract Manager in managing the contract efficiently. This negative effect is magnified due to the limited data being populated and analysed to date with these objectives in mind, as a result of poor awareness and understanding on most stakeholders towards the importance of this issue at hand. The complexity of naval ship maintenance activities coupled with the limited literatures available to date on factors having negative influence on ship availability has created a seemingly impossible task to improve the current situation faced by the contract managers in the implementation of the ISS contract.

The step by step approach in this research would provide all stakeholders with a clearer view to recover from the situation beginning with the identification of the range of DIFs that influence naval ship availability, concentration on the severe or critical DIFs using a Risk Analysis, identification of the severe DIFs' impact to cost, budget, schedule and scope of the contract and finally the development of a mathematical algorithm that provides the opportunity to produce a ship availability-oriented Contract Management System for naval vessels that would provide a solution to systematically tackle the issues mentioned above.

2. AVAILABILITY-ORIENTED CONTRACT MANAGEMENT APPROACH: KEY OBJECTIVES

In accordance to Ford *et al.* (2013) the In-Service phase is considered 70% of the ships through life cost. During In-Service phase, the number of involved stakeholders will vary as the vessel cycles through tasking, upkeep and regeneration. Prior to the ISS phase, a maintenance contract for the vessel would then be prepared, drafted and negotiated. The overall process of the preparation of the maintenance contract is described in Figure 6.



(a) **(b)**

Figure 6: Maintenance contract preparation and implementation: (a) Maintenance contract preparation (b) Maintenance cycle.

There would be numerous maintenance cycles happening onboard the ship on a daily basis as illustrated in Figure 6, and some may even happen concurrently. There are many factors related to human and equipment (system and machinery) that affects the maintenance of the vessels over the contract period, but there exists insufficient reference or historical records to be analysed.

As with most of other navies especially from the developing world, the RMN PV ISS contract managers (RMN, 2011), only monitor and record equipment and spares related information against cost versus schedule, since the contractual requirements common to most navies are to monitor the start and end of rectification process through defect reports as well as the delivery of spares. Nevertheless, these contract managers as well as the maintainers realize that there exist other factors impacting the availability of the vessels, but they were uncertain on the significance and scale of each issue because it has been considered an unchartered territory in research. Therefore, unless a comprehensive study on these factors is conducted to identify and rank these human and equipment related factors, they will continually never be monitored and recorded. Only when historical records are established, analysis could be conducted to decipher the contents. Ultimately, a Contract Management Control and Monitoring System (ConCaMS) could be developed to assist contract managers and maintainers to maintain the vessels effectively and efficiently. This ConCaMS system which is not available in the market to date, may also be utilized as a Diagnostic Tool in guiding the stakeholders in making critical decisions towards meeting the targeted Ship Operational Availability.

In order to pave the way towards this uncharted area of knowledge, the authors have established a list of Key Objectives to be achieved on this research, as indicated in Table 1.

Table 1: Availability-oriented contract management approach key objectives.

Key Objectives	Description
Objective 1	Development of a Conceptual Model on how the human and equipment related factors affect the maintenance and availability of the vessel over a contract period.
Objective 2	Identification of the best methodology to approach the study.
Objective 3	Development of a Conceptual Model depicting the relationship between Operational Availability (Ao), Maintenance activities and Maintenance Cycles.
Objective 4	Identification of human and equipment related factors (Downtime Influence Factors) affecting Ship Operational Availability.
Objective 5	Ranking of the Downtime Influence Factors (DIFs) from most severe to least severe.
Objective 6	Identifying the Impact of DIFs from Contract and Project Management perspectives, especially on Cost, Time, Quality and Scope.
Objective 7	Development of a Contract Management Control and Monitoring System Spiral.
Objective 8	Development of an Availability-oriented Contract Management Framework.
Objective 9	Development of an Availability-oriented Contract Management Cycle.
Objective 10	Development of an Availability-oriented Contract Management Model.
Objective 11	Improving Availability through Change in Contract Clauses – A suggested Mechanism.
Objective 12	Development of an Availability-oriented Contract Management Control and Monitoring System (ConCaMS).
Objective 13	Development of an Availability-oriented Contract Management Dashboard

2.1 Objective 1: Development of a Conceptual Model on the How the Human and Equipment Related Factors Affect the Maintenance and availability of the Vessel Over a Contract Period

Brainstorming sessions and Focus Group Discussions (FGD) managed to derive some conceptual models on how the human and equipment related factors affect the maintenance and availability of the vessel over the contract period. The model is described in Figure 7 as follows:

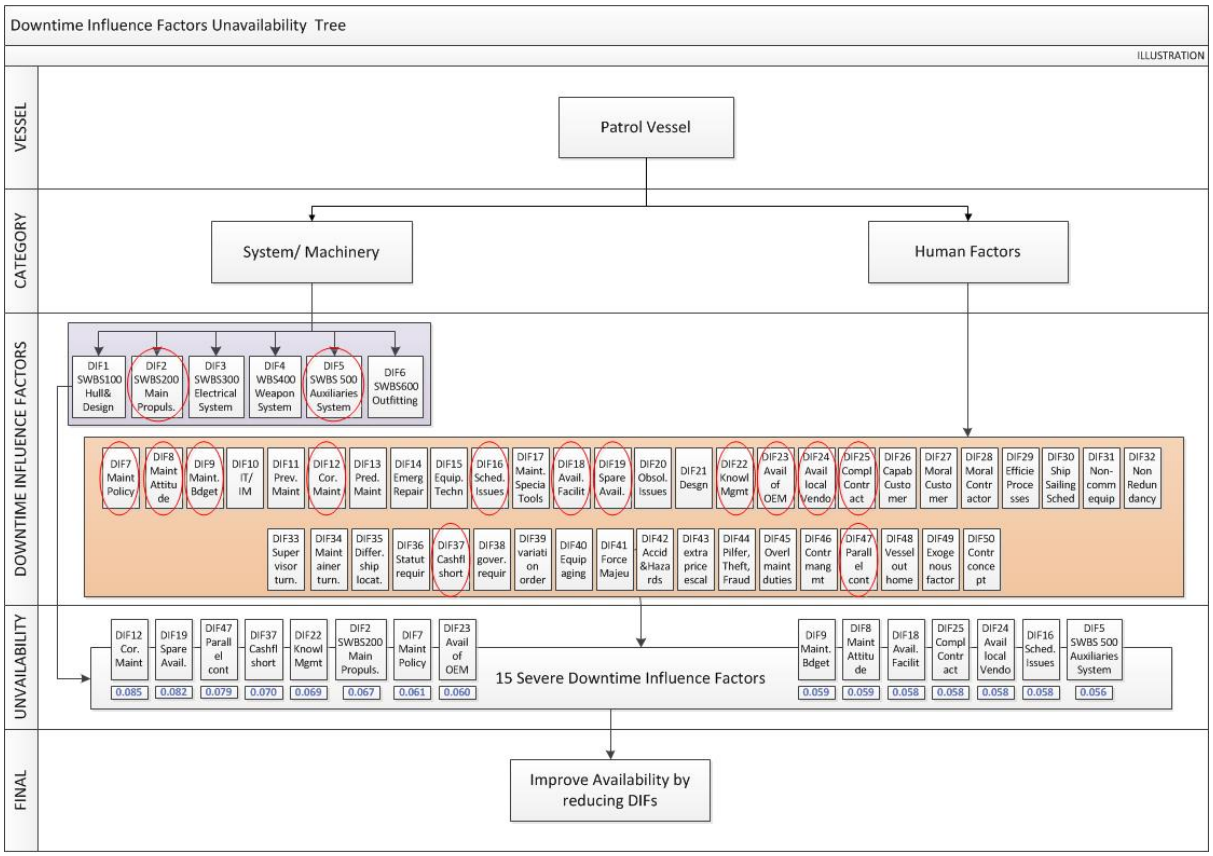


Figure 7: Relationship between human and equipment factors to availability.

2.2 Objective 2: Identification of the Best Methodology to Approach the Study

Since the 1980's efforts have been initiated in studying availability improvement concepts to military assets (GAO, 1982). Various maintenance concepts had been applied by diverse industries worldwide ever since with varying degrees of success. The authors have explored the usage of many methodologies for this research, however Delphi was chosen as the most suitable methodology in line with Skulmoski *et.al* (2007) to explore new concepts within and outside the existing body of knowledge in the field and in accordance to Franklin & Hart (2007) since the complexity of naval ship availability phenomenon is without previous history, a quickly changing event that outdates the literature, and a very complex phenomenon that truly requires experts for understanding it.

2.3 Objective 3: Development of a Conceptual Model depicting the relationship between Operational availability (Ao), Maintenance Activities and Maintenance Cycles

The relationship between Operational Availability (Ao), Maintenance activities and Maintenance Cycles for the RMN PV ISS Contract (RMN, 2011) can be described in Figure 8. This situation is also generically applicable to most of other navy fleets.

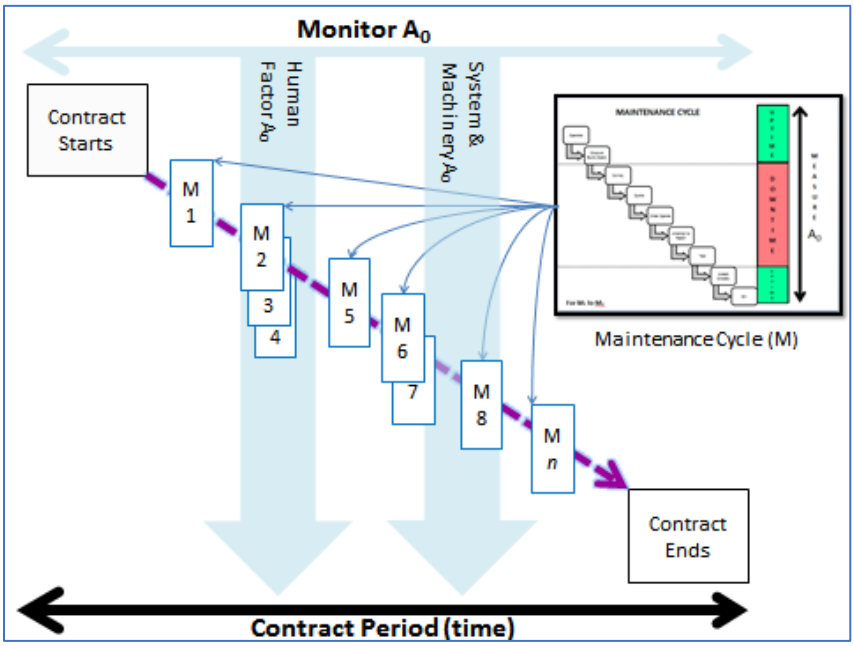
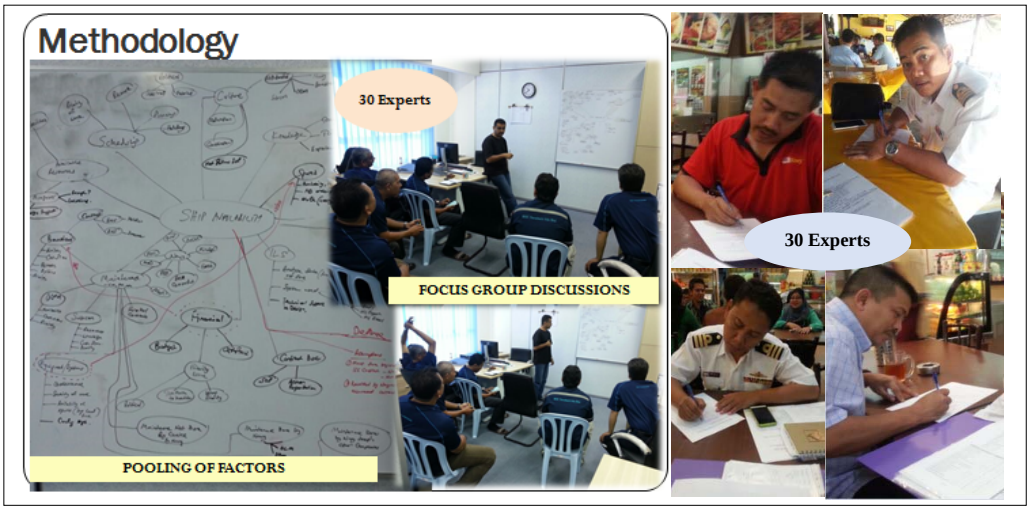


Figure 8: Relationship between Operational Availability (Ao), Maintenance activities and cycles.

2.4 Objective 4: Identification of the Human and Equipment Related Factors (Downtime Influence Factors) Affecting Ship Operational Availability

Prior to this research there was no specific holistic study of all the factors that affect naval maintenance downtime or naval ship availability. Therefore, for this exploratory study, extensive Literature Research across more than 500 literatures were conducted by the authors across various engineering disciplines, and the screening process resulted in close to 200 literatures found applicable to generate the list of factors. The result was used during the subsequent brainstorming session and Focus Group Discussions with 30 experts in Figure 9 from the PV ISS Maintenance organisations and the RMN to reconfirm and pool the variables into relevant groups. The population of interest has been described in this study as experienced, knowledgeable Malaysian Naval ISS experts that have direct involvement in the PV ISS Contract. The total number of experts complying with these criteria was 46. Subsequently, the researcher applied judgmental sampling based on the accessibility of these experts to determine the selected 30 experts. Baker & Edwards (2012) recommended guidance on sampling size for qualitative interviews and stated that saturation is central to qualitative sampling depending on the methodological and epistemological perspective. Meanwhile Adler & Adler (2011) advised sample pool sizes with a mean of 30, though later confirmed that the best answer is to gather data until empirical saturation has been reached since some qualitative researchers have argued that as little as 1 opinion can add value to the area of research. Moreover, good results can be obtained with a homogenous group of experts, with a panel as small as 10 to15 individuals (Adler & Ziglio, 1996).

The result was presented by Al-Shafiq *et al.* (2017a, b) as the identified Downtime Influence Factors (DIFs) that affect naval ship availability for the PVs in Malaysia. The method in identifying the variables is reflected in Figure 10. Subsequently, the authors proceeded with multiple rounds of Mixed Method Sequential Delphi with Snowballing Technique as reflected in Figure 11, moving from 30 Experts to five Top Management Experts as part of the self-validating iterations in the Delphi process.



(a) (b)

Figure 9: Experts participating in Brainstorming and FGD during Delphi Rounds: (a) The PV ISS Maintenance organisation (b) The RMN officers.

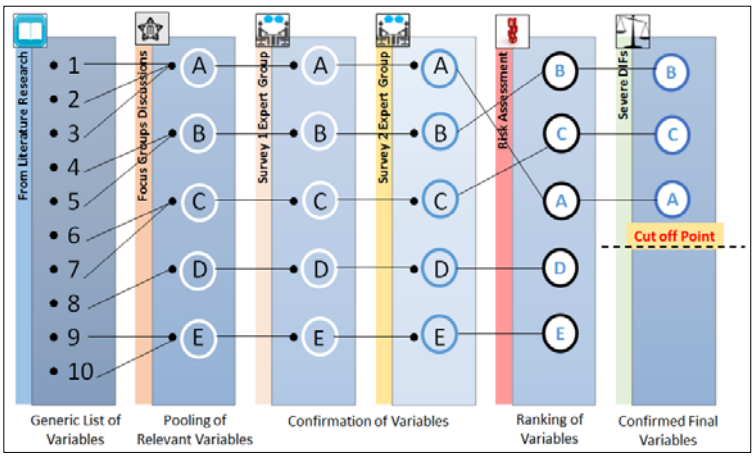


Figure 10: Method of identifying key variables.

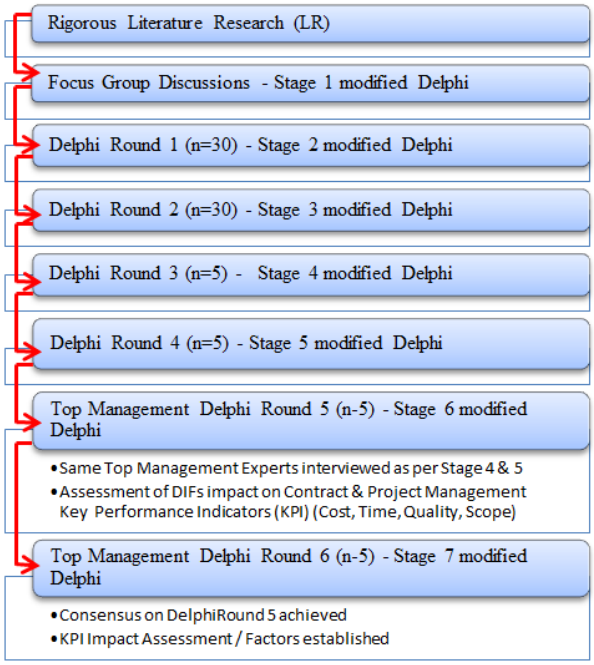


Figure 11: The Delphi rounds.

2.5 Objective 5: Ranking of the Downtime Influence Factors (DIFs) from Most Severe to Least Severe

Al-Shafiq *et al.* (2017a, b) identified via a mixed method sequential modified Delphi approach the Downtime Influence Factors (DIFs) that impact the Naval Availability of the Royal Malaysian Navy (RMN) In-Service Support Contract. In total 35 expert opinions were elicited. The list of 50 DIFs was then prioritised based on a risk management model and 15 Severe DIFs were identified as per Figure 12.

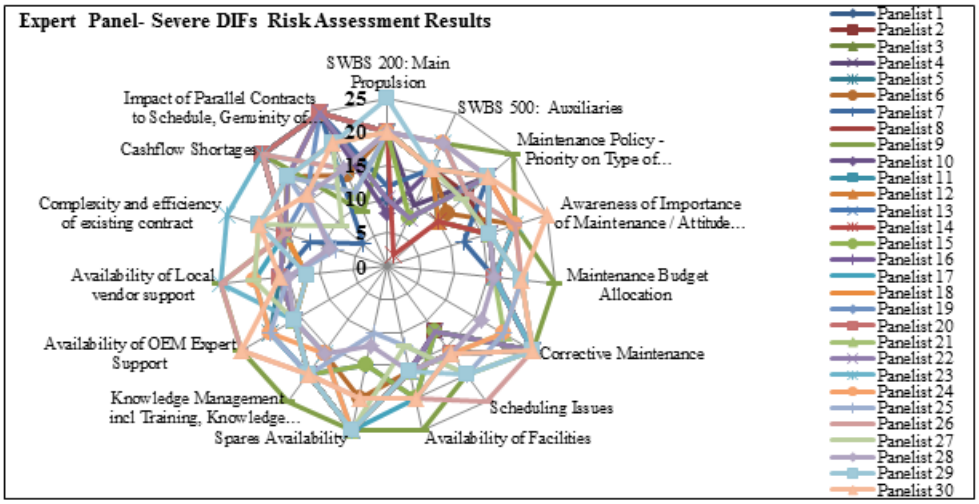


Figure 12: Expert panel – Severe DIFs risk assessment results.

2.6 Objective 6: Identifying the Impact of DIFs from Contract and Project Management Perspectives, Especially on Cost, Time, Quality and Scope.

There is a clear relationship between Project Management (PM) and Contract Management (CM), as well as the relationship of both towards maintenance activities. On the other hand, there is an existing relationship between maintenance activities and availability. Darnall & Preston (2010) describes that PM is complicated because project manager must understand several knowledge areas and develop a variety of tools and technique to successfully manage a project. In a nutshell, PM is focused at managing all aspects of a project to ensure that it can be completed and that the project deliverables are achieved within the main project constraints (Scope, Time, Cost & Quality) which are basically in accordance with the contract. CM is focused at ensuring that terms and commitments agreed in the contract are adhered to. Contract Managers responsibility areas overlap at times with those of a Project Manager, since contract managers are tasked with ensuring that projects are delivered on budget or profitably. Both PM and CM activities for Naval ISS contracts are intrinsically linked via the limiting factors or constraints to the Ship Availability through the DIFs as diagrammatically represented in Figure 13.

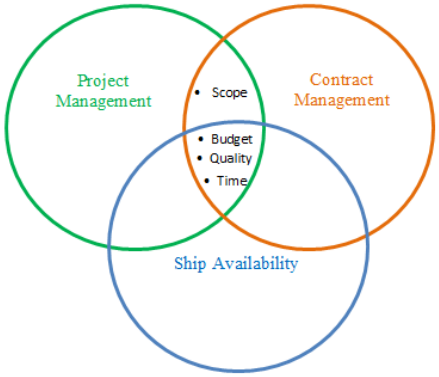


Figure 13: PM, CM and ship availability constraints and impact factors.

A questionnaire was produced and administered to the Top Management Experts in a follow up interview from Stage 5 of Delphi. The objective was to understand the link between the 15 Severe DIFs to the PM Constraints and the CM Objectives. The constraints of "Cost", "Time", "Quality" and Scope" were identified as Key Performance Indicators (KPIs). A three-point rating scales for the effect on each KPI was used as per Table 2. Table 3 contains the consolidated results on PM and CM KPIs.

Table 2: Three-point rating scale to quantify the effect of each KPI.

Cost:	**Time:**	**Quality**	**Scope:**
No Impact (NI)	No Impact (NI)	No Impact (NI)	Fixed
Lower (L)	Shorter Duration (SD)	Better (B)	
Higher (H)	Extended Duration (ED)	Reduced (R)	

Table 3: Severe DIF impact on KPIs.

S/No	Severe DIFs	Top Management Experts					
		KPI	E1	E2	E3	E4	E5
1	SDIF1	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
2	SDIF2	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
3	SDIF3	Cost	NI	NI	NI	NI	NI
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
4	SDIF4	Cost	NI	NI	NI	NI	NI
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
5	SDIF5	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
6	SDIF6	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
7	SDIF7	Cost	NI	NI	NI	NI	NI
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
8	SDIF8	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F

S/No	Severe DIFS	Top Management Experts					
		KPI	E1	E2	E3	E4	E5
9	SDIF9	Cost	H	NI	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
10	SDIF10	Cost	NI	NI	NI	NI	NI
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
11	SDIF11	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
12	SDIF12	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
13	SDIF13	Cost	NI	NI	NI	NI	NI
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
14	SDIF14	Cost	NI	NI	NI	NI	NI
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F
15	SDIF15	Cost	H	H	H	H	H
		Time	S	S	S	S	S
		Quality	B	B	B	B	B
		Scope	F	F	F	F	F

The results of the study were described by the authors in Al-Shafiq *et al*. (2017c) as follows:

a. The improvement of certain Severe DIFs would not have a negative "cost" impact.
b. In addition, the reduction of all 15 severe DIFs will have a positive effect on "time" and "quality". Since "scope" is considered fixed for the ISS contract period there is no impact on scope.
c. The possibility that the negative impact on "costs" to be outweighed by the positive effects on "time" and "quality.
d. The findings confirm that all 15 Severe DIFS have impact on project management and contract management constraints of cost, time, quality and scope. It is also possible to identify whether the impact is positive, negative or neutral.
e. An important finding is that contract managers are now able to pinpoint which DIFs to improve when there are budget or cost limitations.

2.7 Objective 7: Development of a Contract Management Control and Monitoring System (ConCaMS) Spiral

The authors realised the necessity to develop a figure to reflect the steps taken in the study, so a Contract Management Control and Monitoring System Spiral in Figure 14 was produced.

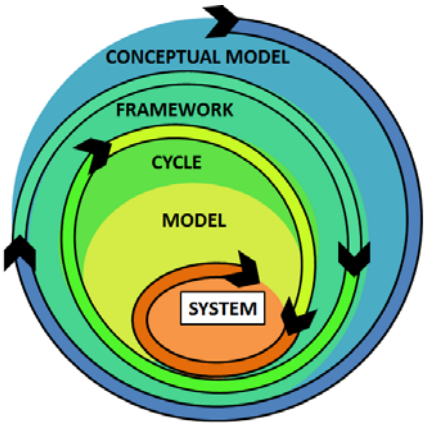


Figure 14: Contract Management Control & Monitoring System (ConCams) development spiral.

2.8 Objective 8: Development of an Availability-oriented Contract Management Framework

An Availability-oriented Contract Management Framework was developed by the authors after taking consideration of all available INPUT from prior research steps, and inserted all requirements for the PROCESS as well as the expected OUTPUT, based on the McGrath (1984) IPO Model. The developed Framework is described in Figure 15.

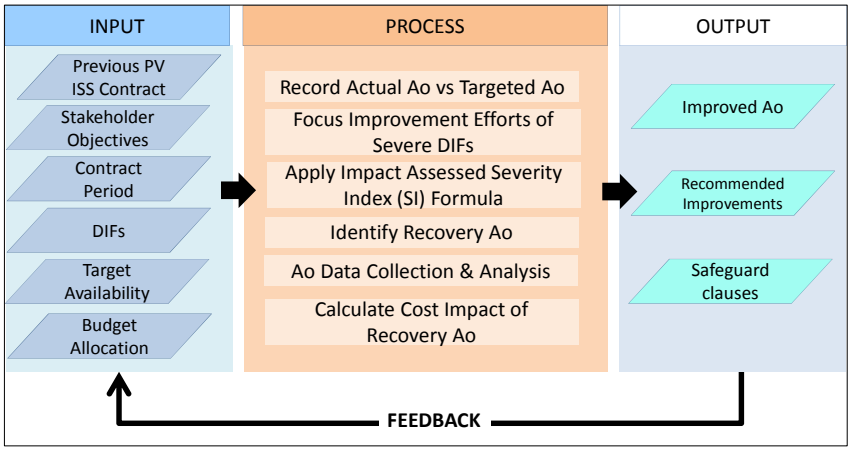


Figure 15: Input-Process-Ouput (IPO) framework for ISS ship availability-based contract.

2.9 Objective 9: Development of an Availability-oriented Contract Management Cycle

During the contract period, there would be uptimes and downtimes for the naval vessels. This downtime includes the maintenance periods (M_1 to Mn) as described in Figure 8. The DIFs would be the factors that influence the downtime, whereby those that have a negative impact over a prolonged period are considered severe DIFs. Ideally, at the end of the 3-year PV ISS contract period, the targeted availability is compared with the actual availability of the vessels, and improvements from "lessons learned" are expected to be implemented in the next contract. An Availability-oriented Contract Management Cycle has been developed by the authors for this purpose following discussions with the Experts as described in Figure 16.

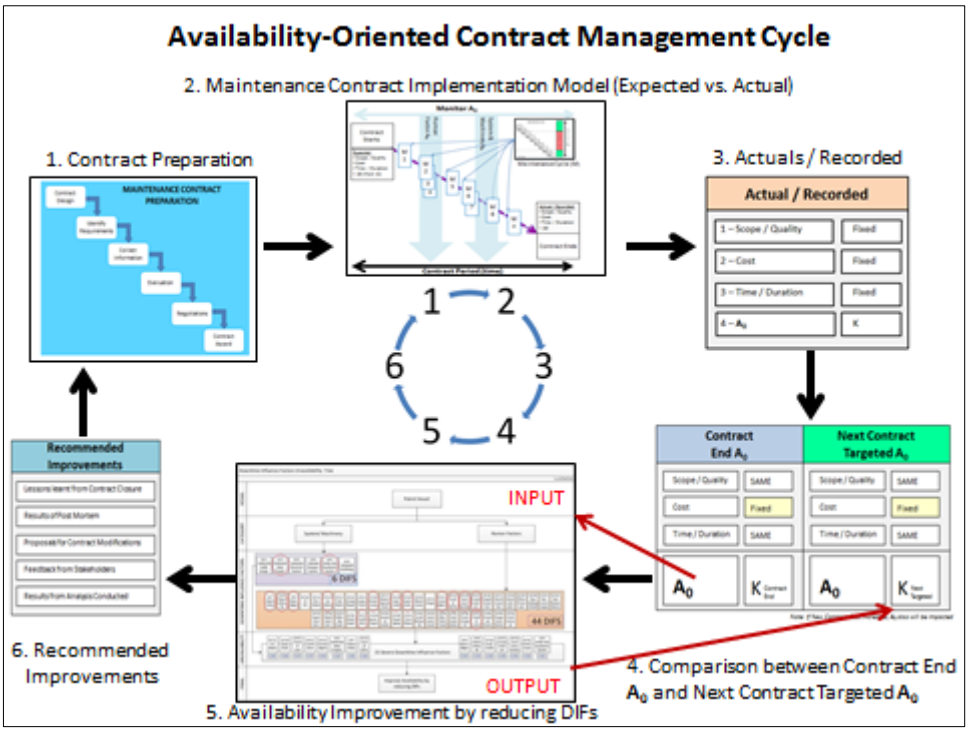
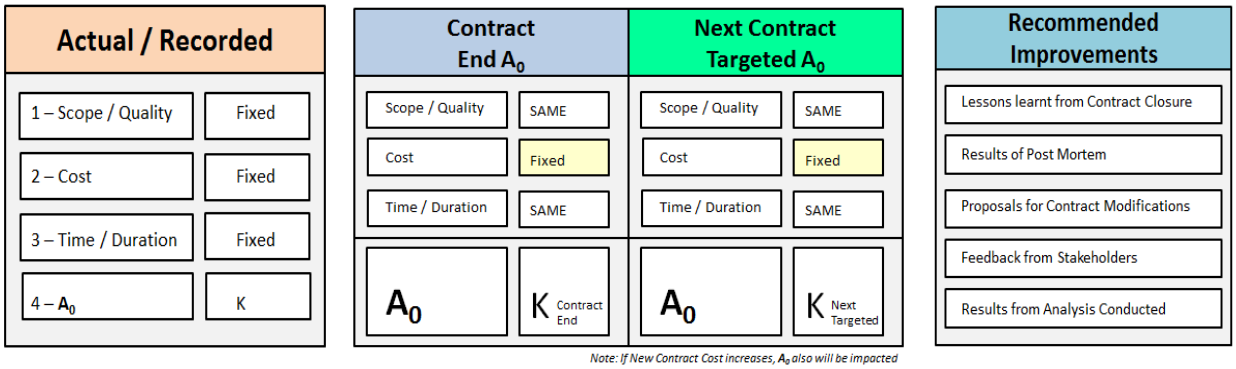


Figure 16: Availability-oriented Contract Management Cycle.

The Key Monitoring Criteria for the Availability-oriented Contract Management Cycle is reflected in Figure 17 and the steps to be followed are described in Steps 1 to 6 in Figure 16:



(a) (b) (c)

Figure 17: Key Monitoring Criteria: (a) Actual vs Recorded (b) Comparison of contract end Ao vs Targeted Ao (c) Recommended improvements.

The Experts who are stakeholders realized that they concentrate mostly on day-to-day operations and kept busy in "everyday fire-fighting culture' that they have never been able to record and analyse the past to improve in the future. Urgencies supersedes importance, and problems become crises. On the other hand, many of the stakeholders agreed that an in-depth research as triggered by the authors is necessary before a concerted effort could possibly be placed in improving the implementation of the PV ISS contract in the future, as they are currently blind and clueless to the root causes as well as the recommended solutions. It is the authors' aim to develop a systematic approach towards managing these real-life and legacy issues through this research.

Equally important is the new concept of "Recovery Availability" or "Recovery Ao" as termed by the authors which shall assist in guiding the Top Management especially the Contract Managers in taking the necessary steps on a daily basis in achieving the targeted availability as opposed to only be able to

discuss during scheduled monthly or quarterly meetings without any substance or evidence. With proper guidance and Top Management buy-in, the Contract Managers shall have the mandate to implement the necessary improvement efforts on resource and time allocations in order to be on the path of recovery towards the Targeted Availability. An example of the calculated availabilities and Recovery Ao based on the developed ConCaMS system is reflected in Figure 18.

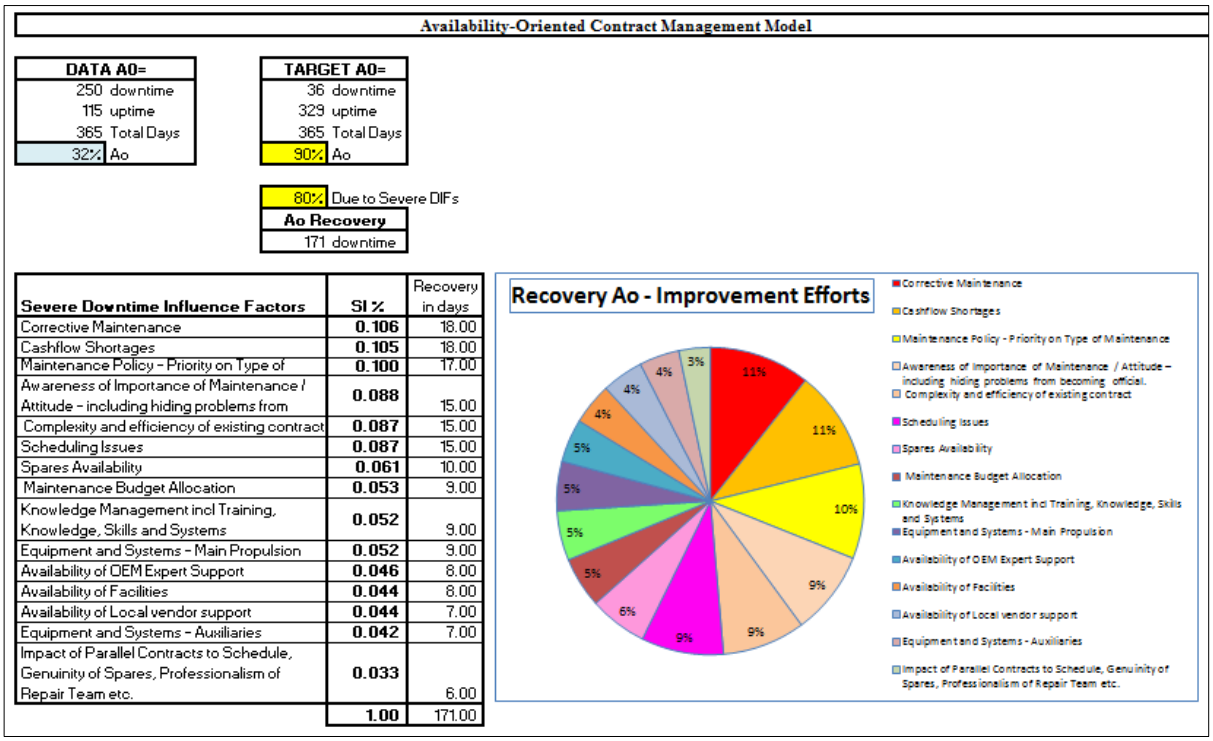


Figure 18: Example of Recovery Ao in the ConCaMS.

2.10 Objective 10: Development of an Availability-oriented Contract Management Model

The 50 DIFs identified in the earlier stages were assessed based via Risk Assessment Method. Qualitatively, risk is proportional to the expected losses that can be induced by a certain accident and to the likelihood of an occurrence. Greater loss and greater likelihood result in an increased overall risk (Ristic, 2013). In engineering, the definition of risk is:

$$\text{Risk} = (\text{Probability of Incident/ Accident}) \ X \ (\text{Losses per Incident / Accident}) \qquad (2)$$

Both panellist groups were requested to identify and rank the DIFs by severity by assigning a value to the probability of the DIF occurring during the contract duration and the Impact the DIF had onto the Availability of Naval Vessel for the ISS Contract by means of a 5-point Likert Scale as per below.

Table 4: 5 Point Likert scale for impact and likelihood.

Impact	
Description	**Rating**
Extreme	5
High	4
Medium	3
Low	2
Negligible	1

Likelihood	
Description	**Rating**
Almost Certain	5
Likely	4
Possible	3
Unlikely	2
Rare	1

After identifying the quantity of key measures of DIFs, the experts scoring was referred to determine the DIF Severity Index. The starting point was to identify the importance of each weighting. The cut off point for a Severe DIF was determined as 16 with an availability impact perceived as "High and above" and a probability of occurrence of "Likely and above". A preliminary series of weighted Severity Measures (SM) was developed based on the mean ratings advocated by all the respondents as presented by Al-Shafiq *et al*. (2017d). The weighting for each of the top DIFs was computed using the following equation:

$$W_{SMi} = \frac{M_{SMi}}{\sum_{1}^{15} S_{SMi}} \quad (3)$$

where:

$W_{SM}i$ represents the importance weighting of particular severe DIFs
$MSMi$ represents the mean rating of particular severe DIFs
$\sum SSMi$ represents the summation of the mean rating of the severe DIFs

A composite indicator was developed to evaluate severity of the DIF for a particular contract or project. A Severity Index (SI) was designed which can be represented by the following formula:

$$SI = W_{SM\ (DIF1)} + W_{SM\ (DIF2)} + W_{SM\ (DIF3)} + W_{SM\ (DIF4)} + W_{SM\ (DIF5)} + W_{SM\ (DIF6)} + W_{SM\ (DIF7)} + W_{SM\ (DIF8)} + W_{SM\ (DIF9)} + W_{SM\ (DIF10)} + W_{SM\ (DIF11)} + W_{SM\ (DIF12)} + W_{SM\ (DIF13)} + W_{SM\ (DIF14)} + W_{SM\ (DIF15)} \quad (4)$$

Once the SI had been defined, the PM and CM KPI score was quantified for each of the severe DIFs. The initial algorithm was derived based on the assumption that this is a linear and additive model. Nevertheless, it is only valid to derive a linear and additive model if there is no correlation between the weighted Severe DIFs. Pearson correlation matrix was calculated and analysed for the algorithm development in this study using SPSS to ascertain the linear correlation. The Pearson's correlation coefficient obtained in SPSS was referred to determine whether the linear relationship between Weightage of Severity (WOS) was statistically significant. A statistically significant relationship between two or more WOS represented a challenge and requirement to adjust the SI algorithm to consider the multiplier effect between these factors. A linear correlation or multiplier effect is subsequently singled out and adjusted in the SI.

Downtime Influence Factors to Ship Availability	Initial Rank	Adjusted SI Rank	Initial SI	Adjusted SM & SI
Corrective Maintenance.	1	1	0.085	0.142
Spares Availability.	2	2	0.082	0.082
Impact of Parallel Contracts.	3	15	0.079	0.022
Cashflow Shortages.	4	3	0.078	0.078
Knowledge Management.	5	4	0.070	0.070
Equipment and Systems – Propulsion.	6	5	0.069	0.069
Maintenance Policy and Priority.	7	6	0.067	0.067
Availability of OEM Expert Support.	8	7	0.061	0.061
Maintenance Budget Allocation.	9	8	0.060	0.060
Awareness of Importance of Maintenance / Attitude.	10	9	0.059	0.059
Availability of Facilities.	11	10	0.059	0.059
Availability of Local Vendor Support.	12	11	0.058	0.058
Complexity and Efficiency of Existing Maintenance Contract.	13	12	0.058	0.058
Scheduling Issues.	14	13	0.058	0.058
Equipment and Systems – Auxiliaries	15	14	0.056	0.056
TOTAL			1.000	1.000

$$SI = 0.142\ X\ Corrective\ Maintenance + 0.082\ X\ Spares\ Availability + 0.022\ X\ Impact\ of\ Parallel\ Contracts + 0.078\ X\ Cashflow\ Shortages + 0.070\ X\ Knowledge\ Management + 0.069\ X\ Equipment\ and\ Systems:\ Main\ Propulsion + 0.067\ X\ Maintenance\ Policy + 0.061\ X\ Availability\ of\ OEM\ Expert\ Support + 0.075\ X\ Maintenance\ Budget\ Allocation + 0.059\ X\ Awareness\ of\ Importance\ of\ Maintenance\ \&\ Attitude + 0.059\ X\ Availability\ of\ Facilities + 0.058\ X\ Availability\ of\ Local\ Vendors + 0.058\ X\ Complexity\ and\ efficiency\ of\ existing\ contracts + 0.042\ X\ Scheduling\ issues + 0.056\ X\ Equipment\ and\ Systems:\ Auxiliaries$$

Figure 20: Impact assessment adjusted SI.

A preliminary series of weighted Severity Measures (SM) was developed based on the mean ratings advocated by the 35 respondents. The weighting for each of the top 15 SMs was computed according to Equation 4. Only two instances of linear correlation or multiplier effect were found. These were singled out and adjusted in the SI. The resulting Impact Assessment (IA) SI formula is described as per Figure 20.

2.11 Objective 11: Improving Availability through Change in Contract Clauses – A suggested Mechanism

Based on authors' experience, during the naval ship ISS maintenance contract preparation and negotiation stage, neither the RMN nor the Subcontractor is aware of any mechanism or model to simulate possible outcomes of the ISS Contract to be signed. As a result, the ISS Contracts continue to be awarded based on legacy contract terms and clauses. There has been no improvement due to the lack of studies being carried out on improving the contract clauses as well as the contract clauses' relevancy towards the dictated Ship Availability. In the case of the RMN PV ISS Contract, the contract contains a total of 58 Clauses.

A possible approach in improving the Availability is by identifying which clauses have a direct impact on Availability, i.e. Availability Subset clauses. The proposed mechanism is to cross-tabulate the totality of the PV ISS Contract clauses against the 15 Severe DIFs identified in earlier research stages. Thereon each clause is carefully analysed and dissected in terms of the likelihood that a change in the clause would impact the said DIF. The clauses are rated as either "Not Relevant" (NR), 1 as "Relevant" and editions required to clauses and "0" for Relevant but no editions required.

An example is Clause 1 Definitions of Terms. There are certain terms that if defined explicitly with the corresponding action it can guide and prompt contract stakeholders to improve availability. i.e. Defining Beyond Economical Repair (BER) with the corresponding action that a spare part classified as BER requires an immediate notification to be sent to the GOVERNMENT. Another example is specifying that a minimum stock as per the suggested preventive maintenance plan must be met in order to avoid spare parts unavailability. The researchers have identified a total of 32 clauses out of 58 clauses for which the clause formulation could impact the availability throughout the contract period. For this, Figure 21 shows a subset of the findings.

SEVERE DIFs	
DIF S1	SWBS 200: Main Propulsion
DIF S2	SWBS 500: Auxiliaries
DIF S3	Maintenance Policy - Priority on Type of Maintenance
DIF S4	Awareness of Importance of Maintenance / Attitude – including hiding problems from becoming official.
DIF S5	Maintenance Budget Allocation
DIF S6	Corrective Maintenance
DIF S7	Scheduling Issues
DIF S8	Availability of Facilities
DIF S9	Spares Availability
DIF S10	Knowledge Management incl Training, Knowledge and Skills
DIF S11	Availability of OEM Expert Support
DIF S12	Availability of Local vendor support
DIF S13	Complexity and efficiency of existing contract
DIF S14	Cashflow Shortages
DIF S15	Impact of Parallel Contracts to Schedule, Genuinity of Spares, Professionalism of Repair Team etc

NR=	Not Relevant
YES = 1	Needs to be improved
NO=0	Relevant but no editions

CONTRACT CLAUSES		DOWNTIME INFLUENCE SEVERITY FACTORS															
		1 DIF1	2 DIF S2	3 DIF S3	4 DIF S4	5 DIF S5	6 DIF S6	7 DIF S7	8 DIF S8	9 DIF S9	10 DIF S10	11 DIF S11	12 DIF S12	13 DIF S13	14 DIF S14	15 DIF S15	Ao Subset?
CLAUSE 1	DEFINITION OF TERMS		1	1	1	1	1	1	1	1	1	1	1	1	1	1	YES
CLAUSE 2	INTERPRETATION	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 3	REPRESENTATION AND WARRANTY	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 4	SCOPE OF CONTRACT	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	YES
CLAUSE 5	DESCRIPTION OF THE SPARES, MAINTENANCE, ILS AND	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NO
CLAUSE 6	TENURE OF CONTRACT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NO
CLAUSE 7	COSTOF CONTRACT AND STAMP DUTY	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 8	AMENDMENTS TO THE CONTRACT	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 9	GOVERNMENT RIGHTS	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	YES
CLAUSE 10	ORDERING METHOD	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	YES
CLAUSE 11	CONTRACT VALUE	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	YES
CLAUSE 12	PRICES	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	YES
CLAUSE 13	TAXES	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 14	PERFORMANCE BOND	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 15	METHOD OF PAYMENT	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NO
CLAUSE 16	PACKING AND MARKING	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	YES
CLAUSE 17	PACKING AND PRESERVATION	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	YES
CLAUSE 18	BAR CODING	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	YES
CLAUSE 19	TRANSPORTATION	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	YES
CLAUSE 20	INSURANCE	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	YES
CLAUSE 21	DELIVERY PERIOD FOR SPARES	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	YES
CLAUSE 22	SUPPLY OF SPARE PARTS FOR MAINTENANCE	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	YES
CLAUSE 23	TURN AROUND TIME FOR MAINTENANCE AND ILS	0	1	1	0	0	1	1	0	0	0	1	1	0	0	1	YES
CLAUSE 24	DELIVERY PERIOD FOR TRAINING	1	1	0	1	0	1	1	1	0	1	1	1	0	0	0	YES

Figure 21: Availability oriented approach – impacted clauses.

The proposed mechanism requires a significant amount of contract stakeholder engagement and feedback to corroborate the proposed changes to the clauses. Contracts are typically signed at the beginning of a 3-year contract period and will not be amended until the next contract period. As such any proposed changes would require to be incorporated into the new contract. Due to the time constraints of the study parameter, the intention of the authors was to "pave the way" for future research to validate the proposed mechanism. An example of the Contract Clause Flow Mechanism to improve the impacted clauses is reflected in Figure 22.

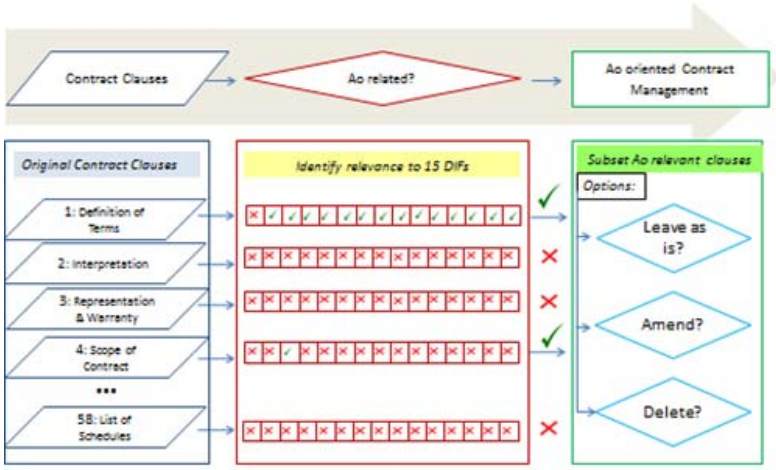


Figure 22: Contract clause flow mechanism.

2.12 Objective 12: Development of an Availability-oriented Contract Management Control and Monitoring System (ConCaMS)

The step by step approach in the ConCaMS Spiral with associated objectives as reflected in Figure 23 would provide all stakeholders with a clearer view of the steps taken from Objective 1 to Objective 13 for the purpose of achieving the target of improving ship operational availability. This includes development of conceptual models, identification of DIFs, ranking of DIFs using Risk Analysis methodology, Identification of the DIFs that impact to cost, budget, schedule and scope of the contract, the development of a mathematical algorithm resulting in the Severity Index (SI), all the way to the development of the Availability-oriented Framework, Model and System.

This would provide all stakeholders including the contract managers the tool to systematically plan, calculate, diagnose, project, and manage the contract implementation during and after the contact period with a firm control of all factors that impact the ship availability. This shall also enable the Top Management of organisations to use this tool to compare contact performance between similar contracts albeit having some differences between them. To date, there has not been any suitable tool that is generally being able to assist in conducting contract performance benchmarking especially on naval ship ISS maintenance contracts. Two contracts with different Budget, Time, Quality and Scope could still be compared by using Ship Availability as the determining criteria using the ConCaMS System.

This Availability-oriented Approach, is a breakthrough that would eliminate the previous real-life issues of contract manager's inability to use any guide or model or mechanism to measure and control risks during the implementation of the contract, which has a snowballing effect in another blind preparation of future contracts. A display of the ConCaMS output is reflected in Figure 24.

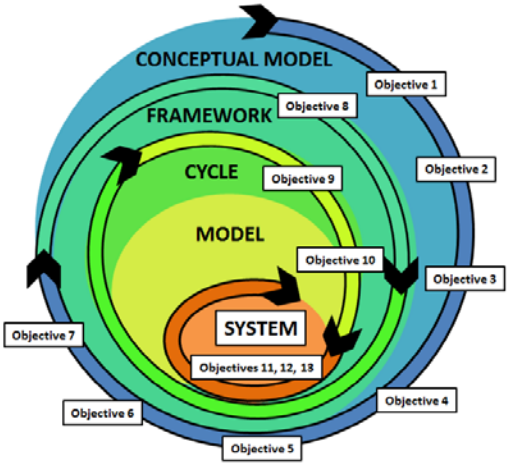


Figure 23: ConCams Development Spiral with associated objectives.

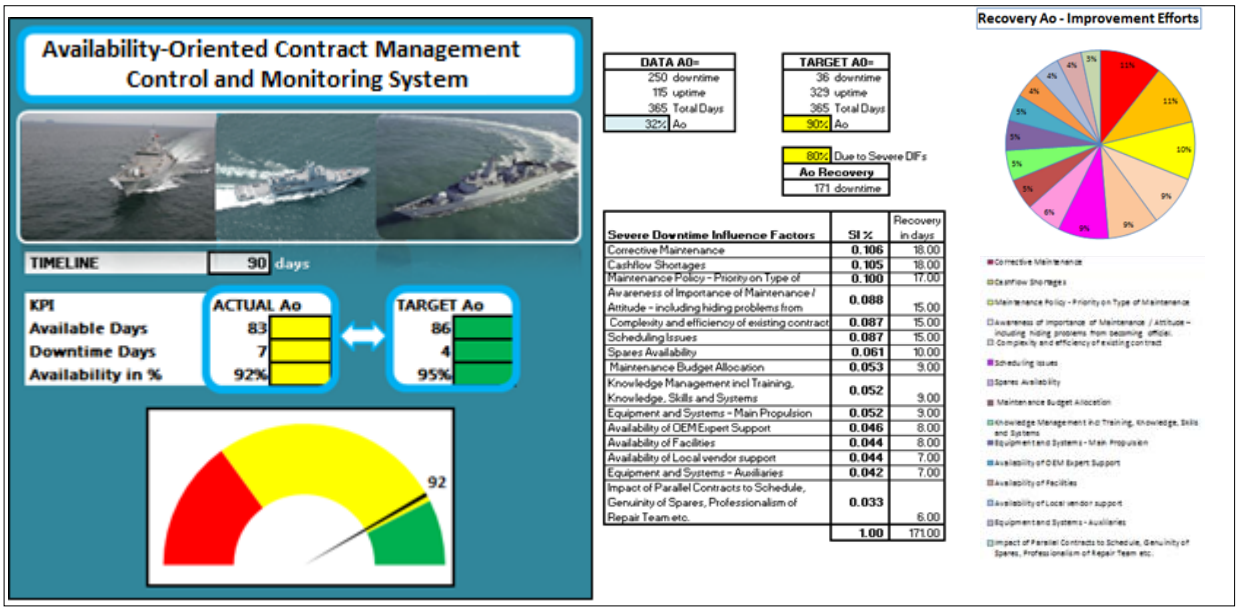


Figure 24: ConCaMS Display output.

2.13 Objective 13: Development of an Availability-oriented Contract Management Dashboard

For the benefits of future research, a template of an Availability-oriented Contract Management Dashboard has been developed by the authors. The dashboard was developed with feedback from the Experts and confirmed by Top Management of the RMN Planning as well as the RMN Strategic Management Department as logical and reasonable method in daily collection of data onboard every vessel in the future. The data collection would enable the ISS Contract Managers from the private sector and the RMN to better analyse the impact of DIFs on the availability of navy vessels, and make any necessary improvements when compared to the published results of the current study by the authors using Expert opinions.

The Dashboard is also Availability-oriented, enabling the Contract Manager to monitor the availability status of each vessel, also the combined availability status of the fleet, with simple

indicators highlighting the daily Actual and compounded Actual versus Targeted Availabilities, with a calculated Recovery Availability (Recovery Ao) figure displayed for reference. The dashboard shall also be able to record possible additional DIFs that have not been discovered previously in the current research. An example of the Dashboard Input and Ouput screens are as per Figure 25.

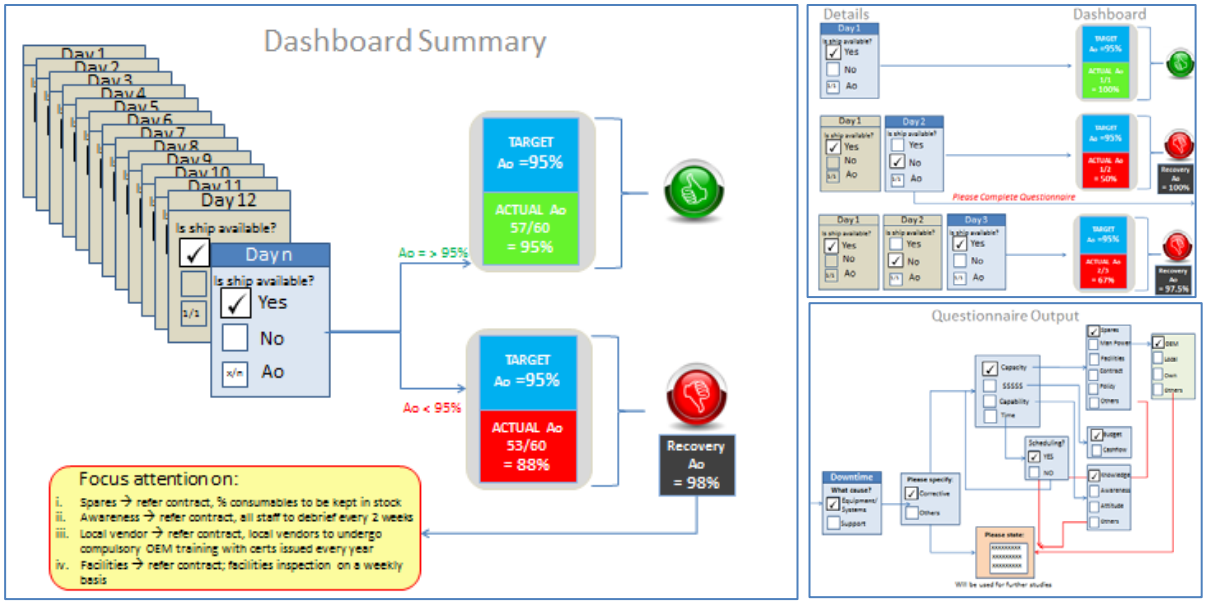


Figure 25: Dashboard input and output screen examples.

3. CONCLUSION & RECOMMENDATION FOR FUTURE STUDY

The ultimate goal of the authors in this study was to explore the possibility of producing a system or mechanism that could assist in improving Naval Ship Operational Availability, even though the authors realize that the situation was very complex in nature. Due to the limited research available previously, improvement efforts could not be placed precisely in tackling issues involving human and equipment related factors impacting ship availability. The challenging journey began when the authors realized that the "huge step" towards demystifying the complex naval issue in improving ship availability begins with "a tiny step" in identifying the factors impacting the ship availability. As elaborated above, the authors continued to broaden the horizon on available knowledge by progressively evolving through the "ConCaMS Development Spiral" achieving various levels of progress on each of the 13 Key Objectives researched in this paper. The ConCaMS Spiral with the labelled location of the various objectives would further assist policymakers and stakeholders of various organisations to develop their respective action-plans.

The authors have exhaustively researched and screened though more than 700 literatures of possible factors affecting maintenance from various engineering disciplines during this study and found that there has not been any discovery of a "one-size fits all solution" towards this complex naval ship availability issue. This research therefore provides a valuable contribution to the body of knowledge towards improving Naval Ship Availability. Nevertheless, due to the time, resources and financial constraint involved in this exploratory but highly specialized research in naval ship maintenance that spanned over 5 years, and in order for the results to remain current for the partial fulfilment of the Doctorate in Mechanical Engineering, the authors have concluded the research by paving the way for more focused future research in all of the areas covered in the 13 Objectives described in this paper.

One of the major obstacles of the research is the necessity to implement an acceptable verification and validation methodology for the "immense" number of variables concerning a complex asset such as the naval ships. This is a tremendous task when only limited data have been collected by various organisations. The same issue on verification and validation applies for the findings by researchers on

recommended amendments to contract clauses, as the time-consuming Delphi Methodology would not work as the availability of the Experts could not be guaranteed for an extended amount of time. Future studies could also utilise other methods including Operational Research tools and techniques so that results could be used to supplement or confirm the findings of the current research

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of this work by the Ministry of Higher Education Malaysia through the MyPhD scholarship for tuition fee, UTM and UPNM for academic guidance, STRIDE for technical guidance, as well as BHIC and RMN for access to data and relevant personnel in enabling this study in partial fulfilment of a Doctorate in Mechanical Engineering.

REFERENCES

Adler, P.A. & Adler, P. (2011). *The Tender Cut: Inside the Hidden World of Self-Injury*, NYU Press, New York.

Adler, M. & Ziglio E. (1996). *Gazing into the Oracle - The Delphi Method and its Application to Social Policy and Public Health*. Jessica Kingsley Publishers.

Al-Shafiq, B.A., Mohd Zamani, A., Sunarsih, Mohd Najib, A.G., Ubaidah, M.A., Abdullah, A.B., Nurhanani, A.A. (2017a). Measuring severity of downtime influence factors to naval ship operational availability – a Delphi study. *2017 Asia Int. Multidiscipl. Conf. (AIMC 2017)*, 1-2 May 2017, Skudai, Johor Bahru, Malaysia.

Al-Shafiq, B.A, Mohd Zamani, A., Sunarsih, Mohd Najib, A.G., Ubaidah, M.A., Abdullah, A.B., Nurhanani, A.A (2017b). Severity of downtime influence factors impacting naval ship operational availability – A five stage Delphi consensus procedure with snowballing technique. *7th IEEE Int. Conf. Control Syst., Comput. Eng. (ICCSCE 2017)*, 24-26 November 2017, Penang, Malaysia.

Al-Shafiq, B.A, Mohd Zamani, A., Sunarsih, Mohd Najib, A.G., Ubaidah, M.A., Abdullah, A.B. & Nurhanani, A.A (2017c). Impact of severe downtime influence factors on operational availability of naval ships – From the contract and project management perspectives. *5th Int. Conf. Exhib. Sustainable Ener. Adv. Mater. (ICE-SEAM 2017)*, 16-19 October 2017, Melaka, Malaysia.

Al-Shafiq, B.A, Mohd Zamani, A., Sunarsih, Mohd Najib, A.G., Ubaidah, M.A., Abdullah, A.B., Nurhanani, A.A (2017d). Development of a downtime influence factor severity index for improvement of naval ship availability: a simple approach for the Malaysian patrol vessel in-service support contract. *7th IEEE Int. Conf. Control Syst. Comput. Eng. (ICCSCE 2017)*, 24-26 November 2017, Penang, Malaysia.

Baker, S.E. & Edwards, R. (2012). *How Many Qualitative Interviews Is Enough? Expert Voices and Early Career Reflections on Sampling and Cases in Qualitative Research*. National Centre for Research Methods (NCRM), Southampton, UK.

Barringer, H.P. (1997). Life cycle cost & reliability for process equipment. *8th Ann. Ener. Week Conf. Exhib.*. American Society of Mechanical Engineering and PennWell Publishing, Tulsa, Oklahoma.

Blanchard, B.S. & Fabrycky, W.J. (1998). *Systems Engineering and Analysis*, Prentice Hall, Upper Saddle River, New Jersey.

Canadian Navy (2012). Available online at: http://publications.gc.ca/collections/collection-2012/dn-nd/db3-22-2001-eng.pdf. (Last access date: 27 January 2017).

Darnall, R.W. & Preston, J.M (2010). *Project Management from Simple to Complex*. The Open University of Hong Kong, Hong Kong.

Dattaa, P.P. & Roy, R. (2010). Cost modelling techniques for availability type service support contracts: A literature review and empirical study. *CIRP J. Manuf. Sci. Tech.*, **3**: 142-157.

Dattaa, P.P. & Roy, R. (2009). Cost modelling techniques for availability type service support contracts: A literature review and empirical study. *CIRP Indusr. Prod. Service Syst. (IPS2) Conf.*, 1-2 April 2009, Cranfield, UK.

Dell'Isola, A. & Vendittelli, A. (2015). Operational availability (Ao) of warships: A complex problem from concept to in service phase. *IEEE Metrology for Aerospace (MetroAeroSpace)*, 26-32.

Directorate of Maritime Strategy, Canada. (2001). *Leadmark: The Navy's Strategy for 2020*. Ottawa, Ontario, Canada.

Donkelaar, A.V. (2017). *Improving the Operational Availability of the Ships of the Royal Netherlands Navy*. Master's Thesis, Delft University of Technology, Delft, Netherlands.

Ford, G., McMahon, C. & Rowley, C (2013). Naval Surface Ship In-service Information Exploitation. *Procedia CIRP, 2nd Int. Through-Life Eng. Services Conf.*, Elsevier, pp. 92-98.

Franklin, K.K. & Hart, J.K. (2007). Idea generation and exploration: Benefits and limitations of the policy Delphi research method. *Innovative Higher Educ.*, **31**: 237–246.

GAO (1982). *Factors Limiting the Availability of F-15 Aircraft at the 1st Tactical Fighter Wing*. US General Accounting Office (GAO), Washington DC.

GAO (2015), *U. S. Navy Force Structure: Sustainable Plan and Comprehensive Assessment Needed to Mitigate Long-Term Risks to Ships Assigned to Overseas Homeports*. US General Accounting Office (GAO), Washington DC.

Henry, R. & Bil, C. (2015). Sustainment management in the Royal Australian Navy. *Proc. 22nd ISPE Inc. Int. Conf. Concurrent Eng.*, IOS Press, 249-254.

Inozu B. (1996). *Reliability, Availability and Maintainability (RAM) Database of Ship Operations Cooperative Program*. PN 1996. Research Project No. 95-18.

Lazakis, I., Turan, O., & Aksu, S. (2010). Increasing ship operational reliability through the implementation of a holistic maintenance management strategy. *Ships Offshore Struct.*, **5**: 337–357.

Marais, K.B., Rivas, J., Tetzloff, I.J. & Crossley, W.A. (2013). Modeling the impact of maintenance on naval fleet total ownership cost. *Syst. Conf. 2013 (SysCon 2013), IEEE Int.*, pp. 801-808.

McGrath, J.E. (1984). *Groups: Interaction and Performance*. Prentice-Hall, New Jersey.

Na,H., Yi, L.,Wang, Y., Liu, J., Bo, Z. & Lv, X. (2012). Research on the mean logistic delay time of the development phrass. *Physics Procedia Int. Conf. on Med. Physics Biomed. Eng. 2012 (ICMPBE 2012)*, **33**: 375-379.

Olivier J.P., Balestrini-Robinson, S. & Briceño, S. (2014). Approach to capability-based system-of-systems framework in support of naval ship design. *8th Ann. IEEE Syst. Conf.*,pp. 388-395.

Olivier J. P., Balestrini-Robinson, S. & Briceño, S. (2012). Ship cost-capability analysis using probabilistic cost modeling and hierarchical functional decomposition methodologies. *11th Int. Naval Eng. Conf. Exhib. (INEC)*, Institute of Marine Engineering, Science and Technology, London, UK.

Paik, S. A. (2014). Study on the case study and evaluation methodology of operational availability for a naval ship using OT&E data. *J. Korea Inst. Mil. Sci. Tech.*, **17**:471-478.

Pascual, R., Meruane, V. & Rey, P.A. (2006). On the effect of downtime costs and budget constraint on preventive and replacement policies. *Reliability Eng. Syst. Safety*, **93**: 144-151.

Ristic, D. (2013). A tool for risk assessment. *J. Safety Eng.*, 3, **3** :121-127.

RMN (Royal Malaysian Navy) (2011). *RMN Patrol Vessel - In-Service Support (ISS) Contract*. Royal Malaysian Navy (RMN), Malaysia.

SIA (2018). *The Collins Class: Submarine Institute of Australia*. Available online at: http://www.submarine institute.com/submarines-in-australia/the-Collins-Class.html. (Last access date: 2 February 2018).

Skulmoski, G.J., Hartman, F.T. & Krahn, J. (2007). The Delphi method for Graduate Research. *J. Infor. Tech. Educ.*, **6**: 1-21.

Tomkins, C. (2012). *Transforming Warship Support: Class Output Management*. RUSI Defence Systems, London, UK.

Xiao-Feng, D., Yu-Jiong, Gu, & Kun, Y. (2008). Study on intelligent maintenance decision support system using for power plant equipment, *IEEE Int. Conf. Auto. Log.*, 1-3 September 2008, Qingdao, China, pp. 96-100.