DEFENCE S&T TECHNICAL BULLETIN

VOL. 10 NUM. 1 YEAR 2017 ISSN 1985-6571

CONTENTS

STRIDE

SCIENCE & TECHNOLOGY RESEARCH INSTITUTE FOR DEFENCE

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. 10, Num. 2) is expected to be published in November 2017. The due date for submissions is 22 August 2017. **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.

APPLICATION OF OPTICAL TECHNIQUES TO DETECT CHEMICAL AND BIOLOGICAL AGENTS

Gaudio Pasqualino[1,2,*], Gelfusa Michela[1], Murari Andrea[3], Pizzoferrato Roberto[1], Carestia Mariachiara[1,2], Cenciarelli Orlando[1,2], Parracino Stefano[1], Ludovici Gianmarco[2], Gabriele Jessica[2], Gabbarini Valentina[2], Di Giovanni Daniele[1,2], Rossi Riccardo[1], Jean Francois Ciparisse[1], CarloBellecci[2,4]&Andrea Malizia[5]

[1]Department of Industrial Engineering, University of Rome Tor Vergata, Italy
[2]International Master Courses in Protection Against CBRNe Events, Department of Industrial
[3]Consorzio RFX-Associazione EURATOM ENEA per la Fusione, Italy
Engineering and School of Medicine and Surgery, Via del Politecnico 1, Italy
[4]CRATIScrl - Consorzio per la Ricerca e le Applicazioni di Tecnologie Innovative, Università degli Studi della Calabria, Italy
[5]Department of Biomedicine and Prevention, University of Rome Tor Vergata, Italy

[*]Email: gaudio@ing.uniroma2.it

ABSTRACT

Chemical or biological contamination, due to natural or man-made disasters, represents a severe concern for safety and security of people, and of the environment. Chemical agents (CAs) and biological agents (BAs) are commonly used for a number of civilian and military applications, and can be deployed as a weapon with terroristic purposes. Therefore, it is necessary to develop specific systems aimed at preventing or reducing the consequences of the spread of these agents. To this end, the authors have developed optical systems to detect (LIDAR) and identify (DIAL) CAs, and to detect BAs (fluorescence technique). These systems and technologies will be presented in this work together with the analysis and discussion of the results obtained.

Keywords:*Light detection and ranging (LIDAR);differential absorption of light (DIAL); fluorescence; chemical agents (CAs);biological agents (BAs).*

1. INTRODUCTION

Nowadays the perception of risks has completely changed as compared to that in the past. The geographical boundaries are thin, so too the geo-political ones. The social boundaries have totally disappeared leaving space to a world with different roles as compared to the past. Globalization and the simplicity to access information on the one hand has improved technologies and quality of life, but on the other hand, has created more opportunities to create terror. The cases of the dispersion of toxic gases in the past and in recent history (Okumura *et al*., 1998; Singer *et al*., 2005; Varney *et al*., 2006), the DAESH progression in the Arabic countries (Hansen-Lewis & Shapiro, 2015; Fromson & Simon, 2015; Akbarzadeh, 2015) connected to the actual risk of toxin dispersion in water, Ebola diffusion through aerial transportation (Cenciarelli *et al.*, 2014, 2015a,b), and radioactive dispersion due to failures in fission or fusion devices (Gallo *et al*., 2012; Cacciotti *et al.,* 2014; Malizia *et al.,* 2014; Koninhs *et al.,* 2015; Malins *et al.*, 2016) are all examples of how the new merge between the well-known risks together with the new "dangerous ideas" have exponentially increased the uncertainties of human health all around the world.

The University of Rome Tor Vergata, in collaboration with the Faculty of Engineering and Faculty of Medicine and Surgery, is facing these new problems with innovative working

approaches devoted to the use the expertise developed in classic scientific disciplines in an unconventional way. Currently several technologies have been investigated in order to develop systems based on optical sensing of chemical agents (CAs) and biological agents (BAs). BA and CA contamination are important not only for security reasons but also for safety issues. As a matter of fact, chemical contamination in actuality is a real menace, with the DAESH scaring the world by threatening the contamination of water. Biological epidemic spreads are relatively easy thanks to the worldwide aerial connection. The risks of chemical or biological contaminations can also be found in hospitals or in private factories and it is necessary to guarantee, in all these areas, a continuum monitoring to reduce the probabilities and the magnitude of those events.

The market is full of products that are able to detect and identify the CAs for both chemical warfare agents (CWA) and toxic industrial components or toxic industrial materials (TIC/TIM), and to identify BAs (such as biological warfare agents or similar), but these technologies work only in short range (chemical) or in the laboratory, after automatic or manual sample collection (biological). The current commercialized tools are only able to detect the putative variation of the environmental biological or chemical background in a defined confined area, providing an on/off response type. In this work, the authors go over simple detection though optical techniques, demonstrating how they are working to have a quasi-real time detection and/or identification of BAs and CAs that is fundamental during operation in a contaminated scenario, such as one characterized by an accidents, or terroristic or war events.

2. MATERIALS AND METHODS

This section is devoted to describing the experimental set-up (together with the methods) developed and used by the authors to detect (LIDAR) and identify (DIAL) CAs, and to detect BAs (fluorescence technique).

2.1 Light Detection and Ranging (LIDAR)

The measurements described in the paper were performed with the LIDAR system developed by the Quantum Electronic Plasma Physics (QEP) research group at the University of Rome Tor Vergata. Furthermore, a ground-based mini LIDAR (Figure 1) station has been built and continuously upgraded at the CRATI s.c.r.l. c/o University of Calabria. The measurements have been performed with a commercial Nd:Yag (neodymium-doped yttrium aluminium garnet; Nd:Y3Al5O12) laser operating in so-called Q-Switching mode. Q-switching, sometimes known as giant pulse formation or Q-spoiling, is a technique by which a laser can be made to produce a pulsed output beam. This technique allows for the production of light pulses with extremely high peak power, much higher than would be produced by the same laser if it were operating in a continuous wave (constant output) mode. Wavelength of 1,064 nm was used during the experimental campaign. This choice satisfies the need for developing a compact system that is robust enough to guarantee continuous operations. This technology has also become relatively standard and thus, it can be procured at reasonable costs. The receiver system is composed of an avalanche photodiode and a telescope, based on the Newtonian configuration. The entire apparatus is installed on a mobile system and thus, constitutes a fully mobile station, which can be used for in field campaigns in any location. The experimental campaign was conducted at Crotone, in the Calabria region in the south of Italy (Bellecci *et al.,* 2009, 2010; Gaudio *et al.,* 2013a, 2014, Gelfusa *et al.,* 2015, Parracino *et al.,* 2016).

Using the LIDAR system, it is possible to detect anomalies in terms of variation of the normal backscattered LIDAR signal sent into the atmosphere through the application of Equation1, which is the LIDAR equation, and Equation 2, which is the Klett inversion (Parracino *et al.*, 2016).

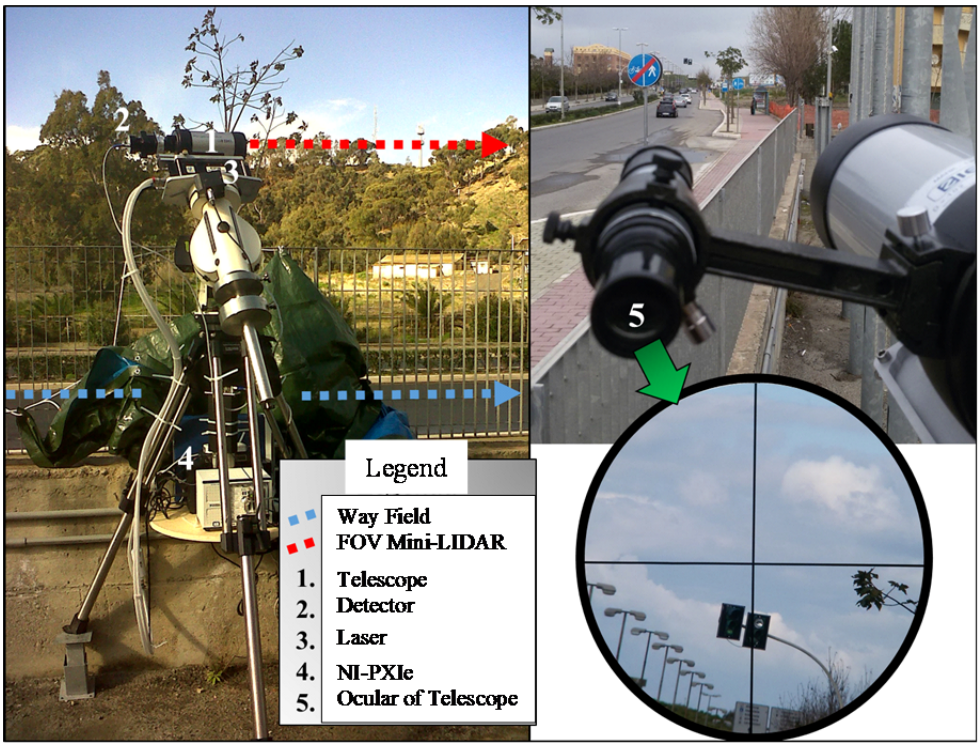


Figure 1:The developed mini-LIDAR system.

$$P_r(\lambda,R) = E_1 \frac{c}{2} \frac{A_r}{R^2} \beta(\lambda,R) C(\lambda,R) \exp\left(-2\int_0^R K(\lambda,R')dR'\right) \quad (1)$$

The LIDAR equation is the way to relate the light power backscattered by the atmospheric target with the signal collected by the LIDAR receiver. In the equation,$P_r(\lambda,R)$ is the backscattered power received from the distance R, at the specific laser operative wavelength λ (because the equation describes the elastic scattering);E_1 is the laser pulse energy, c is the speed of light;$A_r \cdot R^{-2}$ is the acceptance solid angle of the receiver optics with collecting area A_r; and $C(\lambda,R)$ is a function determined by the geometric considerations of the receiver optics, the transmitter (set at 100%) and receiver efficiencies, and the overlap of the emitted laser beam with the field of view of the receiver. The last two terms are known as the optical parameters and are closely linked to the aerosols load. The backscattering coefficient $\beta(\lambda, R)$ is a measure of the scattering in the backward direction (i.e., towards the incident direction, at a scattering angle of 180°) for the light encountering the atmospheric aerosol particles and molecules. The extinction coefficient $K(\lambda, R)$ is a measure of attenuation of the light passing through the atmosphere due to the scattering and absorption by aerosol particles and molecules. Since it is reasonable to assume that molecular backscattering is negligible in the infrared region, and that particulate backscattering dominates in this region, the LIDAR signal gives particle backscatter information directly if laser is operating in this region. By assuming a relationship between relative variation of backscattering coefficient and particulate mass or number density, the LIDAR signal can be used to monitor and track atmospheric particles (Parracino *et al*., 2016).

From Equation 1, and using the Klett inversion method (Klett, 1981), it is possible to extract both the extinction and backscattering coefficients. In this paper, the attention is focused on the particle backscattering coefficient, the primary atmospheric parameter that determines the strength of the LIDAR signal. The far boundary solution is reported into the following equation:

$$\beta(R,\lambda) = \frac{e^{[\mu(R,\lambda)-\mu_m(\lambda)]}}{\frac{1}{\beta_m(\lambda)} + \frac{2}{P_\pi(R,\lambda)}\int_R^{R_m} e^{[\mu(R',\lambda)-\mu_m(\lambda)]}dR'} \quad (2)$$

2.2 Differential Absorption LIDAR (DIAL)

The QEP research group developed two dial carbon dioxide (CO_2) based laser systems. The first one is mounted on a truck. Today, it is used in experimental campaigns all around Italy for pollutants identification in the atmosphere (Gaudio *et al.,* 2011, 2013b; Gelfusa *et al*., 2014a, 2014b). After that, a mini-CO_2 system (Figure 2) was developed to obtain remote, stand-off identification of CWAs, TIMs and TICs. Both the systems work with the DIAL techniques that provide the capability to remotely measure the concentration and spatial distribution of compounds in the atmosphere. The ability to scan the optical measurement beam throughout the atmosphere enables pollutant concentrations to be mapped and emission fluxes to be determined.

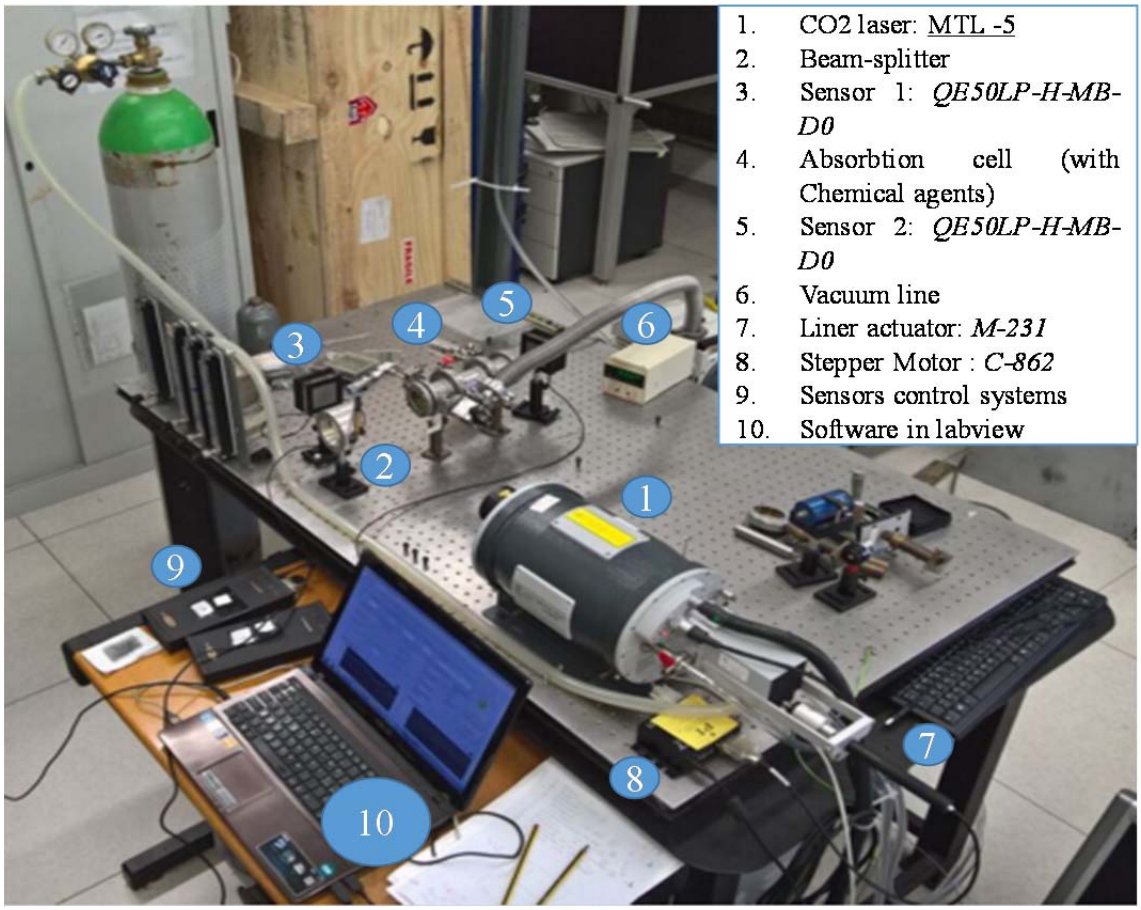


Figure 2: The developed mini-DIAL system.

The experimental set-up is composed of a mini MTL-5 CO_2 TEA laser (1) (a manual tuned laser that usesCO_2 using transverse excited atmospheric pressure) that emits laser beams at different wavelengths obtained by a linear actuator (7) controlled by a stepper motor (8) on a beam-splitter (2) that divides the beam into two beams, one directed to energy meter 1 (3) and one that passes through an absorption cell (4) that contains the CA under investigation. This second beam is partially absorbed by the CA inside (4) and then goes to energy meter 2 (5). Both the sensors are connected to an analog to digital converter (ADC) board (9) to acquire the measurements. The absorption cell is connected to a vacuum line (6) to get different pressure conditions.

2.3 Fluorescence Apparatus

The QEP research group, in collaboration with the Department of Industrial Engineering, is working on an experimental set-up (Figure 3) based on fluorescence techniques to detect BAs (Carestia *et al.*, 2014a, 2014b, 2015a, 2015b). Steady-state photoluminescence measurements on liquid samples in an optical cuvette were performed in a standard laboratory set-up with a photomultiplier optical detector located at the output port of a 25-cm monochromator. The excitation light beam was provided by the monochromatized output of a Hg(Xe) 200-W discharge lamp (that is A Mercury Xenon lamp). All measurements were performed at temperature of 20 ± 2°C. The two wavelengths selected for the excitation of the samples were 355 and 266 nm, referring to the third and fourth harmonics of a Nd:YAG laser source respectively. The experimental set-up is shown in Figure 3. All the components of the experimental apparatus are reported in Table 1. A schematic view of the experimental apparatus is shown in Figure 4.

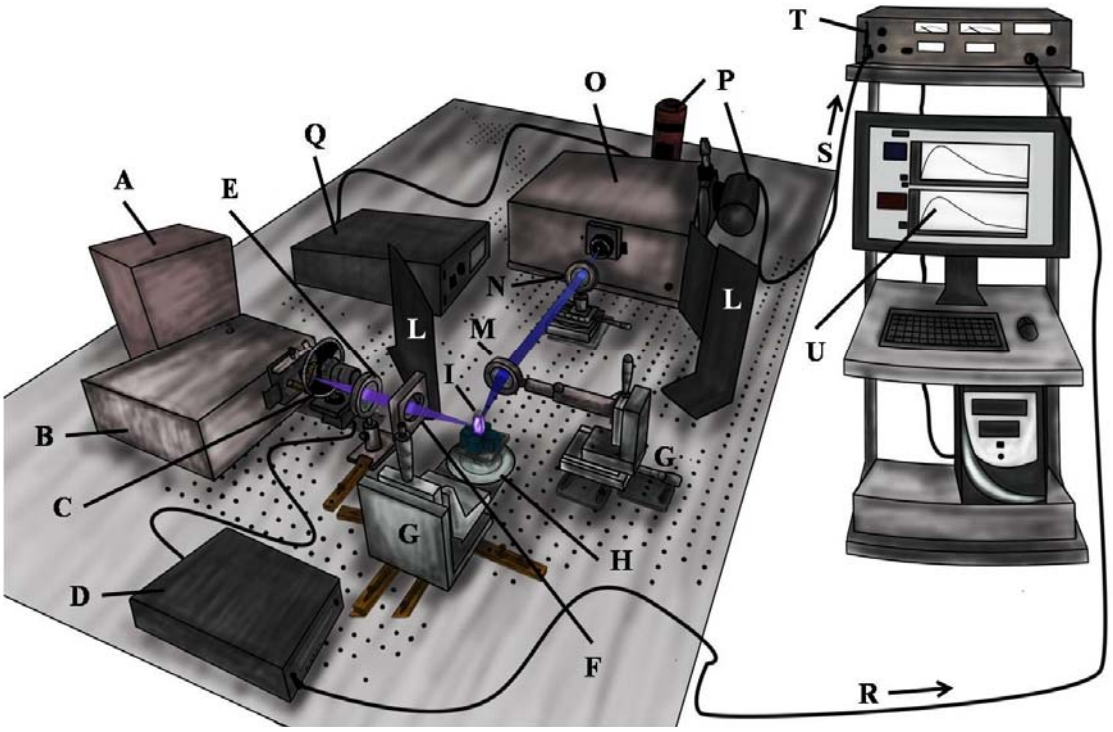


Figure 3: Experimental set-up of the fluorescence technique.

Table 1:Components of the experimental set-up.

A	Hg(Xe) 200 W discharge lamp	**M**	Collecting collimating lens
B	Excitation monochromator	**N**	Collecting focusing lens
C	Excitation modulator (chopper)	**O**	Emission monochromator
D	Chopper power supply	**P**	Photomultiplier
E	Excitation collimating lens	**Q**	Emission monochromator power supply
F	Excitation focusing lens	**R**	Modulation reference
G	Positioning	**S**	Signal
H	Sample holder	**T**	Lock in amplifier
I	Cuvette	**U**	Software analysis bench
L	Light shield		

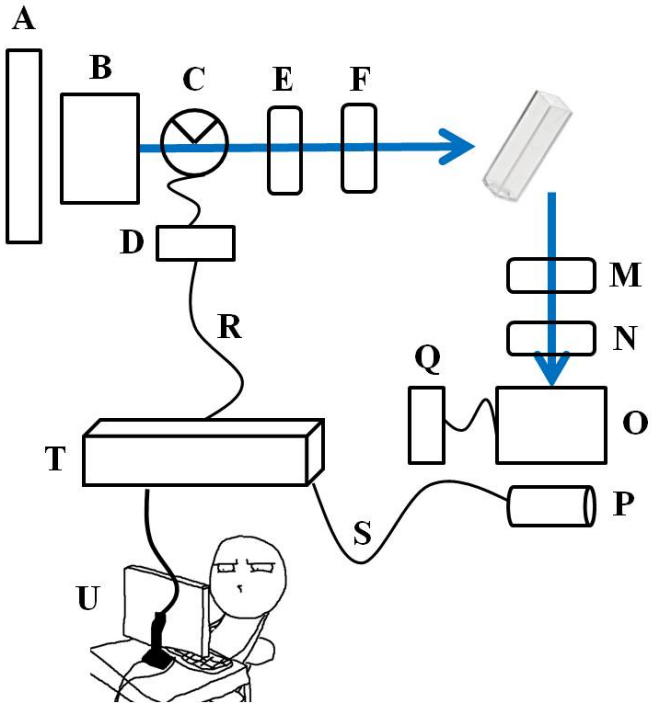


Figure 4. Schematic view of the experimental set-up.

The light emitted by the lamp (A) is focused on the input slit of an excitation monochromator (B), which selects the required wavelength. The monochromatic beam passes through an amplitude modulator (C) and through two lenses, collimating (E) and focusing (F); then it strikes the sample (I). The wavelength intensity distribution of the emitted light (emission spectrum) that passes through two lenses, collimating (M) and focusing (N), is analyzed by the emission monochromator (O). The light intensity is transduced to an electric signal by the photomultiplier (P). The emission monochromator is also necessary to avoid exciting radiation from reaching the detector. This can happen despite the emitted radiation being measured perpendicular to the exciting radiation, due to the scattering phenomena, in particular due to the Rayleigh and Tyndall scattering. The light scattering, as a result of the interaction of the light with the matter, can be defined as elastic (Rayleigh or Tyndall scattering) or inelastic (Raman scattering). In the first case, the scattered light has the same frequency of the incident beam, while in the case of inelastic scattering, the frequencies are different and are present in the sample spectrum. The Rayleigh scattering occurs when the radiation passes through a

transparent medium that illuminates particles with a diameter much lower (less than 1/10) than its wavelength. The scattered light has the same wavelength of the incident light and an intensity proportional to the fourth power of the frequency of the incident light. Tyndall scattering occurs in the presence of colloidal systems in which the particles have a diameter comparable with the wavelength of visible light. The intensity of the scattered light is proportional to the second power of the frequency of the incident radiation. Raman scattering, which is solvent-dependent, occurs when only a very small fraction of the light that is incident on a transparent medium is diffused. (Carestia *et al.,* 2015b)

3. EXPERIMENTS AND RESULTS

This section is devoted to describe the experimental results obtained with the experimental set-ups previously discussed.

3.1 LIDAR

The variations in atmosphere detected with the COmpact LIDAR system (COLI) can be resolved both in time (Figure 5) or space (Figure 6), and visualized through backscattering coefficient maps that essentially give the exact time and position of the detected anomaly. The maximum range reached by the system is 1.5 km, which is quite common. However for reasons of safety and due to the constraints of the experimental field/investigated area, the far range boundary point of LIDAR system was reduced. For these reasons, the two backscattering maps reported in the following figures have been cut from 0 to 500-600m. In spite of these limitations, this improves the accuracy of experimental results, increasing the signal to noise ratio (SNR) level.

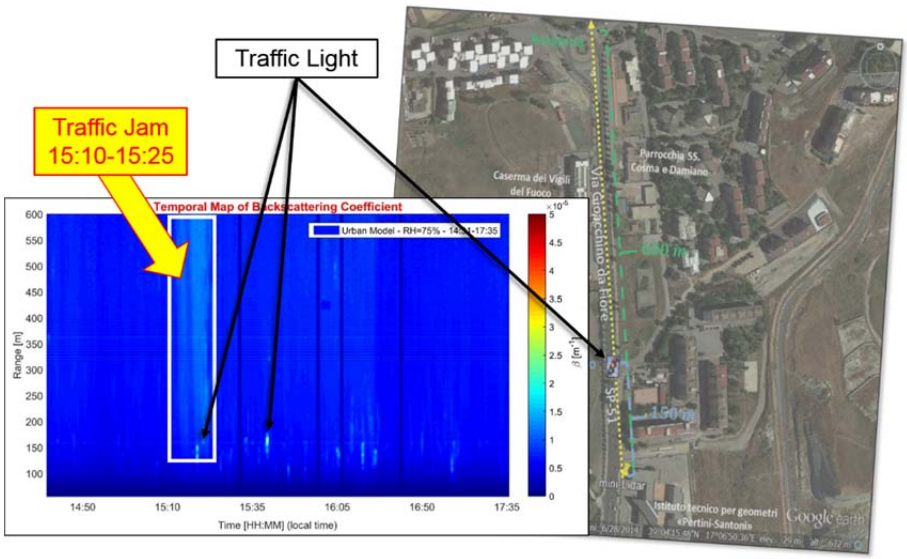


Figure 5: The COLI signal's map resolved in time.

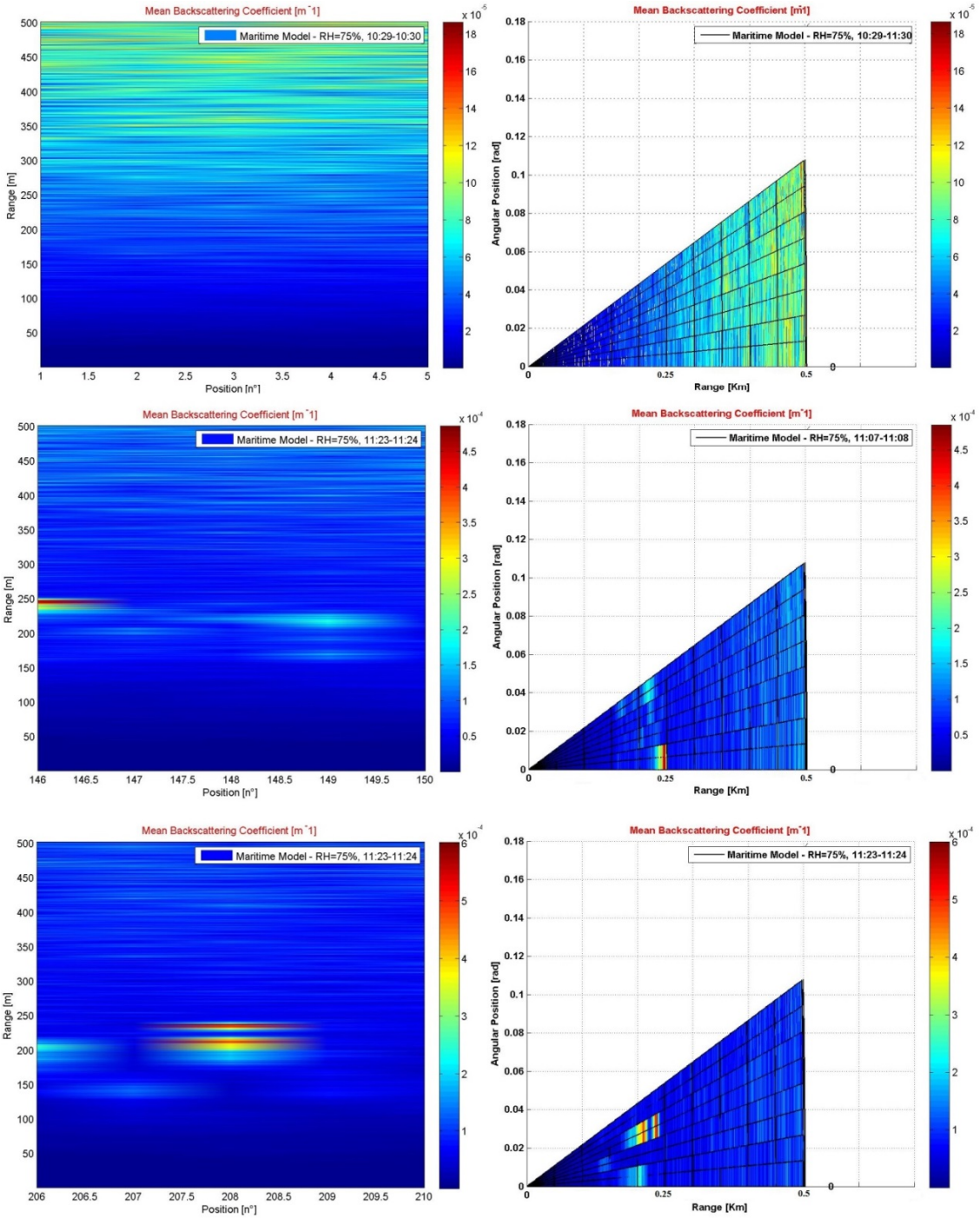


Figure 6: Backscattered maps resolved in space

3.2 DIAL

The CO_2 DIAL systems are used to obtain the absorption spectra, which are the characteristic of each CA, in order to obtain a standoff medium range (0-1 km) identification that can help the decision makers to understand the typology of contamination present in an open or closed environment. An experimental campaign, dedicated to test the DIAL system, was conducted in order to demonstrate the capability of this system to identify CAs, in order to use it for CWA, TIC and TIM dispersion problems. In this experimental campaign, QEP used mainly TIC, some of them are official recognized as CWA simulants. The capability to identify chemicals with the

optical system developed is demonstrated in this section. The experiments were performed not only to demonstrate the capability to identify chemicals through the optical apparatus developed but also to create a database of the absorption spectra of the chemicals to guarantee use on the field with a fast comparison between the spectra acquired and the one on the database.

The toxic component of a chemical weapon is called its CA. Based on their mode of action (i.e. the route of penetration and their effect on the human body), chemical agents are commonly divided into several categories (OPCW, 2017):

- Blister agents
- Blood agents
- Nerve agents
- Riot control agents
- Potential CW Agents
- Mustard agents
- Psychotomimetic Agents
- Toxins

All the experimental measures were conducted in two different conditions:

- Internal cell pressure of 50 mbar (high vacuum): This is a condition with a quasi-unique presence of analyte molecules.
- Internal cell pressure of 1,000 mbar (atmospheric pressure): in this condition, the air has little amount (in percentage) of analyte molecules (JRAMC, 2002).

The composition of air is variable in general depending on the considered quantity of analyte. For a fixed quantity, the ratio between the amount of nitrogen and the amount of oxygen contained in the air remains almost constant thanks to the balance between the consumption and continuous intake of these elements associated with the oxygen cycle and the cycle of nitrogen. Instead, the water vapor and carbon dioxide concentrations are variable. For this reason, they often indicate the properties of the air without water vapor (which is called "dry air"), otherwise it is called "moist air." Dry air at the ground level is composed of approximately 78% nitrogen (N_2), 21% oxygen (O_2), 0.96% argon (Ar) and 0.04% carbon dioxide (CO_2), plus other components in minor amounts, including solid particles in suspension, which constitute the so-called "atmospheric dust." Moist air can contain up to 7% of the volume composed by the water vapor; the percentage of water vapor in the air corresponds to the relative humidity of air and temperature. The maximum value of water vapor that the air can hold at equilibrium corresponds to the saturation conditions; above this maximum value, the water vapor tends to condense spontaneously. This maximum value is related to the temperature changing from values close to 0% (to a temperature equal to -40 ° C), about 0.5% (0 °C), up to 5-7% (around 40 ° C) (JRAMC, 2002).

It was experimentally verified that the variations of the absorption spectra between cell under vacuum and at atmospheric pressure cell (both with analyte) are not relevant. The differences between spectra in dry and moist air conditions are also negligible. In addition, the temperature variations also do not give considerable variation to the measuring spectra if we are in the temperature variation range present in nature.

The absorption spectra are all prepared with pure substance in the cell (black graph in Figure 7) and with the combination of substance+air (red graph in Figure 7) using all the wavelengths available with MTL-5. Figure 7 shows the absorption spectra of acetone that has no specific absorption values with the exception of short wavelengths, in particular 9,225 µm both for vacuum and air conditions and a peak at 9,425 µm just in vacuum conditions, which is probably an error during the acquisition.

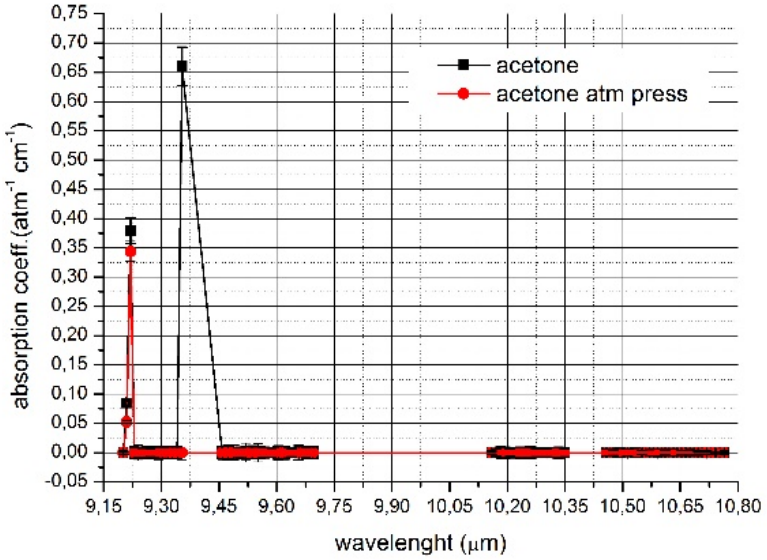


Figure 7:Acetone absorption spectra

3.3 Fluorescence Measurements of BAs

The capability to obtain fluorescence spectra (Figure 8) was investigated by the authors using bacillus spores samples (washed, unwashed and supernatant). These results clearly show the potentiality of this technique merged with statistical methods, such as the universal multi-event Locator (UMEL), a statistical tool based on support vector regression, to improve the ability to discriminate different BAs' emission spectra, to achieve early detection of BAs. It can be used to detect the contamination of surfaces, in liquid samples or in the atmosphere. The experimental and data analysis methods have to be improved in order to reduce the false positives, increase the field of view (FOV) of the system, and to speed up and automate the detection process.

4. CONCLUSION AND FUTURE DEVELOPMENTS

The authors in the last fifteen years have demonstrated that laser-based optical techniques can enable continuous and quasi real-time monitoring of CAs (CWA, TIC, TIM) in the air at medium-long distances. In particular, the capability of the LIDAR systems to detect CAs for the early warning phase and of the DIAL systems to identify at short-medium range certain CAs for the decision making phase has been demonstrated. Using the same approach (optical techniques for standoff detection) but changing the method (fluorescence), the authors assessed its potential for BA detection. The challenges for the future are to optimize the systems by reducing the dimensions and cost, while at the same time improving the performances in terms of reduction of false alarms, increase of robustness and acquisition speed, and increase reliability and determinism.

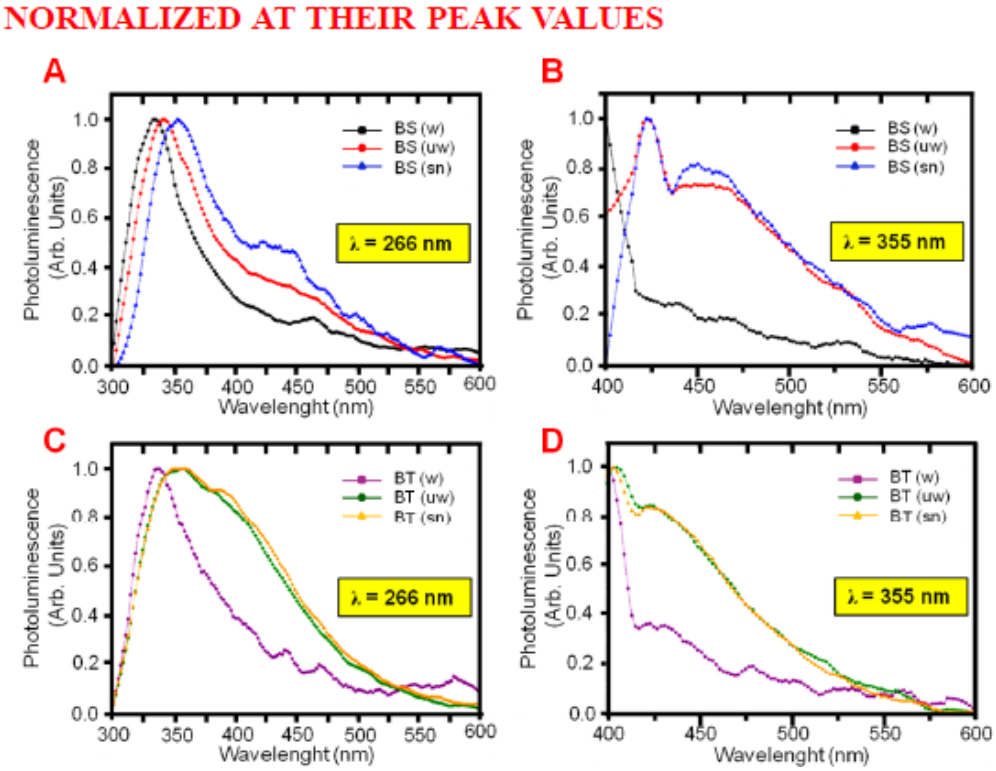


Figure 8: Normalized absorption spectra: *Bacillus subtilis* (A,B); *Bacillus thuringiensis* (C,D).

REFERENCES

Akbarzadeh, S. (2015). Iran and Daesh: The case of a reluctant Shia Power. *Middle East Policy*, **22:**44-54.

Bellecci, C., Gaudio, P., Gelfusa, M., Lo Feudo, T., Malizia, A., Richetta, M. &Ventura, P. (2009). Raman water vapour concentration measurements for reduction of false alarms in forest fire detection. *Proc. SPIE*,**7479**.

Bellecci, C., Gaudio, P., Gelfusa, M., Malizia, A., Richetta, M., Serafini, C. &Ventura, P. (2010).Planetary Boundary Layer (PBL) monitoring by means of two laser radar systems: Experimental results and comparison. *Proc. SPIE*, **7832**.

Cacciotti, I., Aspetti, P.C., Cenciarelli, O., Carestia, M.. Di Giovanni, D., Malizia, A., D'amico, F., Sassolini, A., Bellecci, C. &Gaudio, P. (2014). Simulation of Caesium-137 (137CS) local diffusion as a consequence of the Chernobyl accident using HOTSPOT. *Defence S&TTech. Bull.*, **7:**18-26.

Carestia, M., Pizzoferrato, R., Cenciarelli, O., D'Amico, F., Malizia, A., Gelfusa, M., Scarpellini, D. & Gaudio, P. (2014a). Fluorescence measurements for the identification of biological agents features for the construction of a spectra database. *2014 Fotonica AEIT Ital. Conf. Photon. Technol.*, 12-14 May 2014, Naples, Italy.

Carestia, M., Pizzoferrato, R., Gelfusa, M., Cenciarelli, O., D'Amico, F., Malizia, A., Scarpellini, D., Murari, A., Vega, J. &Gaudio, P. (2014b). Towards the implementation of a spectral data base for the detection of biological warfare. *Proc. SPIE*, **9251**.

Carestia, M., Pizzoferrato, R., Gelfusa, M., Cenciarelli, O., Ludovici, G.M., Gabriele, J., Malizia, A., Murari, A., Vega, J. &Gaudio P. (2015a).Development of a rapid method for the automatic classification of biological agents' fluorescence spectral signatures. *Opt. Eng.*, **54:**114105.

Carestia, M., Pizzoferrato, R., Lungaroni, M., Gabriele, J., Ludovici, G.M., Cenciarelli, O., Gelfusa, M., Murari, A., Malizia, A. &Gaudio P. (2015b).Multispectral analysis of

biological agents to implement a quick tool for stand-off biological detection. *Proc. SPIE*, **9652**.

Cenciarelli, O., Pietropaoli, S, Frusteri, L., Malizia, A., Carestia, M., D'Amico, F., Sassolini, A., Di Giovanni, D., Tamburrini, A.,Palombi, L., Bellecci, C. &Gaudio, P. (2014). Biological emergency management: the case of Ebola 2014 and the air transportation involvement. *J. Microb. Biochem. Technol.*, **6**:247-253.

Cenciarelli, O., Pietropaoli, S., Malizia, A., Carestia, M., D'Amico, F., Sassolini, A., Di Giovanni, D., Rea, S., Gabbarini, V., Tamburrini, A., Palombi, L., Bellecci, C. &Gaudio, P. (2015a).Ebola virus disease 2013-2014 outbreak in West Africa: An analysis of the epidemic spread and response. *Int. J. Microbiol.*, **2015:**769121.

Cenciarelli, O., Gabbarini, V., Pietropaoli, S., Malizia, A., Tamburrini, A., Ludovici G.M., Carestia M., Di Giovanni, D., Sassolini, A., Palombi, L., Bellecci, C. & Gaudio P. (2015b). Viral bioterrorism: Learning the lesson of Ebola virus in West Africa 2013-2015. *Virus Res.*, **210:**318-326.

Fromson, J. & Simon, S. (2015). ISIS: The dubious paradise of apocalypse now. *Survival,***57:**7-56.

JRAMC (Journal of the Royal Army Medical Corps) (2002). Toxic industrial chemicals. *J Royal Army Med Corps*,**148**: 371-381.

Klett J.D. (1981). Stable analytical inversion solution for processing lidar returns.*Appl. Optics,***20:** 211-220.

Gallo, R., De Angelis, P., Malizia, A., Conetta, F., Di Giovanni, D., Antonelli, L., Gallo, N., Fiduccia, A., D'Amico, F., Fiorito, R., Richetta, M., Bellecci, C., Gaudio, P. (2012). Development of a georeferencing software for radiological diffusion in order to improve the safety and security of first responders. *Defence S&T Tech. Bull.*, **6**:21-32.

Gaudio, P., Gelfusa, M., Lupelli, I., Malizia, A., Moretti, A., Richetta, M., Serafini, C. &Bellecci, C. (2011). First open field measurements with a portable CO2 lidar/ dial system for early forest fires detection.*Proc. SPIE*, **8182.**

Gaudio, P., Gelfusa, M., Malizia, A., Richetta, M., Serafini, C., Ventura, P., Bellecci, C., De Leo, L., Lo Feudo, T. &Murari, A. (2013a).New frontiers of forest fire protection: A portable laser system (FfED). *WSEAS T. Environ.Dev.*, **9:**195-205.

Gaudio, P., Gelfusa, M., Malizia, A., Richetta, M., Antonucci, A., Ventura, P., Murari, A. &Vega, J. (2013b).Design and development of a compact Lidar/Dial system for aerial surveillance of urban areas. *Proc. SPIE*, **8894.**

Gaudio, P., Gelfusa, M., Malizia, A., Parracino, S., Richetta, M., Murari, A. &Vega J. (2014). Automatic localization of backscattering events due to particulate in urban areas. *Proceedings of SPIE*,**9244.**

Gelfusa, M., Malizia, A., Parraciono, S., Richetta, M., Bellecci, C., Avolio, E., De Leo, L., Perrimezzi, C. &Gaudio P. (2014a). Detection of pollutant sources in the atmosphere with Lidar/Dial techniques: Results of an experimental campaign in the south of Italy. *2014 Fotonica AEIT Ital. Conf. Photon. Technol.*, 12-14 May 2014, Naples, Italy.

Gelfusa, M., Gaudio, P., Malizia, A., Murari, A., Vega, J., Richetta, M. & Gonzalez S.(2014b). UMEL: A new regression tool to identify measurement peaks in LIDAR/DIAL systems for environmental physics applications, *Rev. Sci.Instrum.,***85:**063112.

Gelfusa, M., Malizia, A., Murari, A., Parracino, S., Lungaroni, M., Peluso, E., Vega, J., De Leo, L., Perrimezzi, C &Gaudio P. (2015). First attempts at measuring widespread smoke with a mobile LIDAR system. *2015 Fotonica AEIT Ital. Conf. Photon. Technol.*, 6-8 May 2015, Turin, Italy.

Hansen-Lewis, J. & Shapiro, J.N. (2015). Understanding the Daesh economy. *Perspect.Terrorism*, **9**: 142-155.

Koninhs, R.J.M, Wiss T.& Benes O. (2015). Predicting material release during a nuclear reactor accident. *Nat. Mater.*, **14:**47–252.

Malins A., Kurikami H., Nakama S., Saito T., Okumura M., Machida M. &Kitamura A. (2016). Evaluation of ambient dose equivalent rates influenced by vertical and horizontal distribution of radioactive cesium in soil in Fukushima Prefecture, *J. Environ. Radioact.*, **151:**38-49.

Malizia A., Lupelli I., Richetta M. Gelfusa M., Bellecci C. and Gaudio P. (2014), Safety analysis in large volume vacuum systems like tokamak: experiments and numerical simulation to analyze vacuum ruptures consequences, *Adv. Mater. Sci. Eng.,***2014:**201831.

Okumura, T., Suzuki, K., Fukuda, A., Kohama, A., Takasu, N., Ishimatsu, S. & Hinohara S. (1998). The Tokyo subway sarin attack: Disaster management, Part 1: Community emergency response, *Acad.Emerg.Med.***5:**613-617.

OPCW (Organisation for the Prohibition of Chemical Weapons) (2017). *Types of Chemical Agent*. Available online at: https://www.opcw.org/about-chemical-weapons/types-of-chemical-agent (Last access date: 22 February 2017).

Parracino, S., Richetta, M., Gelfusa, M., Malizia, A., Bellecci, C., De Leo, L., Perrimezzi, C., Fin, A., Forin, M., Giappicucci, F., Grion, M., Marchese, G., Gaudio, P. (2016). Real-time vehicle emissions monitoring using a compact LiDAR system and conventional instruments: First results of an experimental campaign in a suburban area in southern Italy.View at Publisher. *Opt. Eng.*, **55:** 103107.

Singer, B.C., Hodgson, A.T., Destaillats, H., Hotchi, T., Revzan, K.L. & Sextro, R.G. (2005), Indoor sorption of surrogates for sarin and related nerve agents. *Environ. Sci. Technol.*, **39:**3203-3214.

Varney, S., Hirshon, J.M., Dischinger, P. & Mackenzie, C. (2006). Extending injury prevention methodology to chemical terrorism preparedness: the Haddon Matrix and sarin, *Am. J.DisasterMed.*, **1:**18-27.

HARDWARE EFFICIENT VEIN ENHANCEMENT AND FEATURE EXTRACTION METHOD

Suhaimi Yusoff[1,2,*] & Abdul Rahman Ramli[1]

[1]Department of Computer and Communication System Engineering, Universiti Putra Malaysia (UPM), Malaysia
[2]Electronics Technology Department, UniKL British Malaysian Institute, Malaysia

[*]Email: suhaimi.yusoff@student.upm.edu.my

ABSTRACT

In this paper, an improved method for vein enhancement and feature extraction is proposed. The proposed method is aimed to be implemented in hardware. It uses image resampling to reduce the number of pixels in the vein image, thus reducing the cost of computation. Then, Difference of Gaussian (DoG) and thresholding are used to segment the vein image. Finally, thinning is applied to get single line veins. The experimental results and analysis show that the proposed method speeds up the processing time and is effective to detect veins in images. The method also reduces the use of hardware resources and storage, making it suitable for hardware implementation.

Keywords: *Vein enhancement and feature extraction; image resampling; Difference of Gaussian (DoG); thresholding; thinning.*

1. INTRODUCTION

Nowadays, there are various biometric modalities such as fingerprint, iris, hand geometry, vein and many more. Each of these modalities has its own advantage and disadvantage. Veins have emerged as an important biometric modality because it uses features inside the human body rather than on the surface. This modality is very secure because veins exist inside the human body (NTSC, 2006). Vein biometric can be widely used in various fields such as security, banking, military, automotive and many more where authentication is needed (Iterson, 2004; Fujitsu Limited, 2006; Ennis, 2012).

Many researchers have developed numerous methods to enhance and extract veins for authentication purposes. A comprehensive review on vein enhancement and feature extraction methods was provided by Yusoff *et al.* (2015). Most researchers have focused on using software such as MATLAB or C language for this purpose, with many of these methods not being optimised for hardware implementation. Thus, hardware resources or processing time constraints was not a priority when developing vein enhancement detection methods. Implementation of vein enhancement methods in hardware is challenging due to resource constraints and the complexity of hardware.

Khalil-Hani & Eng (2010, 2011), Khalil-Hani & Lee (2013), Lee *et al*. (2012), Wei & Rosdi (2012) and Pudzs *et al.* (2013) used field programmable gate array (FPGA) to implement vein detection in hardware. FPGA is a device that is widely used for high computational resources due to its parallelism, high-speed processing and highly configurability (Contreras-Medina *et al*., 2010, 2012). Khalil-Hani & Eng (2010) proposed a finger vein biometric system that consists of image pre-processing, acquisition, feature extraction and matching. The image pre-processing includes colour to greyscale conversion, median filter, segmentation, alignment, resize, Gaussian low pass filter, local dynamic thresholding, binary median filter and thinning. The proposed method was implemented in an embedded system using NIOS II. Then, Khalil-Hani & Eng (2011) proposed a hardware accelerator to implement this image pre-processing, while feature extraction and template matching

was processed in a NIOS II embedded CPU. Pudzs *et al*. (2013) implemented a small part of the algorithm, which is a complex matched filter, in hardware using FPGA for palm vein biometric system. Sun *et al* (2011) used a high performance digital signal processing (DSP) processor to implement their algorithm for finger vein in hardware. Their proposed algorithm consists of average filter as noise removal, image segmentation using maximum curvature model, median filter and thinning. Wei & Rosdi (2012) have implemented a band-limited phase-only correlation matching method in hardware using FPGA for a finger vein biometric system.

The vein enhancement and feature extraction method from our previous work (Yusoff *et al*., 2009) is shown in Figure 1. Previously, this method was developed using MATLAB and was not intended to be implemented in hardware. Figure 2 shows the original vein image and the resulting vein image after it was applied using this method.

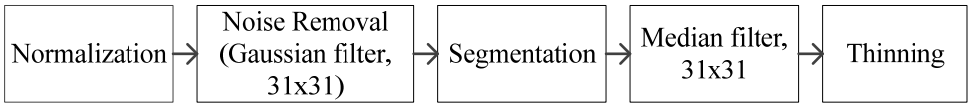
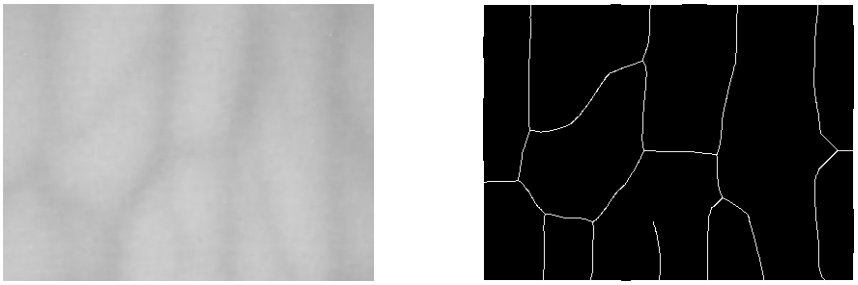


Figure 1: Previous vein enhancement detection method block diagram.



(a) **(b)**

Figure 2: Previous enhancement detection method: (a) Original vein image; (b) Resulting vein image.

It was found that using this method for hardware implementation required large hardware resource and lengthy processing time due to the size of filter masks in the corresponding blocks. The vein image size is 288x384 pixels. The number of pixels in the vein image is 110,592 pixels. In the noise removal block, which is Gaussian filter, the filter mask size is 31x31. To perform the Gaussian filter, for each pixel, there are 961 multiplications, 961 additions and 1 division involved. Hence, completing the Gaussian filter involves enormous multiplications, additions and divisions operation. It also takes a long processing time. Furthermore, there are two more Gaussian filters in the segmentation and a 31x31 median filter involved.

In this paper, we propose a new method for vein enhancement and feature extraction. This method can be effectively implemented in hardware with less hardware resources required. This method is developed using MATLAB. We intend to use a FPGA platform for the implementation in hardware. The rest of the paper is organised as follows. Section 2 gives a detailed description on the proposed method with resulting vein images. Section 3 presents the experiment results for the proposed method with discussion. Section 4 concludes with a brief summary of the proposed method and discusses the future works.

2. VEIN ENHANCEMENT AND FEATURE EXTRACTION METHOD

Vein image acquired from human either palm, finger or wrist is processed to enhance the image and extract the feature to obtain the biometric template, which is then matched against the template in the database for authentication. In this work, the vein image is processed using the proposed method to enhance and extract the feature of the vein image. The proposed method is shown in Figure 3.

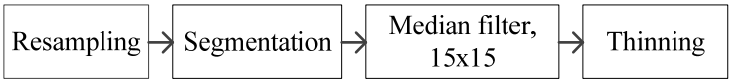


Figure 3: Block diagram of proposed vein enhancement and feature extraction method.

2.1 Image Resampling

Image resampling is known to have the ability to remove noise in images (Liu *et al*, 2010; Liu & Song, 2012; Lu *et al*, 2013). Thus, image resampling is used in the proposed method rather than Gaussian noise remover from our previous work (Yusoff *et al*., 2009). This technique also uses less hardware resources than the Gaussian noise remover that uses a lot of multiplication and division operations. Furthermore, it also significantly reduces the number of pixels in the image, resulting in less processing time and less data storage needed for hardware requirements. For resampling, the vein image has been downsampled to half of the size of the original vein image. Under-resampling will not remove the noise effectively, while over-resampling will remove vital information in the vein image. The vein image size after resampling is 144x192 pixels. The number of pixels is 27,648 pixels. To get the best result, nearest-neighbour, bilinear, box and bicubic interpolations were analysed for the resampling technique. The vein image was also downsampled at different sizes. Peak signal to noise ratio (PSNR) is used to measure the quality of image and noise suppression. PSNR is a quality measurement between the reference image and a degraded image. The higher the PSNR, the better quality of a degraded image compared to the reference image. Using PSNR, different interpolation techniques and downsamples sizes can be compared systematically to identify which produces better results. Table 1 shows the PSNR for each interpolation technique. It is found that bicubic interpolation is the optimal interpolation technique to reduce noise.

Table 1: PSNR of the vein image using different interpolation techniques and downsample sizes.

Downsample Size (with reference to original image size)	Nearest-Neighbour	Bilinear	Box	Bicubic
3/4	55.02	56.41	55.02	57.34
2/3	53.05	54.53	51.62	54.88
1/2	52.76	52.50	52.63	53.80
1/3	50.45	52.15	51.02	52.52
1/4	48.97	50.27	49.45	50.93

In order to determine the resampling downsample size, PSNR cannot be used for measurement because it requires a reference image with similar size. Thus, blind image quality assessment (Gabarda & Cristóbal, 2007) is used to determine the resampling downsample size. It is a method used to measure image quality without a reference or ground truth image. Figure 4 shows the blind image quality measurement for the different downsample sizes. A greater blind image quality measurement value reflects better image quality. From Figure 4, resampling to half of the original vein image size has the highest value, indicating the best image quality. Thus, resampling to half of the original vein image is the optimum resampling downsample size.

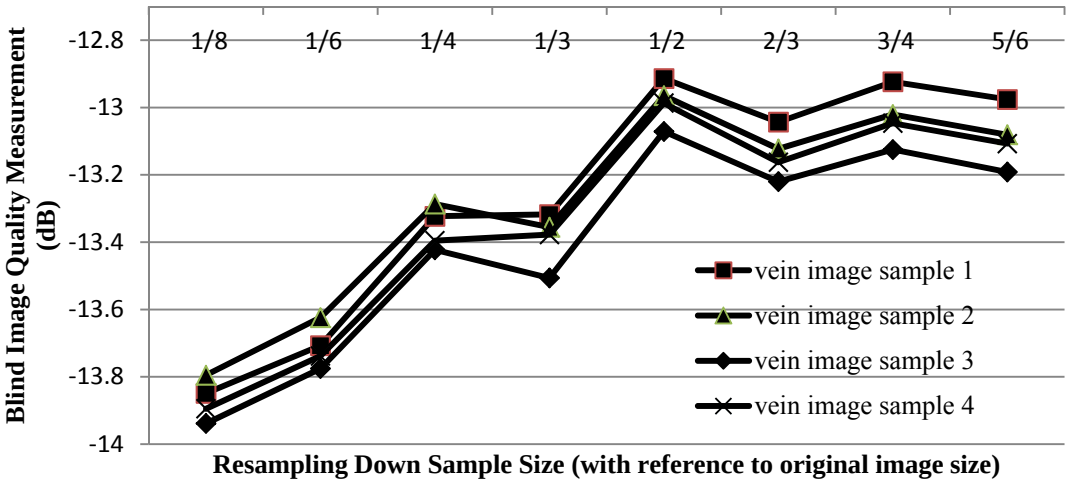


Figure 4: Blind image quality measurement versus resampling downsample size for samples of vein images.

2.2 Image Segmentation

For segmentation, we use difference of Gaussian (DoG) and threshold operation as shown in the Figure 5 to distinguish the veins from background and unwanted noise. DoG is an effective technique for noisy image enhancement. It works by performing two different values of standard deviations on the same image with different blurring radius, then subtracting them to produce the result (Kang *et al*, 2014). The smaller standard deviation will remove high frequency noise from the vein image, while the larger standard deviation will remove the detail including noise. When subtracting, it will suppress the effect of the noise from the vein image. Experiments and analysis were conducted to determine the best filter mask sizes and standard deviations for both Gaussians. There are many combinations of filter mask size and standard deviations for the two Gaussians. For each combination, entropy and standard deviation are measured. Both parameters indicate the quality of the image. Image entropy is a metric which is used to describe the amount of information in an image, while standard deviation is used as noise measurement.

For both parameters, the larger value means better image quality. From the results obtained (Table 2), the combinations that have both maximum entropy and standard deviation are determined. In the proposed method, first image is smoothed using a Gaussian filter with a 3 x 3 mask and standard deviation of 5. The second smoothed image is obtained using Gaussian filter with a 31 x 31 mask and standard deviation of 60. The standard deviation for the first Gaussian is 5, while for the second Gaussian is 60.

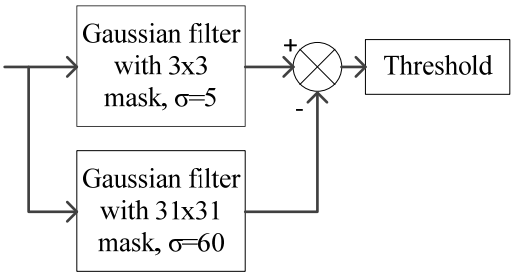


Figure 5: Block diagram for image segmentation.

Table 2: Filter mask sizes and standard deviations for DoG that has both maximum value of entropy and standard deviation.

No.	First Gaussian Filter Mask Size	First Gaussian Standard Deviation	Second Gaussian Filter Mask Size	Second Gaussian Standard Deviation
1	3x3	5	31x31	60
2	3x3	10	31x31	60
3	3x3	20	31x31	60
4	3x3	30	31x31	60
5	3x3	40	31x31	60
6	3x3	50	31x31	60
7	3x3	60	31x31	60

After DoG, thresholding is applied to the image. For each pixel, pixel value of '0' will be suppressed to binary '0', while the pixel is greater than the value of '0' will be changed to binary '1'. Now, the desired vein line will be black or binary '0' and the background will be white or binary '1'.

2.3 Median Filter

After segmentation, the resulting image has unwanted noise where it has some burrs and the edges are not smooth enough. To remove this noise, a 15x15 median filter is used for the vein image. Different median filter sizes were tested to determine the optimums size. According to Ding *et al.* (2005), noise should be removed based on the size of noise. Figure 6 shows the resulting images for median filter sizes of 3x3, 7x7, 11x11, 15x15, 19x19 and 23x23. For median filter sizes of 3x3 to 23x23, the resulting images will still have noise. When the median filter size is 3x3, it cannot remove the noise effectively and there is a lot of noise left in the image. When the median filter size is increased, the noise is reduced. It was found that median filter size of 15x15 removes the noise effectively. Larger median filter sizes remove the desired vein pattern.

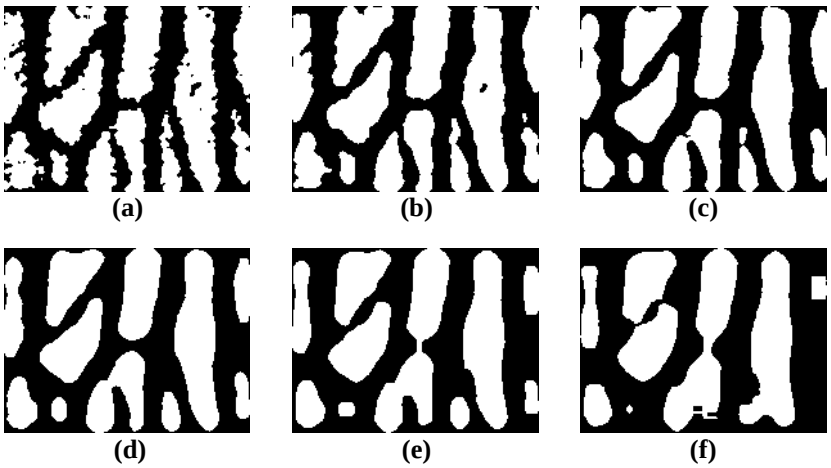


Figure 6: Resulting vein images for different median filter size: (a) 3x3; (b) 7x7; (c) 11x11; (d) 15x15; (e) 19x19; (f) 23x23.

2.4 Thinning

Finally, thinning is applied to reduce the thick lines of the veins to a single line pixel thickness. In this method, the thinning algorithm from Zhang & Suen (1984) is used. This algorithm is fast and requires simple computations. Therefore, this technique is selected for use due to hardware complexity. The desired foreground is set to '1' and the background is '0'. A 3x3 mask size is used in this algorithm as shown in Figure 7.

p9	p2	p3
p8	p1	p4
p7	p6	p5

Figure 7: Arrangement of 3x3 mask for thinning (Zhang & Suen, 1984).

This algorithm consists of two iterations. For first iteration, the point p1 is marked and deleted from the image if it satisfies all the following conditions:

(a) $2 \leq N(p1) \leq 6$
(b) $S(p1) = 1$
(c) $p2*p4*p6 = 0$
(d) $p4*p6*p8 = 0$

where N(p1) is the total number of non-zero neighbours of p1 and S(p1) is the total number of changes of the point value from 0 to 1 in the order of p1, p2, p3,…,p9. Point p1 is not deleted and remains the same if it does not satisfy all the conditions.

For the second iteration, point p1 is marked and deleted from the image if it satisfies all the following conditions:

(a) $2 \leq N(p1) \leq 6$
(b) $S(p1) = 1$
(c) $p2*p4*p8 = 0$
(d) $p2*p6*p8 = 0$

The differences between the first iteration and the second iteration are the conditions for (c) and (d). Both iterations are applied to the image repeatedly until no change is detected in the vein image.

3. RESULTS AND DISCUSSION

After resampling, the number of pixels in the image is 27,648 pixels as compared to 110,592 pixels in original image. This is a reduction of approximately 75% of the original image. The reduction of pixels can reduce hardware storage requirements. In segmentation, although the filter mask sizes are similar as the previous method, the number of pixels in the image has decreased significantly. Instead of having to calculate 110,592 pixels, only 27,648 pixels are calculated. This leads to reduced computation process, thereby reducing processing time.

For segmentation, the first Gaussian filter uses smaller filter mask size than the second Gaussian filter. From the experimental results, the DoG is improved if the mask size of the second Gaussian is greater than first Gaussian. This is the reason why the first Gaussian has 3x3 masks, while the second Gaussian has 31x31 masks. Figure 8 shows the comparison between DoG that uses smaller first Gaussian filter mask and larger first Gaussian filter mask. In Figure 8(c), it is shown that the desired vein pattern is more visible as compared to the vein pattern in Figure 8(b), especially the unclear and

thin vein pattern. Figure 8(f) also shows that the desired vein pattern is more obvious as compared to vein pattern in Figure 8(e).

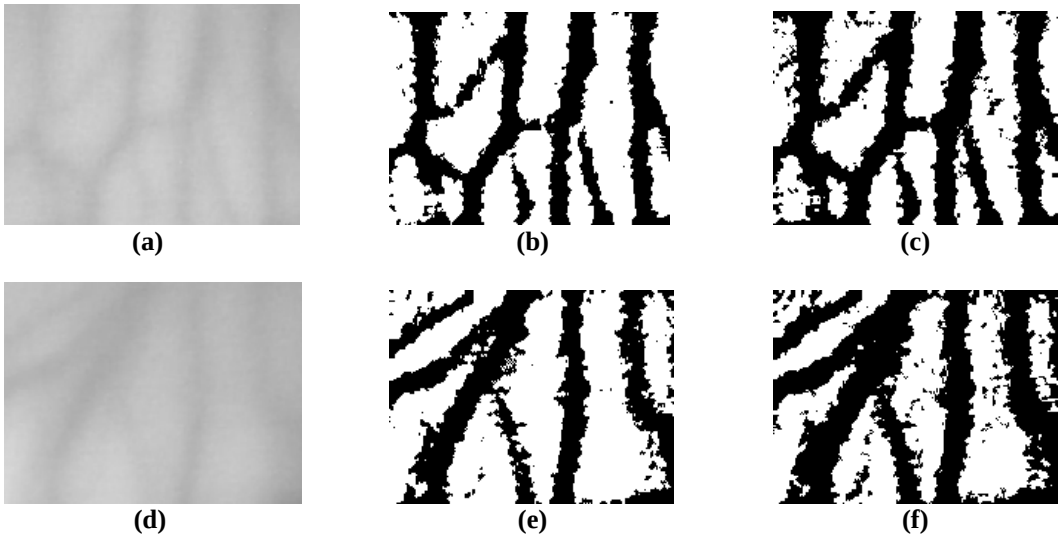


Figure 8: Resulting vein images for DoG that uses smaller first Gaussian filter mask and larger first Gaussian filter mask: (a) Original vein image sample 1. (b) Resulting vein image for sample 1 after being applied with DoG that uses larger first Gaussian filter mask. (c) Resulting vein image for sample 1 after being applied with DoG that uses smaller first Gaussian filter mask. (d) Original vein image sample 4. (e) Resulting vein image for sample 4 after being applied with DoG that uses larger first Gaussian filter mask. (f) Resulting vein image for sample 4 after being applied with DoG that uses smaller first Gaussian filter mask.

Figure 9 shows the resulting vein image for each sub-module for the proposed method. Figure 9(a) is the original vein image. The darker colour is the vein pattern. Figure 9(b) is the vein image after applied with resampling that reduced the number of pixels and also removed the noise. Then, image segmentation is applied to Figure 9(b) to distinguish the desired vein pattern from the surrounding as shown in Figure 9(c). Figure 9(d) shows the vein image after being applied with median filter to remove the noise introduced from image segmentation. Figure 9(e) shows the desired vein pattern after applied with thinning.

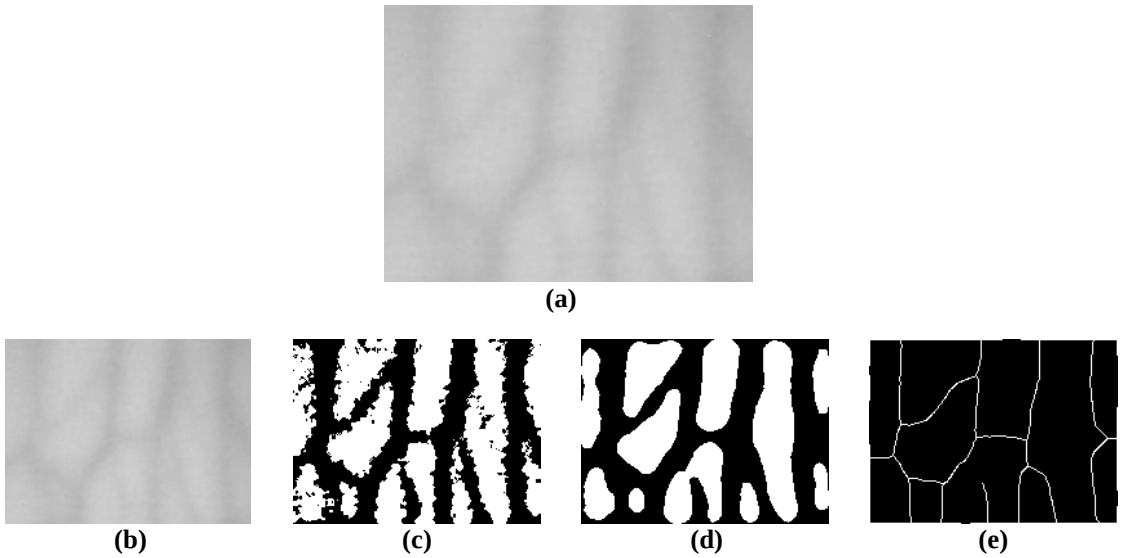


Figure 9: Resulting vein image for each sub-module for the proposed method: (a) Original vein image. (b) Resampling. (c) Segmentation. (d) Median filtering. (e) Thinning.

Figure 10 shows the results for different vein images after applying the proposed method. Figure 10(a), Figure 10(c), Figure 10(e) and Figure 10(g) are the vein image samples, while Figure 10(b), Figure 10(d), Figure 10(f) and Figure 10(h) are the resulting vein image respectively after have been applied with the proposed method. In the vein image samples, the darker colour is the vein pattern. Visibly it is difficult to distinguish between vein pattern and surrounding, especially the unclear and thin vein pattern. After being applied with the proposed method, the vein pattern is enhanced and extracted to a single line vein pattern.

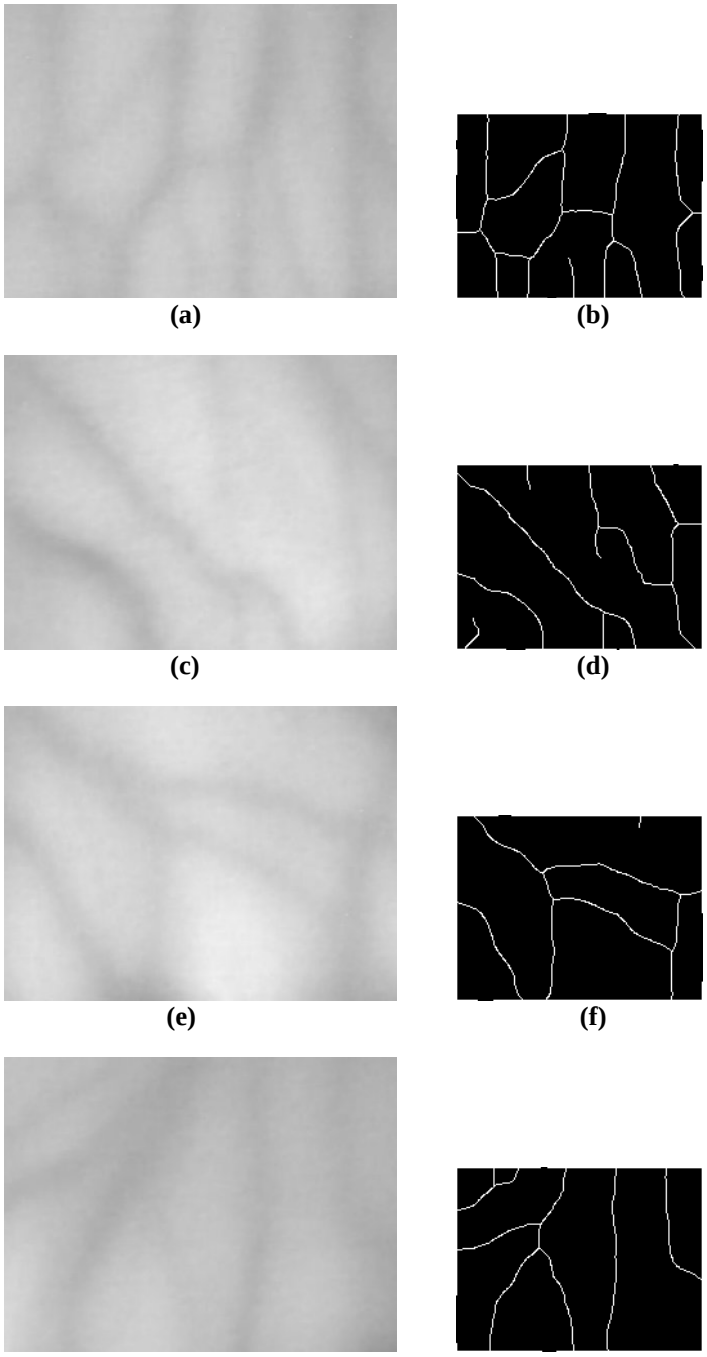


Figure 10: Proposed vein enhancement and feature extraction method results: (a) Original vein image sample 1; (b) Resulting vein image sample 1; (c) Original vein image sample 2; (d) Resulting vein image sample 2; (e) Original vein image sample 3; (f) Resulting vein image sample 3; (g) Original vein image sample 4; (h) Resulting vein image sample 4.

Table 3 shows the average execution time for each image in comparison with the previous method (Yusoff *et al*., 2009). The experiment was conducted using MATLAB on Windows 8.1 with Intel i7-4710HQ (2.5GHz) and 8 GB RAM. It is found that the proposed method provides significantly better execution time than the previous method. The proposed method not only improves execution performance but also uses less hardware resources, which make it suitable for FPGA implementation.

Table 3: The average execution time for vein image samples using the previous and proposed methods.

Image	Previous Method (s)	Proposed Method (s)	Improvement Compared to Previous Method
Sample 1	2.09	0.31	85.17%
Sample 2	2.42	0.29	88.02%
Sample 3	2.01	0.25	87.56%
Sample 4	2.28	0.31	86.40%

4. CONCLUSION

In this paper, an improved vein enhancement and feature extraction method is proposed. Resampling using bicubic interpolation not only removes the noise, but it also significantly reduces the number of pixels in the image. DoG with thresholding effectively segmented the vein image. The second Gaussian filter mask size used is bigger than the first Gaussian filter mask size to improve the segmented image. Median filter size of 15x15 is used to smooth and remove noise introduced from the image segmentation. Then, thinning is applied to obtain the single line veins. All the variables in the proposed method are analyzed and determined to obtain the optimum performance. The results obtained show that the proposed method is effective to detect veins in the images. The proposed method also speeds up the processing time significantly. The resampling technique not only reduces the number of pixels, but also reduces the intensity of computation, hardware resources, processing time and hardware storage. Thus, it is good for hardware implementation. Our future work is to implement the proposed method in FPGA.

REFERENCES

Contreras-Medina, L.M., Osornio-Rios, R.A., Torres-Pacheco, I., Romero-Troncoso, R. de J., Guevara-González, R.G. & Millan-Almaraz, J.R. (2012). Smart sensor for real-time quantification of common symptoms present in unhealthy plants. *Sensors*, **12**: 784–805.

Contreras-Medina, L.M., Romero-Troncoso, R.D. J., Cabal-Yepez, E., Rangel-Magdaleno, J.D.J. & Millan-Almaraz, J.R. (2010). FPGA-based multiple-channel vibration analyzer for industrial applications in induction motor failure detection. *IEEE T. Instrum. Meas.*, **59**: 63–72.

Ding, Y., Zhuang, D. & Wang, K. (2005). A study of hand vein recognition method. *IEEE Int. Conf. Mech. Autom. 2005*, **4**: 2106–2110.

Ennis, J. (2012). Swapping PINs for palms – the potential of biometric technology in retail banking. *Biometric Tech. Today*, **2012**: 8–9.

Fujitsu Limited. (2006). Japanese public library to use palm vein. *Biometric Tech. Today*, **14**: 4.

Gabarda, S., & Cristóbal, G. (2007). Blind image quality assessment through anisotropy. *J. Opt. Soc. Am. A.*, **24**: B42–B51.

Iterson, R. (2004). Vein recognition in Europe. *Biometric Tech. Today*, **12**: 6.

Kang, W., Liu, Y., Wu, Q. & Yue, X. (2014). Contact-free palm-vein recognition based on local invariant features. *PloS One*, **9**: e97548.

Khalil-Hani, M., & Eng, P.C. (2010). FPGA-based embedded system implementation of finger vein

biometrics. *2010 IEEE Symp. Ind. Electr. Appl.*, pp. 700–705.
Khalil-Hani, M., & Eng, P.C. (2011). Personal verification using finger vein biometrics in FPGA-based Systems-on-chip. *Int. Conf. Elec. Electr. Eng.*, pp. II-171-II-176.
Khalil-Hani, M. & Lee, Y.H. (2013). FPGA embedded hardware system for finger vein biometric recognition. *39th Annu. Conf. IEEE Ind. Electr. Soc.*, pp. 2273–2278.
Lee, Y.H., Khalil-Hani, M. & Bakhteri, R. (2012). FPGA-based finger vein biometric system with adaptive illumination for better image acquisition. *2012 Int. Symp. Comput. Appl. Ind. Electr.*, pp. 107–112.
Liu, Z. & Song, S. (2012). An embedded real-time finger-vein recognition system for mobile devices. *IEEE T. Consum. Electr.*, **58**: 522–527.
Liu, Z., Yin, Y., Wang, H., Song, S. & Li, Q. (2010). Finger vein recognition with manifold learning. *J. Netw. Comput. Appl.*, **33**: 275–282.
Lu, Y., Xie, S. J., Yoon, S. & Park, D. S. (2013). Finger vein identification using polydirectional local line binary pattern. *Int. Conf. ICT Convergence*, pp. 61–65.
Pudzs, M., Fuksis, R. & Ruskuls, R. (2013). FPGA based palmprint and palm vein biometric system. *Int. Conf. Biometrics Spec. Interest Group*, pp. 1–4.
Pudzs, M., Fuksis, R., Ruskuls, R., Barkans, D., Eglitis, T. & Greitans, M. (2013). FPGA implementation of CMF for embedded palm biometric system. *Signal Proc. Conf. (EUSIPCO)*, pp. 1–5.
NTSC (National Science & Technology Council) (2006). *Vascular Pattern Recognition*. National Science & Technology Council(NTSC), Washington D.C.
Sun, X., Lin, C.-Y., Li, M.-Z., Lin, H.-W., & Chen, Q.-W. (2011). A DSP-based finger vein authentication system. *2011 Fourth Int. Conf. on Intell. Comput. Technol. Autom.*, pp. 333–336.
Wei, T. C., & Rosdi, B. A. Bin. (2012). FPGA based band-limited phase-only correlation (BLPOC) for finger vein recognition system. *2012 4th Int. Conf. Intell. Adv. Syst.*, pp. 729–734.
Yusoff, S., Ramli, A.R., Hashim, S.J. & Rokhani, F.Z. (2015). Review on vein enhancement methods for biometric system. *Int. J. Res. Eng. Tech.*, **4**: 833–841.
Yusoff, S., Ramli, A. R., Saripan, M. I. & Mashohor, S. (2009). Palm-Dorsa vein detection using webcam. *Int. Adv. Techn. Congr.*, pp. 1–6.
Zhang, T.Y. & Suen, C.Y. (1984). A fast parallel algorithm for thinning digital patterns. *Commun. ACM.*, **27**: 236–239.

DESIGN OF VITERBI DECODER FOR UNDERWATER MARINE RECEIVERS USING MULTI-THRESHOLD NULL CONVENTION LOGIC (MTNCL)

Sudhakar Jyothula[1*] & Abbireddy Uma Maheswari[2]

Department of Electronics and Communication Engineering, Vignan's Institute of engineering for Women, Visakhapatnam, India

[*]Email: sudhakar.jyo @gmail.com

ABSTRACT

In this paper, we develop a Viterbi decoder for underwater ITU- V.34 standard marine receivers. In order to have low power and to accomplish faithful quality of transmission, clockless approaches are found to better alternatives. The underwater message channel is not only band limited, but motionless as well and mostly difficult because of strong distortions. In view of this, underwater communications is the focus of exhaustive research aiming to accomplish higher data rates with optimized delay requirements. Null convention logic (NCL) is the traditional asynchronous design approach with which draw backs of clock distribution, power or delay are overcome. In order to obtain both advantages, we incorporate a multi-threshold approach in the NCL, known as Multi -Threshold NCL (MTNCL). The proposed method of Viterbi decoder is designed using Xilinx tools and is compared with the performance of conventional complementary metal oxide Semiconductor (CMOS) and asynchronous NCL design in terms of power savings.

Keywords: *Asynchronous design; Viterbi decoder; null convention logic (NCL); multi-threshold, marine-receivers.*

1. INTRODUCTION

The public telephone switching network (PSTN) infrastructure was designed for human voice band transmission. As this network carries analogue signals with limited bandwidth, the setup does not allow data with fast data transmission rates (Iwona *et al.,* 2012). Instead of replacing the entire analog communications infrastructure, the modem was devised. The modem's functionality is to adjust digital data to the type of data that is usually carried in a telecommunications channel (Kilfoyle et al., 2000). The first telecommunications modem was devised in Bell laboratories in 1958. Its speed was 0.3kbits/s in an analog telecommunication channel with frequency range of 300-3400Hz. It has the provision of applying complex techniques such as compression and error correction encoding to have better data rates (Texas Inc, 1997).

Underwater communications has been the focus of intensive research aimed at obtaining better data rates and quality of transmission. Digital selective calling (DSC) is used for transport distress alerts from ships as well as for sending acknowledgments of these alerts from a coast station in underwater. The DSC system is a digital synchronous system that uses ten-bit error detecting code. The first seven bits of the ten-bit code are information bits (128 different characters) while the remaining three bits are error detecting bits (providing information of zero bits within each character). For the purpose of error correction, every character is transmitted with twice the time delay that corresponds to the broadcast time of the four following characters that need to be transmitted (Valčić *et al*, 2016). Hence, error correction and power requirements are important issues in the design of marine receivers.

The design of receivers with traditional complementary metal oxide semiconductor (CMOS) methodology leads to issues with clock signals and causes more delay. As asynchronous approaches do not involve clock, these issues can be successfully sorted out (Sudhakar *et al*., 2016). In this paper, we present a novel design methodology for the implementation underwater marine receivers i.e. null convention logic (NCL). Along with this, the multi-threshold concept is incorporated to reap better performance metrics.

2. THEORETICAL BACKGROUND

2.1 ITU V.34 Modem

ITU V.34 is one of the efficient and popular modems for submarine receivers. In a V.34 transmitter (Figure 1) the input bits stream is processed by the scrambler, which is a shift register that changes the sequences 0 and 1 by modulo 2 sum operations. It converts serial errors into simple random errors and nullifies long periodic sequences of 0 and 1, which can produce periodic signal at the transmitter's output. This leads to negative impact on the performance of the receiver's adaptation filters. Pulse distortions and serial errors are eliminated in the scrambler (Iwona *et al*., 2012).

.

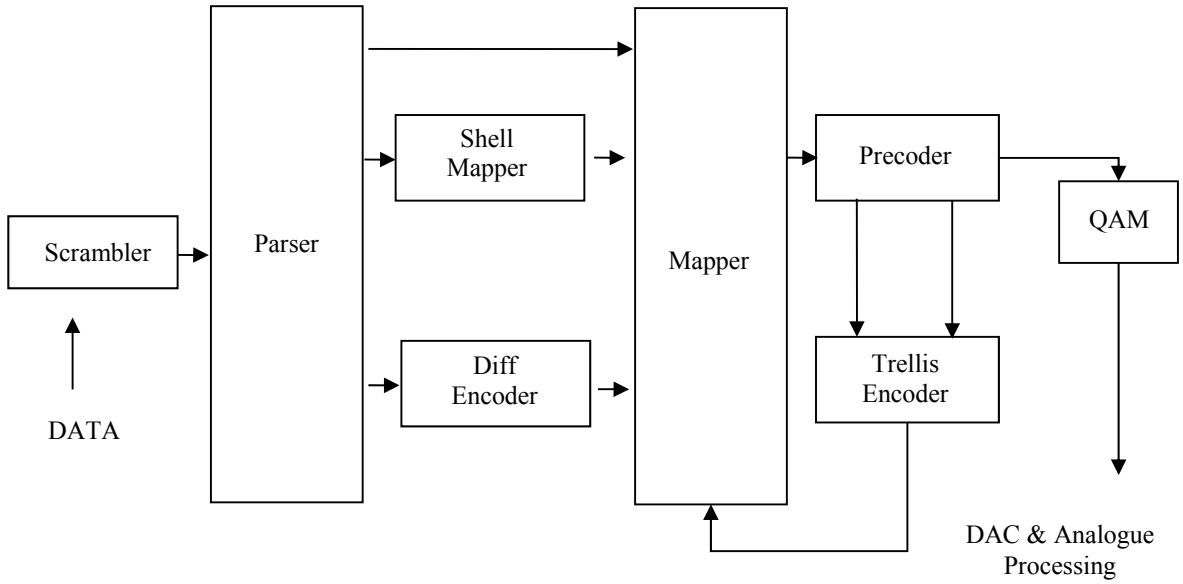


Figure 1: Functional diagram of a V.34 transmitter modem.

The parser divides the block of bits from the scrambler into three groups which are processed further by the shell mapper, differential encoder and mapper. The shell mapper computes ring index integers from the first group of bits. Pre-emphasis was done by the precoder. The trellis encoder generates trellis sequence into the complex signal points. The sequence secures the distances in between the points protecting the resolution of signal power (Iwona *et al*., 2012). In the trellis encoder, the signal points from the past signal points may suffer with errors caused by noise in the channel and inter symbol interference can be detected and corrected by a Viterbi decoder, which is the main theme in this paper.

The functionality of a V.34 receiver is shown in Figure 2. Transmitter and receiver synchronization can be established with clock recovery system. The quadrature amplitude modulator (QAM) modulates the quadrature signal at 1800Hz (Daniel *et al.,* 2000). The equalizer cleans the signal hindrance caused by non linearity in the propagation path. The signal component that was filtered from the signal by the

transmitter's precoder is introduced by the inverse precoder. The Viterbi decoder behaves like an error correction mechanism in underwater receivers. The inverse mapper and inverse selection unit are accountable for performing inverse functions in the transmitter. The descrambler is a shift register that performs the inverse function of scrambler. For any communications channel encoder, channel and decoder are very important (Nupur *et al.* 2015*)*.

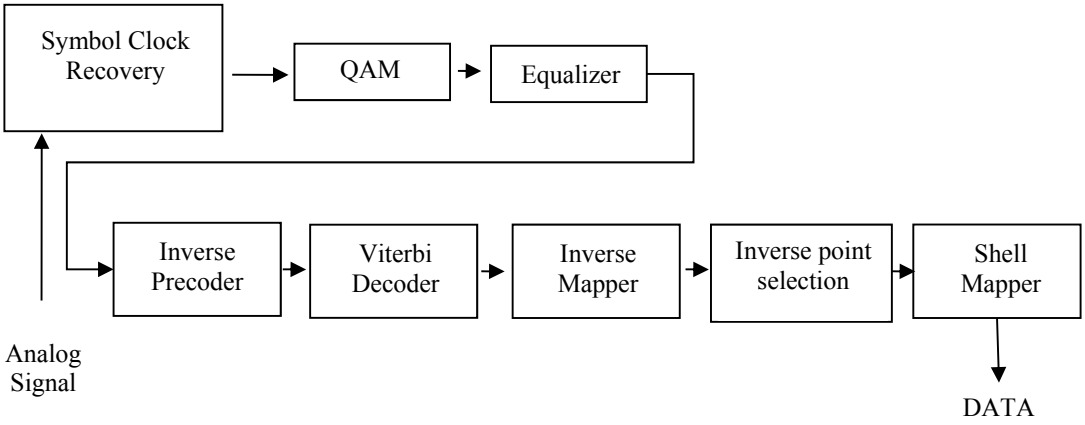


Figure 2: Functional block diagram of a V.34 receiver.

The encoder is used to encode the bit stream that it receives from the source. It consists of multiple Exclusive OR (XOR) gates and multiple-shift registers (Kawokgy *et. al.,* 2007). Figure 3 shows the convolution encoder that consists of two XOR gates and three shift registers for three message bits (*m, m1* and *m2*) (Kalavathi *et al.,* 2012).

$$X1 = m\ XOR\ m1\ XOR\ m2 \quad (1)$$

$$X2 = m\ XOR\ m2 \quad (2)$$

where *X1* and *X2* are the outputs of Adders 1 and 2. The generating sequences of the two adders are *{g0(1),g1(1),…….gm(1)} and {g0(2),g1(2),…..gm(2)}* respectively, i.e., {111} and {101}.The outputs of Adders 1 and 2 are obtained by convolving the generating sequence of adders with message sequences {1000} and {1101} respectively(Wing *et al.,* 2004).

$x_i(1) = \sum_{i=0}^{M} g_i(1) m(i-l) = \{1\ 0\ 0\ 0\}$ (3)

$x_i(2) = \sum_{i=0}^{M} g_i(2) m(i-l) = \{1\ 1\ 1\ 0\}$ (4)

By adding the two adder outputs, the encoder output is= {1 1 0 1 0 1 0 0}.

If the input sequence length is *N,* then the output sequence length will be *2N*.

2.2 Error Correction using Viterbi Decoder

The Viterbi decoding algorithm (Chien *et al.,* 2005*;* Varadpande *et al.,* 2015), proposed in 1967 by Andrew J. Viterbi, is a decoding process for convolutional codes in memory-less noise. The Viterbi algorithm is used to find the most likely noiseless finite-state sequence, given a sequence of finite-state signals that are corrupted by noise. The Viterbi algorithm process is based on the trellis diagram. In Viterbi decoding, a metric is assigned to each surviving path (Habib *et al.,* 2009). This metric is nothing but the discrepancy between the received signal and the decoded signal at a particular node. The metric can be added over a few nodes for a particular path. The metric of a particular path is obtained by adding the individual metrics of the nodes along that path. The surviving path (*Y*) refers to the path of the

decoded signal with the minimum metric. Consider following example of Viterbi decoding (Balamurugan *et al.*, 2015).
Y=11 01 01 00

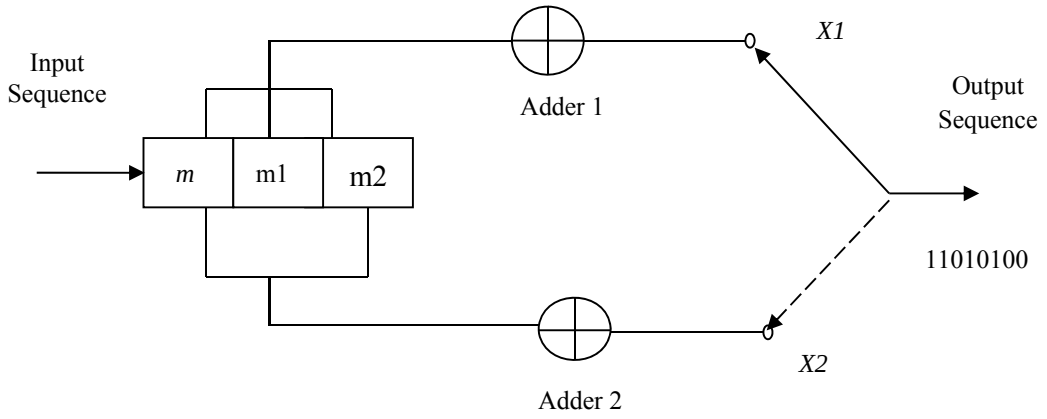


Figure 3: Functional behavior of encoder.

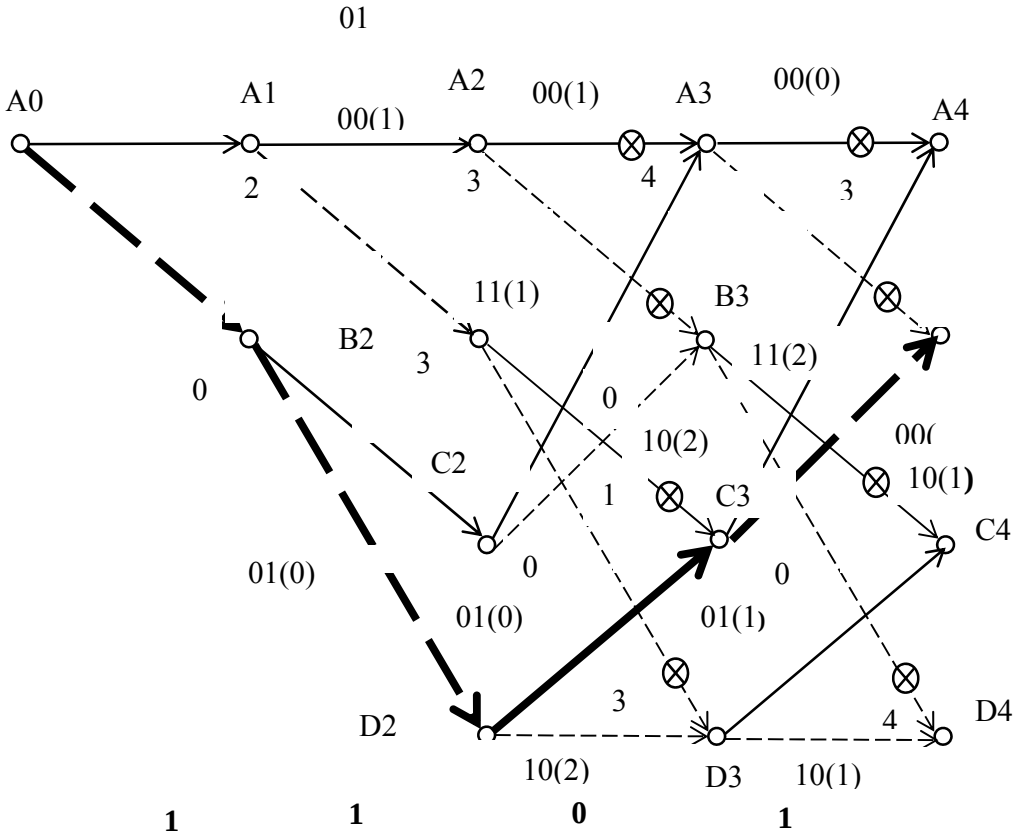


Figure 4: Working principle of Viterbi algorithm without noise (Simon *et al.*, 2012).

Decoding of the first message bit *y* = 11: Let the initial state be A0. From the Trellis diagram it is clear that if the current state is A, then the next state will be either A or B. The two branches are shown from A0. One branch is at next node A1 representing the decoded signal as 00 and the other branch is at B1 representing the decoded signal as 11.The branch from A0 to B1 represents the decoded output as 11, which is same as the received signal at that node. In this particular case, there is no discrepancy between the received and decoded signals. Hence, the metric of that branch is 0. This metric is shown in brackets along with that of branch metric. The metric of the branch from A0 to A1 is 2. The encoded number near a node shows the path metric reaching the node.

Decoding of second message bit *y* = 01: When the next part of the input bits are received at nodes A1 and B1, then four possible next states are A2, B2, C2 and D2.Figure 4 shows all the branches, their decoded outputs and branch metrics corresponding to those decoded outputs. The encoded numbers near A2, B2, C2, and D2 indicate the path metric emerging from A0.For example, the path metric of path A0-A1- A2 is 3 and the path metric of path A0 -B1- D2 is 0.

Decoding of third message bit *y* = 01: While decoding the third message bit, two paths are possible at node A. One path is A0-A1-A2-A3 with metric 4.The other path is A0-B1-C2-A3 with metric 3. According to Viterbi decoding, only one path with lower metric will be retained at that particular node. Four paths with lower metrics are stored in the decoder and the decoding will continue to next received bits.

3. METHODOLOGY

In this paper, we present the architecture of a Viterbi decoder with Multi Threshold Null Convention Logic (MTNCL) approach. The asynchronous MTNCL provides delay insensitive logic operation with significant leakage power and active energy reduction. The MTNCL circuit is also capable of functioning properly under extreme supply voltage scaling down to the sub-threshold region for further power reduction (Sudhakar *et al.,* 2016). The general structure of the MTNCL design is shown in Figure 5.

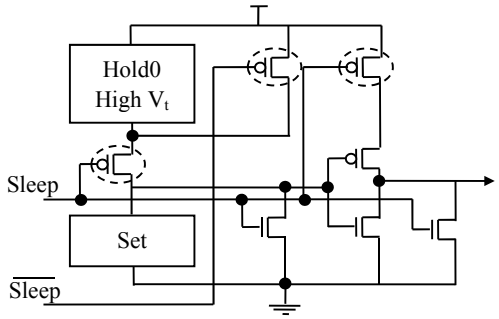


Figure 5: General structure of the MTNCL design (Sudhakar *et al.,* 2016).

The proposed Viterbi decoder (Figure 6) mainly consists of three blocks, which are branch metric unit (BMU), add compare and select unit (ACSU), and survivor memory unit (SMU). The BMU compares the received code symbol with the standard code symbol and counts the number of difference bits (Dholakia, 1994; Jun *et al,*2004.; Meilana *et al.*, 2006; Yun *et al.,*2009). The BMU unit comprises of XOR gate and 3-bit counter. The XOR gate's output works like a clock input to the counter. The ACSU comprises of ripple carry adder, comparator and selector. The ACSU adds the branch metrics (BM) to the corresponding path metrics (PM), compares the new PMs, and then stores the selected PMs in the path metric memory (PMM). At the same time, the ACSU stores the associated survivor path decisions in the SMU (Lang *et al.*, 1997; Arun *et al.*, 2009). The SMU is designed by using four 4-bit SISO shift registers.

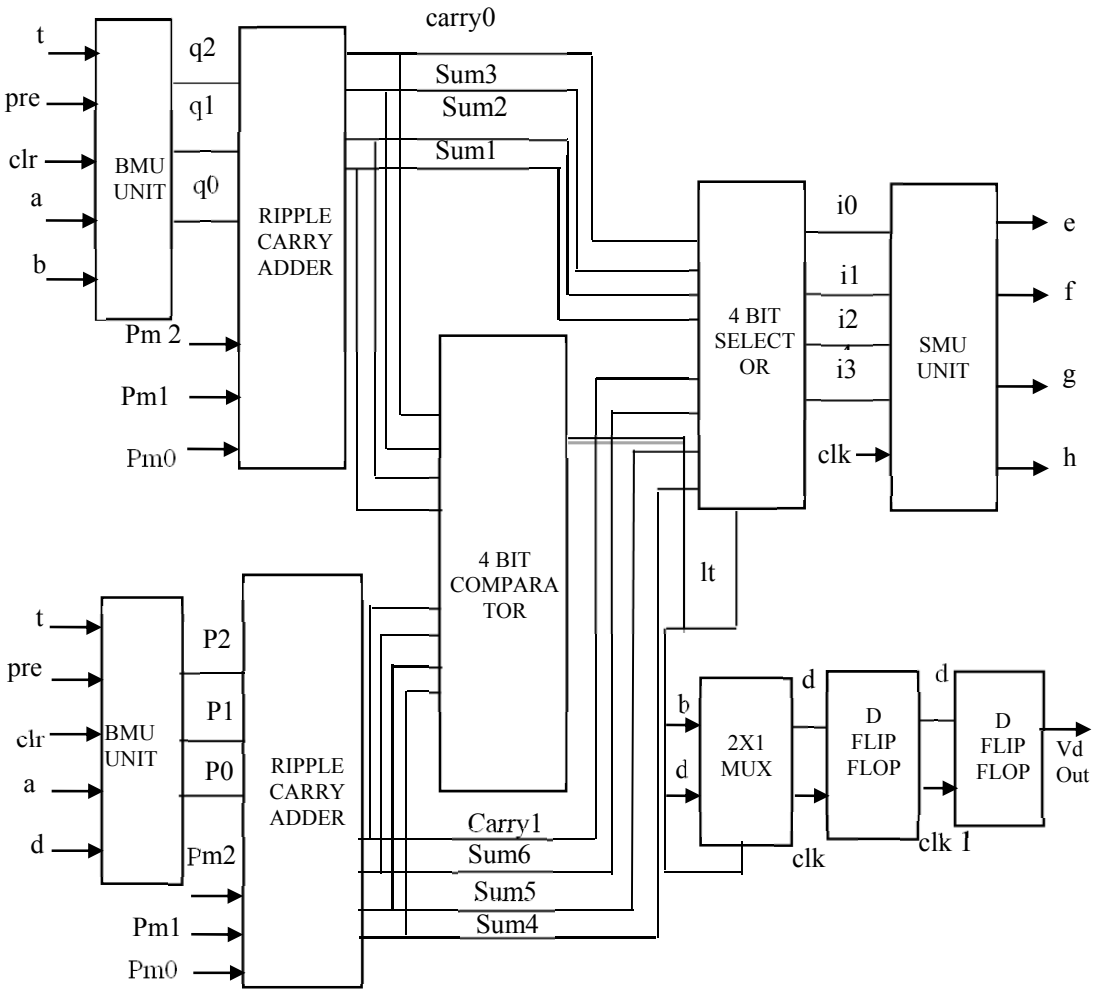


Figure 6: Block diagram of the proposed Viterbi decoder.

4. RESULTS AND DISCUSSION

In this paper, a Viterbi decoder with MTNCL is designed using Verilog code and implemented using Xilinx ISE. The generated register transfer logic (RTL) schematic is shown in Figure 7. Initially, the input is given to the two BMU units, where the counter starts counting. The outputs of the BMU units are given as input to the two ripple carry adders and 4 bit selector unit. The “lt” output of the comparator is given as selection input to the 4-bit selector. Based on the selection input to the 4- bit selector the corresponding output is obtained, which is processed as input to the last block i.e., the SMU block. The Viterbi decoder also consists of 2:1 multiplexer and two D flip flops. The decoded output is achieved at the output of the D flip flops as shown in Figure 8

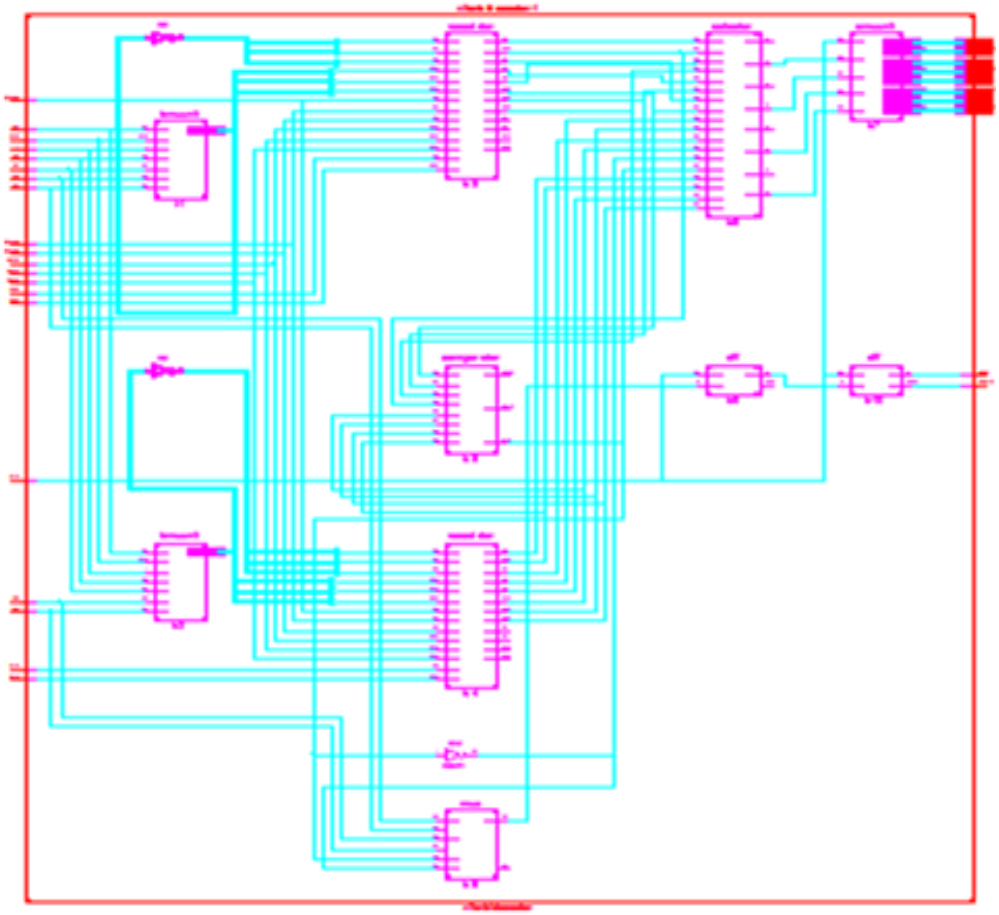


Figure 7: RTL schematic of the Viterbi decoder.

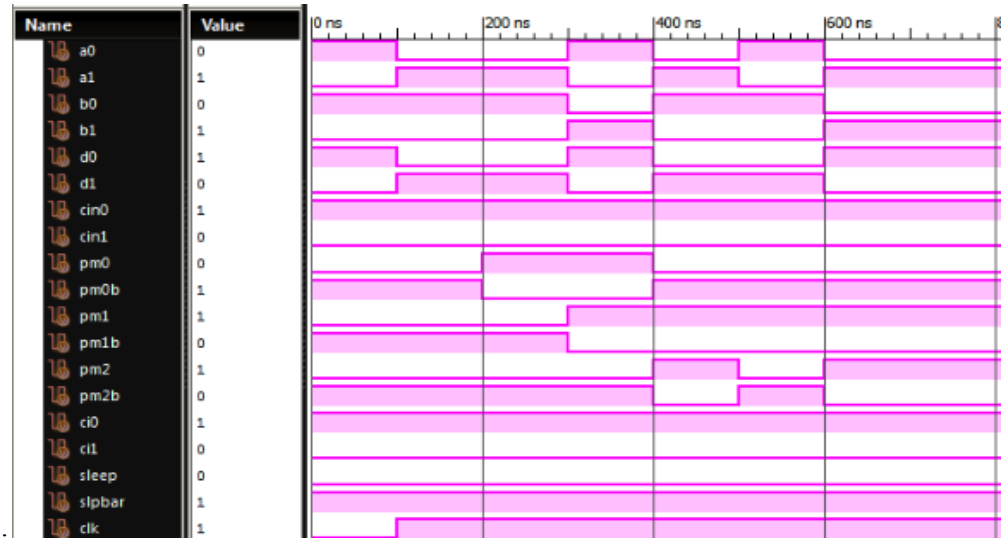


Figure 8: Output waveform of Viterbi Decoder.

Along with functionality verification, the verilog code is synthesized and power dissipation report is generated using XPower Analyzer. The power dissipation obtained with MTNCL is compared with existing implementations - CMOS and NCL (Table 1). It is found that the power consumption for the proposed MTNCL Viterbi decoder is better than the existing approaches. The MTNCL Viterbi decoder requires 0.029 µw, while their counter parts need 0.040 (CMOS) and 0.030 µw (NCL) respectively. With the MTNCL Viterbi decoder we achieved 27.5 % power savings.

Table 1: Comparison of power consumption for various methodologies.

Parameter	CMOS	NCL	Proposed MTNCL
Power (µW)	0.040	0.030	0.029

5. CONCLUSION

This paper has presented the design and implementation of a MTNCL Viterbi decoder. Its streamed input-output, regular architecture and parallel execution favor an FPGA implementation. Applications such as data communications and image processing require more processing power but when the fastest DSP processor is not fast enough- the only alternatives are to add multiple DSP processors or to design custom hardware devices. Multiple DSP processors are expensive, require many components and consume too much power. The performance gain that comes with each additional processor is small when compared to the increase in cost, board space, power consumption, and development time. Custom devices deliver the performance but sacrifice flexibility and require a large engineering investment with no chance to recover from mistakes. FPGAs are the new solution used by many engineers to implement computationally intensive algorithms. Some of the generic parameters are basic decoder specifications, metric size, Trellis window length, number of surviving paths and pipeline depth. In this paper, we demonstrated the improved power consumption of the proposed Viterbi decoder using MTNCL as compared with NCL and CMOS topologies. Both NCL and MTNCL support return-to- zero (RTZ) protocol. We may reap better results for marine receivers with return -to -one (RTO) protocol due to its capability of staying at logic '1' in idle cases. In future, this work may be extended with better encoder rates.

REFERENCES

Arun,C. & Rajamani.V. (2009). A low power & high speed Viterbi Decoder based on Deep pipelined, clock blocking and hazards filtering, Network and System sciences. *International Journal of Comm.*, **6**: 575-582

Balamurugan, V. & Nandhitha, N.M. (2015). Performance Evaluation of BPN Based Viterbi Decoder for Decoding 2-bit and 3-bit Errors. *ARPN Journal of Engineering and Applied Sciences*. **10**:4.

Chien, C.L., Yen, H.S., Hsie, C.C. & Chen, Y. L. (2005). Design of a Power Reduction Viterbi Decoder for WLAN Applications. *IEEE Tran. on Circuits and Systems-I*, Regular papers. **52**.

Habib, I., Paker, O. & Sawitzki, S. (2009). Design Space Exploration of Hard Decision Viterbi Decoding.

Haykin, S. *COMMUNICATION SYSTEMS*,2012

Iwona, B., Henryk, L. & Romuald, M. (2012). Adapting Telecommunications Mod-ems Techniques for use in Underwater Communication Systems. *Hydrostatics*, **6**: 285-292.

Jun, J. K. & Keshhab K. P. (2004) Low Latency Architectures for High Throughput Rate Viterbi Decoder. *IEEE Transactions on Vlsi Systems***12**:642-651

Kalavathi Devi, T. & Venkatesh, C. (2012). Design of Low Power Asynchronous Viterbi Decoder using Asynchronous Techniques. *International Journal of Advances in Engg. & Tech*. **4:** 561-570.

Kawokgy, M. & Salama, A.T. (2007). A Low Power CSCD Asynchronous Viterbi Decoder for Wireless Applications. in Proc. of *International Symposium on Low Power Electronics and Design*, Portland, USA, August 27-29.

Kilfoyle, D.B. & Baggeroer, A.B. (2000). The State of the Art in Underwater Acoustic Telemetry. *IEEE Journal of Oceanic Engg.*.**25**.

Lang, L., Tsui, C.Y. & Cheng, R.S. (1997). Low power soft output Viterbi decoder scheme for turbo code decoding. *IEEE Conference-Paper*, ISCAS '97, New York, USA, 24, pp. 1369-1372.

Meilana. S., Masuri, O. & Edmond Z.. (2006). VLSI Implementation of ½ Viterbi Decoder for IEEE P802.15-3a UWB Communication. *IEEE ICSE2006 Proc.*, Kuala Lumpur, Malaysia, pp. 666-670.

Nupur, R. C. & Surekha, T. K. (2015). Design of Asynchronous Viterbi Decoder using Bundled Data Protocol for Low Power Consumption. *International Journal of Modern Trends in Engg and Research*.

Sudhakar, J., Prasad, A.M. &Ajit Kumar, P. (2016).An Ultra Low Power Self Timed Design Approach with Multi Threshold Variations: MTNCL+.*International Journal of Innovations in Engg & Tech.***06**:273-278

Wing, K. C., Chiu, S. C., Cheong, F. C. & Kong, P. P. (2004). An Asynchronous Sova Decoder for Wireless Applications. *Lectures in Department of Electronics and Communication Engineering*, The Chinese University of Hong Kong, Shatin, Hongkong.

Valčić,S., Zoran, M., & Gulic, M. (2016). Analysis of Advantages and Disadvantages of Existing Maritime Communication Systems for Data Exchange. *Multidisciplinary Scientific Journal of Maritime Research.* **30**:28-37.

Varadpande, P.S. & Mali, M. B. (2015). A Review on Low Power Asynchronous Viterbi Decoder for Wireless Communication. *International Journal of Scientific Research and Education.***03:**06.

Yun, C. T., Do C. H., Weiyi, W., Chung, L. W. & Hongchin L.. (2009).A Memory-Efficient Architecture for Low Latency Viterbi Decoders. *IEEE Transactions*. **06**:335-338

.
.

.

EVALUATION OF THE EFFECT OF GLOBAL POSITIONING SYSTEM (GPS) ANTENNA ORIENTATION ON GPS PERFORMANCE

Dinesh Sathyamoorthy*, Shalini Shafii, Zainal Fitry M Amin, Mohamad Firdaus Ahmad, Asmariah Jusoh & Siti Zainun Ali

Science & Technology Research Institute for Defence (STRIDE), Ministry of Defence, Malaysia

*Email: dinesh.sathyamoorthy@stride.gov.my

ABSTRACT

This study is aimed at evaluating the effect of Global Positioning System (GPS) antenna orientation for three Garmin GPS receivers that use built-in quad helix antennas; GPSmap 60CSx, GPSmap 62Cs and Oregon 550. The study is conducted using GPS simulation, which allows for the tests to be held with various repeatable conditions, as defined by the authors. Readings are taken for GPS antenna orientations of 0 to 345°, at increments of 15°. From the results obtained, it is found that there is degradation of accuracy for orientations of 75 to 120° and 240 to 285°. For the remaining orientations, the accuracy remains constant. This indicates that the quad helix antennas are operating in endfire and backfire modes simultaneously. While this type of design has smaller antenna gain than quad helix antennas that use only endfire or backfire modes, it allows for a more isotropic antenna performance. This study will be extended to evaluate the performance of antennas of a wider range of GPS receivers.

Keywords: Global Positioning System (GPS) antenna orientation; GPS simulation; estimate probable error (EPE); quad helix antenna; endfire and backfire modes.

1. INTRODUCTION

Antennas are a critical part of any Global Navigation Satellite System (GNSS) receiver design and their importance cannot be stated highly enough. Even the best receiver cannot bring back what has been lost due to a poor antenna design. GNSS signals are extremely weak and present unique demands on the antenna. The choice and implementation of the antenna plays a significant role in GNSS performance (Moernaut & Orban, 2009; Gupta, 2016; Hautcoeur *et al*., 2016).

Ideally, a GNSS antenna should have an isotropic response pattern that is independent of its orientation or direction of arrival of GNSS signals. However, there are no ideal antennas in the real world and real antennas do not have an isotropic response pattern. This means that the same signal received at various antenna orientations can result in stronger or weaker signals being presented to the receiver front end. To this end, the evaluation of the effect of GNSS antenna orientation on GNSS performance has received significant attention (D'eon & Delparte, 2005; O'Driscoll *et al*., 2007; Grimm, 2012).

This study is aimed at evaluating the effect of Global Positioning System (GPS) antenna orientation for three Garmin GPS receivers that use built-in quad helix antennas; GPSmap 60CSx (Garmin, 2007), GPSmap 62Cs (Garmin, 2011) and Oregon 550 (Garmin, 2010). These receivers employ the GPS L1 coarse acquisition (C/A) signal, which is an unencrypted civilian GPS signal widely used by various GPS receivers. The signal has a fundamental frequency of 1,575.42 MHz and a code structure which modulates the signal over a 2 MHz bandwidth (DOD, 2001; Kaplan & Hegarty, 2006; USACE, 2011).

The study will be conducted using GPS simulation, which will allow for the tests to be held with various repeatable conditions, as defined by the authors. As the tests are conducted in controlled laboratory environments, they will not be inhibited by unintended signal interferences and obstructions (Aloi *et al*., 2007; Kou & Zhang, 2011; Pozzobon *et al*., 2013). In previous studies, GPS simulation was used to evaluate the vulnerabilities of GPS to radio frequency interference (RFI) (Dinesh *et al*., 2012a, 2014a), multipath (Dinesh *et al*., 2013, 2014b), GPS satellite clock error (Dinesh *et al*., 2015a), varying speeds (Dinesh *et al*., 2015b) and power consumption (Dinesh *et al*., 2016).

2. METHODOLOGY

The apparatus used in the study are an Aeroflex GPSG-1000 GPS simulator (Aeroflex, 2010) and a notebook running GPS Diagnostics v1.05 (CNET, 2004). The study is conducted in STRIDE's mini-anechoic chamber (Kamarulzaman, 2010) to avoid external interference signals and unintended multipath errors. The test setup employed is as shown in Figure 1. Simulated GPS signals are generated using the GPS simulator and transmitted via the coupler. The following assumptions are made for the tests conducted:

i) No ionospheric or troposheric delays
ii) No unintended GPS satellite clock or ephemeris error
iii) No obstructions or multipath
iv) No interference signals.

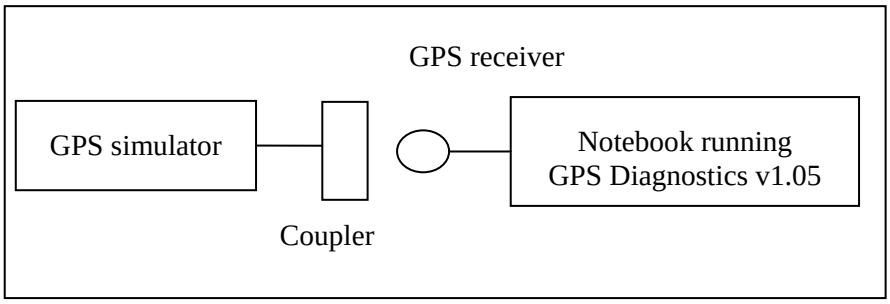


Figure 1: The test setup employed.

The tests are conducted for coordinated universal time (UTC) times of 0000, 0600, 1200 and 1800 for the following coordinates:

i) N 2° 58', E 101° 48', 0 m (Kajang, Selangor, Malaysia)
ii) N 39° 45', W 105° 00', 0 m (Denver, Colorado, USA)
iii) S 16° 55', E 145° 46', 0 m (Cairns, Queensland, Australia)
iv) S 51° 37', W 69° 12', 0 m (Rio Gallegos, Argentina).

The almanac data for the periods is downloaded from the US Coast Guard's web site (USCG, 2016) and imported into the GPS simulator. The GPS signal power level is set at -130 dBm. Readings are taken for GPS antenna orientations of 0 to 345°, at increments of 15°. For each reading, values of estimate probable error (EPE) are recorded for a period of 15 min.

3. RESULTS & DISCUSSION

From the results obtained (Figures 2-4), it is found that there is degradation of accuracy for antenna orientations of 75 to 120° and 240 to 285°. This indicates that for these orientations, the antenna gain is lower, resulting in reduced carrier-to-noise density (C/N_0) levels for GPS satellites tracked by the receivers, which is the ratio of received GPS signal power level to noise density. Lower C/N_0 levels result in increased data bit error rate when extracting navigation data from GPS signals, and hence, increased carrier and code tracking loop jitter. This, in turn, results in more noisy range measurements and thus, less precise positioning (DOD, 2001; Kaplan & Hegarty, 2006; Petovello, 2009; USACE, 2011). For the remaining orientations, the performance remains constant.

These results indicate that the quad helix antennas are operating in endfire and backfire modes simultaneously (Figure 5). While this type of design has smaller antenna gain than quad helix antennas that use only endfire or backfire modes, it allows for a more isotropic antenna performance.

Varying average positional error patterns are observed for the each of the readings. This is due to the GPS satellite constellation being dynamic, causing varying GPS satellite geometry over location and time, resulting in GPS accuracy being location / time dependent (DOD 2001; Kaplan & Hegarty 2006; Dinesh *et al*., 2010).

It should be noted that usage of lower GPS signal power levels, as compared to -130 dBm as used in this study, would result in reduced C/N_0 levels and hence, higher EPE values, as demonstrated in Dinesh *et al*. (2012b).

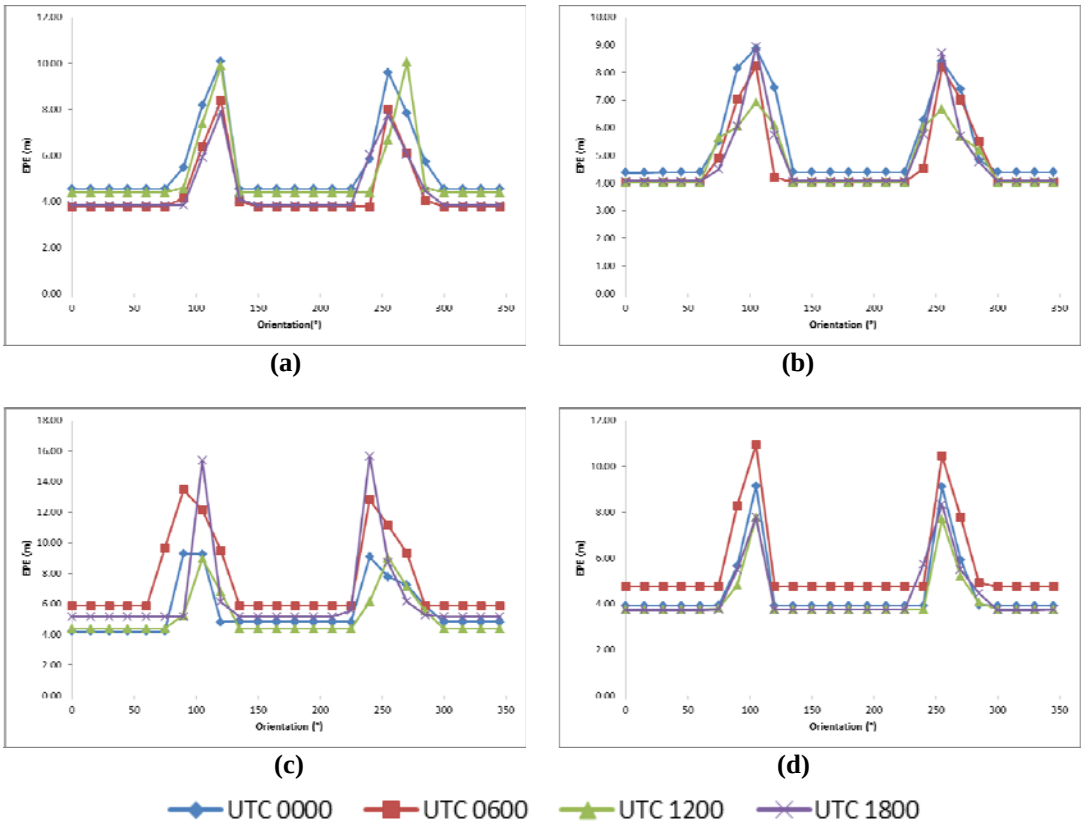


Figure 2: Recorded EPE values for the GPSmap 60CSx receiver: (a) Kajang (b) Denver (c) Cairns (d) Rio Gallegos.

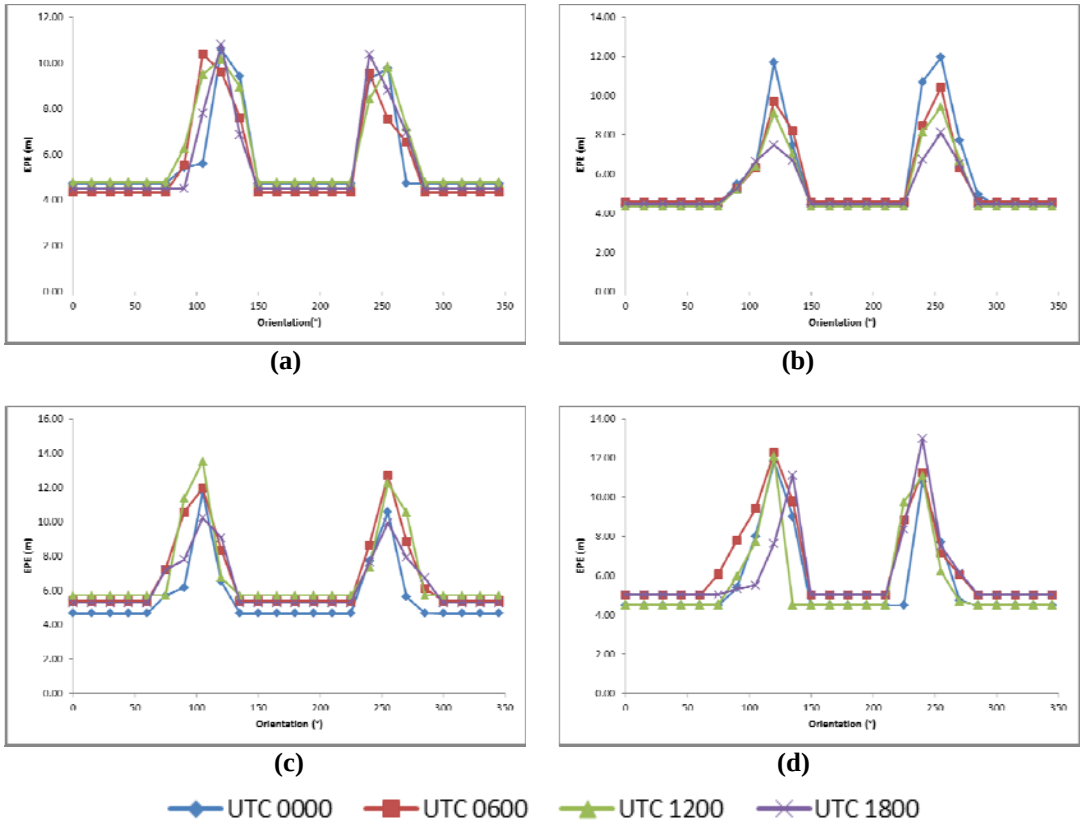


Figure 3: Recorded EPE values for the GPSmap 62Cs receiver: (a) Kajang (b) Denver (c) Cairns (d) Rio Gallegos.

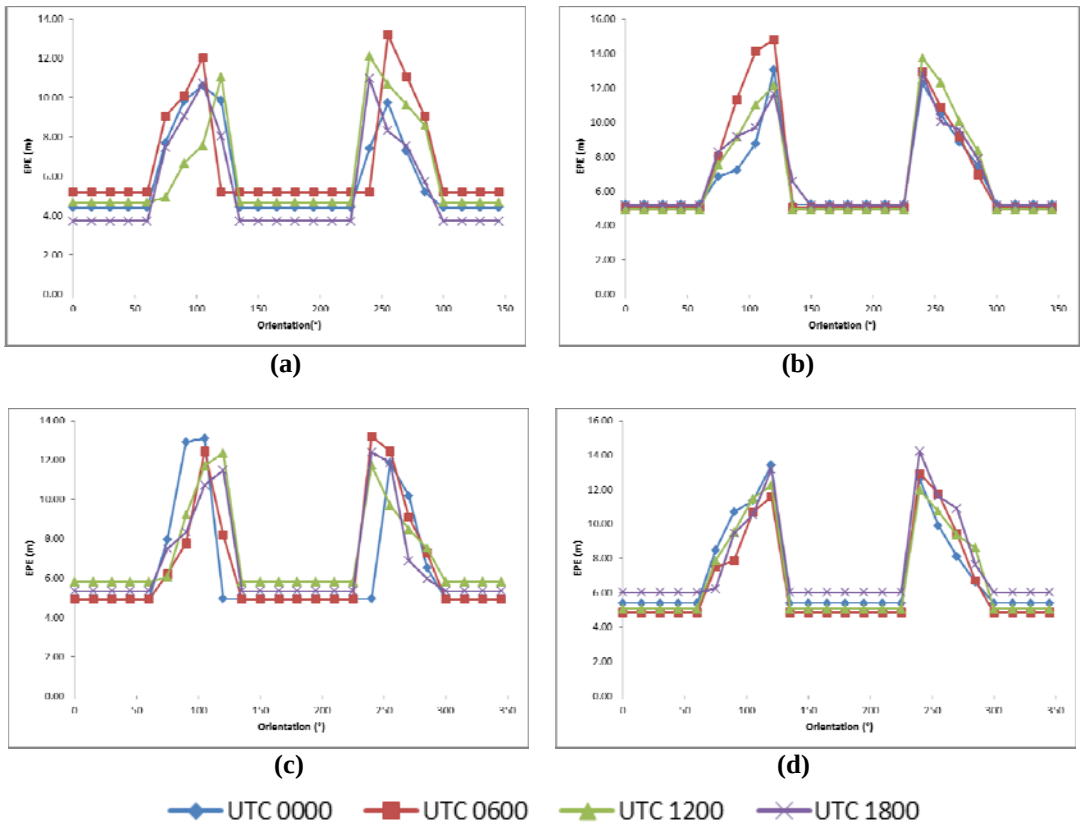


Figure 4: Recorded EPE values for the Oregon 550 receiver: (a) Kajang (b) Denver (c) Cairns (d) Rio Gallegos.

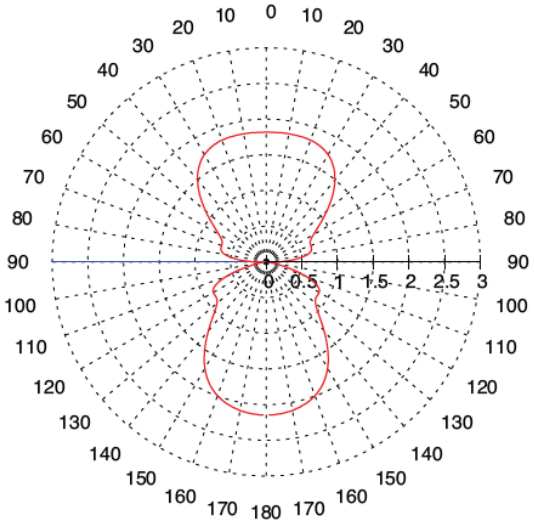


Figure 5: Example of the radiation pattern of a quad helix antenna operating in endfire and backfire modes simultaneously
(Source: Slade, 2015)

4. CONCLUSION

In this study, the effect of GPS antenna orientation on GPS receivers using quad helix antennas was evaluated using GPS simulation. From the results obtained, it was found that there was degradation of accuracy for orientations of 75 to 120° and 240 to 285°. For the remaining orientations, the accuracy remained constant. This indicates that the quad helix antennas are operating in endfire and backfire modes simultaneously. While this type of design has smaller antenna gain than quad helix antennas that use only endfire or backfire modes, it allows for a more isotropic antenna performance. This study will be extended to evaluate the performance of antennas of a wider range of GPS receivers.

REFERENCES

Aeroflex (2010). *Avionics GPSG-1000 GPS / Galileo Portable Positional Simulator*. Aeroflex Inc., Plainview, New York.

Aloi, D.N., Alsliety, M. & Akos, D.M. (2007). A methodology for the evaluation of a GPS receiver performance in telematics applications. *IEEE T. Instrum. Meas.*, **56**: 11-24.

CNET (2004). *GPSDiag 1.0*. Available online at: http://download.cnet.com/GPSDiag/3000-2130_4-4951103.html (Last access date: 31 January 2010).

D'eon, R.G. & Delparte, D. (2005). Effects of radio-collar position and orientation on GPS radio-collar performance, and the implications of PDOP in data screening. *J. Appl. Ecol.*, **42**: 383-388.

Dinesh , S., Wan Mustafa, W.H., Mohd Faudzi., M., Kamarulzaman, M., Hasniza, H., Nor Irza Shakhira, B., Siti Robiah, A., Shalini, S., Jamilah, J., Aliah, I., Lim, B.T., Zainal Fitry, M.A., Mohd. Rizal, A.K., Azlina, B. & Mohd. Hasrol, H.M.Y. (2010). Evaluation of the effect of radio frequency interference (RFI) on Global Positioning System (GPS) accuracy. *Defence S&T Tech. Bull.*, **3**: 100-118.

Dinesh, S, Mohd Faudzi, M. & Zainal Fitry, M.A. (2012a). Evaluation of the effect of radio frequency interference (RFI) on Global Positioning System (GPS) accuracy via GPS simulation. *Defence. Sci. J.*, **62**: 338-347.

Dinesh, S., Mohd Faudzi., M., Rafidah, M., Nor Irza Shakhira, B., Siti Robiah, A., Shalini, S., Aliah, I., Lim, B.T., Zainal Fitry, M.A., Mohd. Rizal, A.K., & Mohd Hasrol, H.M.Y. (2012b). Evaluation of the effect of varying Global Positioning System (GPS) signal power levels on GPS accuracy. *Defence S&T Tech. Bull.*, **5**: 57-71.

Dinesh, S., Shalini, S., Zainal Fitry, M.A. & Siti Zainun, A. (2013). Evaluation of the repeatability of Global Positioning System (GPS) performance with respect to GPS satellite orbital passes. *Defence S&T Tech. Bull.*, **6**: 130-140.

Dinesh, S., Mohd Faudzi, M., Rafidah, M., Nor Irza Shakhira, B., Siti Robiah, A., Shalini, S., Aliah, I., Lim, B.T., Zainal Fitry, M.A., Mohd Rizal, A.K. & Mohd Hasrol Hisam, M.Y. (2014a). Evaluation of the effect of radio frequency interference (RFI) on Global Positioning System (GPS) receivers via GPS simulation. *ASM Sci. J.*,**8**: 11-20.

Dinesh, S., Shalini, S., Zainal Fitry, M.A., Siti Zainun, A., Siti Robiah, A., Mohd Idris, I. & Mohd Hasrol Hisam, M.Y. (2014b). Evaluation of the effect of commonly used materials on multipath propagation of Global Positioning System (GPS) signals via GPS simulation. *Adv. Mil. Tech.*, **9**: 81-95.

Dinesh, S., Shalini, S., Zainal Fitry, M.A., Asmariah, J. & Siti Zainun, A. (2015a). Evaluation of the effect of Global Positioning System (GPS) satellite clock error via GPS simulation. *Defence S&T Tech. Bull.*, **8**: 51-62.

Dinesh, S., Shalini, S., Zainal Fitry, M.A., Asmariah, J. & Siti Zainun, A. (2015b). Evaluation of the accuracy of Global Positioning System (GPS) speed measurement via GPS simulation. *Defence S&T Tech. Bull.*, **8**: 121-128.

Dinesh, S., Shalini, S., Zainal Fitry, M.A., Asmariah, J. & Siti Zainun, A. (2016). Evaluation of trade-off between Global Positioning System (GPS) accuracy and power saving from reduction of number of GPS receiver channels. *Appl. Geomatics*, **8**: 67-75.

DOD (Department of Defence) (2001). *Global Positioning System Standard Positioning Service Performance Standard, Command, Control, Communications, and Intelligence*. Department of Defence (DOD), Washington D.C.

Garmin (2007). *GPSmap 60CSx Owner's Manual*. Garmin International Inc., Olathe, Kansas.

Garmin (2010). *Oregon Series Owner's Manual*. Garmin International Inc., Olathe, Kansas.

Garmin (2011). *GPSMAP C62 Series Owner's Manual*. Garmin International Inc., Olathe, Kansas.

Grimm, D.E. (2012). *GNSS Antenna Orientation Based on Modification of Received Signal Strengths*. PhD Dissertation, ETH Zurich, Zurich.

Gupta, I. (2016). GNSS antennas: The link between satellites and receivers. *Insider GNSS*, **11**: 26-27.

Hautcoeur, J., Johnston, R.H. & Panther, G. (2016). Evolution and revolutionary: The development and performance of the VeraPhase GNSS antenna. *GPS World*, **27**: 42-48.

Kamarulzaman, M. (2010). *Technical Specification for STRIDE's Mini-Anechoic Chamber*. Science & Technology Research Institute for Defence (STRIDE), Ministry of Defence, Malaysia.

Kaplan, E.D. & Hegarty, C.J. (2006). *Understanding GPS: Principles and Applications*. Artech House, Norwood, Massachusetts.

Kou, Y. & Zhang, H. (2011). Verification testing of a multi-GNSS RF signal simulator. *Inside GNSS*, **6**: 52-61.

Moernaut, G.J.K. & Orban, D. (2009). *GNSS Antennas: An Introduction to Bandwidth, Gain Pattern, Polarization and All That*. Available online at: http://gpsworld.com/gnss-systemreceiver-designinnovation-gnss-antennas-8480 (Last access date: 20 June 2016).

Petovello, M. (2009). Carrier-to-noise density and AI for INS / GPS integration. *Inside GNSS*, **4**: 20-29.

Pozzobon, O., Sarto, C., Chiara, A.D., Pozzobon, A., Gamba, G., Crisci, M. & Ioannides, R. (2013). Developing a GNSS position and timing authentication testbed: GNSS vulnerability and mitigation techniques. *Inside GNSS*, **8**: 45-53.

O'Driscoll, C., Izadpanah, A. & Lachapelle, G. (2007). Investigation of the effect of antenna orientation on indoor GPS signal reception. *RIN Nav 2007*, 31 October 2007, London.

Slade, B. (2015). *The Basics of Quadrifilar Helix Antennas*. Orban Microwave Inc., Orlando, Florida.

USACE (US Army Corps of Engineers) (2011). *Engineer Manual EM 1110-1-1003: NAVSTAR Global Positioning System Surveying*. US Army Corps of Engineers (USACE), Washington D.C.

USCG (US Coast Guard) (2016). *GPS NANUs, Almanacs, & Ops Advisories*. Available online at: http://www.navcen.uscg.gov/?pageName=gpsAlmanacs (Last access date: 20 June 2016).

PIEZOELECTRIC TRANSDUCERS IN GUIDED WAVE INSPECTION FOR DEFECT DETECTION IN A STRAIGHT PIPE

Nor Salim Muhammad*, Jong Nyet Nyet, Abd. Rahman Dullah, Ahmad Fuad Ab Ghani & Ruztamreen Jenal

Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka

*Email: norsalim@utem.edu.my

ABSTRACT

The article presents a study on sensitivity of piezoelectric transducers for defect evaluation in a straight pipe. The transducers are placed in configurations of one, two, and four in circumferential direction on aluminum pipe with a circumferential defect. The wave propagation of L(0,2) is used to observe visibility of the defect echo in configurations of one, two, and four piezoelectric transducers. Results from the pulse-echo inspections show that the increase of piezoelectric transducers on the measurement also improved the detectability of the defect in pipe where the measurement at 4 transducers given the significant reflected wave from defect.

Keywords: *Non-destructive testing (NDT); acoustic waves; Lamb wave; guided wave inspection.*

1. INTRODUCTION

Recently, structure health monitoring become important for monitoring of static structures in industries, militaries, and civilian facilities. The failures in storage tanks, pressure vessels, pipes, aircraft structures, and bridges can cause catastrophic damage to the affected communities of a developing country which is not prepared for the maintenance and major services of the facilities. In worst cases, new equipments or facilities are required due to the failure consequences or severity of the failures. In high risk facilities, the failure can cause casualties from explosions or poison gases.

Pipe structures are widely used to transport fluids and gases in petrochemical industries. The structures are also used to control the temperature of the cooling system as found in refinery and power plants facilities. The extensive use of pipe structure has encouraged many studies on the pipe structures compared to the other structures in structure health monitoring. The requirement of health monitoring on pipe structures is also expected to increase proportionally due to the risks of failures in facilities like nuclear fuelled power plants and submarines (Paredes *et al.*, 2012; Koo *et al.*, 2014; Su, 2014; Mike & Migdal, 2014).

Ultrasonic guided wave is one of the robust and effective methods of inspection as the method allows rapid, high accuracy and long distance inspection on pipe structures (Allenye *et al.*, 1997). The method was introduced into industry in 1998, mainly in energy based industry, to inspect pipelines for corrosion, defects and metal degradation (Allenye *et al.*, 2000). Allenye also described the technique as interesting tool as it needs only exposing the area which transducers will be located, hence requiring less insulation removal and digging activity for a buried pipe. Therefore, the use of guided wave technique can significantly reduce the economical aspect of the monitoring program in comparison with other non-destructive testing (NDT) techniques due to its fast screening capability and apart of using smart materials as transducers (Monkhouse *et al.*, 2000; Rose, 2002).

There are several literatures discuss about the difference between conventional inspection method against guided wave inspection method (Demma *et al.*, 2003; Feltcher *et al.*, 2009; Sonia & Fouad, 2010). Dissimilar to standard ultrasonic inspection techniques, utilization of guided waves enable part of pipe structures to be inspected from a single point of access location. Guided wave utilizes ultrasonic waves that transmit along a medium and guided by geometric borders of the medium. The waves are formed by exciting a transducer with high voltage electric pulses which converted to mechanical vibration by a transducer. The waves then propagated through medium under inspection and reflected back when interacts with defects, welds or pipe ends along the wave trajectory. For faults detection in pipes, fundamental torsional mode, T(0,1), and longitudinal mode, L(0,2), are normally used modes. These modes are selected due to non-dispersive behavior over wide frequency bandwidth. Non-dispersive movement means that wave motion do not vary during motion propagation and their phase velocity is not dependent on number of waves.

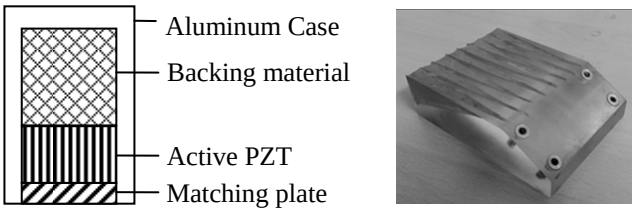
There are few transducers available in guided wave inspections including piezoelectric element, magnetostrictive foil, and electromagnetic acoustic transducer (EMAT). However, transducers of piezoelectric composite materials have the potential to produce large amplitude in excitations and have good sensitivity in detecting reflected waves from defects (Lin & Yuan, 2001; Giurgiutiu, 2003). Their studies also stated that piezo composite materials are formed in the arrangement of piezo ceramic structures with passive polymers like epoxy or other material of active polymers. The element has advantage over general piezo ceramic in term of low acoustic impedance and higher coupling coefficient. Therefore, ultrasonic transducers using piezo composite element has ability to produce high energy of ultrasonic waves in guided wave applications compared to the general piezo ceramic elements. The excitation of high power ultrasonic waves in structures is one of the technique which is used to increase the inspection range on large structures in guided wave inspections. At the same time, the composite transducers also provide high sensitivity to defect detection over the large structures like storage tanks and pipes. However, the cost for the transducers will expand as the structures become larger whereas more transducers are required to generate a specific wave mode in the inspected structures (Allenye *et al.*, 2000). Simultaneously, lack of local expertise on the guided wave technique also contributes to the increase in the total inspection cost.

This paper presents a study on defect inspection in pipes using the developed transducers made of the piezo composite elements. Excitations and measurements of the guided wave are conducted from four types of inspections by using one, two, and four transducers at the end of aluminum pipe to locate a circumferential defect in a straight pipe.

2. EXPERIMENTS

2.1 Preparation of Transducers

The transducers used in this study are developed by using piezoelectric composites of lead zirconate titanate (PZT) from Fuji Ceramics Corp. that designed the active elements to have central frequency approximately at 100 kHz. The structure of the developed transducers is depicted in Figure 1(a) which consists of a matching layer, the PZT active element, and backing material that is made out of tungsten-epoxy material. Figure 1(b) shows the transducer shoe which was designed to excite and receive with predominant to L(0,2) mode in pipes (Salim *et al.*, 2013). Meanwhile, the tungsten-epoxy backing layers were made of tungsten powder at 70% in weight ratio to reduce the internal oscillation of the transducers.



(a) Structure of PZT **(b) Transducer shoe**

Figure 1: Develop PZT transducer.

2.2 Preparation of Defect

Defect on the pipe surface are made by fabricating a 3 mm depth of groove about 100 mm length in circumferential direction with opening groove of 10 mm in width. The actual filed defect was fabricated by using a round metal file as shown in Figure 2. It was placed in the 6 m aluminum pipe at a distance of 5 m from the location of transducers as depicted in Figure 3. The aluminum pipe is made of 6 mm thickness with the inner and outer diameters approximately at 98 mm and 110 mm, respectively.

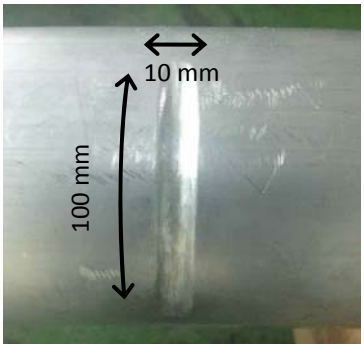


Figure 2: Circumferential defect at 3 mm depth in pipe.

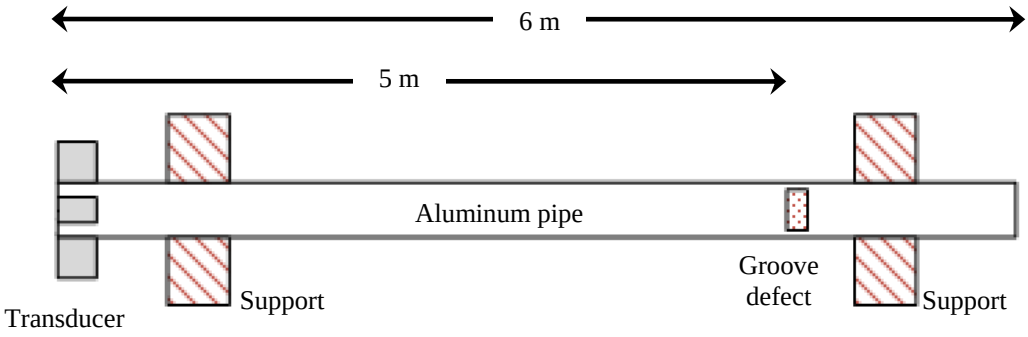
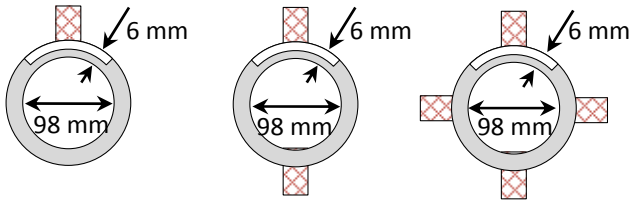


Figure 3: Defect location.



(a) 1 transducer (b) 2 transducers (c) 4 transducers

Figure 4: Transducer arrangements for defect detection in pipe.

At the location of the transducers in Figure 3, the developed transducers are placed on their shoes to form angle beam transducers in configurations of one, two, and four piezoelectric transducers as shown in Figure 4.

2.3 Selection of Wave Mode

Group velocity dispersion curves of the guided wave propagation in 6 mm thickness of aluminum pipe are computed in CIVA computational software for guided wave module developed by the CEA LIST and its partners as plotted in Figure 5 which show the possible propagation waves of L(0,1) and L(0,2) modes. The selection of frequency 100 kHz was made based on the L(0,2) dispersion curve which has small dispersion and large difference in group velocity between L(0,1) and L(0,2) at 100 kHz frequency. At 100 kHz, the plotted curves indicated group velocity of L(0,1) and L(0,2) at about 5200 m/s and 2900 m/s, respectively. The fastest group velocity also indicates shorter travel time of the defect echo which allows it to be decoupled from the complicated wave propagation in pipes which is probably propagating at lower velocities than L(0,2) mode.

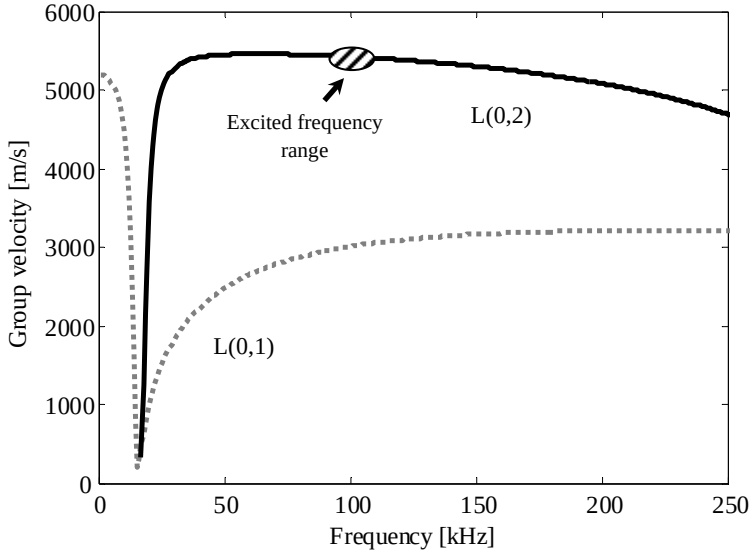


Figure 5: Group velocity dispersion curves of in aluminum pipe at 6 mm thickness.

2.4 Guided Wave Defect Inspection

The measurement system is shown in Figure 6, which consists of a signal generator, power amplifier, diplexer, transducers, preamplifier, data logger, and a host computer to synchronize the measurement. Signal generator (Tektronik AFG3022C) is controlled from a LabVIEW program to synthesize a three cycle of toneburst signal at 100 kHz and power amplifier (NF HAS4052) is used to amplify the

synthesized signal at 40 dB gain. The amplified signal is passed to a diplexer that controls the connection of excitation signal and defect signal. The attached transducers will excite acoustic wave with predominant to L(0,2) mode. The received interaction wave from diplexer is amplified by preamplifier (Ritec PAS-0.1-20) at 20 dB gain. Simultaneous to the excitation of input signal from the function generator, the data logger (NI USB 5133) was triggered by the same LabVIEW program to acquire the signal and transfer the data back to the host computer.

Prior to the test on the defected pipe, the developed transducers were experimented on the non defect pipe to record the signatures of the non defect pipe. The results are compared to the results with defect in pipe for the configurations of one, two, and four transducers as depicted in Figure 4.

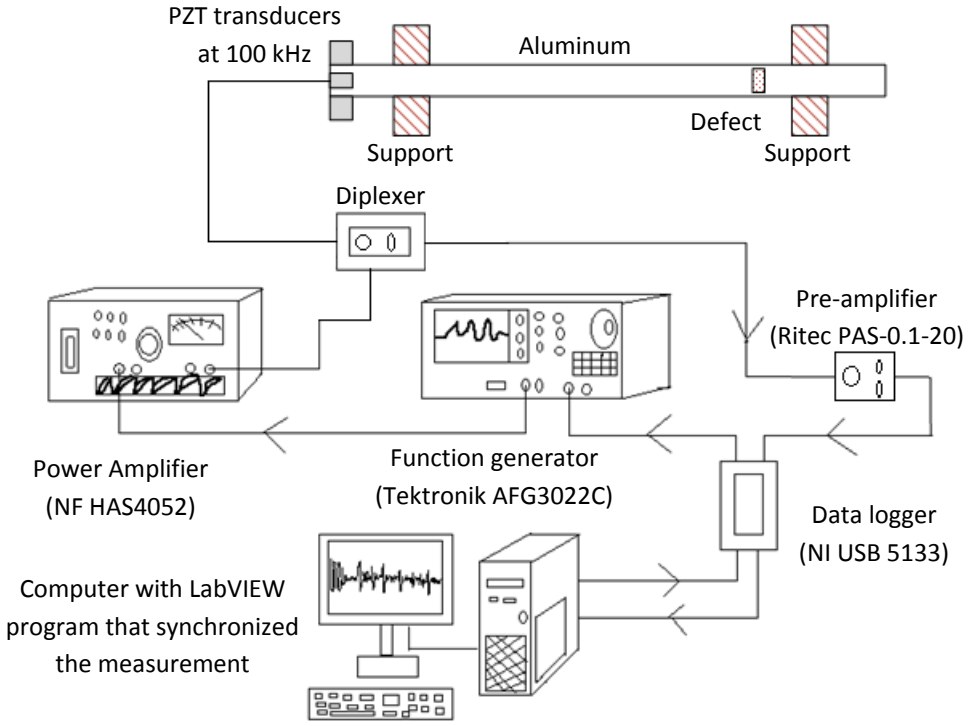
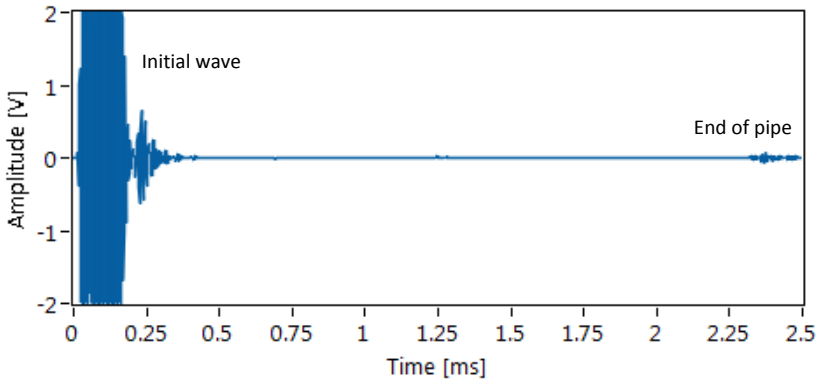


Figure 6: Guided wave pulse-echo inspection in pipe.

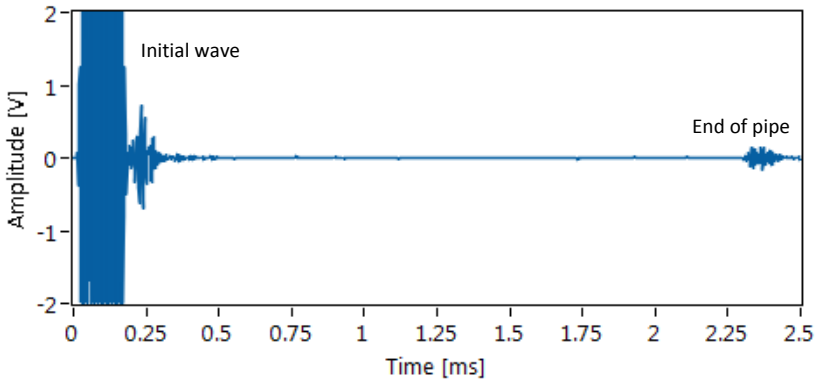
3. RESULTS AND DISCUSSION

Guided wave propagations in non defect aluminum pipe for the configurations of one, two, and four transducers are shown in Figure 7. The time waveforms are recorded in record length up to 2.5 ms to show the reflected wave from the pipe end of the 6 m aluminum pipe. The figures also indicated large oscillation of initial waves up to 0.4 ms for measurement with one, two and four of the developed angle beam transducers. Regardless to the number of transducers used in the measurements, reflections from the pipe end are observed approximately at 2.3 ms for the configurations of one, two, and four transducers. The expected travel time ($t_{\text{end of pipe}}$) of a round trip wave propagation can also be calculated from the total propagation distance ($2l_{\text{pipe}}$) and wave speed of L(0,2) mode ($v_{\text{L(0,2)}}$) as in Eq. 1 where in this experiment the total distance and wave speed are at 12 m and 5200 m/s, respectively. The echoes from the end of pipe also show larger amplitude as the number of transducers increases from one to four transducers which indicates values of 0.07 V, 0.15 V, and 0.30 V at the configurations of one, two, and four transducers, respectively.

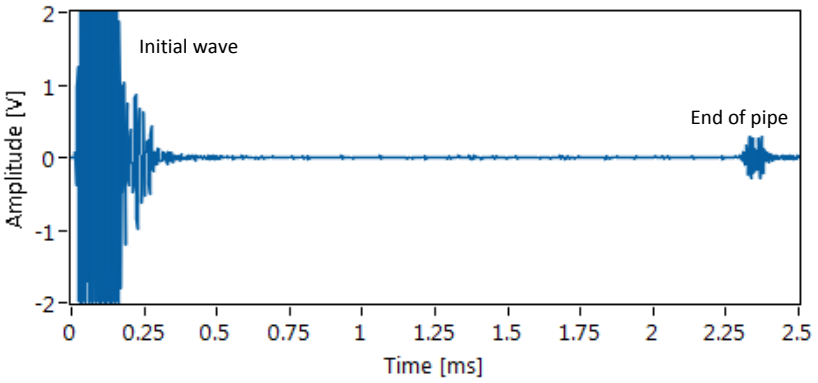
$$t_{\text{end of pipe}} = \frac{2l_{\text{pipe}}}{v_{\text{L(0,2)}}} \tag{1}$$



(a) 1 transducer.



(b) 2 transducers.



(c) 4 transducers.

Figure 7: Wave propagation in the aluminum pipe (without defect).

Guided wave propagations in pipe with the groove defect for the configurations of one, two, and four transducers are shown in Figure 8. The attachment of transducers on the pipes with and without defect can result into different contact conditions between the first and second experiments which can affect into the sensitivity of the transducers. Since there are other possibilities which are contributed to the difference in amplitude level, amplitude comparison are not to be discussed between the non defect and defected pipes.

The results in Figures 7 and 8 show close similarity between the two pipes in the reflections from the pipe end which appeared approximately at 2.3 ms. It also indicates that the presence of defect in the pipe will show reflected wave earlier then travel time for the round trip wave propagation. This agreed to the results from wave propagations in pipe with defect that indicated defect echoes at about 1.93 ms. The expected location of defect ($l_{\text{defect location}}$) can be evaluated from its travel time ($t_{\text{defect echo}}$) and wave speed of L(0,2) mode ($v_{\text{L(0,2)}}$) as in Eq. 2 which suggests a location approximately at 5.02 m from the fixed transducers.

$$l_{\text{defect location}} = \frac{t_{\text{defect echo}} \times v_{\text{L(0,2)}}}{2} \tag{2}$$

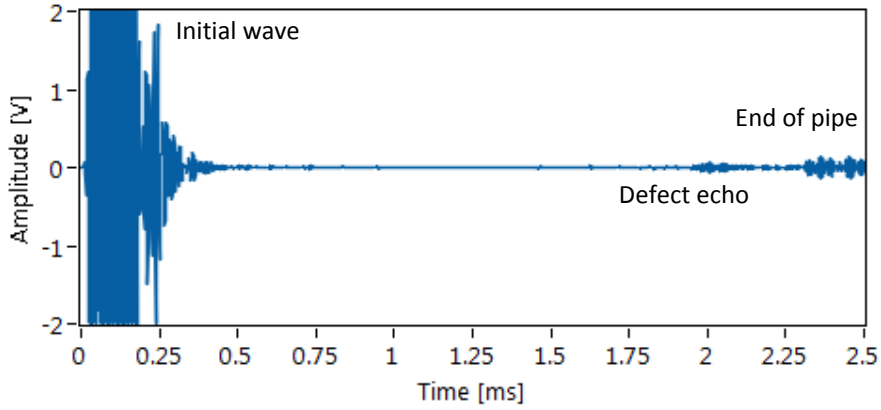
Amplitudes of defect echoes are also obtained at 0.07 V, 0.16 V, and 0.13 V for the configurations of one, two, and four transducers, respectively. The pattern of amplitudes are observed to be complicated than the reflected wave from the end of pipe which show similarity to the pattern in the experiment on pipe without defect. Although result in configuration of four transducers show slightly lower amplitude of defect echo than two transducers, defect echo of four transducer seem to have better time resolution with small dispersion behavior that reflect to the characteristic of L(0,2) at 100 kHz.

Defect echoes obtained from the measurements are zoomed in Figure 8 to detail the difference in the inspection sensitivity when using one, two, and four transducers. The zoomed time waveforms show small amplitude with large dispersion behavior for measurement using one transducer as in Figure 8(a). The use of another transducer on the opposite side on pipe shows higher amplitude level of defect echo but with large dispersion behavior as in Figure 8(b). The use of four transducers in the experiment significantly increases the sensitivity of defect detection from one and two transducers as in Figure 8(c) where the amplitude level increased in small dispersion of wave packets.

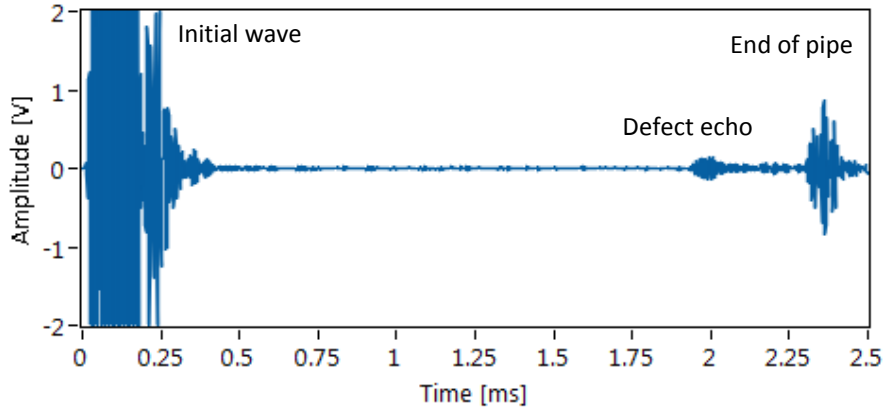
In configuration of four transducers, there are two wave packets which arrived at travel time indicated as t_0 and t_1, respectively. From the measurement data, t_0 and t_1 are recorded at 1.93 ms and 1.97 ms, respectively. The visibility of two separate wave packets at the record time between 1.9 ms to 2.0 ms indicates an increase in time resolution compared to the other configurations with less transducers.

The presence of two wave packets in the time waveform between 1.9 ms to 2.0 ms can be explained in Figures 10 and 11 which possibly caused by the complicated reflections from top and bottom edges of the groove defect (Salim *et al.*, 2009) as in Figure 10 or internal reflection as depicted in the transducer shoe (Jezzine & Lhémery, 2006) in Figure 11. The complicated wave propagation in pipe is difficult to be explained from the experiment results and is required further observation in future studies. However, the increase of transducers in circumferential direction which increased in pairs has proven to has effect in increasing the sensitivity of the defect detection in a straight pipe as stated in (Rose, 1998) and Allenye *et al.* (2000).

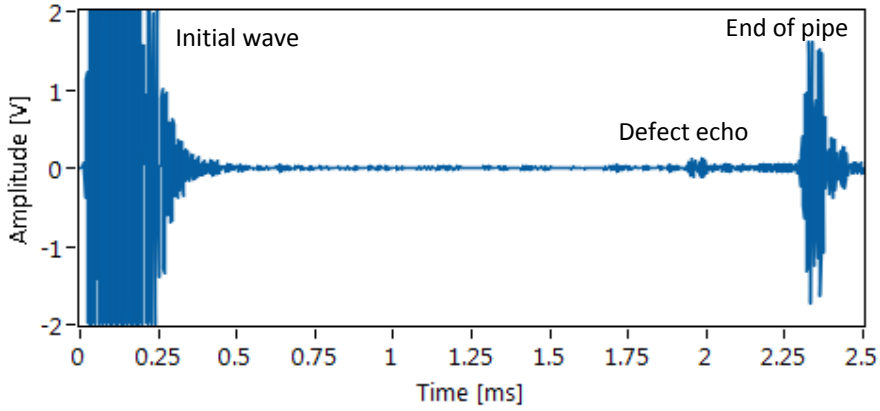
The results of our study may not be as good as studies conducted in developed countries. However, it may provide information on the potential use of guided waves to detect or find defects in the structures. Especially for those who are planning to start a study in a guided wave technique. Although the waves propagate in different wave modes in structures like plates and rails. Similar structures like tubes or hollow cylindrical structures which found in defense facilities may apply the similar technique as in pipe structures for monitoring of the structures degradation (Buderath, 2009).



(a) 1 transducer.

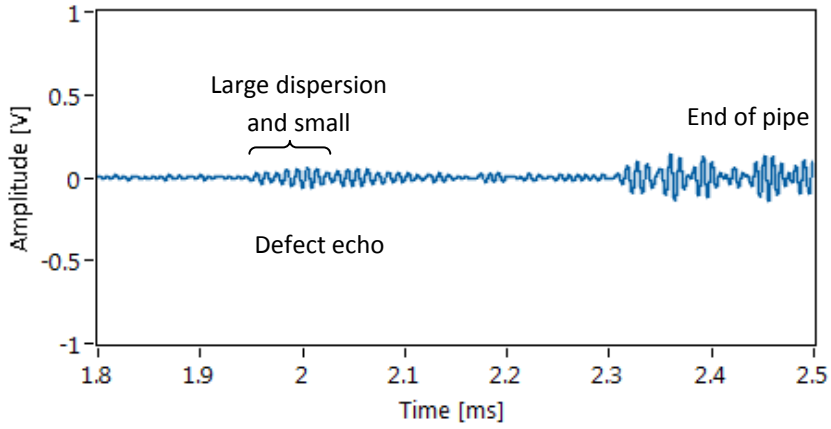


(b) 2 transducers.

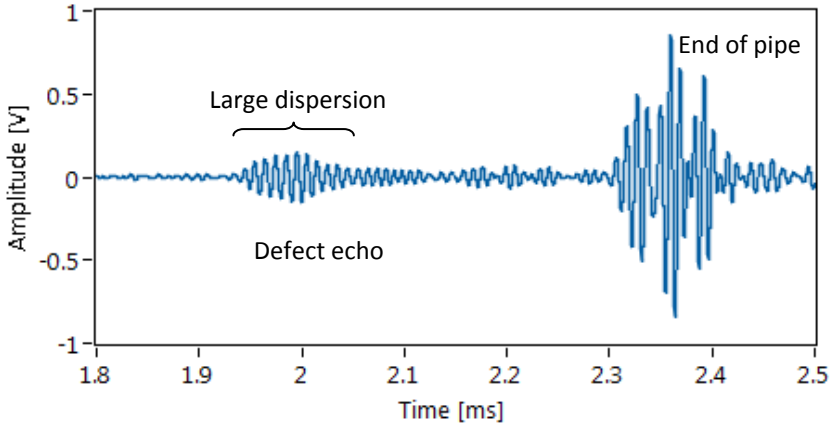


(c) 4 transducers.

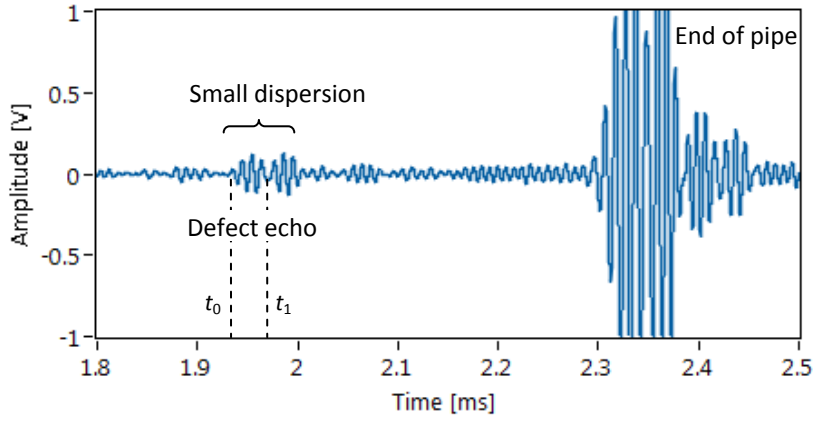
Figure 8: Wave propagation in the aluminum pipe (with 3mm defect depth).



(a) 1 transducer.



(b) 2 transducers.



(c) 4 transducers.

Figure 9: Zoom on time waveforms in the aluminum pipe (with 3mm defect depth).

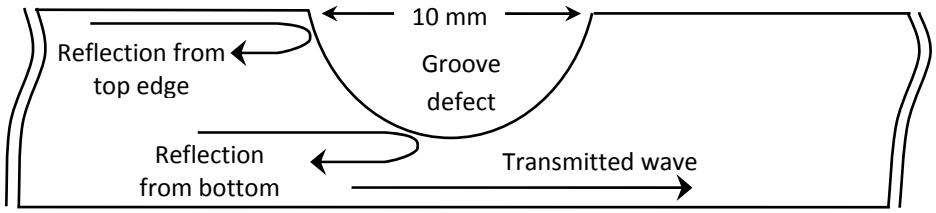


Figure 10: Expected reflection waves around defect.

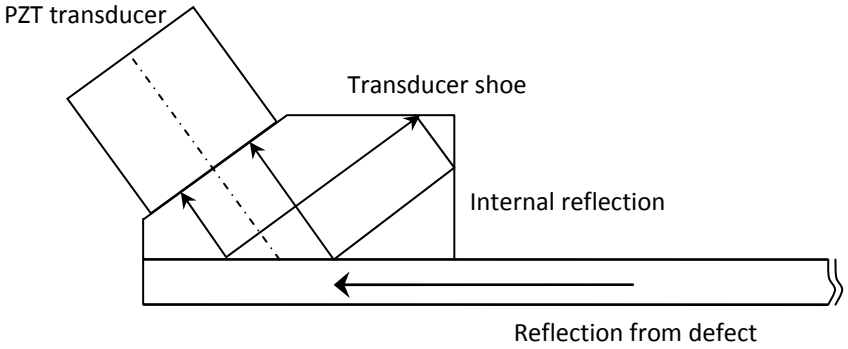


Figure 11: Expected internal reflection waves in transducer shoe.

4. CONCLUSION

Guided wave inspections had been conducted using L(0,2) mode in configurations of one, two, and four in circumferential direction on a straight pipe which made of aluminum pipe with a circumferential defect. The guided wave propagation in pipe show slightly complicated wave propagation in the measured time waveform but the results show that the increase of transducers in circumferential direction which increased in pairs as two and four transducers has improved the sensitivity of the defect detection in the straight pipe.

ACKNOWLEDGMENT

This work is supported by Universiti Teknikal Malaysia Melaka (UTeM) and Malaysia Ministry of Higher Education (MOHE) under research grant ERGS/1/2013/FKM/TK01/UTEM/02/02/E00015, FRGS/2/2013/TK01/UTEM/02/F0171, and FRGS/2/2013/TK01/UTEM/02/F00174.

REFERENCES

Alleyne D.N., Cawley P., Lank A.M. & Mudge, P. J. (1997). The Lamb wave inspection of chemical plant pipework. *Rev. Prog. in Quantitative NDE,*, **16**: 1269-1276.

Alleyne D.N, Pavlakovic B., Lowe M.J.S. & Cawley P. (2000). Rapid, long range inspection of chemical plant pipework using guided waves. *Rev. Prog. in Quantitative NDE*, **20**: 180-187.

Demma, A., Cawley P. & Lowe M.J.S. (2003). The reflection of the fundamental torsional mode from cracks and notches in pipes. *J. Acoust. Soc. Am.*, **114**: 511.

Fletcher S., Lowe M.J.S, Ratassepp M. & Brett C. (2009). Detection of axial cracks in pipes using focused guided waves. *Rev. Prog. Quantitative NDE*, **29**: 56–59.

Giurgiutiu V. (2003). Lamb wave generation with piezoelectric wafer active sensors for structural health monitoring. Proc. SPIE, **5056**: 111-122.

Jezzine, K. & Lhémery, A. (2006). Diffraction effects on ultrasonic guided waves radiated or received by transducers mounted on the section of the guide. *Rev. Prog. in Quantitative NDE*, **25**: 134–141.

Koo Y.H., Yang Y.S. & Song K.W. (2014). Radioactivity release from the Fukushima accident and its consequences: A review. *Prog. Nucl. Energy*, **74**: 61-70.

Lin X. & Yuan F.G. (2014). Diagnostic Lamb waves in an integrated piezoelectric sensor/actuator plate: analytical and experimental studies. *Smart Mater. Struct*., **10**: 907–913.

Mike A. & Migdal A. (2014). *Learning from the Kursk Submarine Rescue Failure: The Case for Pluralistic Risk Management*. Havard Business School, Boston, Massachusetts

Monkhouse R.S.C., Wilcox P.W., Lowe M.J.S., Dalton R.P. & Cawley P. (2000). The rapid monitoring of structures using interdigital Lamb wave transducers. *Smart Mater. Struct*., **9**: 304-309.

Paredes G.E., Batet L., Carrera A.N. & Sugimoto J. (2012). Severe accident analysis in nuclear power plants. *Sci. Tech. Nucl. Install.*, **2012**: ID 430471

Rose J.L. (2002). A baseline and vision of ultrasonic guided wave inspection potential. *J. Press. Vessel Tech.*, **124**: 273 – 282.

Rose J.L., Pelts S.P. & Quarry M.J. (1998). A comb transducer model for guided wave NDE. *Ultrasonics*, **36**: 163-169.

Salim M.N, Hayashi T. & Murase M. (2009). Visualization and modal analysis of guided waves from a defect in a pipe. *Japanese J. Appl. Phys.*, **48**: DOI: JJAP.48.07GD06.

Salim M.N., Jenal R.B. & Ismail M.P. (2013). Reflected guided wave in plate with flaw using CIVA simulation software. *6th Int. Eng. Conf. Ener. Environ.*, pp. 357-363.

Sonia D. & Fouad B. (2010). Propagation of guided waves in a hollow circular cylinder application to non destructive testing. *EDP Sci.*, **6**: DOI:10.1051/epjconf/20100615004.

Su D.T. (2014). Hazard analysis of underwater vehicles from submarine casualties. *Int. J. Sci. Commerce Humanities*, **2**: 68-76.

Buderath M. (2009). Fatigue monitoring in military fixed-wing aircraft. Encyclopedia of Structural Health Monitoring, *John Wiley & Sons*. **4** (97).

STRESS DETERMINATION BY USING OUT-OF-PLANE DEFLECTION WITH SCANNING LASER DOPPLER VIBROMETER

Tino Hermanto, Abd. Rahman Dullah, Ruztamreen Jenal* & Nor Salim Muhammad

Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka

*Email: ruztamreen@utem.edu.my

ABSTRACT

Scanning Laser Doppler Vibrometer is one of the technologies in non-contact measurement that is capable of measuring deflection of structure. This project isto determine stress on plate structure due to vibration with free boundary conditions by using non-contact out-of-plane single head scanning laser doppler vibrometer measurement. The first vibration mode frequency is used to vibrate the plate meanwhile the single head scanning laser vibrometer is used to measure out-of-plane deflection of the plate surface. The measured deflection values are then used to calculate stresses at the plate surface by using the relation of stress with bending formation. Furthermore, direct strain values of the plate have been measured at certain bending formation that corresponded to the bending formation due to the first vibration mode excitation. This experiment method shows that the stress determination by using out-of-plane deflection is possible by using single head scanning laser doppler vibrometer.

Keyword*: Bending stress; Scanning laser doppler vibrometer.*

1. INTRODUCTION

Structural Health Monitoring (SHM) is an application of inspection and characterization of damage for engineering structures. In the past, the main focus of SHM is on military aircraft applications with recent development of non-contact measurement (Goggin *et al*., 2003), the use of laser measurement has become more practical and precise in SHM. The advantages for non-contact laser measurement were summarized as wavefield images, pristine condition of a target structure, and applicable to harsh environments and require less maintenance (Shon*et al*., 2016). In 2015, the UK Military Aviation Authority (MAA) has integrated SHM technology or system into the UK military aircraft (Azzam *et al*., 2015).

Scanning Laser Doppler Vibrometer (SLDV) is used for non-contact measurement, visualization, analysis, and to measure vibration of a structure or component with very high spatial resolution (Vanlanduit *et al*., 2004; Omidi*et al*., 2015; Khalil *et al*., 2016). It determines the operational deflection shapes as easily as taking a video. The SLDV is commonly used in acoustic, architectural, aerospace and automotive structural dynamics analysis (Omidi*et al*., 2015; Kitazawa *et al*., 2016, Shin *et al*., 2017).

The advantage of SLDV over similar measurement devices such as an accelerometer is that the SLDV can be directed at the targets that are difficult to access (Avitabile *et al*., 2010). Over recent years, 3D-SLDV was used to investigate into the accuracy of strain measurement (Wildy *et al*., 2010), the dynamic strain over the surface of a planar structure (Cazzolato *et al*., 2008), and multiaxial fatigue (George *et al*., 2004; 2005; 2006). The dynamic stress distribution between SLDV and strain gage measurement was also compared (Schüssler *et al*., 2011). The result shows that the feasibility of using dynamic in-plane for measurement of dynamic strain field is possible by using 3D displacement (Weisbecker *et al*., 2012). This paper presents the strain and stress measurement based on a single head SLDV and the validation of the calculated strain by strain gauge measurement.

2. EXPERIMENTAL SETUP

Two experimental setups were used in this paper. The measurement of plate vibration by using single head SLDV and measurement of strain from three point bending test.

2.1 Scanning Laser Doppler Vibrometer (SLDV)

The experiment was performed on an aluminium Al-2024 plate specimen with the dimensions of 150 mm x 400 mm with thickness 2 mm. The materials commonly used in aircraft structure such as fuselage structurals, wing tension members, shear webs and ribs and structural areas. It possess good fatigue resistance, average machinability and high strength. The properties are shown in Table 1.

Table 1: Material properties of the Al-2024 plate.

Property	Detail
Density	2780 kg/m^3
Young Modulus	72400 MPa
Poisson Ratio	0.33

Figure 1 shows the primary experiment set-up for the measurement of aluminium plate deflection by using a single head SLDV. The plate is hanged by a rope on the bar to create free boundary conditions at all edges. A mechanical shaker (TIRA GmbH type S 50018, rated force sine 18 N) is attached at bottom corner of the plate and being amplified by a power amplifier (TIRA type BAA 60). The trigger signal for the amplifier is controlled by a signal generator (Tektronix AFG 3022).

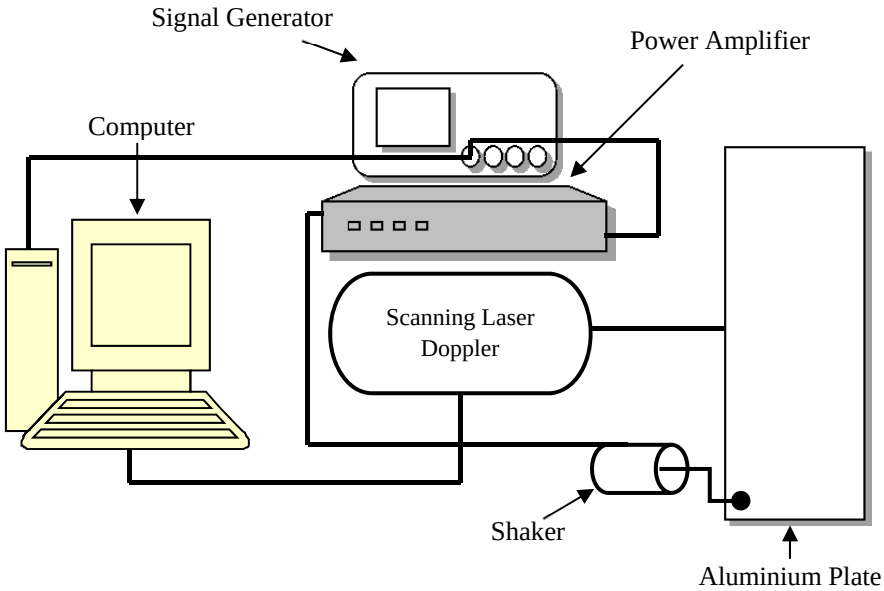


Figure 1: Test setup for modal analysis.

The set-up was further explained in Figure 2. The single head SLDV beam was directed at the center of the plate surface at a distance of around 1370 mm. The setting parameters and configuration of the single head SLDV are shown in Table 2.

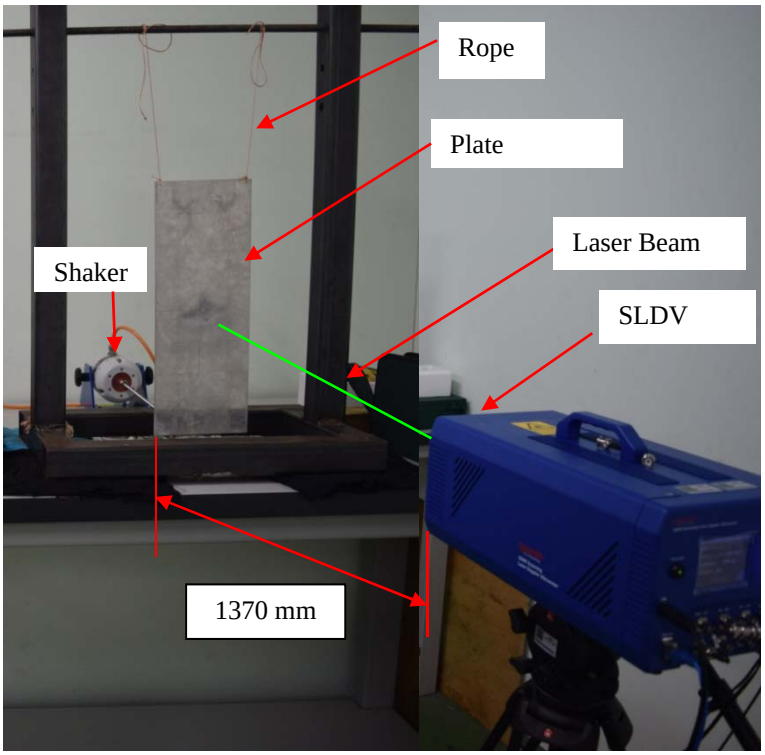


Figure 2: Single head SLDV experimental set-up.

Table 2: Parameters and configuration for the single head SLDV.

Item	Parameter
Scanning area	140 mm x 390 mm
Scanning point	126
Velocity resolution	122.5 mm/s

The output of the laser is recorded through USB and Ethernet cables into a personal computer (PC) and the result was later analyzed by using Vibrometer Laser Scanner (VL Scanner) software. The input signal from signal generator was also recorded in the PC for frequency response analysis.

2.2 Three Point Bending Test

In order to compare the result of plate stress from the SLDV measurement due to vibration, three points bending test was performed as below.

Figure 3 shows the schematic diagram for the test that consists of the aluminium plate specimen being supported by simple supports. The strain gauge (KR 120-3p-23-1L and gauge factor is 2.08) and dial gauge were placed at the middle of the plate. Weights are used to apply load at the center of the plate until the bending deflection value on dial gauge is equal with deflection of point 3 at SLDV measurement. The strain values on strain meter (TML Digital Static Strain meter TC-31K DC) are recorded when the deflection are similar to the result from the SLDV measurement.

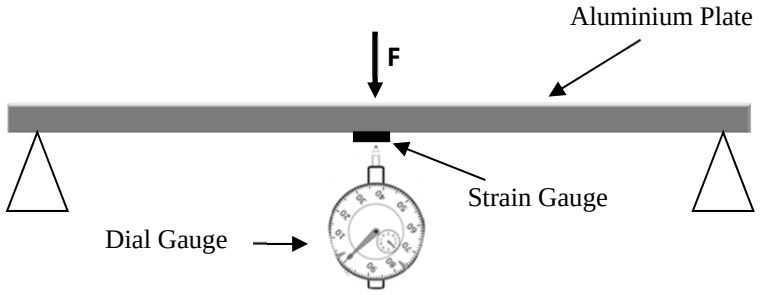


Figure 3: Schematic of the experimental set-up for bending test.

The stress at center of the plate was calculated by replacing at the measured strain is written as

$$\sigma = \mathrm{E}\ \varepsilon \tag{1}$$

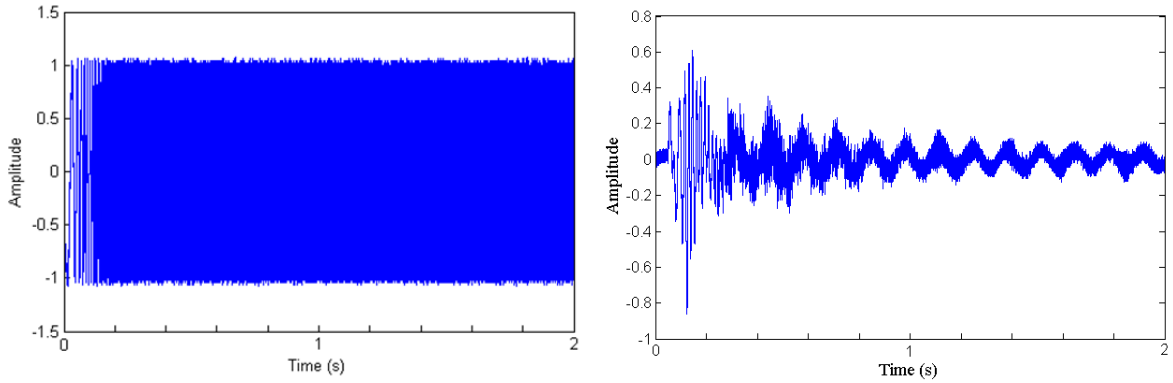
where ,E is Young's modulus of the aluminium plate. Strain (ε) is the amount of deformation of a specimen due to applied force.

3. MODAL ANALYSIS

In order to determine the resonance frequency of the plate, modal analysis was performed on the plate. The experimental set-up is as explained in the previous section with signal generator parameters were set as in Table 3. Figure 4 show an example of the plate response in time domain for the input and output signals. The responses of the plate from the sweep frequency excitation were captured by single head SLDV and recorded into PC.

Table 3: Parameters of signal generator for modal analysis.

Item	Setting value
Sweep frequency	1-2000 Hz
Sweep time	2 s
Amplitude	2 Vpp
Sampling size	100 K sample



(a) **(b)**

Figure 4: (a) input signal and (b) output signal in time domain.

The signals were then converted into frequency domain by using transfer function defined by Matlab as

$$T_{xy} = tfestimate(x, y) = \frac{P_{yx}}{P_{xx}} \tag{2}$$

where T_{xy} is the estimated transfer function of the input signal vector x and output signal vector y. P_{yx} is cross power spectral density of x and y, and P_{xx} is the power spectral density of x.

Figure 5 shows the results of the frequency response for the plate by using single point measurement. The figure shows a few peaks that can be identified as the vibration mode frequency for the plate. The first six vibration modes frequencies from Figure 5 are 87 Hz, 152 Hz, 208 Hz, 258 Hz, 315 Hz and 428 Hz as listed in Table 4.

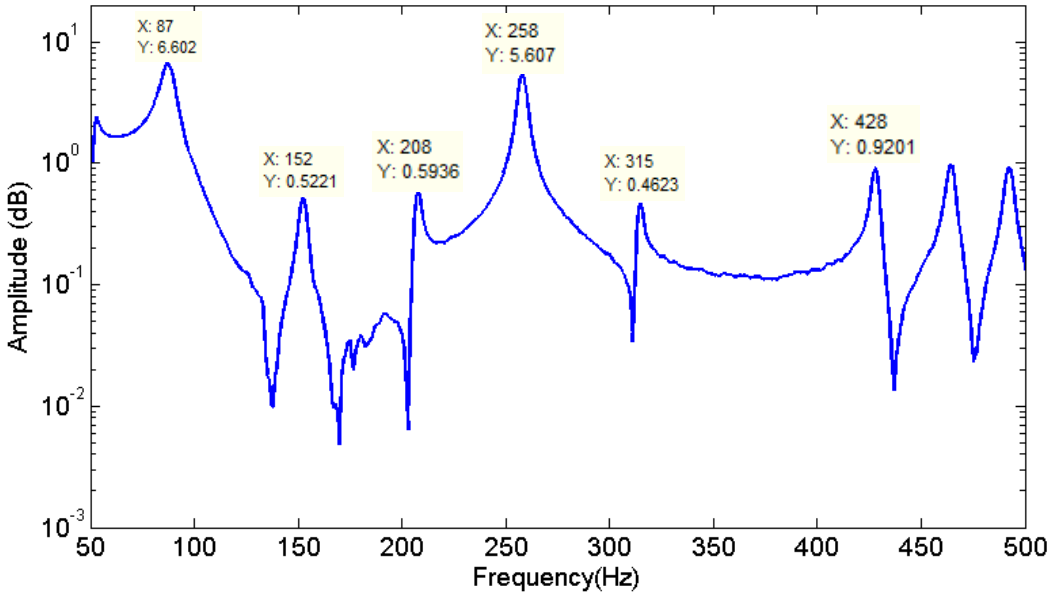


Figure 5: Frequency response function.

Table 4: List of vibration mode frequency for the plate.

Vibration mode	Frequency (Hz)
1	87
2	152
3	208
4	258
5	315
6	428

Figure 6 shows an example of vibration mode shape for the first vibration mode at 87 Hz. The first vibration mode shape shows the maximum deflection happened at the center of the plate and the deflection is reduced at both upper and bottom partsof the plate gradually. This criterion of deflection is the first bending mode for the rectangular plate. The result was obtained by using scanning measurement with resolution of 126 points for 140 mm x 390 mm area.

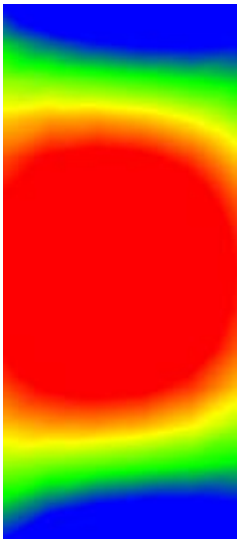
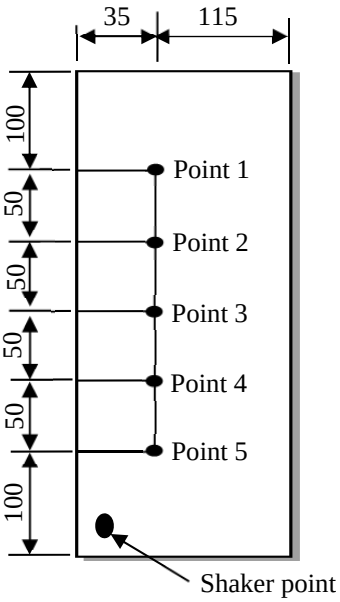


Figure 6: Vibration mode shape for the first vibration mode at 87 Hz.

4. DETERMINATION OF STRESS FROM SLDV MEASUREMENT

Figure 7 shows the location of point 1 to 5 on the aluminium plate for the single point measurement. Point 1 is located at 100 mm from top edge and 35 mm from left edge. The plate was vibrated at bottom left corner of the plate at vibration power level 1 Vpp to 5 Vpp at 87 Hz frequency (1st vibration mode) obtained from the previous experiment.



Note: All dimensions in mm

Figure 7: The points for measurement.

Table 7 shows the result of the single point deflection measurement at point 1 to 5 with 5 Vpp vibration power level. As expected, point 3 shows the highest deflection point at 1.96 mm and gradually decreasesfor the other points. From the table, the plate was deflected as first bending mode shape since it was vibrated by the first vibration mode frequency.

Table 7: Deflection at point 1 to 5 for the plate vibrated at 5 Vpp.

Point of measurement	Deflection (mm)
1	1.2
2	1.48
3	1.96
4	1.44
5	1.28

The deflection at point 3 due to vibration level 1 Vpp to 5 Vpp is shown in Table 8. The result shows that the deflection gradually increased as the excitation power level increaseswith the maximum deflection is 1.96 mm at 5 Vpp.

Table 8: Deflections at point 3.

Excitation power (Vpp)	Deflection (mm)
1	0.56
2	1.04
3	1.44
4	1.76
5	1.96

The stress in the plate is calculated by using Hooke's Law (Beer *et al*., 1992), where:

$$\sigma = -\mathrm{E}\left(\frac{y}{\rho}\right) \tag{3}$$

Here, E is the modulus of elasticity, y is the distance between the plate surface with its neutral axis and ρ is the radius of bending arc or curvature radius. Using this equation,$y$is half of the plate thickness, the ρ can be calculated by using equation (4).

$$\rho = \frac{\delta_0}{2} + \frac{{L_0}^2}{8\delta_0} \tag{4}$$

Where, δ_0 is the deflection at center of the plate surface, L_0 is the length of arcof the plate deformation. Since the plate at all edges is in free boundary condition, the δ_0 is determined by the difference between maximum and minimum deflections. L_0 is the distance between measurements points 1 to 5 which is 200 mm. Meanwhile the δ_0 is the difference between measured deflections at point 3 (middle of plate) with point 1 or 5 that is the lowest deflection as shown in Figure 8.

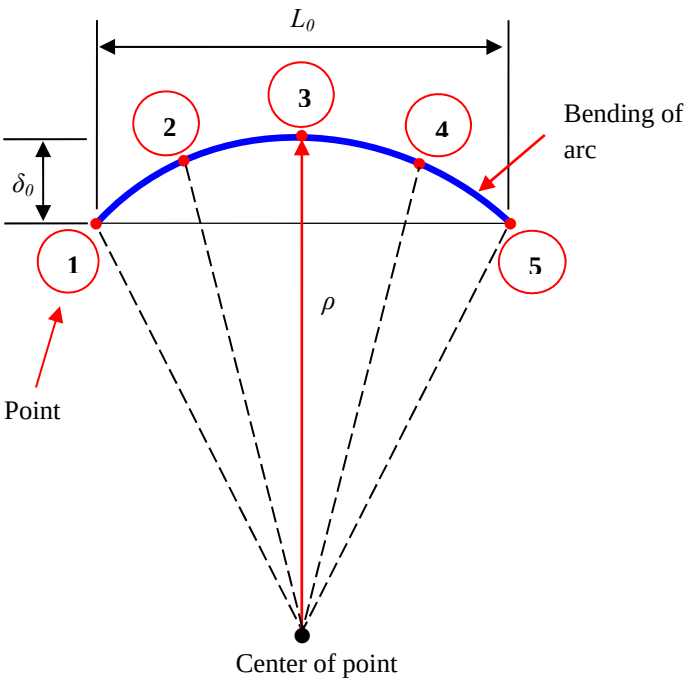


Figure 8: Bending mode deformations.

5. RESULTS AND DISCUSSION

Figure9 shows the deflection at measurement point 1 to 5 with vibration power level 1-5 Vpp. The maximum deflection for allexcitation power levels happened at measurement point 3 which is at the center of the plate. The deflection at all measurement points increased linearly with the increment of the excitation power level. At both left and right of measurement point 3, the deflections were gradually decreasing and resemble a first bending mode shape. The measured deflections at point 3 arc then were used to calculate bending arc and stress by using equations (3) and (4) respectively.

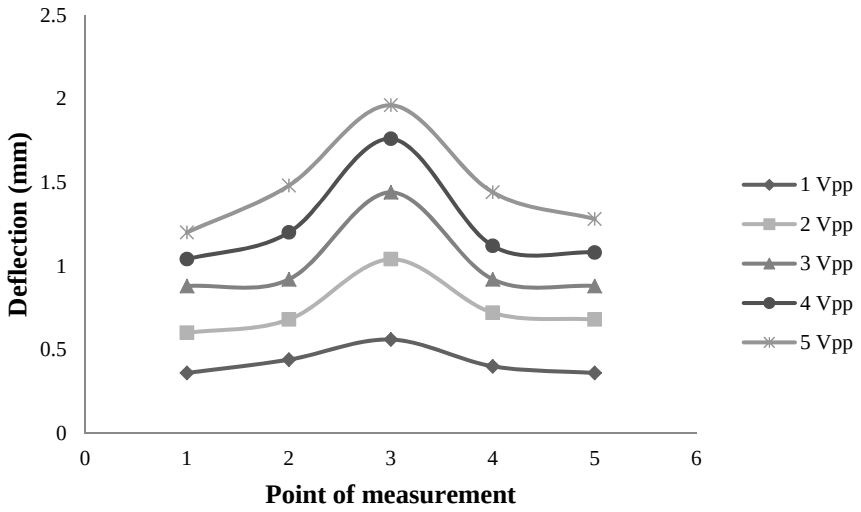


Figure 9: Deflection results at 5 measurement points due to 1st vibration mode frequency excitation excited by various excitation power levels.

By replacing the result from Table 8 into equation (4), Table 9 shows the result of the plate deflection (δ_0) at vibration power level 1 Vpp to 5 Vpp with it corresponding radius of bending arc and stress. As expected the stressesareincreased as the power levels increased andthe maximum stress is 11.00 MPa.

Table 9: The result between deflection (δ_0) and radius of bending arc.

Excitation power (Vpp)	Deflection, δ_0 (mm)	Radius of bending arc, ρ (mm)	Stress (MPa)
1	0.2	25000.1	2.90
2	0.44	11363.9	6.37
3	0.56	8928.9	8.11
4	0.72	6944.8	10.43
5	0.76	6579.3	11.00

Table 10 shows the result of the strain gauge measurement based on reference to single head SLDV deflection reading at center of the plate. The result shows that the strain is gradually increased as the deflection increased where the maximum stress is 11.08 MPa at 1.96 mm deflection.

Table 10: List of stress from the bending test.

Deflection (mm)	Strain ($\mu\varepsilon$)	Stress (MPa)
0.56	**50**	**3.62**
1.04	**98**	**7.10**
1.44	**119**	**8.62**
1.76	**145**	**10.50**
1.96	**153**	**11.08**

The stress versus deflection comparison for the single head SLDV measurement and three point bending test is shown in Figure10. The figure shows the result of stress at the center of the plate (point 3) against deflection values. At low deflection, the difference of the stress between both experiments is 0.24% and the difference is reduced as the deflection increased. At 2 mm deflection, the stress values between both measurements are nearly similar. These results show that the measurement of stress by using out-of-plane deflection measurement with single head SLDV is possible and nearly similar with the measurement by using strain gauge.

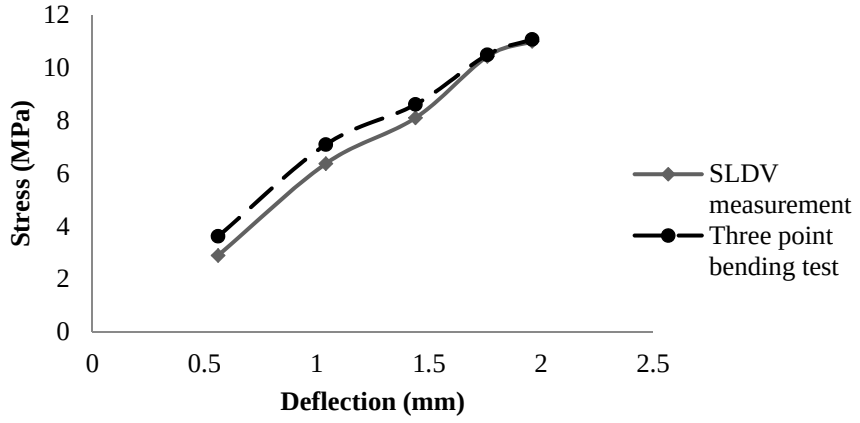


Figure10: Comparison of the stress measured by single head SLDV and calculated based on three point bending test.

The SLDV is very useful and gives precise measurement by using non-contact technology. However, thestudy is only for the first vibration mode stress measurement only. Further analysis is required for the other vibration modes due to the complexity of the plate responses against the vibration modes.

Comparing the result from Weisbecker *et al*., it is shown that the measurement of in-plane strain over the surface is possible. Schüssler *et al*. also shows that the 3D-SLDV is possible to measure strain at high frequency level. However, this study is only limited to measure out-of-plane strain with single vibration frequency or at first vibration mode frequency.

6. CONCLUSION

In summary, a single head SLDV is capable of measuring surface in dynamic condition. The result of stress measured by out-of-plane deflection by using a single headSLDV is nearly similar with result from strain gauge measurement. The single head SLDV is also able to measure strains on the surface of structure. This technique however only limited tothe first bending mode shape of plate.

The single head SLDV is also capable of measuring dynamic of stress on structure precisely without interfering the structure dynamic response and to plot certain of the vibration mode shape by using the scanning capability. This was shown by high agreement between single head SLDV and strain gauge measurements.

However, the single head SLDV measurement has limitation if compared withthe measurement by using 3D-SLDV. It is only limited to out-of-plane stress measurement with single frequency excitation or only limited at first vibration mode frequency excitation. But with minimum cost of measurement, it is possible to measure the surface dynamic stress particularly at certain surfacein the study of SHM.

ACKNOWLEDGEMENT

Authors would like to thank Universiti Teknikal Malaysia Melaka for the Postgraduate Fellowship Scheme awarded to the first author.

REFERENCES

Avitabile, P., Niezrecki, C., Helfrick, M., Warren, C. & Pingle, P. (2010). Noncontact measurement techniques for model correlation. *Sound Vib.,* **44**: 8-12.

Azzam, H., &Mcfeat, J. (2015). Guidance on integrating matured SHM systems into UK military aircraft. *Struct. Health Monit. 2015*, September 2015, Stanford, California.

Beer, F.P., Johnston, E.R. & De Wolf, J.T. (1992). *Mechanics of Materials*. McGraw-Hill Engineering Series. New York.

Cazzolato, B., Wildy, S.J., Codrington, J., Kotousov, A. & Schüssler, M. (2008). Scanning laser vibrometer for non-contact three-dimensional displacement and strain measurements. *Proc. Aust. Acoustical Soc. Conf.*, *Vol. 26*, November 2008, Geelong, Victoria, Australia, pp. 206-214.

George, T.J., Seidt, J., Shen, M.H.H., Nicholas, T. & Cross, C.J. (2004). Development of a novel vibration-based fatigue testing methodology. *Int. J. Fatigue*, **26**: 477-486.

George, T.J.,Shen, M.H.H., Nicholas, T. & Cross, C.J. (2006). A new multiaxial fatigue testing method for variable-amplitude loading and stress ratio. *J. Eng. Gas Turbines Power*, **128**: 857-864.

George, T.J., Shen, M.H.H., Scott-Emuakpor, O., Nicholas, T., Cross, C.J. & Calcaterra, J. (2005). Goodman diagram via vibration-based fatigue testing. *J. Eng. Mater. Tech.*, **127**: 58-64.

Goggin, P., Huang, J., White, E., & Haugse, E. (2003). Challenges for SHM transition to future aerospace systems. *Proc. 4th Int. Workshop Struct. Health Monitor.*, Stanford University, Stanford, CA, pp. 30-41.

Khalil, H., Kim, D., Nam, J. & Park, K. (2016). Accuracy and noise analyses of 3D vibration measurements using laser Doppler vibrometer. *Measurement*, **94**: 883-892.

Kitazawa, S. I., Wakai, E. & Aoto, K. (2016). Evaluation of annealing and double ion beam irradiation by a laser-induced and laser-detected surface acoustic wave diagnostic system. *Radiat. Phys. Chem.*, **127**: 264-268.

Longo, R., Vanlanduit, S., Vanherzeele, J. & Guillaume, P. (2010). A method for crack sizing using laser Doppler vibrometer measurements of Surface Acoustic Waves. *Ultrasonics*, **50**: 76-80.

Omidi, E., Mahmoodi, S.N. & Shepard, W.S. (2015). Vibration reduction in aerospace structures via an optimized modified positive velocity feedback control. *Aerospace Sci. Tech.*, **45**: 408-415.

Park, S.W., Park, H.S., Kim, J.H. & Adeli, H. (2015). 3D displacement measurement model for health monitoring of structures using a motion capture system. *Measurement*, **59**: 352-362.

Schüssler, M., Mitrofanova, M. & Retze, U. (2011) Measurement of 2D dynamic stress distributions with a 3D-Scanning Laser Doppler Vibrometer. *In Modal Analysis Topics*. *Conference Proceedings of the Society for Experimental Mechanics Series.* Springer, New York, NY. vol. 3, pp. 141-151.

Shin, H.J. & Lee, J.R. (2017). Development of a long-range multi-area scanning ultrasonic propagation imaging system built into a hangar and its application on an actual aircraft. *Structur. Health Monitor.*, **16**: 97-111.

Sohn, H., & Liu, P. (2016).Non-contact laser ultrasonics for SHM in aerospace structures. Fuh, G.W. (Ed), *Struct. Health Monitor. (SHM) Aerospace Struct. (2016)*, Elsevier, Cambridge, UK, pp. 325-352.

Vanlanduit, S. & Guillaume, P. (2004). An automatic scanning algorithm for high spatial resolution laser vibrometer measurements. *Mech. Syst. Signal Proc.*, **18**: 79-88.

Weisbecker, H., Cazzolato, B., Wildy, S. Marburg, S, Codrington, J. & Kotousov, A. (2012). Surface strain measurements using a 3D scanning laser vibrometer. *Exp. Mech.*, **52**: 805-815.

Wildy, S.J., Cazzolato, B.S., Kotousov, Andrei G. & Weisbecker, H. (2010). New method for accurate strain measurements utilising a 3D scanning laser Doppler vibrometer. *Proc. 6th Aust. Cong. Appl. Mech.*. Perth, Australia, pp. 738-747.

TECHNICAL TENDER EVALUATION USING ANALYTICAL HIERARCHY PROCESS (AHP)

Nor Hafizah Mohamed[1*], Hendrik Lamsali[2] & Dinesh Sathyamoorthy[1]

[1]Science& Technology Research Institute for Defence (STRIDE), Ministry of Defence, Malaysia
[2]Centre for University-Industry Collaboration (CUIC), Universiti Utara Malaysia (UUM), Malaysia

[*]Email: norhafizah.mohamed@stride.gov.my

ABSTRACT

The selection of competent tenderers in high-value projects and procurements is very critical, as it determines the risks and outcomes of the projects. In some government procurement systems, millions of dollars are at stakes. Unfortunately, a significant number of high-value projects and procurements have failed to achieve the intended goals and standards due to incompetent tenderers. Hence, employing the right decision making approach in the tender evaluation and selection process is vital. This study evaluates the use of analytical hierarchy process (AHP) for technical tender evaluation using a case study. The findings show that AHP is more effective than the present method which employs the scoring and weighted sum model. AHP also provides more accurate results with minimal bias. It is expected that the findings of this study will be beneficial to improve the existing procurement system. In general, the findings may also be applicable and beneficial to other high-value government procurement systems.

Keywords: *Technical tender evaluation; analytical hierarchy process (AHP); pairwise comparison; consistency; priority ranking*.

1. INTRODUCTION

The tender procedure is very complex and involves coordination tasks with different priorities and objectives for each item. Cheng & Li (2004) stated that if no system or technique is in place to accurately assess the most suitable tenderer, project performance can be affected. Technical evaluation of tenderers is one of the main activities of an organisation in acquisitions, whether in the form of services, supplies or works to facilitate the planning of a project. Bias and inconsistency in judgement are inevitable if the technical evaluation depends on intuition, subjective judgment or emotion. Therefore, a transparent, flexible guidance tool to support the assessment of the tenderers is required to produce a more effective evaluation (Mohemad *et al*., 2010; Chua *et al*., 2015).

In Malaysia, difficulties in determining the appropriate tenderers can be seen in the Federal Auditor General's Reports for 2012, 2013, 2014 and 2015, which indicated that a number of tenderers were unable to complete the respective projects or tasks within the stipulated time, the work was done not according to specifications, and the quality of materials supplied was not suitable (National Audit Department, 2012, 2013, 2014, 2015). In addition, "The Red Book" (Putrajaya Committee on GLC High Performance, 2006) highlighted a number of common weaknesses in the procurement system in Malaysia, namely the failure to purchase a product in the right quantity and inaccurate specifications, resulting in increased total cost. Hui *et al*. (2011) reported that the present procurement process is ineffective, resulting in a longer time being taken in the selection of the tenderer and losses for the government.

At present, the scoring and weighted sum model is used in the technical evaluation of tenderers. The Technical Tender Evaluation Committee (*Jawatankuasa Penilaian Teknikal Tender*, JPTT) needs to check and ensure that every specification is complied with by tenderers on the basis of information submitted. Compliance with the specification is checked using two criteria, i.e., either comply or not comply.

There are also instances that when the scoring is calculated from the results obtained, there are two or more tenderers who have scored equal marks. This is mainly because the scoring and weighted sum model does not take into account the weight ratio of each criterion, but is merely a comparison with alternative and criteria only. Therefore, it is possible to provide the same weightage values for all criteria. This causes difficulty to the JPTT to submit a list of tenderers in order of importance and ranked based on the assessment that has been made. This will lead to the tenderer selection process being time consuming, with the assessment sometimes being done twice due to such confusions. In parallel with this problem, the calculation of marks for each assessment is done manually or only by Microsoft Excel. There has been no decision support system used to perform calculations of score points for the technical evaluation.

In addition, pairwise comparisons which compare criteria with criteria, sub-criteria with criteria, and criteria with alternative (tenderer) cannot be conducted using the technique of scoring and weighted sum model. Thus, the evaluation of tenderers is not detailed and rigorous for each of the criteria, sub-criteria and alternatives. If pairwise comparison is made, the calculation is more detailed and more accurate because each member of the JPTT will compare the criteria and give a weightage value to each criterion for comparison. On the other hand, the technique currently used only gives weightage to the comparison criteria with alternative. The technical evaluation carried out only refers to scoring given for a procurement of equipment and the evaluation is also based on consensus among members of the JPTT, whereby sometimes the members have to follow the decision of the chairman. Moreover, when there are a large number of criteria, there is a high probability that the scoring and weighted sum model will provide the same weightage for most criteria (Meredith & Mantel, 2012). Furthermore, there could be problems to distinguish the best choice according to the criteria established. Previous studies have shown that the scoring and weighted sum model cannot stand on its own as a technique that is capable of producing the best choice when it is not supported by other techniques (Padumadasa & Rehan, 2009).

The inadequacies of the existing scoring system show that there is a need for improvements to the use of techniques and procedures for systematic and transparent evaluation of tenderers. From previous studies, the field of operations research (OR) is an area that can support decision making in the evaluation and selection of tenderers. Various OR methods and techniques can support decision making in addressing the complexities of electoral issues, such as multi-criteria decision making (MCDM). This approach enables decision makers to assess various criteria and compare alternatives to achieve specific goals (Pomerol & Barba-Romero, 2000; Büyüközkan, 2004). The main feature of MCDM is its emphasis on judgment of group decision making, determining the objectives (goals) and criteria, estimating the weighted relative importance, and assessing the contributions of each option for each criteria (DCLG, 2009).

Following the above problems, this study suggests improvements in the technical evaluation of the tenderer in the public sector by using an OR method that is more appropriate and effective, namely the analytical hierarchy process (AHP). AHP is a technique that has a hierarchy consisting of goals, criteria or factors, sub-criteria and alternatives, comparative judgments, and synthesis of priorities by combining qualitative and quantitative criteria simultaneously. Gabb & Henderson (1996), Yahya & Kingsman (1999), Tam & Tummala (2001), Shiau *et al.* (2003) and Diabagaté *et al*. (2015) found that AHP provides a fair and open process for the technical evaluation of tenderers. The primary advantage of AHP is its simplicity and transparency in developing an interactive selection model to facilitate decision makers in evaluating tenderers (Anagnostopoulos &Vavatsikos, 2006; Mohammed & Rami, 2015). It is a flexible, easy to use and cost effective method that does not require much time for evaluating and selecting the best tenderers (Palcic & Lalic, 2009). In addition, it is effective and

practical for complex tenderer evaluations, whereby judgements are made using pairwise comparisons, which leads to more precise and concise decisions (Padumadasa & Rehan, 2009). Hence, the application of AHP in evaluation of tenderers provides versatility for various projects through alternatives and multi-criteria (Kwok & Lim 2006; Kwok, 2011). The importance of each criterion can be seen clearly as it is in the form of a hierarchical structure, while the consistency test reduces bias in making decisions (Aruldoss *et al.*, 2013; Chua *et al.*, 2015; Diabagaté *et al.* 2015; Gayathri & Nagaraju, 2016). In this paper, the effectiveness of AHP for technical tender evaluation and selection of the best tenderer is demonstrated using a case study.

2. RESEARCH METHODOLOGY

2.1 Development of AHP Model

Due to the confidentiality of the procurement used in this case study, we cannot indicate the real criteria and alternatives. Therefore, we decided to name the procurement as ABC, the criteria as Criteria 1 to 8, and the alternatives as Tenderers A, B and C. In the technical tender evaluation, the goal is the evaluation and selection of the best tenderer. The identification of criteria is based on the specifications developed by the users. Then appointed JPTT members will provide the judgment for each criterion based on pairwise comparisons of the procurement specifications. As for assessing the consistency of judgments made by the JPTT, consistency tests are done to ensure that the judgments that are made are consistent, in order to shortlist the tenderers by priority. In this study, there are some steps or stages to be followed for determining priorities and the best alternative, as shown in Figure 1.

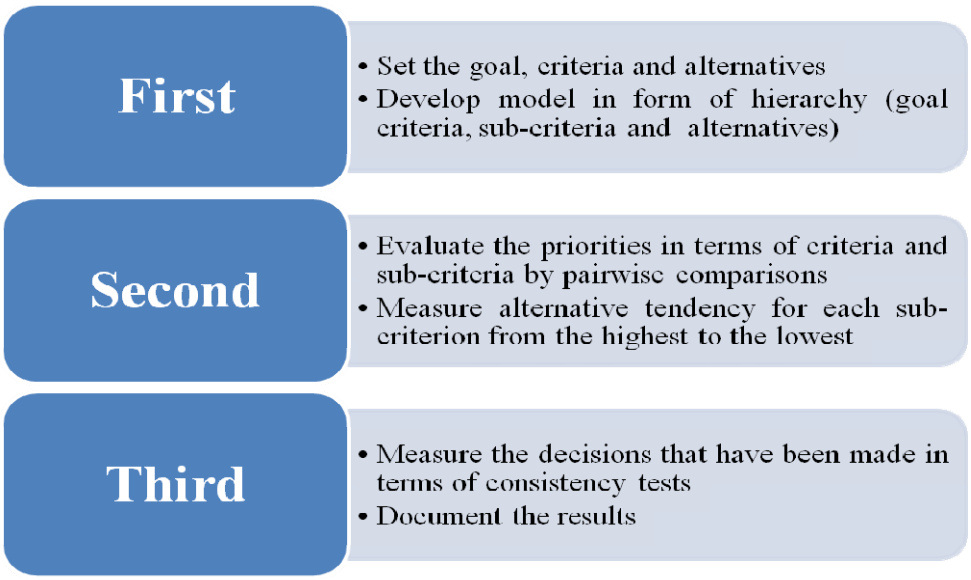


Figure 1: Steps used in this study for prioritising alternatives.

In this study, the development of the AHP model for the technical evaluation of procurement ABC refers to the flow chart as shown in Figure 2, which includes the identification of objectives, criteria and alternatives.

The JPTT members should be experts in areas of the equipment to be tendered. In this case, seven officers have been appointed to evaluate the tenderers. One of the members of the JPTT will be appointed as the chairman, designated as P1, to lead the process of the evaluation, while the other members are designated as P2 to P7. In this evaluation, the JPTT conducts the assessment on the basis of predetermined criteria, of which there were nine main criteria, 22 sub-criteria, 34 sub-sub-criteria and three alternatives (tenderers) (Table 1).

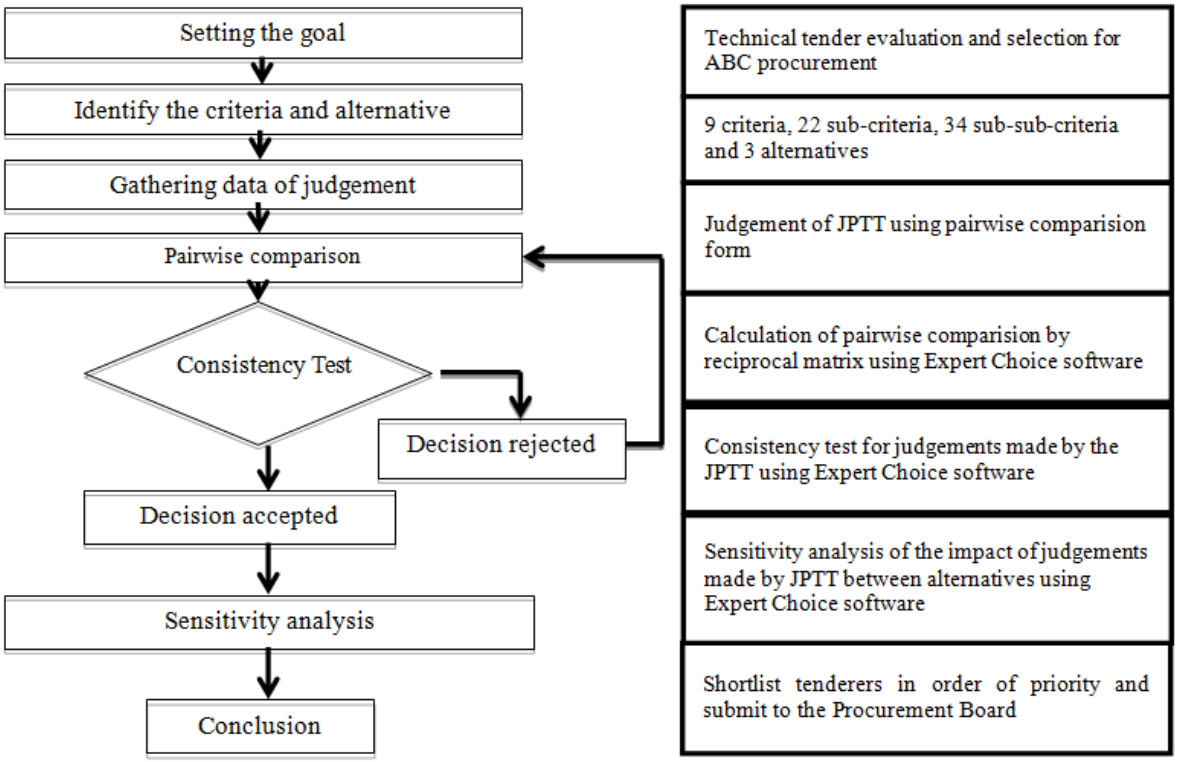


Figure 2: Flow chart of the study.

2.2 Structuring Hierarchy Model

The hierarchy model for this study is as shown in Figure 3, which shows the levels or hierarchy that should be taken into account in the evaluation of tenderers for procurement ABC. It is organised according to a hierarchical model, where the goal is top-ranked, followed by the nine main criteria, the second stage is 22 sub-criteria, the third level is a 34 sub-sub-criteria and the fourth stage is the alternatives.

2.3 Research Instrument, Calculation Method and Consistency Test

The instrument used in this study is the comparison of pairs developed to scale the relative importance that is used in AHP, which is based on the goals and factors to be achieved. The pairwise comparisons form shown in Figure 4 was designed to facilitate the measurement made based on pairwise comparisons (Saaty, 1980; Partovi, 1994; Omkarprasad & Kumar, 2006; Render et al, 2012). Generally this form is divided into two parts, namely pairwise comparisons between criteria and pairwise comparisons between alternatives. Each JPTT member will give a score for the relative importance of each of the two parts for the respective pairwise comparisons.

In this study, each JPTT member needs to assess a total of 255 pairwise comparison items, consisting of 153 comparison items for criteria with criteria and 102 comparison items for criteria with alternatives, by filling out the pairwise comparison form based on criteria, sub-criteria, sub-sub-criteria and alternatives with assessment scores and enter the data into the Expert Choice software for calculation purposes.

Table 1: Structuring of multi-criteria for procurement ABC.

Goal	Criteria	Sub-criteria	Sub-sub-criteria	Alternative (Number of tenderers)
Technical tender evaluation and selection of the best tender for procurement ABC	Criteria 1	• Criteria 1.1	• Criteria 1.1.1 • Criteria 1.1.2 • Criteria 1.1.3 • Criteria 1.1.4 • Criteria 1.1.5 • Criteria 1.1.6	Three Tenderers: Tenderer A, B and C
	Criteria 2	• Criteria 2.1	• Criteria 2.1.1 • Criteria 2.1.2 • Criteria 2.1.3 • Criteria 2.1.4 • Criteria 2.1.5 • Criteria 2.1.6 • Criteria 2.1.7 • Criteria 2.1.8 • Criteria 2.1.9	
		• Criteria 2.2	• Criteria 2.2.1 • Criteria 2.2.2 • Criteria 2.2.3 • Criteria 2.2.4 • Criteria 2.2.5	
		• Criteria 2.3	• Criteria 2.3.1 • Criteria 2.3.2 • Criteria 2.3.3 • Criteria 2.3.4 • Criteria 2.3.5	
		• Criteria 2.4	• Criteria 2.4.1 • Criteria 2.4.2 • Criteria 2.4.3 • Criteria 2.4.4 • Criteria 2.4.5	
		• Criteria 2.5	• Criteria 2.5.1 • Criteria 2.5.2 • Criteria 2.5.3 • Criteria 2.5.4	
	Criteria 3	• Criteria 3.1		
	Criteria 4	• Criteria 4.1 • Criteria 4.2 • Criteria 4.3		
	Criteria 5	• Criteria 5.1 • Criteria 5.2		
	Criteria 6	•Criteria 6.1		
	Criteria 7	• Criteria 7.1 • Criteria 7.2 • Criteria 7.3 • Criteria 7.4		
	Criteria 8	• Criteria 8.1 • Criteria 8.2 • Criteria 8.3 • Criteria 8.4 • Criteria 8.5		
	Criteria 9			

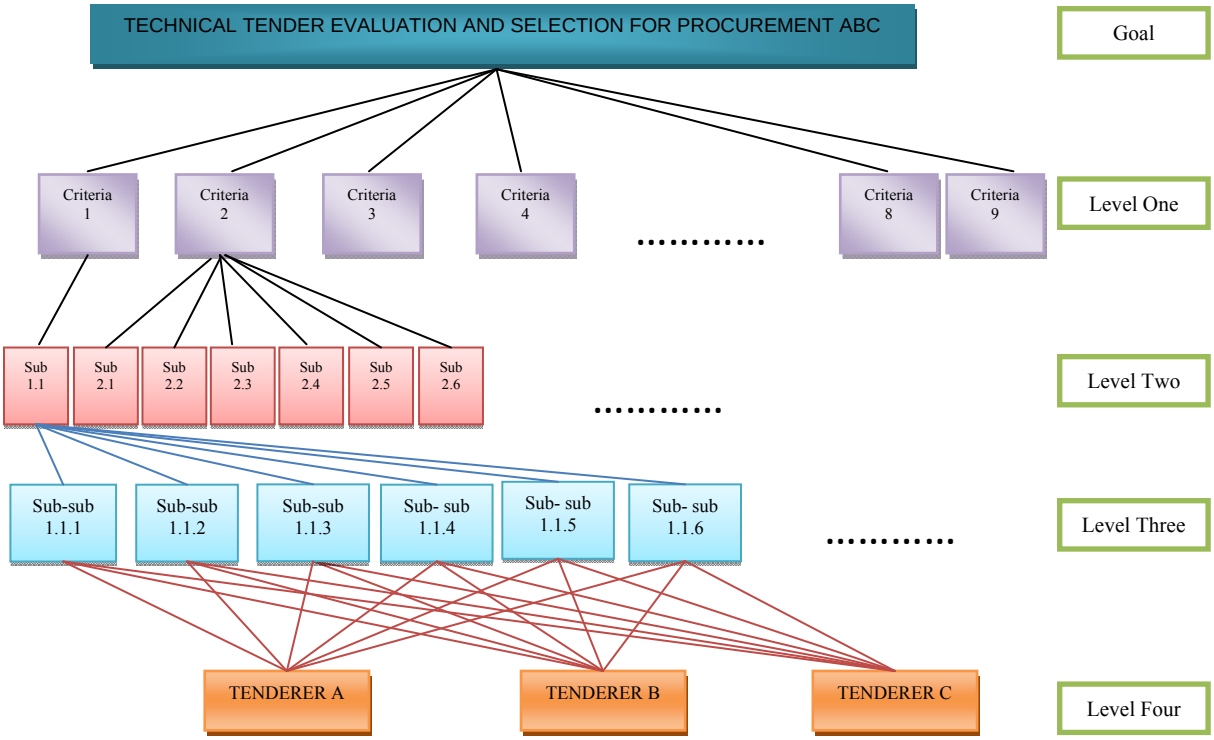


Figure 3: The hierarchy model for technical tender evaluation for procurement ABC.

Criteria 1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Criteria 2
Criteria 1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Criteria 3

Figure 4: Example of form developed for pairwise comparison matrix.

All the data obtained from the form will be prepared in the form of a matrix (Saaty, 1980; Render & Stair, 1992; Saaty, 1988; Saaty & Vargas, 1994; Cabala, 2010) :

$$A=(a_{ij}) \text{ where } i,j=1,2,...,n,$$
$$a_{ii} = 1 \text{ for } i=j,$$
$$a_{ij} = 1/a_{ji} \text{ for } i \neq j \qquad (1)$$

The matrix is known as an eigenvector ($\boldsymbol{\omega}$) with $\boldsymbol{n}$ number of criteria, and each cell in the column and row known as eigenvalue ($\boldsymbol{\lambda}$) or equation ($\boldsymbol{A\omega = \lambda\omega}$). The weights are calculated based on the scale of importance {1-9},where $\boldsymbol{a_{ik} = a_{ij}a_{jk}}$ for i, j and k, or known as reciprocal comparison matrices.

The calculation will be assisted by the Expert Choice software for easy and accurate calculation and analysis. Expert Choice is a software that is devoted to analysis using AHP technique.In any study using AHP,consistency analysis is crucial in ensuring that the results obtained through calculations or software can be adopted. According to Saaty (1980) and Ayag & Ozdemir (2006), in general, the judgements are consistent (*Consistency Ratio*) when:

Consistency Index (CI) = $\frac{\lambda - n}{n - 1}$

where;
n number of item/criteria that compared
λ is $\frac{Consistency\ vector}{n}$

Consistency Ratio (CR) = $\frac{Consistency\ Index\ (CI)}{Random\ Index\ (RI)}$ < 0.10 Consistent (Acceptable)

and $\frac{Consistency\ Index\ (CI)}{Random\ Index\ (RI)}$ > 0.10 Not Consistent (Not Acceptable) (2)

If the value of consistency is more than 0.10, it shows that there is an inconsistency in the judgments made. Therefore, the JPTT members should review the evaluation to determine if any mistakes were made. On the other hand, if the consistency is less than 0.10, it indicates that the judgements made are consistent and acceptable (Saaty, 1980; Render & Stair, 1992; Ayag & Ozdemir, 2006; Cabala, 2010).

3. RESULTS & DISCUSSION

3.1 Pairwise Comparison

Any comparison made is written in the form of a matrix of reciprocity and the assessment is comprised of two parts, namely pairwise comparisons between criteria with criteria and pairwise comparisons between the criteria and alternatives. This comparison is continued from the first stage to the fourth stage. Examples of pairwise comparisons that were made in the form of a matrix are shown in Tables 2 to 5.

The JPTT members simply fill out the comparison matrix. For example, assuming that the ratio between Criterion 1 and Criterion 9 is 4, it is equivalent to the ratio of 4 to 1 (4: 1). Therefore, a comparison of Criterion 9 and Criterion 1 is a 1 to 4 ratio (1: 4) or 1/4. A value of 1 is given to the diagonal elements that show that the criteria are equally important. After consideration of the pairwise comparisons in Tables 2 to 5 made by all the JPTT members, the next step is the calculation of the priority vector for each element in the matrix. In terms of matrix algebra, this is done through eigenvector matrix calculation by adding a given value of each element in the column to get the total column.

Then, in order to normalise each column to be equivalent to a value of 1 or 100%, each element of the column is divided by the total column. Finally, the elements in each row are added and the sum is divided by the number of elements in a row to get the average. For example, Table 6 shows the vector calculation priority for the JPTT chairman for the main criteria, and calculations should proceed to the sub-criteria, sub-sub-criteria and alternatives. Similarly, the judgments made by the remaining JPTT members should be calculated using the same calculation method.

Table 2: Examples of pairwise comparison matrix in the form of reciprocity (9X9) for the first level (main criteria).

Criteria	Criteria 1	Criteria 2	Criteria 3	Criteria 4	Criteria 5	Criteria 6	Criteria 7	Criteria 8	Criteria 9
Criteria 1	1	2	1	2	3	3	3	3	4
Criteria 2	1/2	1	1	2	1	1	2	2	2
Criteria 3	1	1	1	2	2	2	1	2	2
Criteria 4	1/2	1/2	1/2	1	3	3	2	2	3
Criteria 5	1/3	1	1/2	1/3	1	2	2	2	2
Criteria 6	1/3	1	1/2	1/3	1/2	1	2	1	1
Criteria 7	1/3	1/2	1	1/2	1/2	1/2	1	2	2
Criteria 8	1/3	1/2	1/2	1/2	1/2	1	1/2	1	1
Criteria 9	1/4	1/2	1/2	1/3	1/2	1	1/2	1	1

Table 3: Examples of pairwise comparison matrix in the form of reciprocity (4X4) for the second level (sub-criteria).

Sub-Criteria: Criteria 7	Sub Criteria 7.1	Sub Criteria 7.2	Sub Criteria 7.3	Sub Criteria 7.4
Sub Criteria 7.1	1	3	3	3
Sub Criteria 7.2	1/3	1	3	3
Sub Criteria 7.3	1/3	1/3	1	2
Sub Criteria 7.4	1/3	1/3	1/2	1

Table 4: Examples of pairwise comparison matrix in the form of reciprocity (4X4) for the third level (sub-sub-criteria).

Sub-sub-criteria: Criteria 2.5	Sub-sub-criteria 2.5.1	Sub-sub-criteria 2.5.2	Sub-sub-criteria 2.5.3	Sub-sub-criteria 2.5.4
Sub-sub-criteria 2.5.1	1	4	3	2
Sub-sub-criteria 2.5.1	1/4	1	2	3
Sub-sub-criteria 2.5.1	1/3	1/3	1	3
Sub-sub-criteria 2.5.1	1/2	1/2	1/3	1

Table 5: Examples of pairwise comparison matrix in the form of reciprocity (3X3) for the fourth level (alternative and criteria 1.1.6).

	Tenderer A	Tenderer B	Tenderer C
Tenderer A	1	2	1
Tenderer B	1/2	1	2
Tenderer C	1	1/2	1

Table 6: Examples of the vector priority calculation for the JPTT chairman for the main criteria of Procurement ABC.

Criteria	Criteria 1	Criteria 2	Criteria 3	Criteria 4	Criteria 5	Criteria 6	Criteria 7	Criteria 8	Criteria 9	Total Row	Mean
Criteria 1	0.22	0.25	0.15	0.22	0.25	0.21	0.21	0.19	0.22	1.93	0.21
Criteria 2	0.11	0.13	0.15	0.22	0.08	0.07	0.14	0.13	0.11	1.14	0.13
Criteria 3	0.22	0.13	0.15	0.22	0.17	0.14	0.07	0.13	0.11	1.33	0.15
Criteria 4	0.11	0.06	0.08	0.11	0.25	0.21	0.14	0.13	0.17	1.25	0.14
Criteria 5	0.07	0.13	0.08	0.04	0.08	0.14	0.14	0.13	0.11	0.91	0.10
Criteria 6	0.07	0.13	0.08	0.04	0.04	0.07	0.14	0.06	0.06	0.68	0.08
Criteria 7	0.07	0.06	0.15	0.06	0.04	0.03	0.07	0.13	0.11	0.73	0.08
Criteria 8	0.07	0.06	0.08	0.06	0.04	0.07	0.04	0.06	0.06	0.53	0.06
Criteria 9	0.05	0.06	0.08	0.04	0.04	0.07	0.04	0.06	0.06	0.50	0.06
Total Column	1		1	1	1	1	1	1	1		1

However, for the decision by the group, the calculation must be made a priority vector matrix of the combined participants' judgments. The priorities for each criteria and priority ranking criteria will be obtained. The results of these calculations can be ranked according to priority as shown in Table 7, which lists the criteria, sub-criteria and sub-sub-criteria in order of priority or relative importance.

Table 7: Ranking priority results for the criteria, sub-criteria and sub-sub criteria of Procurement ABC.

Ranking	Criteria	Priority	Sub-Criteria	Priority	Sub-Sub-Criteria	Priority
1	Criteria 8	0.169	Criteria 8.1	0.039		0.039
			Criteria 8.2	0.038		0.038
			Criteria 8.4	0.038		0.038
			Criteria 8.3	0.037		0.037
			Criteria 8.5	0.018		0.018
2	Criteria 9	0.158		0.158		0.158
3	Criteria 2	0.146	Criteria 2.1	0.070	Criteria 2.1.3	0.017
					Criteria 2.1.2	0.016
					Criteria 2.1.4	0.010
					Criteria 2.1.1	0.008
					Criteria 2.1.5	0.007
					Criteria 2.1.7	0.004
					Criteria 2.1.8	0.004
					Criteria 2.1.9	0.003
					Criteria 2.1.6	0.002
			Criteria 2.5	0.037	Criteria 2.5.1	0.015
					Criteria 2.5.4	0.009
					Criteria 2.5.2	0.008

Ranking	Criteria	Priority	Sub-Criteria	Priority	Sub-Sub-Criteria	Priority
					Criteria 2.5.3	0.005
			Criteria 2.2	0.022	Criteria 2.2.2	0.007
					Criteria 2.2.1	0.007
					Criteria 2.2.3	0.003
					Criteria 2.2.4	0.003
					Criteria 2.2.5	0.003
			Criteria 2.3	0.010	Criteria 2.3.2	0.003
					Criteria 2.3.1	0.003
					Criteria 2.3.3	0.001
					Criteria 2.3.4	0.001
					Criteria 2.3.5	0.001
			Criteria 2.4	0.006	Criteria 2.4.2	0.002
					Criteria 2.4.1	0.002
					Criteria 2.4.3	0.001
					Criteria 2.4.4	0.001
					Criteria 2.4.5	0.001
4	Criteria 5	0.140	Criteria 5.1	0.097		0.097
			Criteria 5.2	0.043		0.043
5	Criteria 6	0.131	Criteria 6.1	0.131		0.131
6	Criteria 7	0.087	Criteria 7.2	0.040		0.040
			Criteria 7.3	0.022		0.022
			Criteria 7.4	0.013		0.013
			Criteria 7.1	0.012		0.012
7	Criteria 3	0.061	Criteria 3.1	0.061		0.061
8	Criteria 4	0.061	Criteria 4.1	0.035		0.035
			Criteria 4.3	0.016		0.016
			Criteria 4.2	0.010		0.010
9	Criteria 1	0.048	Criteria 1.1	0.048	Criteria 1.1.1	0.012
					Criteria 1.1.3	0.008
					Criteria 1.1.6	0.008
					Criteria 1.1.2	0.008
					Criteria 1.1.5	0.006
					Criteria 1.1.4	0.005
Total		1		1		1

3.2 Consistency Test

For this study, the consistency of the overall judgements of each member and the combined results are as shown in Table 8. The table shows that the value of each member is consistent. The consistency of the combined results is 0.01, which is less than 0.10, and this indicates that the judgements made are consistent and the results obtained are applicable. This proves that the judgments made by the JPTT are really feasible with negligible error.

Table 8: Consistency values of the results.

JPTT Member	Overall Inconsistency
Chairman	0.030
P2	0.030
P3	0.010
P4	0.040
P5	0.030
P6	0.040
P7	0.030
Combined	0.010

3.3 Ranking of Priority for the Tenderers

Based on the results of the AHP analysis, the ranking of the tenderers by priority was determined; Tenderer A (0.370), followed by Tenderer B (0333) and finally Tenderer C (0.296).While each tenderer has advantages over each other, on the whole, Tenderer A has the highest probability value as compared to Tenderers B and C. This shows that Tenderer A is most eligible to be granted the tender. Based on this, the JPTT's decision will be submitted to the Procurement Board for further processing.

From Table 9, we can see that the overall inconsistency is 0.01, which indicates that the decision made by the JPTT is consistent and accurate. This clearly gives the impression that the decision is acceptable and there is no element of bias because every JPTT member makes judgments and evaluates individually although the technical evaluation is carried out in a group. In addition, in order to fine-tune the results of the analysis using AHP, Table 10 shows the rankings of the tenderers on the basis of priority of criteria.

Table 9: Results of priority ranking of the tenderers.

Combined instance - synthesis with respect to:
Goal: Technical tender evaluation and selection of the best tenderer for Procurement ABC

Ranking	Alternative	Priority
1	Tenderer A	0.370
2	Tenderer B	0.333
3	Tenderer C	0.296
	Overall Inconsistency	0.010

Table 10: Results of priority ranking of the tenderers for the criteria.

Combined instance –synthesis with respect to:
Goal: Technical tender evaluation and selection of the best tenderer for Procurement ABC

Ranking	Criteria	Alternative	Priority
1	Criteria 8	Tenderer A	0.333
		Tenderer B	0.333
		Tenderer C	0.333
		Overall Inconsistency	0.000
2	Criteria 9	Tenderer A	0.333
		Tenderer B	0.333
		Tenderer C	0.333
		Overall Inconsistency	0.000
3	Criteria 2	Tenderer A	0.414
		Tenderer B	0.335
		Tenderer C	0.252
		Overall Inconsistency	0.030
4	Criteria 5	Tenderer A	0.356
		Tenderer B	0.322
		Tenderer C	0.322
		Overall Inconsistency	0.000
5	Criteria 6	Tenderer A	0.475
		Tenderer B	0.262
		Tenderer C	0.262
		Overall Inconsistency	0.000
6	Criteria 7	Tenderer A	0.525
		Tenderer B	0.305
		Tenderer C	0.170
		Overall Inconsistency	0.010
7	Criteria 3	Tenderer A	0.333
		Tenderer B	0.333
		Tenderer C	0.333
		Overall Inconsistency	0.000

8	Criteria 4	Tenderer A	0.500
		Tenderer B	0.258
		Tenderer C	0.242
		Overall Inconsistency	0.020
9	Criteria 1	Tenderer A	0.419
		Tenderer B	0.294
		Tenderer C	0.287
		Overall Inconsistency	0.010

4. CONCLUSION

This paper presented a case study on the application of AHP for technical tender evaluation. It is demonstrated that AHP allows for detailed and rigorous comparisons of criteria by the JPTT members, which produces more reliable results. The analysis showed that the overall level of consistency of the judgements of JPTT members are consistent and acceptable. Additionally, AHP can also be used to further refine the results by calculating the consistency individually for each member. Considerations can also be made for each of the criteria through their consistencies. This shows that the results produced by AHP are very specific and detailed. AHP provides a more systematic method of identification of priority ranking for criteria, sub-criteria, sub-sub-criteria and tenderers. The results of this study will be beneficial to improve the existing procurement system. In general, these findings may also be applicable and beneficial to other high-value government procurement systems.

REFERENCES

Anagnostopoulos, K.P. & Vavatsikos, A.P. (2006). An AHP model for construction contractor prequalification. *Int. J. Oper. Res.*, **6**: 333-346.

Aruldoss, M., Lakshmi, T.M. & Venkatesan, V.P. (2013). A survey on multi criteria decision making methods and its applications. *Am. J. Inform. Syst.*, **1**: 31-43.

Ayag, Z & Ozdemir, R.G. (2006). A fuzzy AHP approach to evaluating machine tool alternatives. *Journal of Intelligent Manufacturing*, 17,179-190.

Büyüközkan, G. (2004). Multi-criteria decision making for e-marketplace selection. *Internet Res.*, **14**: 139 – 154.

Cabala, P. (2010). Using the analytic hierarchy process in evaluating decision alternatives. *Oper. Res. Decis.*, **27**: 31-510.

Cheng, E.W.L. & Li, H. (2004). Contractor selection using the analytic network process. *Constr. Manage. Econ.*, **22**: 1021–1032.

Chua, S.J.L., Ali, A.S. & Alias, A.B. (2015). Implementation of analytic hierarchy process (AHP) decision making framework for building maintenance procurement selection: Case study of Malaysian public universities. *Eksploatacjai Niezawodnosc Maint. Reliability*, **18**: 7–18

DCLG (Department for Communities and Local Government) (2009). *Multi-Criteria Analysis : A Manual*. Department for Communities and Local Government (DGLG), London.

Diabagaté, A., Abdellah, A. & Mohamed, E.H. (2015). The choice of the best proposal in tendering with AHP method: Case of procurement of IT master plan's realization. *Int. J. Inform. Tech. Comput. Sci.*, **12:** 1-11**.**

Gabb, A.P. & Henderson, D.E. (1996). *Technical and Operational Tender Evaluations for Complex Military Systems*. Department of Defence, Australia.

Gayathri, S. & Nagaraju, D. (2016). Evaluation and selection of projects using analytic hierarchy process in a manufacturing industrial sector. *ARPN J. Eng. Appl. Sci.*, **11**: 10028-10034.

Hui, W.S., Othman, R., Omar, N.H., Rahman, R.A. & Haron, N.H. (2011). Procurement issues in Malaysia. *Int. J. Public Sector Manage.*, **24**: 567–593.

Kwok, Y.F. (2011). Using Analytic Hierarchy Process for the evaluation of government projects. *Proc. Int. Symp. Analytic Hierarchy Process 2011*, Defence Science and Technology Agency, Singapore.

Kwok, Y.F. & Lim, H.S. (2006). Using analytic hierarchy process with operations analysis in project evaluation: 'Unique F-15SG is perfect for Singapore's fighter project'. *Flight Daily*, 21 February 2006.

Meredith, J.R. & Mantel, S.K. (2012). *A Managerial Approach: Project Management*. 8th Edition. John Wiley & Son, Inc.

Mohammed, B. & Rami, A. (2015). Application of the analytical hierarchy process (AHP) to multi-criteria analysis for contractor selection. *Am. J. Indust. Bus. Manage.*, **5**:581-589.

Mohemad, R., Hamdan, A.R., Othman, Z.A.L.I., Maizura, N. & Noor, M. (2010). Decision support systems (DSS) in construction tendering processes. *Int. J. Comput. Sci. Iss.*, **7**: 35-45.

National Audit Department (2012). *Auditor General's Report 2012*. National Audit Department, Malaysia.

National Audit Department (2013). *Auditor General's Report 2013*. National Audit Department, Malaysia.

National Audit Department (2014). *Auditor General's Report 2014*. National Audit Department, Malaysia.

National Audit Department (2015). *Auditor General's Report 2015*. National Audit Department, Malaysia.

Omkarprasad, S. V. & Kumar, S. (2006). Analytical Hierarchy Process: An overview of applications. *European Journal Operations Research*, 169, 1-29.

Padumadasa, E.U. & Rehan, S. (2009). Investigation in to decision support systems and multiple criteria decision making to develop a web based tender management system. *Int. Symp. Analytic Hierarchy Process 2009,* School of Computing, Asia Pasific Institute of Information Technology Colombo, Sri Lanka.

Palcic, I. & Lalic, B.S. (2009). Analytical Hierarchy Process as a tool for selection and evaluating projects. *Int. J. Sim. Model.*, **8**: 16-26.

Partovi, F.Y. (1994). Determining what to benchmark: An Analytic Hierarchy Process Approach. *Int. J. Oper. Prod. Man*., 14,25-39.

Pomerol, J.C. & Barba-Romero, S. (2000). *Multicriterion Decision in Management: Principles and Practice*. Kluwer Academic Publishers, USA.

Putrajaya Committee on GLC High Performance (2006). *The Red Book- Procurement Guidelines and Best Practices for all the GLCs*. Putrajaya Committee on GLC High Performance, Putrajaya.

Render, B. & Stair, R.M. (1992). *Introduction to management science*. Division of Simon & Schuster, Inc.United States of America.

Render, B., Stair, R.M. & Hanna, M.E. (2012). *Global Edition: Quantitative analysis for management*. 11th Edition, Pearson Education Limited, England.

Saaty, T.L. (1988). *Decision making for leaders*. RWS Publication, Pittsburgh, PA.

Saaty, T.L. & Vargas, L.G. (1994). *Decision Making in Economic, Political, Social and Technological Environments with the Analytical Hierarchy Process*.RWS Publication, Pittsburgh.

Saaty, T.L. (1980). *The Analytic Hierarchy Process*. McGraw Hill, New York.

Shiau, Y.C., Tsai, T.P., Wang, W.C. & Huang, M.L. (2003). Use questionnaire and AHP techniques to develop subcontractor selection system. *19th Int. Symp. Automat. Robot. Constr.*, National Institute of Standards and Technology, Gaithersburg, Maryland.

Tam, M.C.Y. & Tummala, V.M.R. (2001). An application of the AHP in vendor selection of a telecommunications system. *Omega*, **29**: 171-182.

Yahya, S. & Kingsman, B. (1999). Vendor rating for an entrepreneur development programme: A case study using the Analytic Hierarchy Process method. *J. Oper. Res. Society*, **50**: 916-930.

MILITARY SPENDING AND FISCAL SUSTAINABILITY INDICATOR: CASE STUDY FOR MALAYSIA

Wan Farisan Wan Suleiman*, Zulkefly Abdul Karim, Norlin Khalid & Riayati Ahmad

School of Economics, Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), Malaysia

*Email: wantreasury@yahoo.com

ABSTRACT

The recent budget cuts to military spending in Malaysia may cause concern on the readiness of the country to face future security threat; as such if there is a need to increase military spending in the future to improve military readiness, this need should be balance with the objective of ensuring fiscal sustainability. This study aims to assess the fiscal sustainability in Malaysia using the fiscal sustainability indicator approach, mainly the primary gap indicator, and then compare it to the military spending in Malaysia. Using annual data of primary fiscal balance ratio, public debt ratio, real gross domestic product (GDP) growth, real interest rate, real inflation and military spending from 1981 to 2015, four fiscal regimes were identified; regime 1 (1981-1988) unsustainable, regime 2 (1988-1997) sustainable, regime 3 (1998-2013) unsustainable and regime 4 (2014-2015) sustainable. The conclusion of the study suggests that while there is a need for fiscal consolidation through budget cuts in non-productive sectors such as military to ensure fiscal sustainability and long-term economic growth, this should not be at the expense of the government fulfilling its role as the provider of defence as public good.

Keywords: *Military spending; fiscal sustainability indicator; primary gap indicators; fiscal policy Malaysia.*

1. INTRODUCTION

The 2017 Budget announced by the Malaysian Ministry of Finance showed that Malaysia's budgeted military spending is RM 15.1 billion, which declined 14% from the previous year (Ministry of Finance, 2017). This means that Malaysia's military spending has seen the largest drop since 1998, with the 23.6% cut in 1998 remaining the largest. According to the government, the decrease is due to both lower operation and development expenditures with the military only getting new patrol vessels and armoured vehicles. This decline in military spending could affect future capabilities of the military to rise up to the threat of rising security concerns, such as the overlapping claims in South China Sea by several countries (Singh, 2012).

There are many justifications for the policy decision. Even though military spending has been reduced to 1.1% of GDP, it is comparable to some developed economies. This indicates that Malaysia's military spending is relatively restrained, meaning that military spending should be in line with economic development (Qiong & Junhua, 2015). In addition, in order for Malaysia to ensure fiscal sustainability and economic growth, arising from the threat of increasing public debt ratio and persistent fiscal deficits, the country should reduce its government spending and subsequently budget for the military (Caruso & Domizio, 2016). Under such a situation, another justification for the lower military spending is that it would provide growth in spending for social sectors, such as better public healthcare, and consequently increases the quality of life for its people (Dunne, 2015).

In the Malaysian economy, fiscal policy sustainability is essential in ensuring steady economic growth through sound fiscal policy management, which is crucial in attracting investment and stable monetary policy (International Monetary Fund, 2015). Fiscal policy sustainability management in Malaysia is focused on ensuring government borrowing is for development purpose only and the debt ratio does not exceed 55% of the GDP (Ministry of Finance, 2012). The government started the fiscal consolidation programme in 2010 towards achieving fiscal balance by 2020. The government has also set a target to achieve debt ratio of 45% by 2020. The declining oil revenue from the recent sharp drop of oil price in 2015 also poses a challenge for the government in realising the fiscal target set. Nevertheless, it is essential for the government to achieve those targets to avoid any negative impact from the perception of fiscal unsustainability, such as downgrading of the sovereign rating of the country.

This study aims to provide an assessment of the fiscal policy sustainability for Malaysia. Specifically, this study investigates the short, medium and long-term fiscal sustainability using the Malaysian federal government data set by constructing the primary gap indicator (Blanchard *et al*., 1990) as well as exploring the impact of population ageing on the fiscal sustainability.

In this paper, the relationship between military spending and fiscal sustainability for Malaysia is also explored. Most of the previous studies were conducted to determine the relationship between military spending and economic growth as well as external debt (Ross 1990; Tang, 2008; Hirnissa et al., 2009 ;). None of the existing literature on the topic of military spending and the economy has dealt with the fiscal sustainability indicator. To this end, this paper aims to contribute to the studies of fiscal sustainability and military spending in three aspects. First, it will focus on Malaysia, which would contribute to the study on fiscal sustainability in Malaysia using the fiscal sustainability indicator. Second, it will improve previous studies on fiscal sustainability and military spending in the duration of time. Finally, it extends the studies in the current literature of fiscal sustainability and military spending by jointly incorporating the fiscal sustainability indicator and military spending.

2. TRENDS IN MILITARY SPENDING AND FISCAL SUSTAINABILITY

2.1 Trends in Military Spending

Military spending in Malaysia, as in Figure 1 below has shown a declining trend from a high of 3.2% of GDP in 1991 to the lowest level of 1.1% of GDP in 2017. Among the reasons for the declining trend was the slower economic growth from 9.5% in 1991 to 4.2% in 2016. Secondly, any economic crisis would impact the military spending drastically such as the Asian Financial Crisis in 1998 which saw the military spending drop to 1.6% in 1998 from 2.1% in 1997 and the Global Financial Crisis in 2009 which saw the military spending cut from 2.0% in 1999 to 1.5% in 2010. Another reason for the declining spending was the need to embark on a fiscal consolidation path towards budget balance by 2020 from a deficit of 6.7% in 2009 to 3.1% in 2016. The allocation for the ministry reduced drastically by 14% from RM17.3billion in 2016 to RM 15.1 billion in 2017, a reduction of RM 2.3 billion and a decline of 0.4% of GDP. The decline in military spending ratio is the effect of slower economic growth and efforts to consolidate the budget.

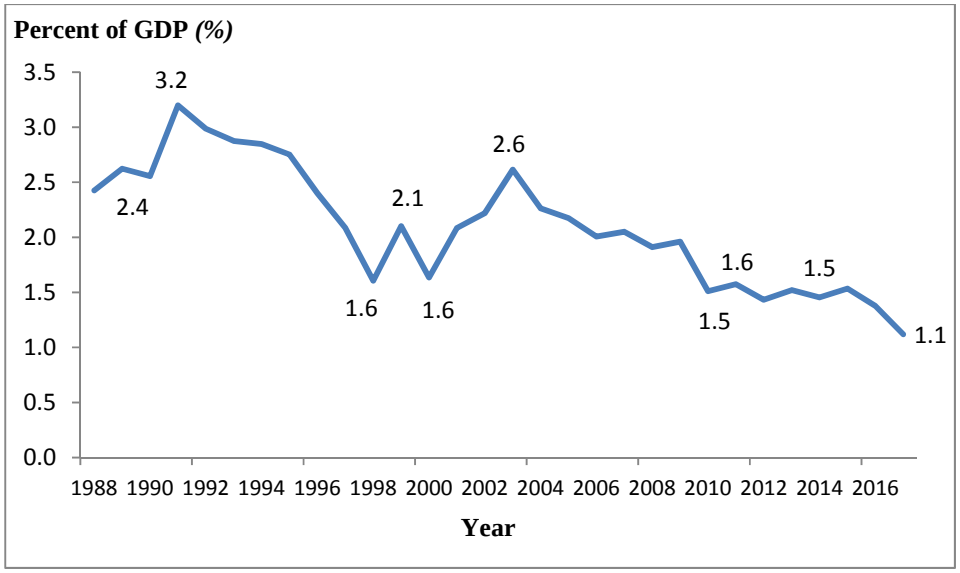


Figure 1: Military spending in Malaysia for 1988-2017 (data from World Bank and Treasury Malaysia).

2.2 Trends in Fiscal Indicators

Deficit consists of fiscal deficit and primary balance deficit, while public debt consists of total value and as percent of GDP. Persistent deficits have been recorded since 1998 with the highest level in 2009 at 6.7% of the GDP after the global financial crisis. Fiscal deficits have shown to be in a declining trend from 2009 to 2017 from 6.7% in 2009 to 3.0% in 2017. On the other hand, the public debt ratio is shown to be in an increasing trend from 2010 to 2015 from 49.6% in 2010 to 54.5% in 2015, almost reaching the limit of 55%, before being expected to decline to 52.5% in 2016 and then to 51.8% in 2017. The persistent fiscal deficits contributes to the increasing public debt ratio and increasing value of public debt from just RM 229 billion in 2005 to almost RM 700 billion in 2017, and the increasing debt ratio from just 42.1% in 2005 to 51.8% in 2017 (Ministry of Finance, Malaysia).

Malaysia's public spending on military consists of the budget allocation to the Ministry of Defence, and comprises of operating and development expenditure. Over the period of 2005 to 2017, total military spending as a percentage of GDP varied from 2.2% in 2005 to 1.1% in 2017. Of the total government expenditure, military spending accounted between 9.5% in 2005 to 5.8% in 2017. Consequently, the percentage of change of military spending also varied from 15.1% in 2011 to -13.7% in 2017. Thus, the fiscal challenge is how to sustain the current fiscal sustainability indicator in terms of deficit and public debt if military spending has to be increased to meet future security concerns.

3. REVIEW OF THE LITERATURE

There have been many studies that investigate the issue of impact of military spending to the Malaysian economy, mainly on growth, budget and debt. For studies on growth, several results can be derived based on growth rate of military spending and economy, and the relationship between the two variables. Ross (1990) observed that Malaysia's growth rate is lower than the growth rate of the military spending. Hirnissa *et al*. (2009) could not find any meaningful relationship between military spending and economic growth. In contrast, Khalid & Mohd Noor (2015) found that military spending has a positive and significant result on growth. On the other hand, on the issue of what determines military spending, it was

discovered that military spending is determined by economic growth (Frederiksen, 1991). Meanwhile, Harris (1986) concluded that military spending is determined by previous year budget. For the study on debt, Ross (1990) indicated that high level of external indebtedness will reduce Malaysia's military spending. To the authors' best knowledge; there has been no study that investigates the relationship between military spending and fiscal sustainability indicator, which will be the focus of this study.

Fiscal sustainability can be defined as public debt ratio that is non-increasing or the ability of the government to pay its debts without any increase in its tax rate or reduction in its spending rate (Chalk & Hemming, 2000). A significant issue in any evaluation of fiscal policy sustainability is the appropriate identification of assessment method for fiscal sustainability. Previous studies have indicated four main approaches to the assessment of fiscal sustainability. First, some studies, for example, Hamilton & Flavin (1985) and Uctum & Wickens (1996) used unit root test for testing stationarity for fiscal variables, while Elliott & Kearney (1988), Cuddington (1997) and Afonso (2004) used the cointegration test to determine the existence of long-run relationship between public expenditure and public revenue to determine the sustainability of fiscal policy. However, the problem with both methods is that it is generally weak against alternatives close to unit root, producing projection of variables such as economic growth based on its current and lagged values, as well as not providing quantitative measure of sustainability (Burnside, 2005). In fact, in reality, many alternatives could be close to unit root, future economic growth could be different than projection based on the current and lagged values, and quantitative measures of fiscal sustainability is needed by fiscal policy makers. This means that any changes in future economic growth would result in changes in fiscal sustainability. To solve the problem of alternative close to unit root, a number of empirical studies have used alternative approaches, such as fiscal reaction function in assessing fiscal sustainability. For example, Bohn (1998), Ghatak & Sanchez-Fung (2006), Stoian & Roman (2006), De Mello (2008) and James & Baba (2015) assessed fiscal sustainability by estimating the fiscal reaction function that explains how fiscal sustainability is achieved when the primary balance to GDP ratio has a positive function to public debt to GDP ratio. Using fiscal reaction function would enable to forecast the reaction of government towards an increase in debt ratio, with fiscal sustainability is achieve when the government will increase its primary balance when there is an increase in public debt ratio.

Another approach to the assessment of fiscal sustainability is by using the fiscal sustainability indicator methodology, which provides quantitative measures of fiscal sustainability for policy makers. Two of the main fiscal sustainability indicators are the primary gap indicator (Blanchard *et al*., 1990) and recursive algorithm (Croce & Ramon, 2003). Examples of research using the primary gap indicator are Blanchard *et al*. (1990) in the Organisation for Economic Co-operation and Development (OECD) countries, Langenus (2006) in selected European countries, and Aristovnik (2014) in Eastern European and former Soviet Union. For example, Blanchard *et al*. (1990) introduced primary gap indicators, which consist of short, medium and long term indicators, and found that most of the advanced economies are not fiscally sustainable. They used the medium term to include future economic fluctuations and long term to identify the impact of population ageing on fiscal sustainability. This fiscal indicator is able to assess future fiscal sustainability from future fiscal challenges.

This study uses the fiscal sustainability indicator approach in assessing fiscal policy sustainability for three reasons. First, the fiscal sustainability indicator approach would provide us with quantitative measures of fiscal sustainability, which could assist the fiscal authorities to set the tax rate and spending decisions after knowing the fiscal sustainability situation. Second, using the primary gap indicator, the government would be able to identify the impact of population ageing on fiscal sustainability and come up with appropriate policies, especially on public healthcare and public pension reform. It is generally believed that fiscal sustainability would be impacted differently to change in the population demography according to age (young and old) and current size of old age public expenditure (small and large). Therefore, an understanding how fiscal sustainability would be impacted by population ageing is a pertinent issue to investigate. For example, Blanchard *et al*. (1990) argued that population ageing would

impact fiscal sustainability of developed economies, because developed economies have a large share of old age population and a generous social protection system for the elderly. Therefore, impact of population ageing on an economy depends on the size of its elderly population as well as the generosity of the social protection system. Specifically, developing economies that tend to have less generous social protection systems are being affected less as compared with developed economies that have a more generous social protection system and thus, have less impact on its fiscal sustainability

Most of the studies done to assess the impact of population ageing on fiscal sustainability were conducted for the developed economies. A recent study done by Narayana (2014) to assess the impact of population ageing on fiscal sustainability in India used the generational accounting approach and revealed that India's long term fiscal policy is not sustainable and requires increase in tax rate to satisfy the intergenerational equilibrium. Other studies, such as Édes & Morgan (2014) and Morgan & Kawai (2013), also identified population ageing as one of the challenges for sustainable fiscal policy in emerging Asia. For Malaysia, a study done by Selvaratnam *et al*. (2009) examined the impact of ageing population on the public expenditure using the historical trend analysis and concluded that population ageing would increase the future public expenditure. Abdullah *et al*. (2012) also assessed the fiscal sustainability of Malaysia using the cointegration test of fiscal sustainability indicator and output, and identified that cointegration exists, indicating fiscal sustainability.

4. METHODOLOGY

This study uses the primary gap indicator approach as the fiscal sustainability indicator.

4.1 Primary Gap Indicator (Blanchard *et al*., 1990)

Primary gap states that if there is a gap between sustainable primary balance and current primary balance, then the fiscal policy is not sustainable. The first step of understanding fiscal sustainability is to understand government intertemporal budget constraint. This is basically the change in new issuance of government debt, which must be equal to the overall deficit of the period. The government intertemporal budget constraint is written as follows in nominal terms:

$$\Delta B_t = G - T + iB_{t-1} \quad (1)$$

where on the left side of the equation, ΔB_t is change in nominal value of debt or new debt created, which depends on the right side of the equation, which consist of G, which is the total government spending before debt payments minus taxes T or revenue plus iB_{t-1}, which is the total interest payments from nominal interest rate (ii) multiplied by the previous debt (B_{t-1}). The government spending minus total revenue is also called primary deficit and denoted by *D*. Using the data from 2015, where *G* is RM 232 billion, T is RM 219 billion, I is 4.17% and (B_{t-1}) is RM 582 billion, while ΔB_t is RM 37 billion or equal to the total overall deficit.

When the government intertemporal budget constraint is written in terms of ratio to nominal GDP to capture growing economy, the Equation 1 becomes:

$$\Delta b = g - t + (r - \theta)b = d + (r - \theta)b_{t-1} \quad (2)$$

Equation 2 states that change in the ratio of nominal debt to GDP (Δb) on the left side of the equation is equal to right side of the equation, which consist of the ratio of nominal government spending minus debt payment over nominal GDP (g) minus the ratio of nominal total revenue over nominal GDP (t) plus ratio

of interest payments, which consist of growth adjusted real interest rate from real interest rate (r) minus real GDP growth rate (θ) multiple with previous debt to GDP ratio b_{t-1}. The ratio of government spending (g) minus ratio of government revenue (t) is denoted as d or ratio of primary deficit to GDP.

Using the data from 2015 (Ministry of Finance, Malaysia), where g is 20.0%, t is 18.9%, while r is 2.0%, (θ)is 5.0% and b_{t-1}. is 52.7% and (Δb)is 1.1%. This should be similar with the difference between public debt ratio in 2015 of 54.5% and in 2014 of 52.7%. However, the difference between those two years is 1.8%, almost similar with the result from the equation.

After understanding the concept of government intertemporal budget constraint, the definition of fiscal policy sustainability can be constructed, which is a fiscal policy that would return the debt to GDP ratio back to its initial level (b_0). The first step is to construct the changes to debt from initial debt to any time given. Therefore, the debt to GDP ratio at any given time (n) is equal to:

$$b_n = b_0 exp(r-\theta)n + \int_0^n d_s \, exp(r-\theta)(n-s)ds \qquad (3)$$

The debt ratio at any time of n is denoted by b_n and is equal to the initial debt ratio b_0 growing at exponential rate of difference of real interest rate and real growth rate multiplied by n ($exp(r-\theta)n$) plus the total sum of value of primary deficits (d_s) growing at similar rate to initial debt.

In order to better understand Equation 3, it could also be used to forecast future public debt ratio. For example, the debt ratio in 2025 could be forecasted using 2015 data as the base year (b_0) or initial debt is the public debt ratio in 2015, which is 54.5%, while the growth adjusted interest rate $(r-\theta)$ is -0.03 that is real interest rate (r) of 2% (from nominal interest rate of 4% minus inflation rate of 2%) minus θ is real GDP growth of 5%, while n is 10 years and d_s is the primary deficit in 2015, which is 1.19%. Substituting all the variables into Equation 3 would produce a result of 28.20%. This result could be interpreted as in ten years; our public debt ratio would drop from 54.5% in 2015 to only 28.20% in 2025 within ten years with all variables in 2015 remaining constant.

In order to derive the fiscal sustainability equation from Equation 3, three steps are required. First, pre-multiplying both sides of Equation 3 with $exp-(r-\theta)$ which would discount both sides to time zero, resulting in:

$$\int_0^n d_s exp-(r-\theta)s \, ds = -b_0 + b_n exp-(r-\theta)\, n \qquad (4)$$

Second, taking the limit of Equation 4 as n goes to infinity would produce the definition of sustainability. As discussed earlier, for a fiscal policy to be sustainable, its debt ratio b_n needs to return to initial level b_0 and as n goes to infinity the discounted value of debt equals to zero:

$$\lim_{n\to\infty} b_n - (r-\theta)n = 0 \qquad (5)$$

Step 3 would combine Equations 4 and 5 to produce the second fiscal policy sustainability equation, which is:

$$\int_0^n d_s \, exp-(r-\theta)s \, ds = -b_0 \qquad (6)$$

Equation 6 states that fiscal policy is sustainable when the total sum of the future value of discounted ratio primary deficit to GDP is equal to the negative value of the current level of debt to GDP. In other words,

the current level of debts must be equal to the discounted value of total primary balance expected to incur in the future. This would mean that a government that currently has a debt outstanding needs to achieve primary budget surpluses that are large enough to satisfy Equation 6.

In order to arrive at the sustainable tax rate *t**, *d* or deficit to GDP in Equation 6 is replaced with *g* - *t* and solving for constant sustainable tax rate t^*t^* would result in:

$$t^* = (r-\theta)\left[\left(\int_0^\infty (g)exp-(r-\theta)s\,ds\right)+b_0\right] \quad (7)$$

Sustainable tax rate or ratio t^* is equal to annuity value future expected government spending (g) plus the difference between ex ante interest rate and growth rate $(r-\theta)$times current debt b_0.

4.1.1 Index of Fiscal Sustainability

If sustainable tax ratio is more than current tax ratio (*t**- *t*)>0, then there is a need for adjustment either by increasing future tax rate or reducing spending. If this happens, then the fiscal policy is determined as fiscally unsustainable. The size of adjustment needed is the gap between sustainable tax rate and the current tax rate (*t**- *t*).

The above equation is for infinite time horizon. For producing the finite time horizon, Equation 7 is manipulated to derive the finite sustainable tax rate or *(t*-t).*

$$t_n^* = (r-\theta)[b_0+[(1-exp-(r-\theta)n]^{-1}]\left[\int_0^n g\,exp-(r-\theta)s\,ds\right] \quad (8)$$

Equation 8 would state that the sustainable tax rate must be able to cover $(r-\theta)b_0$, which is the amount needed to ensure that the debt ratio remains constant without primary deficit and also cover the discounted value of total government spending in period time *0* and time *n* or average government spending during the period.

When *n* goes to infinity t_n^* converges to t^*, similar to Equation 7. When *n* goes to zero, the sustainable tax rate becomes $t_0^* = g+(r-\theta)b$. Then, the index of fiscal sustainability becomes:

$$t_0^* - t = g - t + (r-\theta)b = d + (r-\theta)b \quad (9)$$

The index of fiscal sustainability or fiscal sustainability indicator of sustainable tax rate gap($t^* - t$) is equal to primary deficit (*d*) plus the difference between real interest rate and real growth rate multiple with the debt to GDP ratio. From Equation 9, three type of indicators can be constructed based on time horizon; (1) short term gap (one year), (2) medium term gap (five years) and (3) long term gap (more than five years).

For example, taking the data in 2015, to determine the index of fiscal sustainability, the first step would be to calculate the sustainable tax rate t_0^*, which is the tax rate that would equal the government spending before interest payment (20.0%), plus the difference between real interest rate and real growth rate (-0.03%) multiplied with the previous debt ratio (52.7%), which is 20.5%. The second step would be to subtract with actual tax ratio in 2015 (18.9%), which would produce the result of 1.1%, which is similar to the primary deficit in 2015 of 1.1%.

4.1.2 Fiscal Sustainability Indicator: Short-Term, Medium-Term and Long-Term Gap

From Equation 9, the short-term gap is derived as primary balance (d) plus growth adjusted interest rate multiplied with the previous period public debt, as follows:

$$d + (r - \theta)b_0 \quad (10)$$

Medium-term gap, which is derived from Equation 8 is derived as follows:

$$[(average\ over\ next\ 5\ years\ of\ g) + avearge\ 5\ years\ of (r - \theta)b_0\ ] - t \quad (11)$$

Long-term gap is similar with medium term gap, except rather than five years, the period is longer such as 25 years.

$$[(g\ at\ long\ term\ period\ such\ as\ 25\ years)\ \ + long\ term\ period\ period\ ; e.g.\ 25\ years\ (r - \theta)b_0\ ] - t \quad (12)$$

5. EMPIRICAL RESULTS

5.1 Primary Gap Indicators

The short and medium-term gap estimates are shown graphically in Figure 2, and against military spending to GDP. The chart shows the dramatic turnaround in both fiscal sustainability indicators since the Asian financial crisis in 1998, with the short term and medium term gaps turned from negative gap indicating fiscal sustainability to positive gap indicating fiscal unsustainability in 1998. Military spending also registered the lowest level since 1982 at 1.6% in 1998. Both short and medium term gaps continued in the positive gap until 2003, when the medium term gap became negative at 0.8% but the short term gap remained positive at 2.3%, while the military spending was at the highest level since 1998 at 2.6%. The short term gap then returned to the negative gap in 2006 at 0.5% but the medium term gap became positive at 1.8%. Both short and medium gaps became positive since 2007 and remained until 2011.

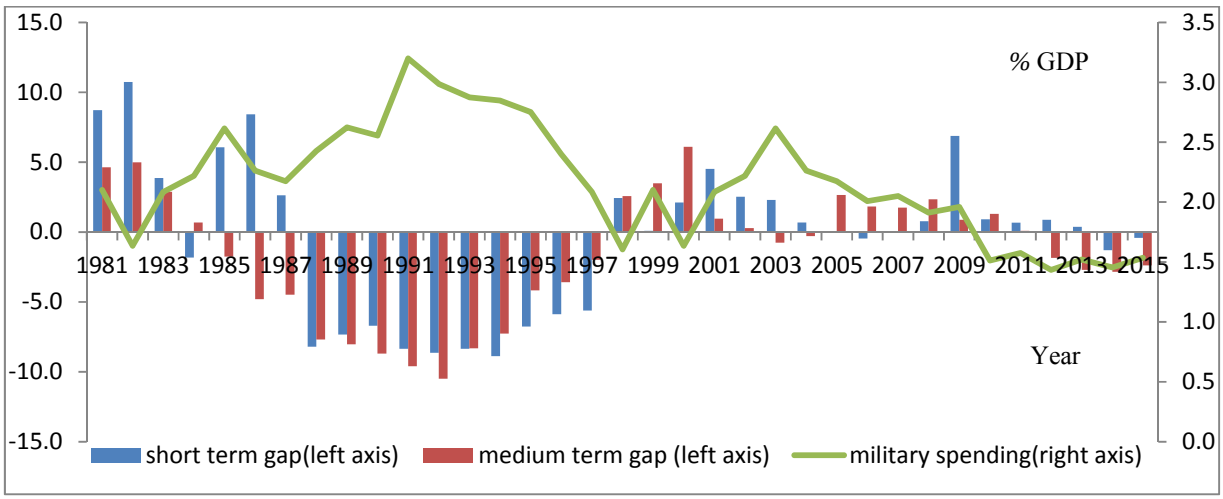


Figure 2: Fiscal sustainability indicator and military spending in Malaysia (percentage of GDP)

The global financial crisis increased the positive short term gap to the highest level since 1986 in 2009 at 6.9%. It also caused military spending to drop sharply from 2.0% in 2009 to 1.5% in 2010, and has been in a declining trend from 2.6% in 2003 to 2.0% in 2009. In 2012, medium term gap became negative for the first time since 2004 at 1.9%, indicating fiscal sustainability, and a small positive for short term gap, indicating small fiscal unsustainability at 0.9%. Both the short and medium term gaps became negative in 2014 for the first time since 1998 at 1.3 and 2.8% respectively. In 2015, both gaps had recorded smaller negative gap at 0.4% for short term gap and 2.4% for medium term gap. The chart also shows that Malaysia's fiscal sustainability has improved since 2009, when the gap had reduced since 2010 and returned to negative gap since 2014. Military spending from 2011 to 2015 was between 1.6% in 2011 and 1.5% since 2013, having declined from 2.0% in 2009. This could indicate that efforts to improve fiscal sustainability through fiscal consolidation since 2010, especially by budget cuts such as in military spending, had succeeded with the fiscal sustainability indicators of short and medium term gaps, with both becoming negative since 2014, indicating fiscal sustainability.

From Table 1, the two periods where short-term and medium-term gap were negative, indicating fiscal sustainability, were in the periods of 1986 to 1990 and 1991 to 1995. During these periods, the overall and the primary fiscal balances were in surplus. Short-term gaps were sustainable for four periods of 1986 to 1990, 1991 to 1995, 1996 to 2000 and 2011 to 2015, as certain years in this period recorded primary fiscal surplus. The most recent period of 2011 to 2015 showed the most improved performance of fiscal sustainability for both short (-0.1) and medium term gap (-2.0) as the first period of negative gap since the period of 1991 to 1995. In all the periods under study, except for 1986 to 1990, the difference in real growth rates and interest rates has always been positive, indicating higher real growth rate as compared to real interest rates. The highest difference of positive (6.2) was registered in the period of 1991 to 1995, which is also the period of the most sustainable fiscal policy. On the other hand, the only period where negative differential was recorded (-1.8) was in the period of 1986 to 1990, which is the period where the highest interest rates were recorded. In terms of military spending, the period of 1991 to 1995, which is also the most sustainable period, has the highest average of military spending of 2.9% of GDP, according to the National Defence Policy, the budget for military spending should consider the affordability as the basis for any arms procurements as the government could afford for higher military spending during this period. The most recent period of 2011 to 2015 has the lowest average military spending at 1.5% of GDP, but during this period, the fiscal sustainability indicators both short term and medium term returned to fiscal sustainability for the first time since the period of 1991 to 1995. The periods where economic recession occurred such as in 1985 and 2009 were also where fiscal policy is not sustainable as in this two periods both short term and medium term gap were positive. Fiscal stimulus were implemented in 1998, 2001 and 2009 to overcome the slower economic growth and this has resulted in fiscal policy to become unsustainable expect for the period where 1998 was included, the fiscal policy was sustainable due to the strong economic performance in the years before 1998.

Table 1: Alternative indicators of fiscal policy sustainability: Malaysia 1981-2015.

variables / Years	Public Debt Ratio (a)	Real GDP Growth Rate * (b)	Real Interest Rate (c)	Difference (b)- (c)	Short term gap	medium term gap	Military Spending
Average							

1981-1985	69.9	5.2	4.0	0.8	5.5	2.3	2.1
1986-1990	93.3	4.0	5.8	-1.8	-2.2	-6.7	2.4
1991-1995	56.4	9.5	3.3	6.2	-8.2	-8.0	2.9
1996-2000	35.2	5.0	4.5	1.5	-1.4	1.3	2.0
2001-2005	43.4	4.8	4.6	0.2	2.0	0.6	2.3
2006-2010	44.2	4.5	2.2	2.3	1.6	1.6	1.9
2011-2015	51.8	5.3	1.8	3.5	-0.1	-2.0	1.5

Source: BNM website* and author's calculation

The comparison between short-term and medium term gaps shows that there are major differences, except for the recent period of 2006 to 2010, where the both gaps were equal at 1.6%. This in contrast to the findings in Blanchard *et al*. (1990), where OECD countries' short term and medium term gaps did not show major differences. This can be attributed to the high volatility in growth rates for Malaysia as compared to OECD countries in the periods concerned.

The long term gap for Malaysia in this study used the time equivalent of 25 periods. Only long term gaps in 2015 are constructed based on projections of spending for the period of 2016 to 2040. The long term challenge that has been identified that could impact long term sustainability of fiscal policy is population ageing in Malaysia. Population ageing forecast is obtained from the Statistics Department of Malaysia (2012). Simulated pension spending was constructed with the assumption that the ratio of public pension expenditure to GDP changes in accordance with population ageing. The growth of healthcare spending is based on Malaysia following the trend of OECD countries from 2020 in terms of growth of public healthcare spending from 2020 to 2040. For the period from 2015 to 2020, the projection would follow the projection by the Ministry of Health, Malaysia for the Eleventh Malaysian Plan.

Table 2 indicates that Malaysia would see an increase in total spending ratio to GDP by 3.4% due to the impact of population ageing on public pension and healthcare. This is because the share of population aged 65 to the total population and above is expected to increase from 5% in 2010 to 11% in 2040.

Table 2: Future growth in general government non-interest spending (as percentage of nominal GDP).

	Non-interest spending as a percentage of GDP in 2015	**Ratio of share non interest spending in GDP in 2040 in share in 2015 accounting potential growth in**	
		Pension Spending	**Pension and healthcare spending**
Malaysia	20.00	1.40	3.40

Table 3 reflects the implied long-term gap for Malaysia, at positive 2.5% for pension spending compared to medium term gap of negative 1.1%, but when included with healthcare spending, the gap becomes positive at 4.5%, which indicates the amount of fiscal reforms that need to be undertaken in the public

pension and healthcare sector to ensure sustainability of the public healthcare and pension. This number specifies the size of adjustment either in taxes or spending that has to be changed to overcome the fiscal implications of ageing. This assumption is based on the continued growth rate of 5% from 2015 to 2040.

Table 3: Sustainability of fiscal policy in the long-run long term gaps based on projected growth of general spending on pensions and healthcare (as percentage of nominal GDP).

	Pensions only[a]	**Gap based on growth in pension and healthcare[b]**
Malaysia	2.5	4.5

[a]Defined as the difference between, on the one hand, general government receipts required on average over the current and next 25 years in order to return debt ratio to its initial level and on the other current receipts, taking into account the potential growth of public pensions associated with rising old-age dependency ratio.
[b]Same as footnote (a) but taking into account the effects of ageing on public health care spending.

6. SUMMARY AND CONCLUSION

The results of the study can be summarised as follows. First, economic recession has a negative impact on the fiscal sustainability in Malaysia due to lower economic growth and the fiscal stimulus needed to overcome the economic challenges. In general, the government will reduce military spending during periods of fiscal unsustainability. Second, population ageing is expected to have an impact on fiscal sustainability in Malaysia due to increasing costs of public healthcare. Third, the Malaysian government fiscal consolidation, such as through budget cuts in military spending, has been shown to be successful in reducing the deficit but less successful in reducing the debt ratio.

In addition, this paper contributes to the current research on the subject of defence and fiscal sustainability by analysing the relationship of military spending with the fiscal sustainability indicator over the period of 1981 to 2015. This study also investigates the impact of population ageing on the long term fiscal sustainability in Malaysia. This paper finds that; firstly, economic crisis will adversely impact the fiscal sustainability indicator as well as military spending. Secondly, fiscal consolidation through budget cuts in military spending has been able to improve the fiscal sustainability indicator performance, hence the fiscal sustainability condition in Malaysia. Thirdly, the population ageing has a significant effect on future fiscal sustainability in Malaysia, through expected increase in pension spending and healthcare expenditure. Finally, the sustainability of fiscal policy will increase the military spending, while unsustainability of fiscal policy will reduce military spending.

Some policy implications from the findings obtained from this study:

a) Reallocation of resources from military spending to other sectors, such as education, must be subject to a strategic security environment facing Malaysia. Any reduction in military spending should be based on informed decisions as regard to the current strategic security environment and to maintain Malaysia strategic interest and sovereignty. There should be a balance between fiscal sustainability and the need to provide defence and security. Fiscal consolidation should not be at the expense of growth and security, through large budget cuts to education and military spending. For military spending, there is a need to ensure that the military readiness is not being affected by budget cuts which could impact areas such as maintaining military assets that have expensive life cycle cost that includes maintenance and military personnel and must be able to be deployed to

address any security threats to the country immediately. The government need to consider full service life cost as the main criteria of any military procurement to determine its affordability and enhance the capability of the military.

b) The government needs to undertake painful fiscal reforms on public pension and healthcare to avoid unmanageable future burden on the government spending due to the impact of population ageing on fiscal policy. The deferment of fiscal reforms for the two sectors could lead to future large fiscal unsustainability, which could result in fiscal measures that could adversely impact economic growth and quality of life of the people.

REFERENCES

Abdullah, H., Muszafarshah, M.M. & Dahalan, J. (2012). An empirical study on fiscal sustainability in Malaysia. *Int. J. Academic Res. Bus. Soc. Sci.*, **2**: 72–90.

Afonso, A. (2004). Fiscal sustainability : The unpleasant European case. *FinanzArchiv: Public Finance Anal.*, **61**: 19-44

Aristovnik, A. (2014). Fiscal sustainability in the Eastern European and the Former Soviet Union. *SSRN*, 1–16.

Blanchard, O., Chouraqui, J.-C., Hagemann, R.P., & Sartor, N. (1990). The sustainability fiscal policy : New answers to an old question. *OECD Econ. Stud*.

Bohn, H. (1998). The behavior of U.S. public debt and deficits. *Quar. J. Econ.*, **113**: 949–963.

Burnside, C. (2005). *Fiscal sustainability in theory and practise A Handbook*. World Bank, Washington D.C.

Caruso, R. & Domizio, M.D. (2016). Military spending and budget deficits: the impact of US military spending on public debt in Europe (1988-2013). *Defense Peace Econ.*, 1–33.

Chalk, N., & Hemming, R. (2000). Assessing fiscal sustainabilities in theory and practice (81 No. 0). *IMF Working Paper*.

Cuddington, J.T. (1997). Analysing the sustainability of fiscal deficits in developing countries. *Int. Financ.*, *9706001*, *E*.

De Mello, L. (2008). Estimating Fiscal Reaction Function: Case of debt sustainability in Brazil. *Appl. Econ.*, **3**: 271–284.

Dunne, J.P. (2015). Economic effects of military expenditure in developing countries. *Peace Dividend*, 439–464.

Édes, B.W. & Morgan, P.J. (2014). Managing fiscal sustainability and aging in emerging Asia. *Public Pol. Rev.*, **10**: 319-348.

Elliott, G. & Kearney, C. (1988). *The Intertemporal Government Budget Constraint and Tests for Bubbles*. Reserve Bank of Australia, Australia.

Frederiksen, P.C. (1991). Economic growth and defense spending: Evidence on causality for selected Asian countries. *J.Philipp. Dev.*, **13***:* 131–147.

Ghatak, S., & Sanchez-Fung, J. R. (2006). Is fiscal policy sustainable in developing economies? *VATT Discuss. Papers*, **384**: 31.

Hamilton, J. D., & Flavin, M. A. (1985). On the limitations of government borrowings, A framework for empirical testing. *Amer. Econ. Rev.*, **76**: 808–819.

Harris, G. (1986). The determinants of defence expenditure in the ASEAN Region. *J. Peace Res.*, **23**: 41–49.

Hirnissa, M. T., Habibullah, M. S., & Baharom, A. H. (2009). Military expenditure and economic growth in Asean-5 Countries. *J. Sustain. Dev.*, **2**: 192–202.

International Monetary Fund. (2015). Fiscal policy and long-term growth. *IMF Policy Paper*, June: 1–184.

James, K. S., & Baba, I. (2015). Fiscal sustainability in the Ghanaian economy : A fiscal reaction function

approach. *Asian Online J.*, **2**: 16–20.
Khalid, M.A. & Mohd Noor, Z. (2015). Military expenditure and economic growth in developing countries: Evidence From system GMM estimates. *J. Emerg. Trends Econ. Manage. Sci.*, **6**: 31–39.
Langenus, G. (2006). Fiscal sustainability indicators and policy design in the face of ageing. *SSRN Elec. J.*, **102**: 719–749.
Ministry of Finance, M. (2012). Public Sector Finance. In *Economic Report 2012/2013* (pp. 122–148). Percetakan Nasional Berhad.
Ministry of Finance, M. (2017). *2017 Estimated Federal Expenditure*. Available online at: http://www.treasury.gov.my/index.php/en/budget/estimated-federal-expenditure.html (Available online at 22 February 2017).
Morgan, P. J., & Kawai, M. (2013). *Long-term issues for fiscal sustainability in emerging Asia*. *Public Pol. Rev.*, **9**: 751-770.
Narayana, M.R. (2014). Impact of population ageing on sustainability of India' s current fiscal policies : A generational accounting approach. *J. Econ. Ageing*, **3**: 71-83.
Qiong, L., & Junhua, H. (2015). Military expenditure and unemployment in China. *Proc. Econ. Financ.*, **30**: 498–504.
Ross, A. L. (1990). Growth debt and military spending in Southeast Asia. *Contemporary Southeast Asia*, **11**: 243–264.
Selvaratnam, D. P., Idris, N. A., Abu Bakar, N., & Ong, B. K. (2009). Kesan peningkatan jangka hayat di Malaysia. *Persidangan Kebangsaan Ekonomi Malaysia (PERKEM IV)*, **1**: 305–315.
Singh, A. (2012). South China Sea disputes: Regionalissue, global concerns. *Maritime Affairs*, **8**: 116–135.
Statistic Department of Malaysia. (2012). *Population Projections*. Statistic Department of Malaysia, Malaysia.
Stoian, A. & Roman, M. (2006). Fiscal sustainability based on reaction function : Case study. **6**: 19.
Tang, C.F. (2008). Defence expenditure and economic growth in Malaysia: A reassessment using bounds and modified Wald Tests. *CFAI J. Public Financ.*, **6**: 45–51.
Uctum, M., & Wickens, M. (1996). *Debt and Deficits Ceilings and Sustainability of Fiscal Policies: An Intertemporal Analysis*. Federal Reserves Bank of New York, New York.

MILITARY SPENDING, ECONOMIC GROWTH AND STRUCTURAL INSTABILITY: CASE STUDY FOR MALAYSIA

Wan Farisan Wan Sulaiman*, Zulkefly Abdul Karim, Norlin Khalid & Riayati Ahmad

School of Economics, Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), Malaysia

*Email: wantreasury@yahoo.com

ABSTRACT

The present study provides new empirical evidence on the relationship between military spending and economic growth using a case study of Malaysia for the period of 1960-2015. Using the military spending and growth nexus, the new econometric method of bootstrap rolling window (BSRW) approach was used to test the possibility of structural breaks. By applying the full sample bootstrap Granger causality tests, the Granger causal link is found between military spending and gross domestic product (GDP), but parameter instability tests show that the estimated vector autoregressive (VAR) models to be unstable. However, using a BSRW estimation procedure, evidence is found of unidirectional Granger causality from growth to military spending in just 2015. This indicates that standard Granger causality tests, which neither account for structural breaks nor time variation, may be invalid.

Keywords: *Military spending; economic growth; bootstrap; time varying causality*.

1. INTRODUCTION

Military spending can have an impact on the economic growth of a country through security of the country and building of an industrial complex (Aye *et al*., 2014). The recent budget cuts for military spending in Malaysia by 14% and military spending over GDP at just 1.1% could impact the development of the defence industry and subsequently the goal of having high end military technology and economic spin off to civilian sectors (Balakrishnan, 2008). This could hinder the transformation of the nation into an innovation led high income developed nation (Akoum, 2016). There is a lack of consensus among economist on the impact of military spending on economic growth, whether it is positive, negative or neutral (Dunne & Tian, 2013). However, most of the studies indicated that there is a negative effect of military spending on economic growth. Thus, a good understanding of the relationship between military spending and economic growth is crucial for policy makers if they are to ensure that military spending would not have an adverse effect on economic growth. This is because military spending could have a positive impact on economic growth through Keynesian demand side or negative impact through the neoclassical supply side. Specifically, expansion in military spending according Keynesian could increase the demand for military goods and new military jobs, and increase growth. In contrast, expansion in military spending according to neoclassical supply side economics could reduce growth through crowding out effect in investment as well as diverting economics resources from productive sector (Deger & Smith, 1983).

The aim of this paper is to provide empirical evidence on the relationship between military spending and economic growth in a developing economy (Malaysia). Specifically, this study aims to investigate the causal relationship between military spending and economic growth using a bootstrap rolling window

Granger non-causality approach to detect any structural change during the period that could make full sample period Granger causality become invalid.

The significant contribution of this study differs from previous work in two ways. First, in Malaysia there have been only a few studies (for example, Frederiksen, 1991; Tang, 2008; Hirnissa *et al*., 2009;) that have examined the relationship between military spending and economic growth but none of these studies have identified any parameter instability or structural change. Therefore, this study improves upon previous studies by identifying parameter instability by applying the parameter instability test and structural break using the bootstrap rolling window Granger non-causality approach. There are two possible reasons for using the parameter stability test and bootstrap rolling window approach in identifying structural break. First, parameter instability would be able to determine if either the short or long run parameters are not stable, which would make the full sample Granger causality test invalid. Second, the bootstrap rolling window approach would able to identify structural breaks, which would make full the sample Granger causality test invalid and determine which period there is structural break (Aye *et al*., 2014).

Second, this study also provides a case study of the dynamics of the relationship between military spending and economic growth within a developing country in Asia, which is Malaysia, and also investigates the possibilities of existence of any structural breaks in the relationship, using the most advanced econometric method. This is important because not only Malaysia does play an important role within the ASEAN region, but also the history of military build-up during the communist insurgency from 1948-1990, and post-insurgence and economic crisis military spending decline present a unique pace of change in military spending due to the impact of economic crisis (Hirnissa *et al*., 2009; Kheng, 2009; Abd Rahman, 2013). This study will provide more details on how the military spending adjust to changes in the economy by observing structural breaks within the relationship.

The remainder of the paper is as organised as follows. In Section 2, describes the theories and literature on military spending and economic growth. Next, in Section 3, the data and empirical model used is outline, while in Section 4, the results are presented. Finally, in Section 5, the summary and concluding remarks are discussed.

2. LITERATURE REVIEW

An important issue in the study of the relationship between military spending and economic growth is the lack of consensus on the impact of military spending and economic growth. Previous studies, for example Benoit (1978), Looney (1988) and Halicioglu (2004), have shown that there is a positive relationship between military spending and economic growth. However, the problem with these results are that they depends on the variables used, such as investment or export type of countries (developed, developing or less developed), period used (either during conflict or peace period) as well as the method applied to analyse. The results are also not conclusive as in the studies, certain countries in same study of various countries, such as resource rich countries, have positive impact while less resource restrained countries have negative impact (Looney, 1988). This means that generalisation of the results across countries could not be done as it depends on the circumstances of the countries.

In contrast, Biswas & Ram (1986), Deger & Smith (1983) and Shahbaz & Shabbir (2010) have found out that there is a negative relationship between military spending and economic growth. This result could be explained through the neoclassical model, where an increase in military spending will reduce economic growth due to the crowding out effect where resources used in military spending could be utilized in more productive sectors of the economy, such as healthcare or education. Furthermore, Dunne (2015) has indicated that there is no significant relationship between military spending and economic growth.

It is generally believed that individual countries' economic growths react differently to military spending according to the level of development. Therefore, understanding how a developing country's growth, for example Malaysia, reacts to military spending is an interesting issue to investigate as well as important to the policy makers. For example, Frederiksen (1991), by comparing five ASEAN countries including Malaysia, found out that there is relationship from growth to military spending but not the other way. In another study, Habibullah & Law (2008) investigated the relationship for selected Asian countries using the panel error correction method and showed that there is no relationship. On the other hand, the study by Tang (2008) on only one country of Malaysia suggested that military spending and economic growth has a negative relationship using the bound test and a bilateral relationship using the modified Wald (MWALD) test. The study conducted by by Hirnissa *et al*. (2009) using bound test and dynamic ordinary least square (OLS) found that there is no relationship between military spending and economic growth, which would support the findings by Habibullah & Law (2008). To overcome the problem of not relating the relationship analysis to growth model, Khalid & Mohd Noor (2015) use the augmented Solow growth model on 35 developing countries, including Malaysia, and applied generalised method of moments (GMM) estimation to indicate that military spending has a negative effect on economic growth. Currently, there is a lack of studies that investigate the relationship between military spending and economic growth using the bootstrap rolling window approach, which would overcome the issue of structural breaks in the relationship.

3. DATA AND EMPIRICAL MODEL

3.1 Data

The empirical analysis employs annual data on GDP and military spending for Malaysia over the period of 1960-2015. The data for real GDP is obtained from World Development Indicators (WDI) from World Bank. All data for military spending in nominal value obtained from Stockholm International Peace Research Institute (SPIRI), and then changed into ratio form by dividing with nominal GDP obtained from Statistics Department Malaysia (DOS). Both the military spending ratio and real GDP are transformed to logarithmic form and denoted as military spending and growth. All data are transformed into log levels as shown in Figure 1.

Table 1: Description of variables (annual data for 1960-2015:Number of periods (years)-55).

Variable	Description	Source
ME	Military Spending	SIPRI (2015)
RGDP	Real GDP	WDI (2015)
GDP	Nominal GDP	DOS

The high value of log of military spending for the period of for the period of 1965-1982, as shown in Figure 1, can be attributed to the period when Malaysia started to increase its military spending at the end of 1960 due to the termination of the defence agreement with the UK and to defend the country against Indonesia during the confrontation (1963-1966) and Communist Party threat (1948-1990). Malaysia's military burden (military spending as a share of GDP) reached a high of 5.4% in 1971 and was at 1.5% in 2015.

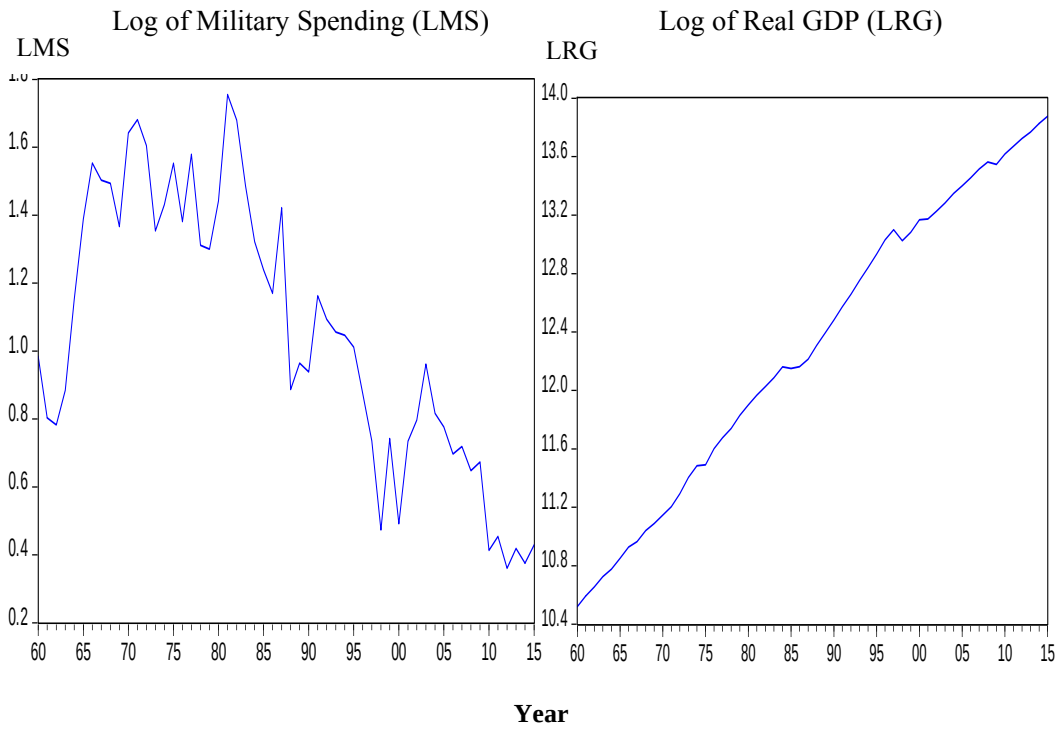


Figure 1: Log of military spending and real GDP.

The purpose of this paper is to analyse the causal relationship between military spending and GDP for Malaysia. The method that is usually used for this purpose is the Granger noncausality approach based on a bivariate vector autoregressive (VAR) framework with the understanding that two variables are not causal if adding another variable does not improve the prediction of the first variable. Among the test statistics under the Granger noncausality approach are standard causality test statistics for joint restriction and standard asymptotic properties, such as Wald, likelihood ratio (LR) and Lagrange multiplier (LM) statistics. However, these test statistics cannot be used when the underlying time series data is nonstationary, as these test statistics may not have standard asymptotic distributions (Sims *et al*., 1990; Toda & Phillips, 1993). In order to overcome this problem, Toda & Yamamoto (1995) proposed a modified Wald test by estimating an augmented VAR model with integrated at first order (*I*(1)) variables to obtain a standard asymptotic distribution for the test. However, this Wald test was proven by Shukor & Mantolos (2004) using Monte Carlo simulations as not having the correct size in small and medium size samples. To improve the size and power, Shukur & Mantalos (2000) suggested to use the residual bootstrap (RB) method critical value. This method has been proven by various Monte Carlo simulations studies (Shukur & Mantalos, 2000; Hacker & Hatemi-J, 2006) to have better performance as compared to other standard asymptotic tests regardless of cointegration properties. Based on this reason, this paper will use the RB based modified cointegration test (CR) test to examine the causal relationship between military spending and output for Malaysia.

In order to show the RB based modified LR causality test in this study, the following bivariate *VAR*(p) process is considered:

$$y_t = \emptyset_0 + \emptyset_1 y_{-1} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t \quad t = 1,2 \ldots T \tag{1}$$

where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})$ is a white noise process with zero mean and covariance matrix Σ and p is the lag order of the process. In the empirical section, optimal lag length is determined by the Schwartz information criteria (SIC). This study will split y_t in Equation 1 into two subsectors, military spending (y_{t1}) and GDP (y_{t2}).

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \emptyset_{10} \\ \emptyset_{20} \end{bmatrix} + \begin{bmatrix} \emptyset_{11}(L) & \emptyset_{12}(L) \\ \emptyset_{21}(L) & \emptyset_{22}(L) \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{12} \\ \varepsilon_{22} \end{bmatrix} \quad (2)$$

where $\emptyset_{ij}(L)=\sum_{k=1}^{p+1} \emptyset_{ijk} L^k, i,j = 1,2$ and L is the lag operator defined as $L^k X_k = X_{t-k}$. Based on Equation 2, the null hypothesis that military spending does not Granger cause GDP (H_0^M) can be tested by imposing zero restriction $\emptyset_{1,2}=0$ for $i = 1,2 \ldots p$, i.e., military spending does not have predictive content, or is not caused for GDP if the joint zero restrictions under the null hypothesis are not rejected:

$$H_0^M = \emptyset_{12,1} = \emptyset_{12,2} = \cdots \emptyset_{12,p} = 0 \quad (3)$$

Similarly, the null hypothesis that GDP does not Granger cause military spending (H_0^Y) can be tested by imposing zero restriction $\emptyset_{2,1}=0$ for $i = 1,2 \ldots p$, i.e., GDP does not have predictive content, or is not caused for military spending if the joint zero restrictions under the null hypothesis are not rejected:

$$H_0^Y = \emptyset_{21,1} = \emptyset_{21,2} = \cdots \emptyset_{21,p} = 0 \quad (4)$$

The Granger causality test in Equation 3 and 4 can be linked to military spending led growth (MLG) nexus and growth led military spending (GLM) nexus hypotheses in two ways. First, under the narrow definition, rejection of H_0^M but not H_0^Y establishes evidence in favour of MLG. On the contrary, if the null hypotheses under H_0^Y is rejected but not H_0^M , then this supports the GLD hypothesis. Second, a broader definition of causality of public debt for GDP can be adopted. Under the broader definition, evidence in favour of the MLG hypothesis is established if H_0^M is rejected, or both H_0^M and H_0^Y is rejected. Analogously, rejection of H_0^Y or rejection of both H_0^M and H_0^Y establishes evidence in support of the GLD hypothesis. The narrow definition requires unidirectional causality running from one variable to another. While the broader definition allows bidirectional causality, in this study, the causality inference is made under both narrow and broader definitions.

The direction and sign of causality between military spending and output have significant implication for fiscal policy. If the causality is unidirectional and the direction of causality is running from military to GDP with positive impact, which would support MLG, then it may be possible to boost economic growth with higher military spending. Several literatures stress the role of military spending in achieving economic growth (Benoit, 1978; Narayan *et al.*, 2007). In the case of bidirectional causality between military spending and output, second round effect from GDP to military spending will reinforce impact on growth. On the other hand, if a unidirectional causality exists running from GDP to military spending in support of the GLD hypothesis, military spending policies will be ineffective and policy makers need to concentrate first on policies that promote output growth. This case may be true for countries at high levels of military spending, where military spending may also result in lower long term economic growth as more government spending needs to be allocated to pay for debt expenses.

The assumption in the Granger non-causality test is that the parameters used in the VAR model are constant over time. This assumption cannot not hold if there is structural change and Granger (1969) strongly pointed out that parameter non-constancy has become a crucial issue in empirical studies.

Therefore, this paper will conduct tests for short-run and long-run parameter stability to check for constancy. For the short-run test, this paper will use the supremum (sup) -LR, mean-LR and exponential (exp)-LR tests developed by Andrews (1993). All these short-run tests are based on Lagrange multiplier statistics, which test the null hypothesis of parameter stability against structural change in each sample time. The sup-LR test is used to evaluate for swift regime change while the mean-LR and exp-LR tests are used to evaluate for stable relationship over time. One of the requirement of these short-run tests is to trim the sample data by 15% from both ends of the sample and calculate the fraction of the sample [0.15, 0.85] (Andrews, 1993).

For the long-run test, it will use the fully modified ordinary least squares (FM-OLS) estimator developed by Park & Phillips (1988) to estimate the parameters of cointegration regressions. FM-OLS will then be used in the limiting condition (LC) test (Nyblom, 1989; Hanson, 1992) to check the stability of long-run parameters. The LC test is also a test of cointegration when the data series are integrated in the first order.

The presence of structural change in other studies is included in the estimation using methods such as dummy variables and sample splitting, but doing so would expose to pre-test bias. To overcome the issue of parameter non-constancy and pre-test bias in model estimation, this paper will adopt a bootstrap estimation. The impact of structural change is examined by using the rolling window causality test, which is based on the modified bootstrap test. Structural changes may create shifts and patterns that change over time. The problem of structural changes and parameter non-constancy in full sample is overcome by applying the bootstrap causality test to the rolling window subsamples for $t = \tau - l + 1, \tau - l, \ldots, \tau = l, l + 1, \ldots, T,$ where l is the size of the rolling window (Balcilar *et al.*, 2010). Rolling window is a technique that uses a fixed- length moving window, such as 15 years, sequentially from the beginning to the end of the sample by adding one observation from ahead and dropping one from behind. For example, the first output is from 1985 as it uses data from 1970 to 1985, while for 1986 uses data from 1971 to 1986, and so on until 2015. This technique is used for two reasons: first, as time changes the relationship between the variables also changes, and second, to capture the instability in different sub-samples caused by structural changes.

To examine the causal relationship between the military spending and real GDP, this paper will use the bootstrap rolling-window approach in four steps. First, we use the unit root tests proposed by Dicky & Fuller (1979) and Phillips & Peron (1988) to test the stationarity of the two variables. Second, the full-sample temporal (Granger) causality test is applied to identify if any relationship between the two variables exists. Third, the short-run and long-run parameter stability tests are conducted. The short-run tests of sup-LR, Mean-LR and Exp-LR are used to check parameter constancy in the short-run. The long-run tests of for LC parameter stability and cointegration test (Johansen, 1991) using FM-OLS estimator are used to determine long-run constancy and cointegration relationship between the two variables. Finally, the rolling VAR regressions are estimated and the Granger causality test is conducted using a 15 year window.

4. EMPIRICAL ANALYSIS AND RESULTS

In this section, the procedure mentioned above is applied to the GDP and military spending series. The results of the augmented Dicky Fuller (ADF) and Phillips-Peron (PP) unit root tests including intercept as well as an intercept and trend in the test regression are reported in Table 1. This unit root tests are conducted to test if there is a need to differentiate the time series data to make it stationary. The null hypothesis of non-stationarity could not be rejected for GDP and military spending at the 5% significance level, but could for the first differences, implying both series are integrated at the first order, i.e. *I*(1). Testing for common stochastic trend, which implies a cointegrating relationship between GDP and military spending, is done using Johansen's (1991) maximum likelihood cointegration method. An

optimal lag order of one is suggested based on the Akaike-information criterion (AIC) cointegrated result, which is based on the trace and maximum eigenvalue statistics as reported in Table 2. The null hypothesis of no cointegration is rejected at 5% significance level, suggesting a long-run relationship between GDP and military spending.

Table 1: Results of the unit root tests.

	ADF		PP	
Series	**Constant**	**Constant and Trend**	**Constant**	**Constant and Trend**
Level				
Military Spending	-1.597	-3.092	-1.555	-2.805
GDP	7.113	1.114	7.336	1.390
First Difference				
Military Spending	-6.163***	-6.493***	-8.566***	-13.696***
GDP	-1.267	-7.284***	-4.344***	-7.284***

*Significance at 10%. **Significance at 5%. ***Significance at 1%.

Table 2: Results of the multivariate cointegration tests.

Series	H_0	H_1	**Trace Statistics**	**Maximum Eigenvalue**
Military Spending	r=0	r>0	33.357	22.677
GDP	r≤1	r>1	10.680	10.680

One-sided test of the null hypothesis (H_0) that the variables are not cointegrated against the alternative (H_1) of at least one cointegrating relationship. The critical values are taken from MacKinnon (1996) with 5% critical values equal to 15.49 for testing $r = 0$ and 3.84 for testing $r \leq 1$ for the trace tests. The corresponding values for the maximum eigenvalue tests are 14.26 and 3.84.

Table 3 shows the estimation results for an optimal lag order of one, as indicated by the AIC and Wald statistics, and bootstrap LR full sample Granger causality test in the VAR model. The null hypothesis that military spending does not Granger cause GDP is not rejected for both tests at any significance level while the null hypothesis that GDP does not Granger cause military spending is rejected at 5% significance level for the Wald test and 10% significance level for the LR test, indicating that GDP Granger causes military spending. For the full sample of 1960-2015, there is evidence of a long-run relationship between military spending and economic growth.

Table 3: Results of the full sample Granger causality tests.

	H_0: Military Spending does not Granger cause GDP		H_0: GDP does not Granger cause military spending	
	Statistics	***p*-value**	**Statistics**	***p*-value**
Wald test	0.385	0.825	9.014	0.010**
Bootstrap LR test	1.067	0.288	6.657	0.060*

To investigate whether the long-run or short-run relationship results is valid, parameter stability tests are conducted, which consists of sup-LR, exp-LR, mean-LR for the short-run relationship, and Lc for the long-run relationship, as shown in Table 4. The results for Lc test from Hansen instability test indicate that both military spending and GDP have no long-run stable parameters or no long-run cointegration between the two variables. For the short-run parameter stability test, only the sup-LR statistics are significant indicating that there is short-run parameter instability against one-time sharp shift in parameter, while the other short-run parameter stability tests of Exp-LR for gradual regime shift and Mean-LR for the Martingale process are insignificant indicating no parameter instability. Using these findings, we can conclude that there is instability in the short-run and long-run parameters of the VAR model, with the results of the full sample Granger causality test being invalid, and supports for the conduct of the bootstrap rolling window regression approach to further investigate the cause of the instability. The purpose of conducting the parameter stability test is to determine the presence of structural change in the relationship between economic growth and military spending. The bootstrap rolling window regression would investigate in more detail through performing causality test using the residual-based bootstrap method by changing the subsample by adding one observation from the front and dropping from the end. The window is the size of sample in one regression is denoted by l, while the full sample length is denoted by T, providing T-l sequence of causality test.

Table 4: Results of the parameter stability tests.

	Military Spending		**GDP**	
	Statistics	***p*-value**	**Statistics**	***p*-value**
Sup-LR	2.884	0.264	5.321	0.004***
Exp-LR	0.865	0.131	0.860	0.131
Mean-LR	1.548	0.124	1.314	0.177
LC	0.826	<0.01***	0.679	<0.01***

*Significance at 10%. **Significance at 5%. ***Significance at 1%. p-values are calculated using 1,000 bootstrap repetitions.

The boostrap *p*-values of the rolling test statistics and the impact of each series on the other are shown graphically in Figures 2-5, with the horizontal axes showing the final observable in each of the 15 year rolling window. Figure 2 presents the bootstrap *p*-values of the rolling statistics, testing the null hypothesis that military spending does not Granger cause GDP for Malaysia from 1975 to 2015, and

shows that the null hypothesis is not rejected for all periods at 10% significance level, which is comparable with the study by Aye *et al*. (2014) for South Africa from 1966 to 2010, whereby it is is not rejected except for 1973, 1982 and 2000-2002. Figure 3 shows the bootstrap estimates of the sum of the rolling window coefficients for the impact of military spending on the GDP for Malaysia at 10% significance level, which shows positive predictive power for all periods except 1989 to 1994, as compared to the study on South Africa, whereby the results suggest that military spending has a positive predictive power for GDP during the 1973-1975 subperiod but negative predictive power during the 1998-2005 subperiod (Aye *et al*., 2014).

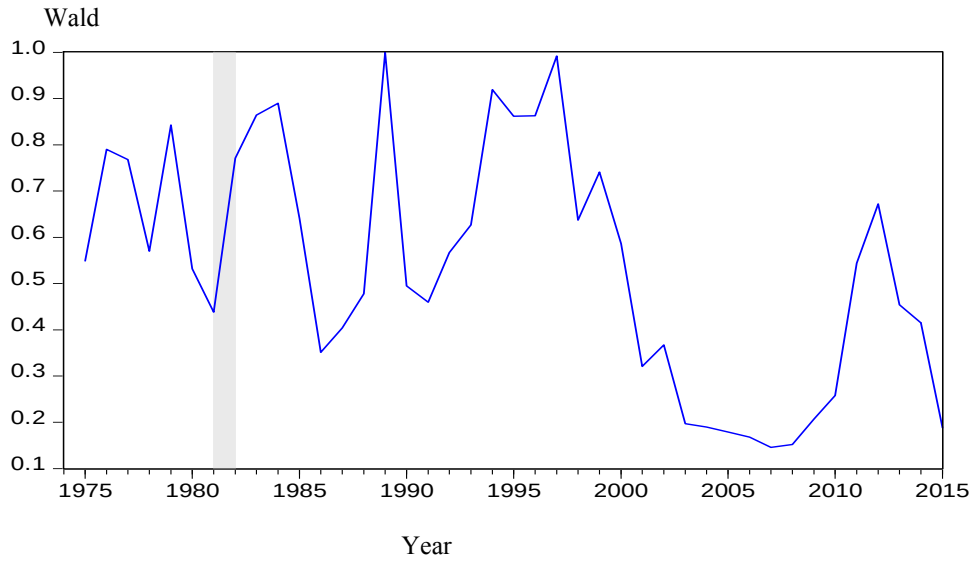


Figure 2: Bootstrap *p*-values of the LR test statistics to test the null hypothesis that military spending does not Granger cause GDP.

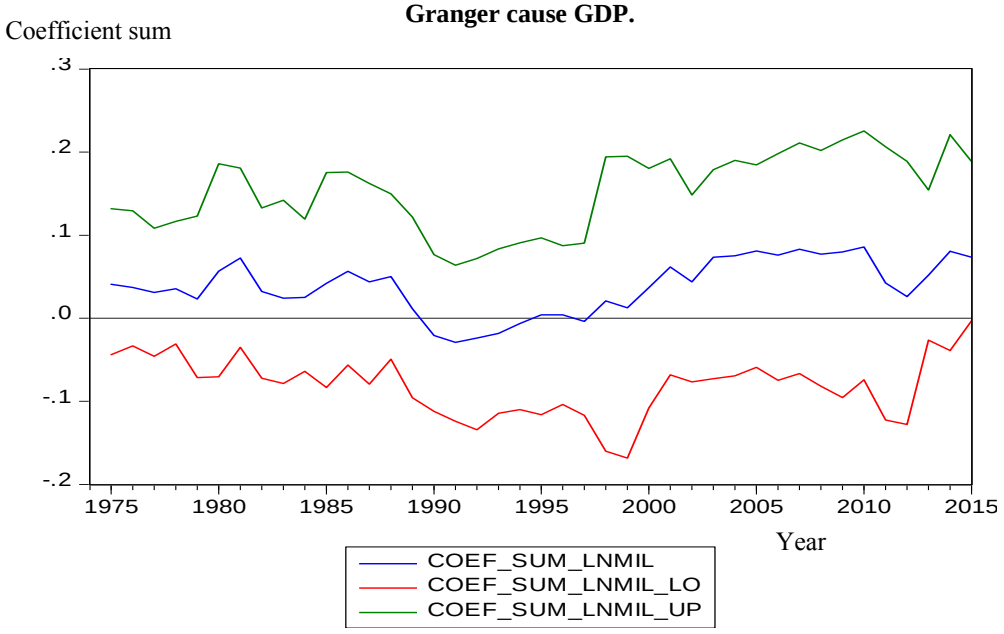


Figure 3: Bootstrap estimate of the sum of rolling window coefficients for the impact of military spending on GDP.

Figure 4 shows the bootstrap *p*-values of the rolling test statistics to test the null hypothesis that GDP does not Granger cause military spending as non-causality tests are evaluated at 10% significance level. The figure shows that the null hypothesis is rejected at 10% significance level for only the period of 2015. All the economic crisis periods in 1985, 2001 and 2009 does not have an impact on military spending. This is in contrast to the study of South Africa that shows that GDP does not Granger cause military spending except for the periods in 1981, 1995 and 1997 (Aye *et al*., 2014).

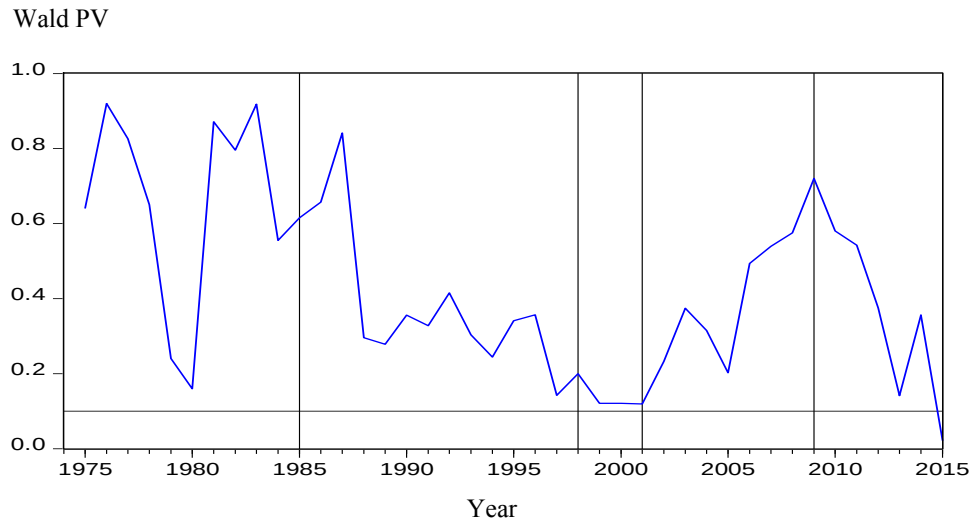


Figure 4: Bootstrap *p*-values of the LR test statistics to test the null hypothesis that GDP does not Granger cause military spending.

Figure 5 shows the bootstrap estimates of the sum of the rolling coefficients for the impact of GDP on military spending and suggest that GDP has negative predictive power for military spending in the 2015 subperiod with coefficient of -0.6. Following the fiscal consolidation since 2010, which probably reflects the reduction in military spending due to the decline in government spending and allocation to other sectors, such as healthcare. GDP has a positive predictive power for 1975-1977 and 1981-1983, and negative predictive power for 1978-1981 and 1984-2015. This is comparable with the South African study by Aye *et al*. (2014) which shows that GDP has positive predictive power for, 1975, 1977 and 1979-1984, and negative predictive power from 1995-2001 subperiods.

Overall, the bootstrap rolling window Granger causality results provide some support for the hypothesis of unidirectional causality from economic growth to military spending especially in 2015. This is in contrast to the results using the full sample, which establish predictive contact from GDP to military spending. These findings point to the fact that using the standard Granger causality tests, which neither accounts for structural breaks or time variation in the relationship between economic variations, may be incorrect (Aye, 2015).

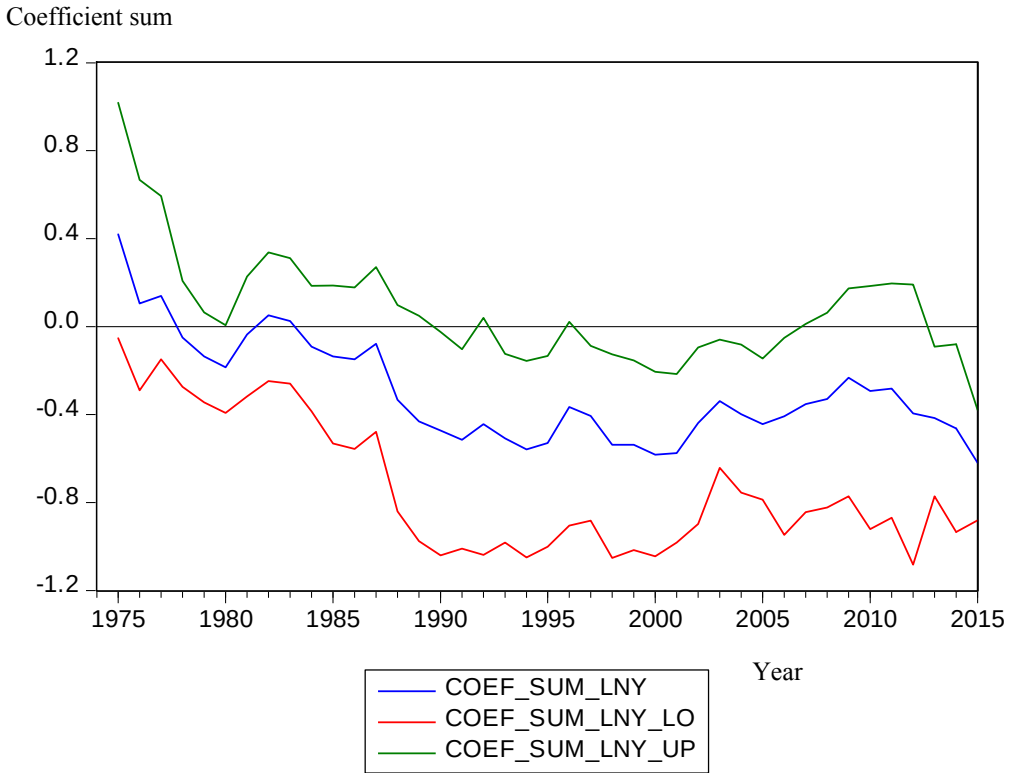


Figure 5: Bootstrap estimate of the sum of rolling window coefficients for the impact of GDP on military spending.

5. CONCLUSION

This paper contributes to the increasing literature on military spending and economic growth, by providing a case study of Malaysia and considering the possibilities of structural breaks in the relationship, using techniques that allow inference on whether or not the series are integrated-cointegrated. The full sample bootstrap Granger non-causality test suggested significant relationship from economic growth to military spending. The parameter stability tests found that estimated VAR to be unstable, suggesting that the inference may be invalid. Allowing for structural change by using bootstrap rolling window estimation, it was found that economic growth has negative predictor power for military spending in the 2015 period. This result supports unidirectional causality, which suggests that economic growth may have had a negative effect on military spending in earlier periods.

REFERENCES

Abd Rahman, S.A. (2013). *History of Special Operations Forces in Malaysia*. Naval Postgraduate School, Monterey, California

Akoum, I. (2016). Research, Development and Innovation in Malaysia: Elements of an Effective Growth Model. *Asian Econ. Financ. Rev.*, **6**: 390–403.

Andrews, D.W.K. (1993). Test of parameter instability and structural change with unknown change point. *Econometrica*, **61:** 821–856.

Aye, G.C. (2015). Causality between financial deepening and economic growth in Nigeria: Evidence from a bootstrap rolling window approach. *J. Econ., Bus. Manage.*, **3**: 795–801.
Aye, G.C., Balcilar, M. & Dunne, J.P. (2014). Military expenditure , economic growth and structural instability : a case study of South Africa. *Defense Peace Econ.*, **25**: 1–17.
Balakrishnan, K. (2008). Defence industrialisation in Malaysia : Development challenges and the revolution in military affairs. *Secur. Chall.*, **4**: 135–155.
Balcilar, M., Ozdemir, Z.A. & Arslanturk, Y. (2010). Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window. *Energy Econ*, **32**: 1398–1410.
Benoit, E. (1978). Growth and defense in developing countries. *Econ. Dev. Cult. Change*, **26**: 271–280.
Biswas, B. & Ram, R. (1986). Military expenditures and economic growth in less developed countries: An augmented model and further evidence. *Econ Dev Cult Change*, **34**: 361–372.
Brzoska, M. (2006). Trends in global military and civilian research and development (R&D) and their changing interface. *Proc. Int. Semin. Defence Financ. Econ.*, pp. 289–302.
Chu, A. C. & Lai, C.-C. (2009). *Defense R&D: Effects on Economic Growth and Social Welfare*. Munich Personal RePEC Archive, University Library of Munich, Germany.
Deger, S. & Smith, R. (1983). Military expenditure and growth in less developed countries. *J. Conflict Resolut.*, **27:**, 335–353.
Dunne, J.P. (2015). Economic effects of military expenditure in developing countries. *The Peace Dividend*, Emerald Group Publishing Ltd., Bradford, UK, pp. 439–464.
Dunne, J.P. & Tian, N. (2013). Military expenditure and economic growth: A survey. *Econ. Peace Secur. J.*, **8**: 5–11.
Dunne, P. & Braddon, D. (2008). *Economic Impact of Military R&D*. School of Economics, Bristol Business School ,University of the West of England, Bristol.
Frederiksen, P.C. (1991). Economic growth and defense spending: Evidence on causality for selected Asian countries. *J. Philippine Dev.*, **18**: 131–147.
Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37**: 424–438.
Habibullah, M.S. & Law, S. (2008). *Defense Spending and Economic Growth in Asian Economies: A Panel Error-Correction Approach*. MPRA Munich Personal RePEc Archive, University Library of Munich, Germany .
Hacker, R.S. & Hatemi-J.A. (2006). Tests for causality between integrated variables using asymptotic and bootstrap distributions: theory and application. *Appl. Econ.*, **38**:, 1489–1500.
Halicioglu, F. (2004). Defense spending and economic growth in Turkey: An empirical application of new macroeconomic theory. *Rev. Middle East Econ. Financ.*, **2**: 193–201.
Hanson, B.E. (1992). Tests for parameter instability in regressions with I(1) processes. *J.Bus. Econ. Statists*, **10**: 321–335.
Herrera, R. & Gentilucci, E. (2013). Military spending, technical progress, and economic growth: A critical overview on mainstream defense economics. *J. Innovation Econ. Manage.*, **2**: 13–35.
Hirnissa, M.T., Habibullah, M.S., & Baharom, A.H. (2009). Military expenditure and economic growth in Asean-5 countries. *J. Sustainable Dev.*, **2**: 192–202.
Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, **59**: 1551–1580.
Khalid, M.A. & Mohd Noor, Z. (2015). Military expenditure and economic growth in developing countries: Evidence from system GMM estimates. *J. Emerg. Trends in Econ. Manage. Sci.*, **6:** 31–39.
Kheng, C.B. (2009). The communist insurgency in Malaysia , 1948-90 : Contesting the nation-state and social change. *New Zealand J. Asian Studies*, **11**: 132–152.
Looney, R. E. (1988). Military expenditures and socio-economic development in Africa : A summary of recent empirical research. *J. Mod. Afr. Stud.*, **26**: 319–325.
Narayan, P.K., Singh, B., & Campus, G.C. (2007). Modelling the relationship between defense spending and economic growth for the Fiji Islands. *Defense Peace Econ.*, **18**: 391–401.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *J Amer. Statistical Assc.,* **84**: 223–230.
Park, J.Y. . & Phillips, P.C.B. (1988). Statistical inference in regressions with integrated processes : Part 1. *Economet. Theor.*, **4**: 468–497.
Peled, D. (2001). *Defense R&D and Economic Growth in Israel: A Research Agenda*. Samuel Neaman Institute for Advanced Studies in Science and Technology, Haifa, Israel.
Shahbaz, M. & Shabbir, M.S. (2010). Military spending and economic growth in Pakistan: New evidence from rolling window. *Ekon. Istraživanja*, **25**: 144–159.
Shukur, G. & Mantalos, P. (2000). A simple investigation of the Granger-causality test in integrated-cointegrated VAR systems. *J. Appl. Stat.*, **27**: 1021–1031.
Sims, C.A., Stock, J.H. & Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica*, **58**: 113–144.
Tang, C.F. (2008). Defence expenditure and economic growth in Malaysia: A reassessment using bounds and modified Wald tests. *CFAI J. Public Financ.*, **6**: 45–51.
Toda, H.Y. & Phillips, P.C.B. (1993). Vector autoregressions and causality. *Econometrica*, **61**: 1367–1393.
Toda, H.Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *J. Econometrics*, **66**: 225–250.