

Jordan Journal of Mechanical and Industrial Engineering (JJMIE)

JJMIE is a high-quality scientific journal devoted to fields of Mechanical and Industrial Engineering. It is published by The Jordanian Ministry of Higher Education and Scientific Research in corporation with the Hashemite University.

EDITORIAL BOARD

Editor-in-Chief

Prof. Mousa S. Mohsen

Editorial board

Prof. Bilal A. Akash
Hashemite University

Prof. Adnan Z. Al-Kilany
University of Jordan

Prof. Ayman A. Al-Maaitah
Mutah University

Prof. Moh'd A. Al-Nimr
Jordan University of Science and Technology

Prof. Ali A. Badran
University of Jordan

Prof. Naseem M. Sawaqed
Mutah University

Assistant Editor

Dr. Ahmed Al-Ghandoor
Hashemite University

THE INTERNATIONAL ADVISORY BOARD

Abu-Qudais, Mohammad
Jordan University of Science & Technology, Jordan

Abu-Mulaweh, Hosni
Purdue University at Fort Wayne, USA

Afaneh Abdul-Hafiz
Robert Bosch Corporation, USA

Afonso, Maria Dina
Institute Superior Tecnico, Portugal

Badiru, Adedji B.
The University of Tennessee, USA

Bejan, Adrian
Duke University, USA

Chalhoub, Nabil G.
Wayne State University, USA

Cho, Kyu-Kab
Pusan National University, South Korea

Dincer, Ibrahim
University of Ontario Institute of Technology, Canada

Douglas, Roy
Queen's University, U. K

El Bassam, Nasir
International Research Center for Renewable Energy,
Germany

Haik, Yousef
United Arab Emirates University, UAE

Jaber, Jamal
Al- Balqa Applied University, Jordan

Jubran, Bassam
Ryerson University, Canada

Kakac, Sadik
University of Miami, USA

Khalil, Essam-Eddin
Cairo University, Egypt

Mutoh, Yoshiharu
Nagaoka University of Technology, Japan

Pant, Durbin
Iowa State University, USA

Riffat, Saffa
The University of Nottingham, U. K

Saghir, Ziad
Ryerson University, Canada

Sarkar, MD. Abdur Rashid
Bangladesh University of
Engineering & Technology, Bangladesh

Siginer, Dennis
Wichita State University, USA

Sopian, Kamaruzzaman
University Kebangsaan Malaysia, Malaysia

Tzou, Gow-Yi
Yung-Ta Institute of Technology and
Commerce, Taiwan

EDITORIAL BOARD SUPPORT TEAM

Language Editor	Publishing Layout
Dr. Wael Zuraiq	MCPD. Osama AlShareet

SUBMISSION ADDRESS:

Prof. Mousa S. Mohsen, Editor-in-Chief
Jordan Journal of Mechanical & Industrial Engineering,
Hashemite University, PO Box 330127, Zarqa, 13133 , Jordan
E-mail: jjmie@hu.edu.jo



Hashemite Kingdom of Jordan



Hashemite University

Jordan Journal of Mechanical and Industrial Engineering

JJMIIE

An International Peer-Reviewed Scientific Journal

<http://jjmie.hu.edu.jo/>

ISSN 1995-6665

Jordan Journal of Mechanical and Industrial Engineering (JJMIE)

JJMIE is a high-quality scientific journal devoted to fields of Mechanical and Industrial Engineering. It is published by The Jordanian Ministry of Higher Education and Scientific Research in corporation with the Hashemite University.

Introduction: The Editorial Board is very committed to build the Journal as one of the leading international journals in mechanical and industrial engineering sciences in the next few years. With the support of the Ministry of Higher Education and Scientific Research and Jordanian Universities, it is expected that a heavy resource to be channeled into the Journal to establish its international reputation. The Journal's reputation will be enhanced from arrangements with several organizers of international conferences in publishing selected best papers of the conference proceedings.

Aims and Scope: Jordan Journal of Mechanical and Industrial Engineering (JJMIE) is a refereed international journal to be of interest and use to all those concerned with research in various fields of, or closely related to, mechanical and industrial engineering disciplines. Jordan Journal of Mechanical and Industrial Engineering aims to provide a highly readable and valuable addition to the literature which will serve as an indispensable reference tool for years to come. The coverage of the journal includes all new theoretical and experimental findings in the fields of mechanical and industrial engineering or any closely related fields. The journal also encourages the submission of critical review articles covering advances in recent research of such fields as well as technical notes.

Guide for Authors

Manuscript Submission

High-quality submissions to this new journal are welcome now and manuscripts may be either submitted online or mail.

Online: For online submission upload one copy of the full paper including graphics and all figures at the online submission site, accessed via E-mail: jjmie@hu.edu.jo. The manuscript must be written in MS Word Format. All correspondence, including notification of the Editor's decision and requests for revision, takes place by e-mail and via the Author's homepage, removing the need for a hard-copy paper trail.

By Mail: Manuscripts (1 original and 3 copies) accompanied by a covering letter may be sent to the Editor-in-Chief. However, a copy of the original manuscript, including original figures, and the electronic files should be sent to the Editor-in-Chief. Authors should also submit electronic files on disk (one disk for text material and a separate disk for graphics), retaining a backup copy for reference and safety.

Note that contributions may be either submitted online or sent by mail. Please do NOT submit via both routes. This will cause confusion and may lead to delay in article publication. Online submission is preferred.

Submission address and contact:

Prof. Mousa S. Mohsen, Editor-in-Chief
Jordan Journal of Mechanical & Industrial Engineering,
Hashemite University,
PO Box 330127, Zarqa, 13133, Jordan
E-mail: jjmie@hu.edu.jo

Types of contributions: Original research papers

Corresponding author: Clearly indicate who is responsible for correspondence at all stages of refereeing and publication, including post-publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Full postal addresses must be given for all co-authors.

Original material: Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher. Authors found to be deliberately contravening the submission guidelines on originality and exclusivity shall not be considered for future publication in this journal.

Supplying Final Accepted Text on Disk: If online submission is not possible: Once the paper has been accepted by the editor, an electronic version of the text should be submitted together with the final hardcopy of the manuscript. The electronic version must match the hardcopy exactly. We accept MS Word format only. Always keep a backup copy of the electronic file for reference and safety. Label the disk with your name. Electronic files can be stored on CD.

Notification: Authors will be notified of the acceptance of their paper by the editor. The Publisher will also send a notification of receipt of the paper in production.

Copyright: All authors must sign the Transfer of Copyright agreement before the article can be published. This transfer agreement enables Jordan Journal of Mechanical and Industrial Engineering to protect the copyrighted material for the authors, but does not relinquish the authors' proprietary rights. The copyright transfer covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microfilm or any other reproductions of similar nature and translations.

PDF Proofs: One set of page proofs in PDF format will be sent by e-mail to the corresponding author, to be checked for typesetting/editing. The corrections should be returned within 48 hours. No changes in, or additions to, the accepted (and subsequently edited) manuscript will be allowed at this stage. Proofreading is solely the author's responsibility. Any queries should be answered in full. Please correct factual errors only, or errors introduced by typesetting. Please note that once your paper has been proofed we publish the identical paper online as in print.

Author Benefits

Page charge: Publication in this journal is free of charge.

Free off-prints: Three journal issues of which the article appears in along with twenty-five off-prints will be supplied free of charge to the corresponding author. Corresponding authors will be given the choice to buy extra off-prints before printing of the article.

Manuscript Preparation:

General: Editors reserve the right to adjust style to certain standards of uniformity. Original manuscripts are discarded after publication unless the Publisher is asked to return original material after use. If online submission is not possible, an electronic copy of the manuscript on disk should accompany the final accepted hardcopy version. Please use MS Word for the text of your manuscript.

Structure: Follow this order when typing manuscripts: Title, Authors, Affiliations, Abstract, Keywords, Introduction, Main text, Conclusions, Acknowledgements, Appendix, References, Figure Captions, Figures and then Tables. For submission in hardcopy, do not import figures into the text - see Illustrations. For online submission, please supply figures imported into the text AND also separately as original graphics files. Collate acknowledgements in a separate section at the end of the article and do not include them on the title page, as a footnote to the title or otherwise.

Text Layout: Use double spacing and wide (3 cm) margins. Ensure that each new paragraph is clearly indicated. Present tables and figure legends on separate pages at the end of the manuscript. If possible, consult a recent issue of the journal to become familiar with layout and conventions. All footnotes (except for table and corresponding author footnotes) should be identified with superscript Arabic numbers. To conserve space, authors are requested to mark the less important parts of the paper (such as records of experimental results) for printing in smaller type. For long papers (more than 4000 words) sections which could be deleted without destroying either the sense or the continuity of the paper should be indicated as a guide for the editor. Nomenclature should conform to that most frequently used in the scientific field concerned. Number all pages consecutively; use 12 or 10 pt font size and standard fonts. If submitting in hardcopy, print the entire manuscript on one side of the paper only.

Corresponding author: Clearly indicate who is responsible for correspondence at all stages of refereeing and publication, including post-publication. The corresponding author should be identified with an asterisk and footnote. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Full postal addresses must be given for all co-authors. Please consult a recent journal paper for style if possible.

Abstract: A self-contained abstract outlining in a single paragraph the aims, scope and conclusions of the paper must be supplied.

Keywords: Immediately after the abstract, provide a maximum of six keywords (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible.

Symbols: All Greek letters and unusual symbols should be identified by name in the margin, the first time they are used.

Units: Follow internationally accepted rules and conventions: use the international system of units (SI). If other quantities are mentioned, give their equivalent in SI.

Maths: Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

References: All publications cited in the text should be presented in a list of references following the text of the manuscript.

Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given.

List: Number the references (numbers in square brackets) in the list in the order in which they appear in the text.

Examples:

Reference to a journal publication:

- [1] M.S. Mohsen, B.A. Akash, "Evaluation of domestic solar water heating system in Jordan using analytic hierarchy process". *Energy Conversion & Management*, Vol. 38, No. 9, 1997, 1815-1822.

Reference to a book:

- [2] Strunk Jr W, White EB. *The elements of style*. 3rd ed. New York: Macmillan; 1979.

Reference to a conference proceeding:

- [3] B. Akash, S. Odeh, S. Nijmeh, "Modeling of solar-assisted double-tube evaporator heat pump system under local climate conditions". 5th Jordanian International Mechanical Engineering Conference, Amman, Jordan, 2004.

Reference to a chapter in an edited book:

- [4] Mettam GR, Adams LB. How to prepare an electronic version of your article. In: Jones BS, Smith RZ, editors. *Introduction to the electronic age*, New York: E-Publishing Inc; 1999, p. 281-304

Free Online Color: If, together with your accepted article, you submit usable color and black/white figures then the journal will ensure that these figures will appear in color on the journal website electronic version.

Tables: Tables should be numbered consecutively and given suitable captions and each table should begin on a new page. No vertical rules should be used. Tables should not unnecessarily duplicate results presented elsewhere in the manuscript (for example, in graphs). Footnotes to tables should be typed below the table and should be referred to by superscript lowercase letters.

PAGES	PAPERS
236 – 245	A Finite Element Calculation of Stress Intensity Factors of Cruciform and Butt Welded Joints for Some Geometrical Parameters <i>A. M. Al-Mukhtar, S. Henkel, H. Biermann, P. Hübner</i>
246 – 251	An Intelligent Opportunistic Maintenance (OM) System: A Genetic Algorithm Approach <i>M. S. Samhoury, A. Al-Ghandoor, R. H. Fouad, S. M. A. Ali</i>
252 – 259	Prediction of Surface Roughness in Turning Using Adaptive Neuro-Fuzzy Inference System <i>B. Sidda Reddy, J. Suresh Kumar, K. Vijaya Kumar Reddy</i>
260 – 267	Wavelet Decomposition for the Detection and Diagnosis of Faults in Rolling Element Bearings <i>J. Chebil, G. Noel, M. Mesbah, M. Deriche</i>
268 -271	Acoustic Diagnosis Technique for Machine Condition Monitoring <i>M. A. Nawafleh, N. Al-Kloub</i>
272 – 279	Vibration Analysis of Plates With Spot Welded Stiffeners <i>S. M. Nacy, N. K. Alsaheb, F. F. Mustafa</i>
280 – 293	Optimal Forward Kinematics Modeling of Stewart Manipulator Using Genetic Algorithms <i>A. Omran, M. Bayoumi, A. Kassem, G. El-Bayoumi</i>
294 - 305	An Intelligent Machine Condition Monitoring System Using Time-Based Analysis: Neuro-Fuzzy Versus Neural Network <i>M. Samhoury, A. Al-Ghandoor, S. A. Ali, I. Hinti, W. Massad</i>
306 - 311	Effect of Exhaust Gas Recirculation (EGR) on the Performance and Emission Characteristics of Diesel Engine with Sunflower Oil Methyl Ester <i>K. Rajan, K. R. Senthilkumar</i>
312 – 321	The Use of Rational Design in the Development of an Improved Plug Tool for the Rotary Tube Widening Process <i>S. D. Al-Shobaki, A. K. A. Al-Dahwi, R. H. Fouad</i>

A Finite Element Calculation of Stress Intensity Factors of Cruciform and Butt Welded Joints for Some Geometrical Parameters

A. M. Al-Mukhtar ^{a,*}, S. Henkel ^a, H. Biermann ^a, P. Hübner ^b

^a Institute of Materials Engineering, Technische Universität Bergakademie Freiberg, Gustav-Zeuner-Straße 5, Germany

^b Fachhochschule Mittweida, Mittweida, Germany

Abstract

With welded joints, stress concentrations occur at the weld toe and at the weld root, which make these regions the points from which fatigue cracks may initiate. To calculate the fatigue life of welded structures and to analyze the progress of these cracks using fracture mechanics technique requires an accurate calculation of the stress intensity factor SIF. The existing SIFs were usually derived for one particular geometry and type of loading. In this study, the finite element method (FEM) was used to calculate the SIF. The stress intensity factors during the crack propagation phase were calculated by using the software FRANC2D, which is shown to be highly accurate, with the direction of crack propagation being predicted by using the maximum normal stress criterion. In the current work, a new analytical approach for the weld toe crack in cruciform welded joints has been used. The SIF results from FRANC2D were compared with those from the International Institute of Welding-IIW, and literature. A good correlation was obtained and the work results have bench marked which made it possible to use FRANC2D to simulate different weld geometries. The results of these comparisons are shown and the agreement is clearly well.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Stress Intensity Factor; FRANC2D; Welded Joints; Cruciform Welded Joint; Butt Welded Joint; IIW.

1. Introduction

In the application of fracture mechanics to fatigue problems, an accurate determination of the stress intensity factor (SIF) for the case under analysis is necessary to use the Paris' equation to calculate the fatigue life.

With fillet welded joints, stress concentrations occur at the weld toe and at the weld root, which make these regions the points from which fatigue cracks may initiate [1]. Therefore Shen and Clayton [2] stated that all the cracks were found to be initiated at the weld end toe, the maximum stress concentration site.

Table 1 shows some of welded joints used in this work. Discarding the major weld defects, fatigue cracks originate from the weld toe, and then propagate through the base material, or from the weld root, and then propagate through the weld throat.

Toe cracks have been considered because it is easier to observe with the naked eye as well as with dye penetration tests, they are often found in many important engineering welded structures. Moreover, there is a high stress concentration located at this point [4-5].

For fracture mechanics treatments, in spite of the fact that several SIF handbooks have been published, it is still

difficult to find solutions adequate to many welded configurations. This is mainly due to a wide variety of complex welded geometries and loading systems.

Thus, the derivation of SIF even for one type of weldment, requires detailed analysis of several parameters such as plate thickness, weld thickness, weld angle, weld toe radius and loading system.

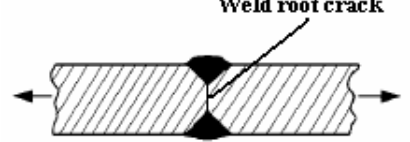
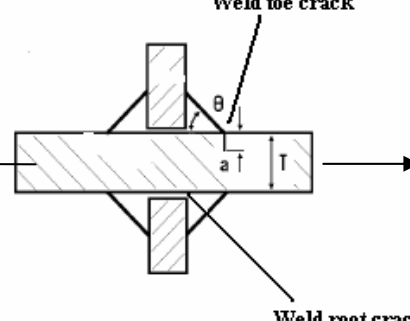
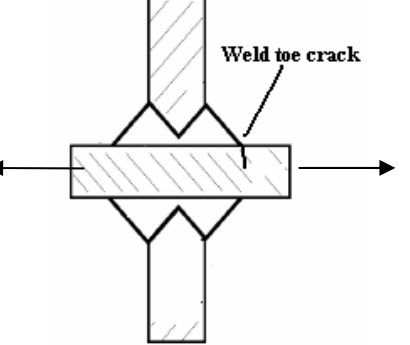
Stress intensity factors are inevitable parameter, which must be determined in fracture mechanics methods. These factors describe the fatigue action at a crack tip in terms of crack propagation.

By characterizing stable macroscopic crack growth using SIF range ΔK , it is possible to predict crack growth rate of a crack under cyclic loading, and hence the number of cycles necessary for a crack to extend from some initial size, i.e., the size of pre-existing crack or crack-like defects, to a maximum permissible size just before catastrophic failures [6].

In this work the SIFs of cruciform joints and butt welded joints failing from weld toe have been calculated using the two-dimensional Fracture Analysis Code program FRANC2D. These results are compared with solutions from International Institute of Welding (IIW) [3], and solutions from literature.

* Corresponding author. a.almukhtar@yahoo.de.

Table 1. Fatigue cracking of some welded joints [3].

 Weld toe crack	Transverse loaded butt weld (X-groove or V-groove) ground flush to plate, toe crack.
	Transverse butt weld with weld reinforcement, toe crack.
 Weld root crack	Transverse partial penetration butt weld. Lack of penetration (LOP) considered being as a root crack.
 Weld toe crack Weld root crack	Cruciform joint or T-joint, fillet welds or partial penetration K-butt welds, toe and root crack.
 Weld toe crack	Cruciform joint or T-joint, K-butt welds, full penetration, toe crack.

2. Specimen Modelling

The cruciform welded joint with equal attachment and main plate thickness ($B/T=1$), is shown in Figure 1. The non-load carrying cruciform joint is made from the welding of stiffener plate (attachment plate) perpendicular to main plate (loading plate). Sheet thickness was taken as 12 mm. Because of geometrical symmetry joint, only quarter or half of the joint was considered during the calculation. The initial crack length, a_i located at weld toe, and h the weld leg length over main plate side, where $w=2h+B$. The weld leg length on the attached plate side is S , as shown in Figure 1. Here, ρ is the weld toe radius.

Cruciform welded joints can be classified into load carrying and non-load carrying cruciform joints. In the latter fatigue cracks usually occur at the weld toe, where the load acted at the end of main plate. By contrast, in the former, cracks starts from the lack of penetration (LOP) where the load acted at the end of attached plate. The same classification can be used for others welded joints.

Butt welded joints have the thickness t , equal to 10 mm. The weld bead height (H) and the weld bead width (W) were 2 and 10 mm respectively, see Figure 2, θ ($\theta=180-\alpha$) is the weld toe angle.

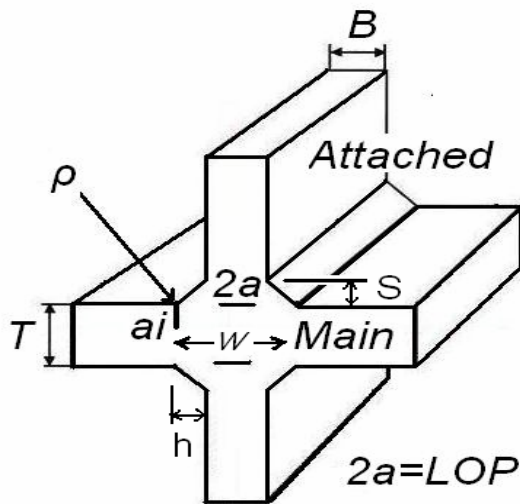


Figure 1. Cruciform fillet welded joint.

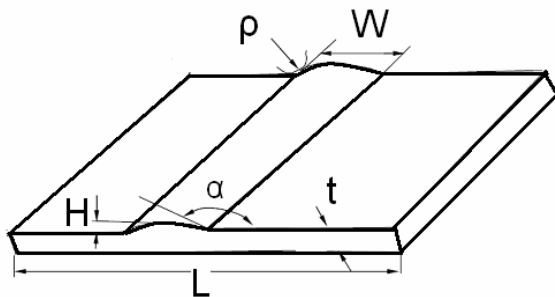


Figure 2. Weld geometry parameters at butt welded joint.

3. Finite Element Analysis

The finite element program FRANC2D have used in this study. The program was developed by the Cornell Fracture Group from Cornell University, USA [7].

3.1. Assumption

The analysis was undertaken based on the assumption of an initially isotropic elastic material for both the base and its weld metal, in which a crack was subsequently allowed to form and grows according to a fracture criterion.

In fracture mechanics, the fatigue strength of a welded joint is not primarily governed by the strength of the base material of the joining members, hence, the governing parameters in fracture mechanics are mainly the local and global geometry of the joint only, i.e., fatigue strength is known to be closely related to the precise geometrical discontinuity of the welded joint [6]. Therefore the similar material as a quantity joint between the weld metal and a base metal has been assumed in the simulation.

3.2. Mesh Description and Boundary Conditions

Figure 3 shows one of the FE meshes used in the present study, which comprises 8 noded quadrilateral elements. The boundary conditions of cruciform and butt joints are shown also in Figure 3. However, there are other possibilities to use mesh and numbers of nodes in FRANC2D program [8 - 9].

Boundary conditions were shown as the fixed sides for half or quarter models that used in this study. One end of the models was supported in the x -direction and a uniform distribution stresses were applied at the other end. To prevent the model from performing rigid body motions, one node on the side where the model is supported in the x -direction, is also locked in the y -direction, as shown in Figure 3.

3.3. Symmetry

If the loading and the cracks are not considered, the cruciform and butt joints are symmetrical about both x - and y -axes. However, this is not the case when the crack is presented at weld toe. The model becomes unsymmetrical.

Quarter symmetry for cruciform joints has been used with 8 noded quadrilateral elements. Moreover, the joint can be simplified to a half model as shown in Figure 3. The difference in the SIFs solutions between the simplified half model and a quarter models for cruciform joint is shown in Figure 4. The differences are quite small.

The comparison between half and complete butt weld joints is shown in Figure 5.

The quarter and half models are therefore having a reasonable approximation of the cruciform and butt joints to estimate the SIFs of these joints.

3.4. Material Properties

The material used in the present study was an extra high strength hot rolled steel with the minimum yield strength 550 MPa and the tensile strength minimum 600 MPa and maximum 760 MPa, respectively. The fatigue simulated with applied loading such that the maximum stress was maintained constant at 200 N/mm² and 104 N/mm² for cruciform and butt weld joints respectively. Values of Poisson's ratio ν and the modulus of elasticity E were chosen as 0.293 and 210 kN/mm² respectively. Experimentally, many structures are optimized by the choice of high strength steel with the very reason for this choice is to allow for higher stresses and reduced dimensions taking benefit of the high strength material with respect to the yield criterion.

4. Solution Procedure

4.1. Mesh Generation

A mesh generating program CASCA, which is distributed with FRANC2D, was used to create the initial mesh configuration for FRANC2D simulations. Other mesh generating programs can also be used, provided a translator is available to convert the mesh description to the FRANC2D *.inp format [8, 9]. The procedure for creating a mesh is very straightforward, as illustrated in Figure 6. To begin with, the problem outline is first created in Figure 6a, followed by the division of sub-regions within the problem boundary. Prior to assigning the type of elements to each of the sub-regions, the boundaries for all sub-regions are divided into the required number of segments, Figure 6b. The resulting mesh for the present simulation is shown in Figure 6c.

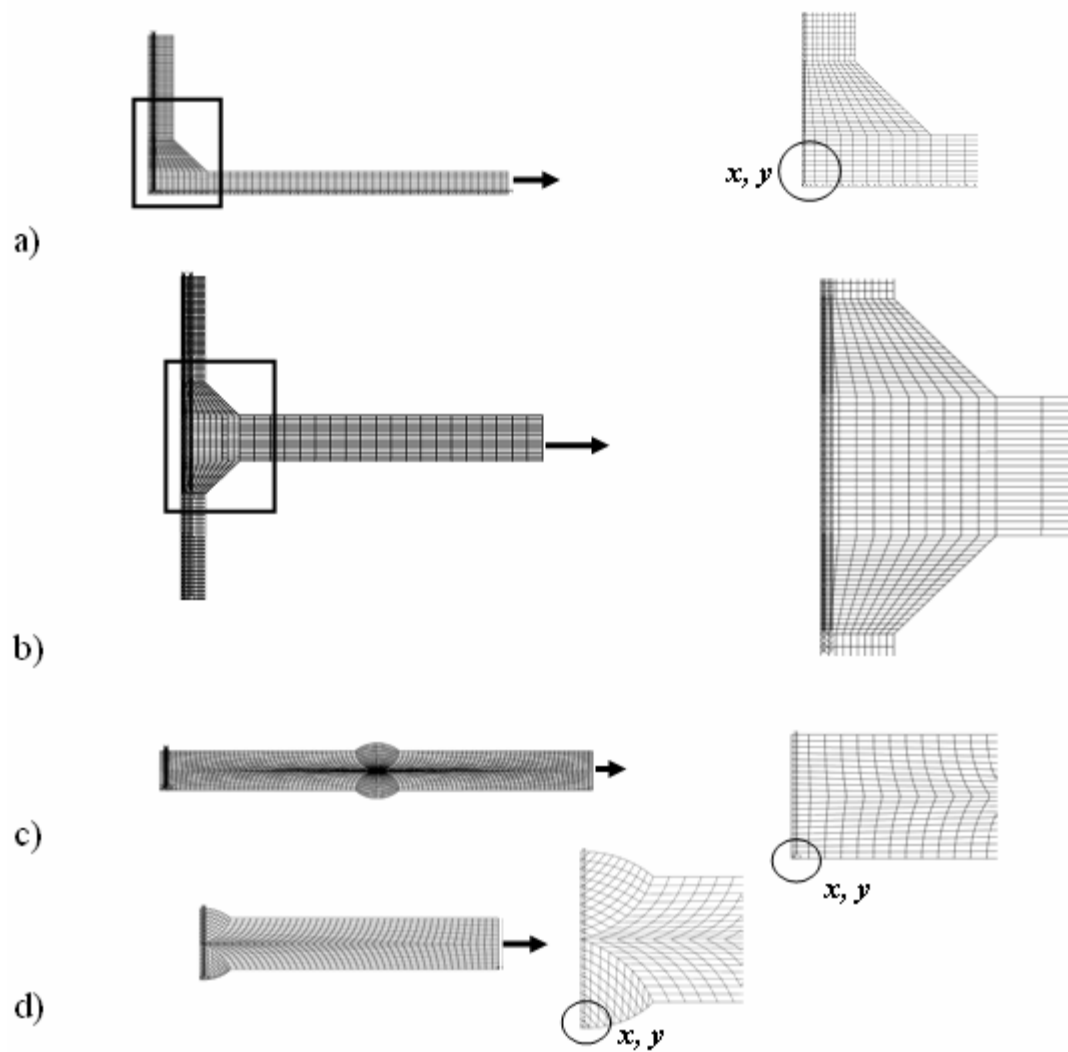


Figure 3. Mesh configuration and boundary conditions for notch cases, a) Quarter cruciform model, b) half cruciform model, c) complete butt model, d) half butt model. The edges of the models were locked in both degrees of freedom.

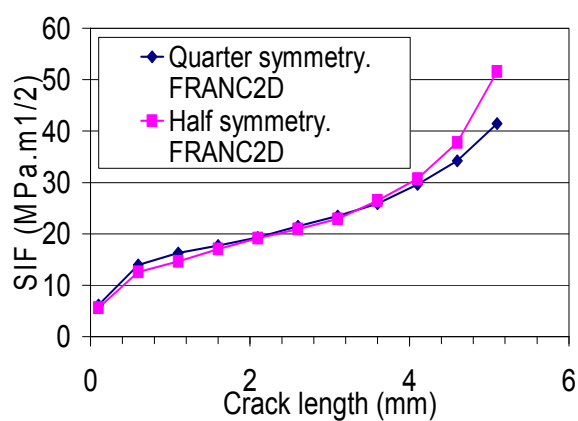


Figure 4. Comparison between half and quarter cruciform joints.

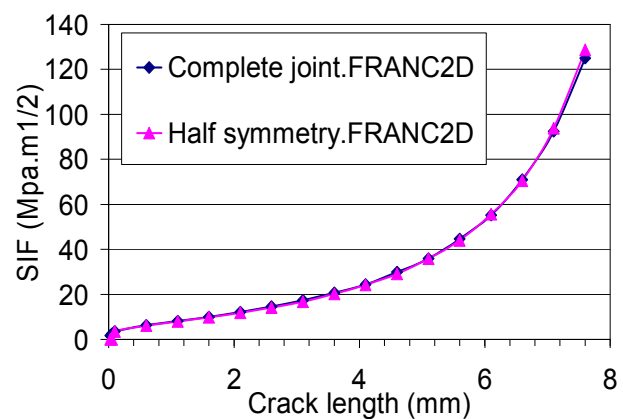


Figure 5. Comparison between half and complete butt joints.

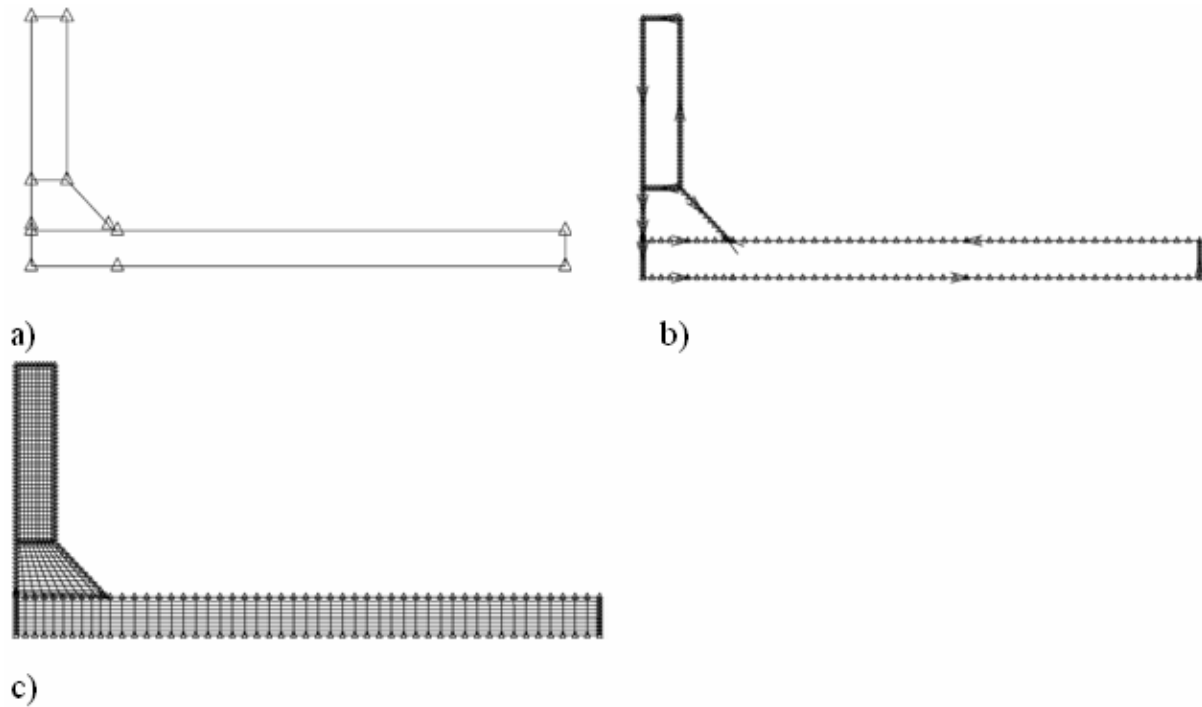


Figure 6. Mesh generating procedure, a) Geometry with subdivided areas, b) Subdivided area into segments, c) Mesh generation (CASA) [7].

4.2. Material Model

A linear elastic material model coupled to a linear elastic discrete fracture model was used in the analysis. It is possible to use an elasto-plastic material for the steel weldments, which is available in the program. But for this analysis based on fracture mechanics for which linear elastic fracture theory is considered reasonable, it was decided to restrict the analysis of SIF solutions to the consideration of a linear material behavior.

4.3. Crack Propagation

A crack could propagate both from the weld toe and from the weld root (LOP). All calculations of the present work will be made on cracks propagating from the weld toe.

In order to study the capability of the program to simulate crack growth, an initial non-cohesive edge crack was placed at the weld toe of the fillet welded joints of the weld-base material interface, where it was predicted that critical tensile stresses would occur. The existence of crack-like imperfections in the welded joint is normally considered to eliminate the so-called crack initiation stage of fatigue life. Therefore, most of fatigue assessment of welded joint is concentrated on the crack growth stage of the fatigue life. Having specified the location of the crack, the program was able to predict the direction in which the crack would propagate. Prior to performing the analysis, it was necessary to specify the magnitude of crack increment and also the number of steps over which the crack would propagate. In the present study, a crack increment (Δa) of 0.5 mm was chosen for all cases. The crack growth was simulated over suitable steps of increment according to plate thickness. Moreover, in this study, the crack path was not pre-selected, but crack direction was allowed to change according to the maximum tangential stress

criterion [6]. Moreover, the auto-mesh was carried out automatically.

5. Stress Intensity Factor Calculation

The expression for the SIF is:

$$K_I = \sigma (\pi a)^{1/2} f(b) \quad (1)$$

The applied stress is σ and $f(b)$ is the correction factor. The stress intensity correction factor $f(b)$ is a function of the crack length to thickness of propagation plate ratio. The analytical expression for $f(b)$ can be of the tangent or secant type. It is not too obvious how to choose between the tangent and secant expression for estimating the value of parameter $f(b)$, although the "secant" version appears to be a bit more straightforward [10].

$$f(b) = \left(\sec \frac{\pi a}{2t} \right)^{1/2} \quad (2)$$

where the term $\left(\frac{\pi a}{2t} \right)$ is given in radians.

Some modifications have been carried out on geometrical correction factor as shown for cruciform joint in the next section.

The FE crack growth program FRANC2D was used to calculate the opening mode SIF using fracture mechanics approach. The influence of K_I on fatigue crack growth was based on the maximum tangential stress criterion by Erdogan and Sih [11]. This criterion assumes that the predicted propagation path of the fatigue crack is perpendicular to the maximum principal stress and the crack grows under opening mode. The equivalent opening

mode (I) SIF which is used in crack growth predictions is calculated according to the same criterion.

5.1. Cruciform Welded Joints

For an elliptical crack at the weld toe of a fillet welded joint, the range of the SIF, ΔK can be written as [12]:

$$\Delta K = \frac{M_k Y_u}{\phi_0} \Delta \sigma \sqrt{a} \quad (3)$$

M_k is the stress concentration magnification factor, $\Delta \sigma$ is the nominal tensile stress range applied on the main plate and a , is the crack depth [6]. ϕ_0 is the complete elliptical integral defined as:

$$\phi_0 = \int_0^{\pi/2} \left[1 - \left(1 - \frac{a^2}{c^2} \right) \sin^2 \phi \right]^{1/2} d\phi \quad (4)$$

where ϕ is parametric angle of ellipse.

In this analysis the M_k -factor functions are based on continuous edge cracks, hence, the crack aspect ratio is zero, $a/2c = 0$, and $\phi_0 = 1$. The correction term Y_u for a double-edge crack in a plate under tensile loading, Eq. 5, given by Brown and Srawley [13] was applied:

$$Y_u = 1.98 + 0.36 \left(\frac{2a}{T} \right) - 2.12 \left(\frac{2a}{T} \right)^2 + 3.42 \left(\frac{2a}{T} \right)^3, \quad 0 < \frac{2a}{T} < 0.95 \quad (5)$$

T , is the main plate thickness. For fracture mechanics treatments Maddox [12] introduced the concept of M_k as a magnification of the SIF, which would be present for a crack of the same geometry but without the presence of the weld. Further work on M_k values for cracks at weld toes has been carried out by Lie [14], Thurlbeck [15] and Bowness and Lee [16].

The formula of M_k from IIW has been used, where IIW introduced a systematic set of formulae for M_k values for different welded joints [3]:

$$M_k = C \left(\frac{a}{T} \right)^k \quad M_k \leq 1 \quad (6)$$

$$C = 0.8068 - 0.1554 \left(\frac{S}{T} \right) + 0.0429 \left(\frac{S}{T} \right)^2 + 0.0794 \left(\frac{h}{T} \right) \quad (7)$$

$$k = -0.1993 - 0.1839 \left(\frac{S}{T} \right) + 0.0495 \left(\frac{S}{T} \right)^2 + 0.0815 \left(\frac{h}{T} \right) \quad (8)$$

where h and S are the weld leg length on main and attachment plate side respectively, T is the main plate thickness and a is the initial crack length from the weld toe.

By substituting Eq. 6 and Eq. 5 in Eq. 3, SIF can be calculated for cruciform joint fail from the weld toe.

For fillet welds, the high stress concentration in weld toe is presented due to the fact that these locations rely to be sound and usually weldment contains flaws and crack-like defects. Therefore, the presence of weld toe radius inevitably will reduce these concentrations of stresses near the weld toe.

5.2. Butt Weld Joint

For the validation of the prediction results, the range of SIF, ΔK for the butt welded specimens was calculated using the following empirical Eq. 9 shown below [17]:

$$\Delta K = \Delta \sigma \sqrt{\pi a} \left(1.12 - 0.23 \left(\frac{a}{t} \right) + 10.55 \left(\frac{a}{t} \right)^2 - 21.72 \left(\frac{a}{t} \right)^3 + 30.39 \left(\frac{a}{t} \right)^4 \right) \quad (9)$$

where $\Delta \sigma$ is the stress range, a , the crack length, t the thickness of the plate (see Figure 2). The results from Eq. 9 were compared with those obtained using FRANC2D.

6. Results and Discussion

6.1. Stress Distribution and Crack Growing

Due to stress concentration and cyclic loading, cracks may initiate and grow in the vicinity of the welds during service life even if the applied stresses are well below the yield limit. The weld toe and the weld root have high stress concentration which makes these regions the easier points from which fatigue cracks may initiate [1].

Figure 7 shows the maximum and minimum principal stress distributions for uniform load distributions of cruciform joint. Figure 8 gives the principal stresses distribution for the case of an edge crack. The stress near the weld toe was 1103 MPa which is higher than the tensile strength of materials.

The site and curved crack growth paths of continuous toe cracks were taken into account as shown in Figure 9. Normally this kind of non-load carrying welded attachments always fails at the weld toe as shown in Figure 10.

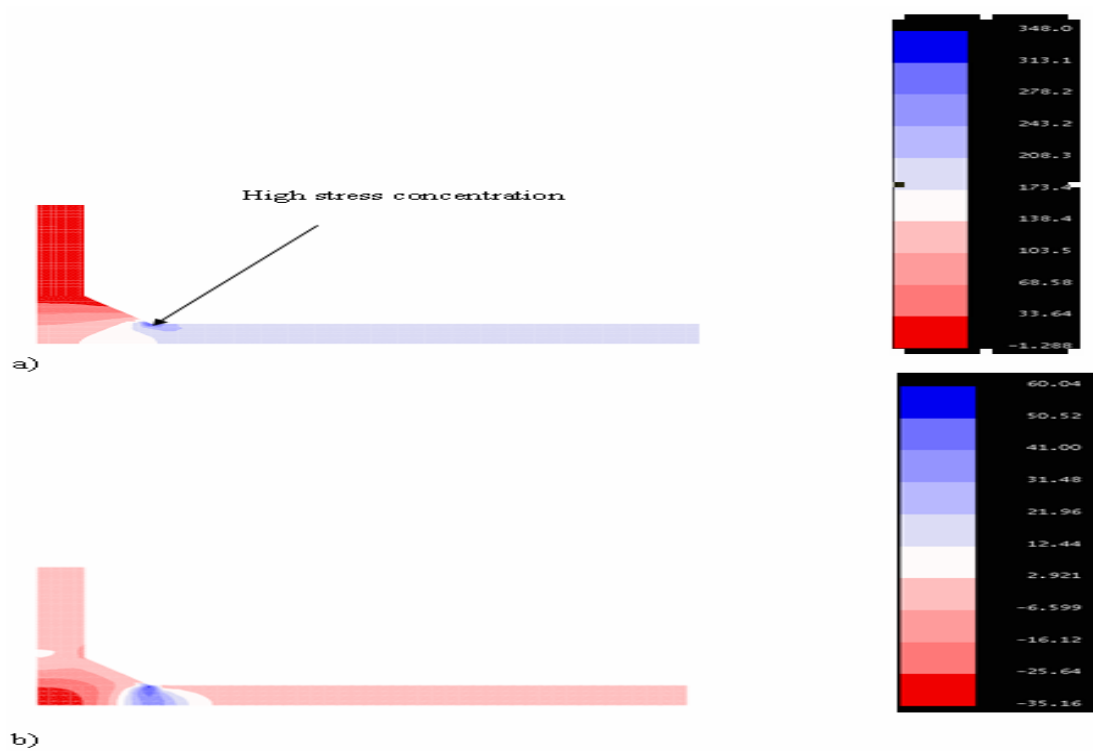


Figure 7. a) Maximum and b) Minimum principal stress distribution for cruciform joint (N/mm^2); the stress concentration sites shown for applied load equal 200 N/mm^2 .

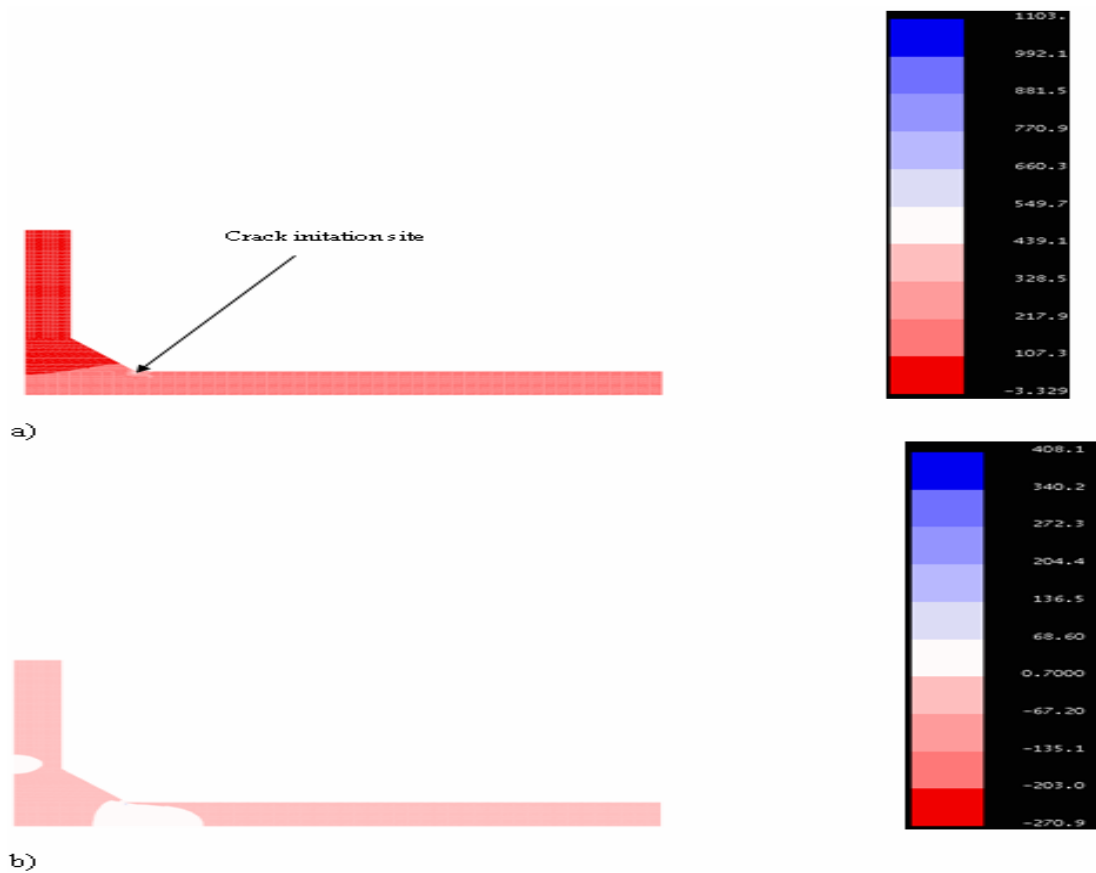


Figure 8. a) Maximum and b) Minimum principal stress distribution (N/mm^2) in existence of edge toe crack ($a_t = 0.1 \text{ mm}$).

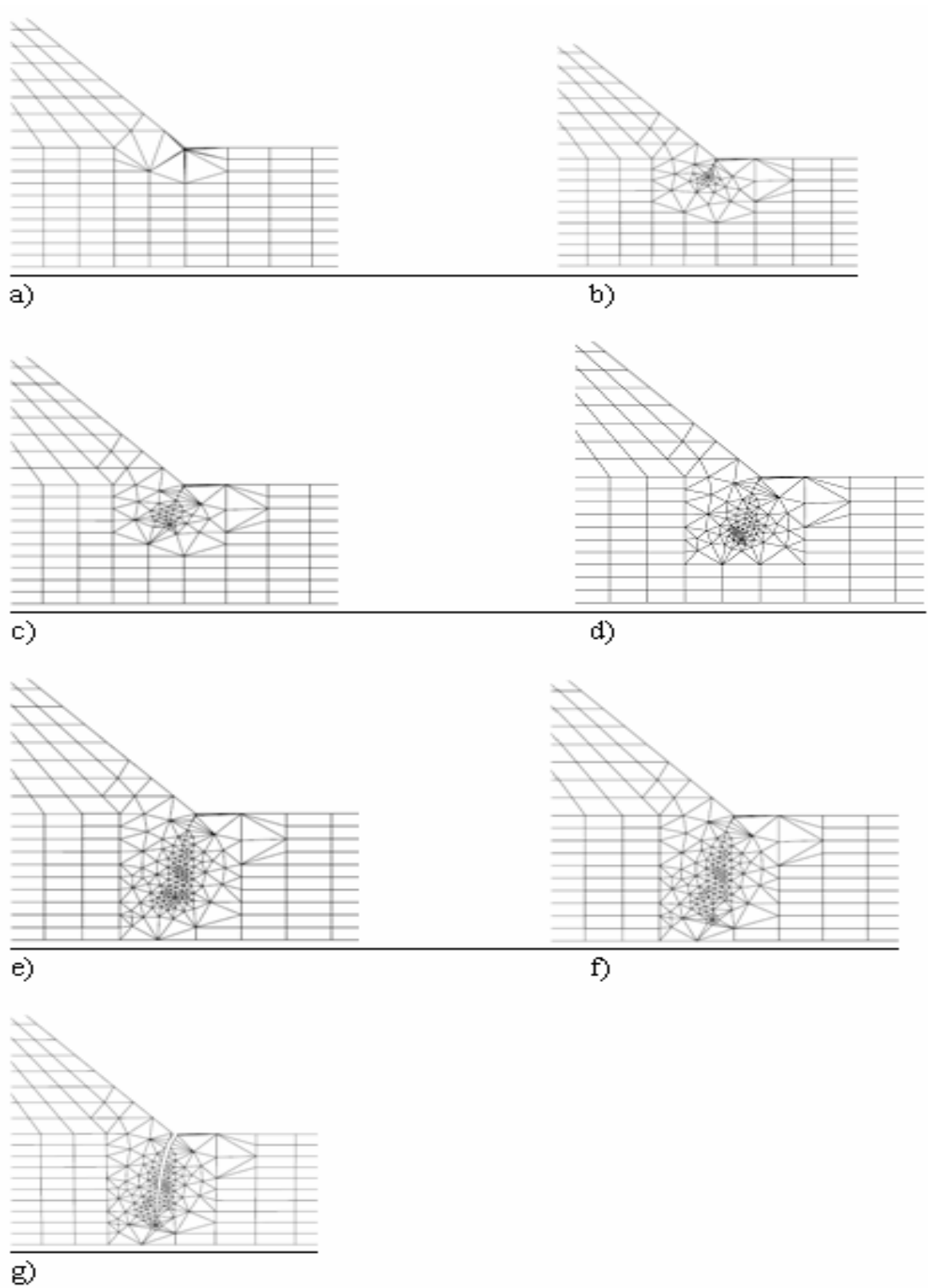


Figure 9. Sequences of crack propagation steps from the weld toe of cruciform welded joints (FRANC2D).

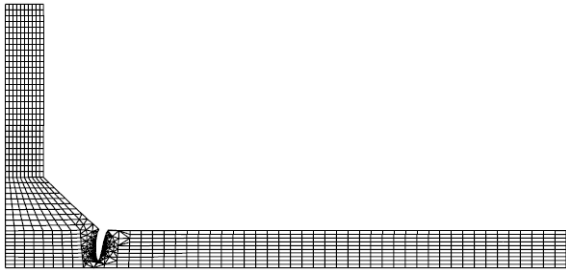


Figure 10. Crack deformation for cruciform weld joint (FRANC2D).

6.2. Butt Weld

The same steel was used for the butt weld joint. Welded toe crack has initiated and grown through the sheet thickness of 10 mm.

Figure 11 shows the modelling and boundary conditions used for butt weld to calculate SIFs. Figures 12 and 13 show the crack initiation and propagation respectively. The final crack deformation is shown in Figure 14.

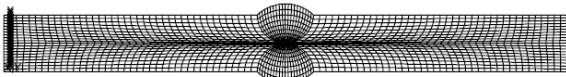


Figure 11. Double side butt weld (FRANC2D).

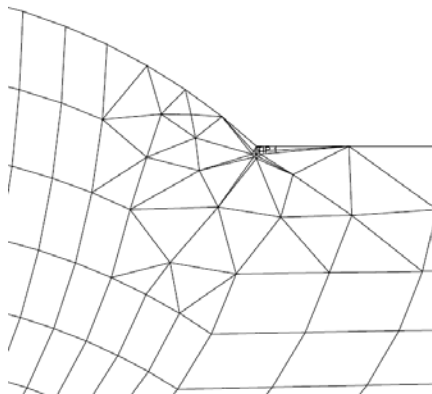


Figure 12. Screenshot from FRANC2D crack tip at weld toe, $a_i = 0.1$ mm.

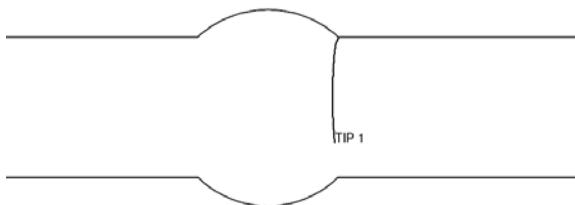
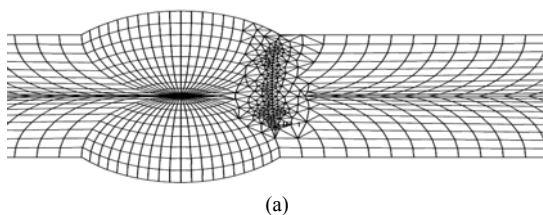


Figure 13. Screenshot from FRANC2D of the crack growth direction.



(a)



(b)

Figure 14. Screenshot from FRANC2D of the crack growth direction with mesh, a) Final crack propagation, b) Crack deformation and failure of butt welded joint.

6.3. Results of SIF Calculations

Cruciform joint and butt welded joint are the most conventional joints used in engineering structures. Accurate predictions of crack shape changes and fatigue lives require accurate SIF estimates because of the power-law nature of the Paris crack growth law. The determination of the SIF for the two dimensional cruciform and butt welded joints has been carried out and compared with empirical solutions using linear elastic finite element analysis.

6.3.1. Cruciform Joint

The results of the analytical study for the cruciform joints with equal main and attached plate thickness are plotted in Figure 15, based on the correction terms Eq. 5, and M_K , Eq. 6 together in Eq. 3. The results compared with the direct calculations of SIF as a function of crack length from FRANC2D. The results have been benchmarked as shown in Figure 16.

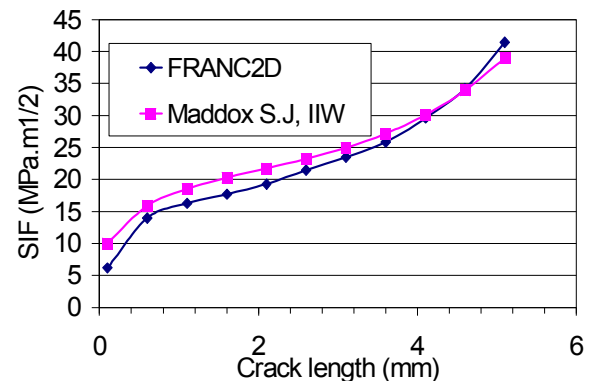


Figure 15. SIF as a function of crack length of cruciform joint fails from the weld toe compared with FEM.

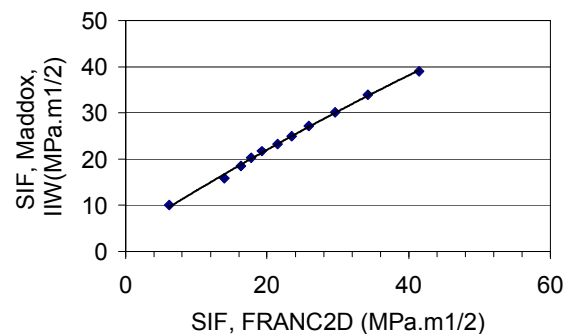


Figure 16. Comparison between analytical solutions and FEM.

6.4. Butt Weld

The comparison of results obtained for the butt weld is shown in Figure 17. It can be seen that the results from FEM (FRANC2D software) is close to the results obtained

from Eq. 9. FRANC2D appears to be more realistic and it compares well with empirical equations.

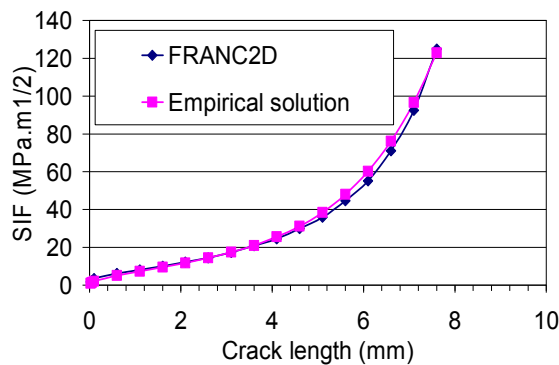


Figure 17. Comparison of FRANC2D and empirical equation of butt welded joint from ref [17].

7. Conclusions

The appropriate solutions for most welded joints are still difficult to find. Therefore SIF at of welded joints have been evaluated using the FEM.

The influence of the weld geometry was incorporated in the solution using FE analysis. The assumption was used that in the as-welded condition the crack remains open (mode-I) during the loading cycle due to the tensile residual stresses caused by welding are high enough. Therefore, the SIFs range corresponding to the nominal stress range is effective and independent of the *R*-ratio of nominal stresses. The FRANC2D software and quadrilateral elements were used to calculate the SIF of the joints from elementary of fracture mechanics.

This program has the ability to analyse the cracked body and describes the singularity a head of the crack tip. Thus, it can be concluded that for specific crack propagation, the SIF can be calculated under mode I-loading conditions. To demonstrate the efficiency of these calculations, two joints were investigated namely cruciform and butt weld joints. In the present study, an analytical approach for toe crack in cruciform joints has been developed.

The comparisons between this analytical approach and FEM have been agreed well.

Therefore, it can be advised to use the current FEM as combine with fracture mechanics to find appropriate SIFs solutions.

References

- [1] K. Motarjemi, A.H. Kokabi, A.A. Ziaie, S. Manteghi, F.M. Burdekin, "Comparison of stress intensity factor for cruciform and T welded joints with different attachment and main plate thickness". Engng Fract Mech, Vol. 65, No.1, 2000, 55–66.
- [2] W. Y. Shen, P. Clayton, "Fatigue of fillet welded A515 steel". Engineering fracture mechanics, Vol. 53. No.6, 1996, 1007-1016.
- [3] Hobbacher A. Recommendations for fatigue design of welded joints and components. International institute of welding-IIW/IIS. IIW document XIII-2151-07 / XV-1254-07; 2007.
- [4] Lindqvist J. Fatigue strengths thickness dependence in welded constructions. Master Thesis, Borlänge, Sweden; 2002.
- [5] Karlsson N., Lenander P.-H. Analysis of fatigue life in two weld class systems. Master Thesis in solid mechanics. Department of Mechanical Engineering, Linköping University SE-581 83 Linköping, Sweden; 2005.
- [6] T. Nykänen, G. Marquis, T. Björk, "Fatigue analysis of non-load carrying fillet welded cruciform joints". Engineering fracture mechanics. Vol. 74, No.3, 2007, 399-415.
- [7] Cornell Fracture Group, <<http://www.cfg.cornell.edu>>
- [8] Wawrzynek P., Ingraffea A. FRANC2D: A Two-Dimensional Crack Propagation Simulator. Tutorial and User's Guide, Version 2.7, NASA contractor report 4572; 1994.
- [9] Iesulaure E. FRANC2D/L: A crack propagation simulation for plane layered structures. Version 1.5 user's guide. Cornell University. Ithaca, New York.
- [10] Shukla A. Practical fracture mechanics in design. 2nd ed. Taylor and Francis Group LLC; 2005.
- [11] F. Erdogan, G.C. Sih, "On the crack extension in plates under plane loading and transverse shear". Journal of Basic Engineering ASME, Vol. 85, 1963, 519–527.
- [12] S.J. Maddox, "An analysis of fatigue cracks in fillet welded joints". International Journal of Fracture, Vol. 11, No. 2, 1975, 221-243.
- [13] Brown WF., Srawley JE. Plane strain crack toughness testing of high strength metallic materials. ASTM STP 410, ASTM, West Conshohocken, PA; 1966.
- [14] Lie ST. The influence of geometrical parameters on the fatigue strength of fillet welds using boundary element and fracture mechanics methods. Ph.D. Thesis, University of Manchester, Institute of Science and Technology (UMIST); 1983.
- [15] Thurlbeck SD. A fracture mechanics based methodology for the assessment of weld toe cracks in tubular offshore joints. Ph.D. Thesis, University of Manchester, Institute of Science and Technology (UMIST); 1991.
- [16] D. Bowness, M.M.K. Lee, "Stress intensity factor solutions for semi-elliptical weld-toe cracks in T-butt geometries". Fatigue Fract Engng Mater Struct., Vol. 19, No. 6, 1966, 787-797.
- [17] Broek D. Elementary engineering fracture mechanics. 3rd ed. Martinus Nijhoff publishers; 1982.

An Intelligent Opportunistic Maintenance (OM) System: A Genetic Algorithm Approach

M. S. Samhouri^{*}, A. Al-Ghandoor, R. H. Fouad, S. M. A. Ali

Department of Industrial Engineering, Hashemite University, Zarka, Jordan

Abstract

Complex systems like aircrafts, space shuttles, nuclear power stations, and some complicated process industries operate under high reliability and safety requirements due to the complicated technology involved and hazardous consequences to the larger community in case of failures. The maintenance regime of complex systems most often consists of a variety of maintenance strategies, like preventive maintenance, corrective maintenance, condition-based maintenance and so on. Opportunistic or opportunity-based maintenance (OM) gives the maintenance staff an opportunity to replace or repair those items, which are found to be defective or need replacement in the immediate future, during the maintenance of a machine or component. This work presents an intelligent method of how to decide whether a particular item requires opportunistic maintenance or not, and if so how cost effective this opportunity-based maintenance will be when compared to a probable future grounding. This maintenance strategy is considered important when dealing with complex systems that contain expensive items with hard lives with condition-based maintenance (CBM) strategies. Genetic algorithms (GA) are employed to decide whether opportunistic maintenance is cost effective or not. An example of applying opportunistic maintenance strategy in process industry is used to describe the methodology for genetic algorithms.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Keywords-Opportunistic Maintenance; Genetic Algorithms.

1. Introduction

Modern engineering systems, like process and energy systems, transport systems, offshore structures, bridges, pipelines are designed to ensure successful operation throughout the anticipated service life, in compliance with given safety requirements related to the risk posed to the personnel, the public and the environment.

Unfortunately, the threat of deteriorating processes is always present, so that it is necessary to install proper maintenance measures to control the development of deterioration and ensure the performance of the system throughout its service life. This requires decisions on what to inspect and maintain, how to inspect and maintain, and when to inspect and maintain. These decisions are to be taken so as to achieve the maximum benefit from the control of the degradation process while minimizing the impact on the operation of the system and other economical and safety consequences.

Engineers are always on the look out for ways of reducing system down time and increasing availability, without compromising on required level of system reliability. The ultimate objective of any maintenance regime is to maintain the system functionality to the maximum extent possible with optimum tradeoffs between the down times and cost of maintenance, avoiding any

hazardous failures. Opportunistic maintenance works out to be the perfect remedy, which utilizes the opportunity of system shutdown or module dismantle to perform any maintenance required in the immediate future and saves a substantial amount of system down-time.

In [1], the use of a genetic algorithm program for analyzing the optimal opportunity-based maintenance problem for real-sized systems, was investigated. They analyzed the performance of the genetic operators with a generation replacement genetic algorithm, using a hypothetical system consisting of 50 maintenance-significant parts, and they paid special attention to the sensitivity of solutions to the maximum number of maintenance groups considered by the genetic algorithm. They also found that better solutions were identified for larger numbers of groups but increasing complexity costs more in terms of the computer time required.

A simulation model for opportunistic maintenance strategies was presented in [2]. They proved that this automated model has a considerable improvement on the performance of the opportunistic maintenance strategies. In [3], a new approach to reliability-centered maintenance (RCM) using the concepts of soft life and hard life to optimize the total maintenance cost, was proposed. The proposed model was applied to find the optimal maintenance policies in the case of military aero-engines using Monte Carlo simulation. This case study showed a potential benefit from setting soft lives on relatively cheap

^{*} Corresponding author. samhouri@hu.edu.jo

components that can cause expensive, unplanned engine rejections.

An opportunistic maintenance policy for a continuously-monitored multi-unit series system with integrating imperfect effect into maintenance activities was developed in [4]. The simulation results implied that the proposed policy was better than the policy to maintain the system units separately. In [5], an opportunistic maintenance policy for a multi-component damage shock model with stochastically dependent components, was proposed. They utilized the coupling method to obtain stochastic maintenance comparisons on failure occurrences under different model parameters. In [6], An opportunistic preventive maintenance (PM) scheduling algorithm for the multi-unit series system based on dynamic programming, was introduced.

Opportunistic maintenance is a systematic method of collecting, investigating, preplanning, and publishing a set of proposed maintenance tasks and acting on them when there is an unscheduled failure or repair "opportunity". In this strategy, preventive maintenance activities are combined with corrective ones as soon as a certain technical and economical conditions are satisfied. Opportunity-based maintenance strategy involves several nonlinear variables which affect the total cost of maintenance that should be optimized to result in a cost-effective decision on maintenance actions. Genetic algorithms (GA) are particularly well-suited to solving problems where the space of all potential solutions is truly huge and too vast to search exhaustively in any reasonable amount of time. A genetic algorithm is a search technique used in computing to find exact or approximate solutions to optimization and search problem. It is categorized as a global search heuristics. It is a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology, such as inheritance, mutation, selection, and crossover.

This work is based on a real problem of improper maintenance strategy of a process industry (potash production). Actually the exploitation of the equipment maintenance records is often a weak point within an operations management organization. Inexistence of proper computerized maintenance system, lack of competences to properly handle maintenance data, or reduced knowledge in advanced maintenance processing techniques, are common problems to solve in order to benefit from the historical record of failures and maintenance operations carried out at certain equipment. In this work, the maintenance records analysis is used to provide critical information from past experience to improve current maintenance process in this Potash processing industry. Genetic algorithms techniques of time and cost analysis are used to build an intelligent maintenance system to predict whether the opportunity-based maintenance strategy is cost effective or not. GA-based opportunistic maintenance technique was applied on one of the critical production units which is the dryer.

This paper is organized as follows: section II will define the main concept of opportunity-based maintenance

strategy. Genetic algorithms (GA) technique will be described in section III. The problem formulation of GA-based opportunity maintenance strategy will be presented in section IV. A hypothetical example on GA-based opportunistic maintenance system will be given in section V, and the last section is to conclude.

2. Opportunistic Maintenance (OM) Strategy

Opportunistic maintenance can be defined as a systematic method of collecting, investigating, preplanning, and publishing a set of proposed maintenance tasks and acting on them when there is an unscheduled failure or repair "opportunity" [1]. Opportunistic maintenance can be thought of as a modification of the run-to-fail maintenance management philosophy. An opportunistic maintenance strategy is proposed to maintain a production line consisting of k non identical processors and without intermediary stocks. Operational characteristics of processors are degraded with usage. In this strategy, preventive maintenance activities are combined with corrective ones as soon as a certain technical and economical conditions are satisfied.

Generally, there are two main purposes for applying opportunistic maintenance: 1. to extend equipment lifetime or at least the mean time to the next failure whose repair may be costly. It is expected that this maintenance policy can reduce the frequency of service interruption and the many undesirable consequences of such interruption, and 2. to take advantage of the resources, efforts and time already dedicated to the maintenance of other parts in the system in order to cut cost.

Opportunistic maintenance consists of opportunistic replacement policies and opportunistic build policies. Replacement policies specify which parts to remove when an opportunity arises. Build policies specify which parts should be taken from the spares inventory to replace the parts removed according to the opportunistic replacement policies. Both policies should be used to reduce future maintenance requirements [3]. The opportunistic maintenance may be divided into two categories: 1. Age related, and 2. Non-age related. Fig. 1 shows the main categories of opportunistic maintenance.

- Non-age related opportunistic maintenance: the maintenance of those items, which failed before, but went undetected until the module's strip. These are the items, which are inaccessible unless the modules containing them are completely dismantled and whose failures do not cause system failure.
- Age-related opportunistic maintenance: this category contains three sub-categories as follows: 1. Hard life, 2. Soft life, and 3. Degradation.
 1. Hard life: is defined as the age of the component, at or by which the component has to be replaced.
 2. Soft life: is the age of the component after which it will be rejected the next time one of the modules containing it is recovered.

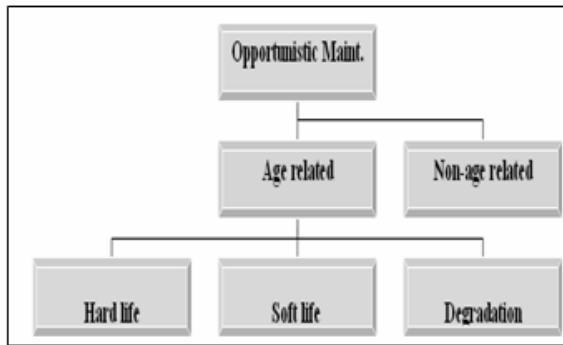


Figure 1. Categories of opportunistic maintenance strategy.

3. Degradation: failure mechanisms are monitored through condition monitoring devices and components are repaired or replaced once the condition deteriorates to a critical level.

In opportunity-based maintenance of a certain component, three important variables need to be collected and analyzed as follows:

- The remaining life of component or sub-module under consideration.
- The cost of down time that will occur if one decides to wait until the component has exhausted its useful life.
- The cost of risk involved or the probability of failure.

In such cases, an optimization model that carries out a comparative analysis of cost of remaining useful life and cost of down time for a group of components that will reach their hard lives within a small period of time will result in a list of components that should be replaced or repaired through opportunistic maintenance. Genetic algorithms (GA) are well-suited to carry out such an optimization task.

3. Genetic Algorithms (GA) Technique

Genetic algorithms are search algorithms based on the mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search [7]. GA are considered one of the most powerful searches and optimization algorithms because GA are conducted from a population of points rather than a single point, thus increasing the exploratory capability of GA. In addition, GA lend themselves naturally to implementation in parallel processing environments leading to the ability to exploit newer technologies in this domain, thus achieving faster computational times. Moreover, GA work with a direct coding of the parameter set rather than the parameters themselves, so, it is suitable for discontinuous, high dimensional and multi-nodal

problems. The mechanics of a simple genetic algorithm are surprisingly simple, involving nothing more complex than copying strings and swapping partial strings. A simple GA that yields good results in many practical problems is composed of three operators: 1. Reproduction, 2. Cross-over, and 3. Mutation.

The reproduction operator may be implemented in algorithmic form in a number of ways. Perhaps the easiest is to create a biased roulette wheel, where each current string in the population has a roulette wheel slot sized in proportion to its fitness.

Crossover is the process of combining information from two parents of strings, such that two children strings have a resemblance to each parent.

Mutation operator plays a secondary role in the simple GA. The frequency of mutation to obtain good results in empirical GA studies is on the order of one mutation per thousand position transfer. In the simple GA, mutation is the occasional random alteration of the value of a string position. Fig. 2 shows the flow diagram of the GA process.

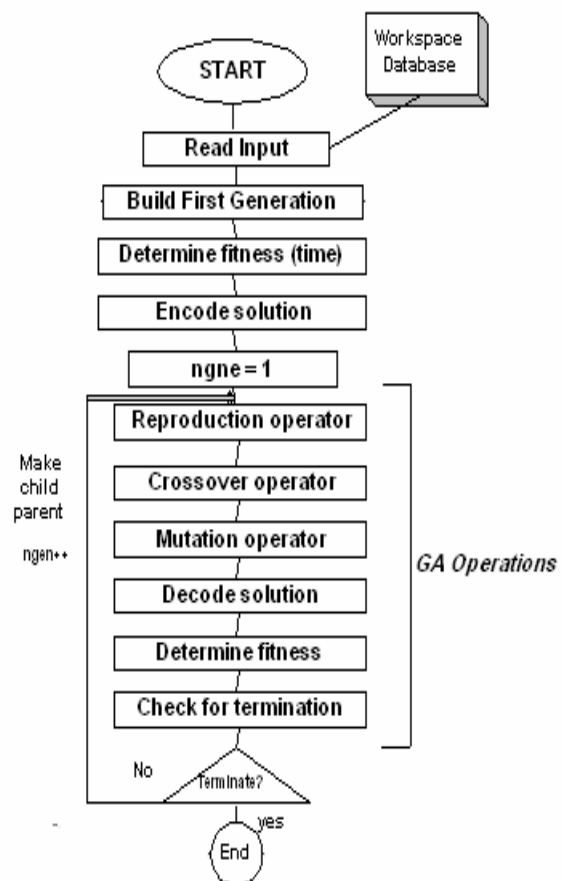


Figure 2. Genetic algorithms (GA) flowchart.

4. GA-based Opportunistic Maintenance: (Problem Formulation)

Using opportunistic maintenance where complex systems require periodic replacements of expensive parts is beneficial and considerably important. But, there are numerous factors to be considered while deciding on whether a particular component should be replaced or repaired when an opportunity arises. Genetic algorithms are used, as an optimization tool to compare the cost of premature replacement with the cost of downtime if grounded for the sole purpose of replacement.

The main factors or variables that may affect the final decision about the application of opportunistic maintenance could be described as follows:

- Remaining life cost (*RLC*): For an item, it is the product of number of hours or cycles remaining (*RL*) and the cost of item (*CI*) per hour/cycle as given in (1).

$$RLC = (RL).(CI) \quad (1)$$

- Down time cost (DTC): This is a very tricky component and completely depends on the complexity of the system, whether the downtime is planned or unplanned, and a lot of other similar factors which are specific to the environment under consideration.
- Unplanned down time cost (UDTC): It is the cost due to unplanned failure and it includes the direct and the indirect costs.
- Risk cost (RC): This is the risk involved in letting the components function until their complete useful life is utilized. The risk cost is given in (2), where (HF) is the hazard function, and (SDC) is the secondary damage cost.

$$RC = HF.(UDTC + SDC) \quad (2)$$

- Hazard function (HF): this is known as the failure rate, hazard rate, or force of mortality. Hazard function $h(x)$ is the ratio of the probability function $p(x)$ to the survival function $s(x)$. Hazard function (HF) is given in (3).

$$HF = \frac{p(x)}{s(x)} \quad (3)$$

- Secondary damage cost due to failure (SDC): the cost which can be obtained with the help of failure modes and critical analysis.
- Unit price: It is the price per finished product unit or service unit.

In this study, the genetic algorithms (GA) toolbox of Matlab 7.0 was used in order to take the right decision whether to repair or replace the components under study. A scattered crossover function and a uniform mutation was adopted. A fitness function (i.e., objective function) was formulated in order to quantify the optimality of a solution (i.e., chromosome in genetic algorithms) so that a particular chromosome may be ranked against all the other chromosomes. Optimal chromosomes, or at least chromosomes which are more optimal, are allowed to breed and mix their datasets by any of several techniques, producing a new generation that will (hopefully) be even better. Also, the fitness function is a way to describe the dynamics of gene frequencies in populations of reproducing

individuals. The fitness function measures the potential for reproductive success of any individual in a given environment. The objective function was formulated for this problem and given in (4), where (*CLR*) is cost of lost revenue, (*DT*) is the downtime in hours, (*RC*) is the risk cost, (*RLC*) is the remaining life cost, and (*Z*) is the fitness value (i.e., total cost of maintenance in Dollars).

$$Z = (CLR).(DT) + (RC) - (RLC) \quad (4)$$

The value of *Z* in (4) gives an indication of the decision; whether to perform the opportunistic maintenance on the studied components or not. In this work, if the fitness function (*Z*) yields a positive value, it will be represented by number one and the decision will be "perform opportunistic maintenance on the component", otherwise if the value of *Z* is negative, it will be represented by number zero, and the decision will be "do not perform opportunistic maintenance on the component".

Since the final decisions; whether to do opportunistic maintenance or not; depend on the value of *Z*, it was computed by first evaluating the total losses if the component continues running until failure occur (i.e., (*CLR*).(*DT*)+(RC) as in (4)). Then, the remaining life cost (*RLC*) was subtracted from the total losses in order to compare which is bigger, the total losses or the cost of remaining life. If the losses were bigger than the (*RLC*), then *Z* is positive and the decision is to perform opportunistic maintenance, because having positive value for *Z* means that repairing or replacing the component during the scheduled maintenance for the other components is better than replacing or repairing it when it fails. Otherwise, if the losses were smaller than the (*RLC*), then *Z* will give a negative value, and the decision is not to perform opportunistic maintenance.

One may notice that the down time (*DT*) and (*RLC*) for all items or components in the same group (as shown in section V), are the same. Therefore, the only independent variables in (4) are the (*CLR*) and (*RC*). These two variables will be the main GA fitness function variables, and will be denoted as follows: *X1*: Cost of lost revenue (*CLR*), and *X2*: Risk cost (*RC*). Therefore, the final GA fitness function will be as shown in (5).

$$Z = (DT).X1 + X2 - (RLC) \quad (5)$$

5. GA-Based Opportunistic Maintenance System: (A Hypothetical Example)

The final fitness function (i.e., objective function) as formulated in (5) contains two constants (i.e., *DT* and *RLC*). These two parameters remain constant for all components in the same group. In (5), *X1* and *X2* are the variables which will enter the GA computational loop, as shown in Fig. (2), in order to optimize the total cost of maintenance (i.e., fitness value *Z*).

In this work, the final formulated GA-based system was applied to an example from process industry (i.e., potash production plants). The data which was collected from the potash plant is given in Table I, and contains values for the following variables: i) the down time (*DT*) in hours, ii) production loss in tons (as an indicator of the cost of revenue lost), and iii) the remaining life cost (*RLC*)

in dollars, for 17 components of a rotary dryer machine, for years 2004, 2005, and 2006. Fig. 3 shows an inside view of a normal rotary dryer, whereas Fig.4 shows an inside view of a faulty rotary dryer.

Table I shows the ungrouped data for years 2004 – 2006. These data points were then grouped into seven different groups depending on the downtime (DT) and remaining life cost (RLC) which both remain constant for all components in the same group. The final grouped data for years 2004 to 2006 is shown in Table II.



Figure 3. The inside view of a normal rotary dryer.

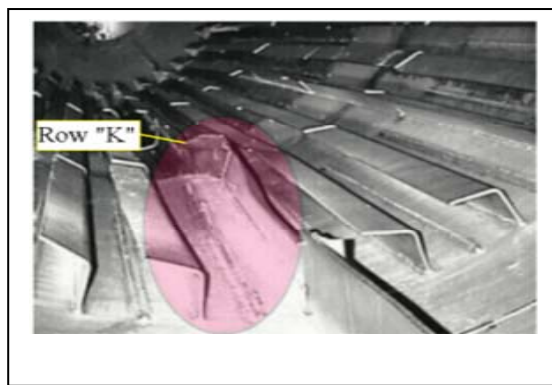


Figure 4. The inside view of a faulty rotary dryer.

Table I. Ungrouped Data for Years (2004 – 2006).

Year	Downtime (DT) (Hours)	Production Loss (tons)	Remaining life Cost (RLC) (\$)
2004	6	988.2	592920
2004	16	2635.2	4216320
2004	18	2964.6	5336280
2004	18	2964.6	5336280
2004	12	1976.4	2371680
2004	24	3952.8	9486720
2004	50	8235	4117500
2005	6	988.2	592920
2005	8	1317.6	1054080
2005	8	1317.6	1054080
2005	24	3952.8	9486720
2005	50	8235	4117500
2006	6	988.2	592920
2006	8	1317.6	1054080
2006	16	2635.2	4216320
2006	24	3952.8	9486720
2006	24	3952.8	9486720

Table 2. Grouped Data Based on Downtime (DT).

Group Number	Downtime (DT) (Hours)	Production Loss (tons)	Remaining life Cost (RLC) (\$)
1	6	988.2	592920
	6	988.2	592920
	6	988.2	592920
2	8	1317.6	1054080
	8	1317.6	1054080
	8	1317.6	1054080
3	12	1976.4	2371680
4	16	2635.2	4216320
	16	2635.2	4216320
5	18	2964.6	5336280
	18	2964.6	5336280
6	24	3952.8	9486720
	24	3952.8	9486720
	24	3952.8	9486720
	24	3952.8	9486720
7	50	8235	4117500
	50	8235	4117500
	50	8235	4117500

The grouped data of Table II was then fed into the genetic algorithms (GA) process in order to identify the desired outputs of this model which are represented by the fitness values (Z). As described earlier, the equation of the fitness function differs from group to group. For example, the fitness function for (group 1) is given in (6).

$$Z = 6X1 + X2 - 592920 \quad (6)$$

Different groups have different fitness functions depending on their downtimes (DT). Each fitness function (Z) was then given to the GA operators with initial population of 130, and terminates at 100 generations with a uniform crossover technique. The final outputs from the GA process are the best fitness values for each group, which could be negative or positive values. If the best (Z) from GA is negative, the decision is (zero) or not to perform opportunistic maintenance for this group of components. Otherwise, the decision is (one) and opportunistic maintenance strategy should be applied. Fig. 5 and Fig 6 are samples of the GA outputs for groups 4 and 7, respectively. As shown in Fig. 5 and Fig. 6, the points at the bottom of the plot denote the best fitness values, while the points above them denote the averages of the fitness values in each generation.

In Fig. 5, the best fitness value (Z) is (-4.2167×10^6) , which indicates a negative total cost of opportunistic maintenance, and the decision is "not to perform opportunistic maintenance". As noticed in Fig. 5 and Fig. 6, the first generation's fitness value was very low, and while the number of generation increases, the fitness value improves, which means that when the generation increases, the fitness value converges into the optimal value. The best fitness value improves more slowly in later generations whose populations are closer to the optimal point.

The final fitness values and decisions generated by the GA are given in Table III. As indicated by Table III, the components of groups 1 and 2 are the only components which have positive best fitness values, therefore; the decision is "to perform opportunistic maintenance" on group 1 and 2, but "not to perform opportunistic maintenance" on group 3, 4, 5, 6, and 7.

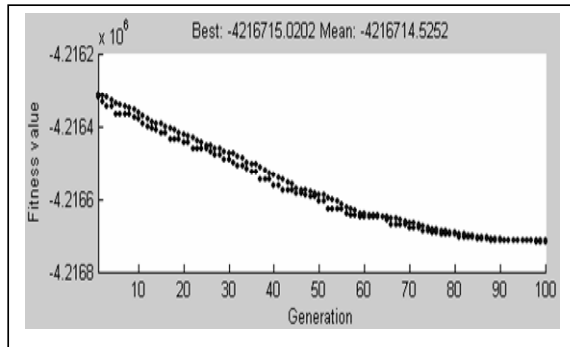


Figure 5. The best and average fitness values Vs. GA generation for group 4.

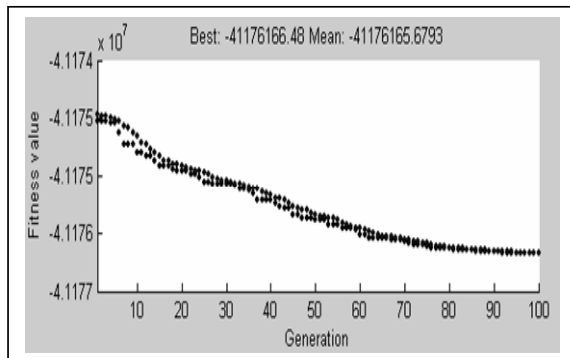


Figure 6. The best and average fitness values Vs. GA generation for group 7.

Table 3. The Final GA-Generated Fitness Values and Decisions

Group Number	Downtime (DT) (Hours)	Best Fitness Value (Z)	Decision
1	6	$+ 0.593 \times 10^6$	1
	6		
	6		
2	8	$+ 1.054 \times 10^6$	1
	8		
	8		
3	12	$- 2.372 \times 10^6$	0
4	16	$- 4.217 \times 10^6$	0
	16		
5	18	$- 5.336 \times 10^6$	0
	18		
6	24	$- 9.487 \times 10^6$	0
	24		
	24		
	24		
7	50	$- 41.176 \times 10^6$	0
	50		
	50		

6. Conclusions

In this work, genetic algorithms (GA) were adopted as an optimization tool to identify the maintenance cost variables that optimize the total cost of maintenance in

order to decide if opportunistic maintenance is favorable or not. The main conclusions of this work could be summarized as follow:

- based approach to opportunistic maintenance is a necessary procedure before deciding on whether to perform opportunistic maintenance strategy or not. This approach optimized the total cost of maintenance and gave an accurate indication about the economic of repairing or replacing a certain component under opportunistic maintenance strategy.
- Genetic algorithms for opportunistic maintenance are a novel application. GA technique is very suitable for the problem of opportunistic maintenance where the variables interactions complicate the problem. The final decision on whether to perform opportunistic maintenance or not depending on a minimum total cost gives the maintenance department an opportunity for considerable savings on the total maintenance expenses.
- The success of genetic algorithms in optimizing the cost of opportunistic maintenance suggests the use of this intelligent technique in many other industrial fields, particularly in maintenance and safety.
- It is recommended to improve the accuracy of the GA fitness function by considering more maintenance variables like the direct and indirect costs of maintenance tasks. This will guarantee GA convergence into better optimal solution.

References

- [1] A. Savic, G. Walters, J. Knezevic, "Optimal, opportunistic maintenance policy using genetic algorithms, 2: analysis". *Journal of Quality in Maintenance Engineering*, Vol. 1, No. 3, 1995, 25-34.
- [2] J. Crocker, U. Kumar, "Age related maintenance versus reliability centered maintenance: a case study on aero-engines". *Journal of Reliability Engineering and Systems Safety*, Vol. 67, No. 2, 2000, 113-118.
- [3] O. Mohamed-Salah, A-K. Daoud, G. Ali, "A simulation model for opportunistic maintenance strategies". 7th IEEE International Conference on Energy Technologies and Factory Automation, Vol. 1, 1999, Barcelona, Spain, 703 – 708.
- [4] X. Zhaou, L. Xi, J. Lee, "A dynamic opportunistic maintenance policy for continuously monitored systems". *Trans. J. of Quality Management Engineering*, Vol. 12, No. 3, 2006, 294 – 305.
- [5] L. Cui, H. Li, "Opportunistic maintenance for multi-component shock models". *Journal of Mathematical Methods of Operations Research*, Vol. 63, No. 3, 2006, 180 – 191.
- [6] X. Zhaou, L. Xi, J. Lee, "Opportunistic preventive maintenance scheduling for a multi-unit Series system based on dynamic programming". *International Journal of Production Economics*, 2008, In Press.
- [7] Goldberg D. *Genetic algorithms in search, optimization, and machine learning*. New York: Addison Wesley; 1994.

Prediction of Surface Roughness in Turning Using Adaptive Neuro-Fuzzy Inference System

B. Sidda Reddy ^{a,*}, J. Suresh Kumar ^b, K. Vijaya Kumar Reddy ^b

^a Department of Mechanical Engineering, R. G. M. Engineering College, Nandyal-518 501. A.P, India.

^b Department of Mechanical Engineering, J.N.T.U. College of Engineering, Hyderabad. A.P, India.

Abstract

Due to the extensive use of highly automated machine tools in the industry, manufacturing requires reliable models for the prediction of output performance of machining processes. The prediction of surface roughness plays a very important role in the manufacturing industry. The present work deals with the development of surface roughness prediction model for machining of aluminum alloys, using adaptive neuro-fuzzy inference system (ANFIS). The experimentation has been carried out on CNC turning machine with carbide cutting tool for machining aluminum alloys covering a wide range of machining conditions. The ANFIS model has been developed in terms of machining parameters for the prediction of surface roughness using train data. The Experimental validation runs were conducted for validating the model. To judge the accuracy and ability of the model percentage deviation and average percentage deviation has been used. The Response Surface Methodology (RSM) is also applied to model the same data. The ANFIS results are compared with the RSM results. Comparison results showed that the ANFIS results are superior to the RSM results.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Adaptive Neuro-Fuzzy; Surface Roughness Prediction; Turning.

1. Introduction

The aluminum alloys are used in various engineering applications like structural, cryogenic, food processing, oil and gas process industries etc. because of light weight and high tensile strength. The quality of the surface plays a very important role in the performance of the turning as a good quality turned surface significantly improves fatigue strength, corrosion resistance, or creep life. Surface roughness also affects several functional attributes of parts, such as contact causing surface friction, wearing, light reflection, heat transmission, ability of distributing and holding a lubricant, load bearing capacity, coating or resisting fatigue. Therefore the desired finish surface is usually specified and the appropriate processes are selected to reach the required quality [1]. To achieve the desired surface finish, a good predictive model is required for stable machining. The number of surface roughness prediction models available in literature is very limited [2]. Most surface prediction models are empirical and are generally based on experiments in the laboratory. In addition, it is very difficult in practice, to keep all factors under control as required to obtain reproducible results [3]. Taraman [4] used Response Surface Methodology for Prediction of surface roughness. Hasegawa et al., [5] conducted 3⁴ factorial designs to conduct experiments for the surface roughness prediction model. They found that the surface roughness increased with an increase in cutting

speed. Sundaram and Lambert [6-7] considered six variables i.e. speed, feed, depth of cut, time of cut, nose radius and type of tool to monitor surface roughness. Mital and Mehta [8] conducted a survey of surface roughness prediction models developed and factors influencing surface roughness. They found that most of the surface roughness prediction models developed for steels. Generally these models have a complex relation ship between surface roughness and operational parameters, work materials and chip breaker types. Salah Gasim Ahmed [9] developed an empirical surface roughness model for commercial aluminum, based on metal cutting results from factorial experiments. The model includes the feed, depth of cut and spindle speed. Dilbag Singh and P. Venkateswara Rao [10] conducted experiments to determine the effects of cutting conditions and tool geometry on the surface roughness in the finish hard turning of the bearing steel (AISI 52100) using mixed ceramic inserts made up of aluminum oxide and titanium carbide with different nose radius and different effective rake angles as cutting tools. They found that the feed is the most dominant factor determining the surface finish followed by nose radius and cutting velocity. Li Zhanjie [11] used Radial Basis Function network to predict surface roughness and compared with measured value and the result from regression analysis. Chen Lu and Jean-Philippe Costes [12] considered three variables i.e., cutting speed, depth of cut and feed rate to predict the surface profile in turning process using Radial Basis Function (RBF). They found that the RBF networks have the advantage over Back Propagation Networks (BPN). In the present work

* Corresponding author. bsrrgmecet@gmail.com.

the adaptive neuro-fuzzy model has been developed for the prediction of surface roughness. The predicted and measured values are fairly close to each other. The developed model can be effectively used to predict the surface roughness in the machining of aluminum alloys within the ranges of variables studied. The ANFIS results are compared with the RSM results. Comparison results showed that the ANFIS results are superior to the RSM results.

2. Aluminum Alloy Material

The work material used for the present investigation is aluminum alloy 6082 cylindrical work pieces. The chemical composition and physical properties of the material used in this work is given in Table 1 and Table 2.

3. Adaptive Neuro Fuzzy Inference Method

The fuzzy logic and fuzzy inference system (FIS) is an effective technique for the identification and control of complex non-linear systems. Fuzzy logic is particularly attractive due to its ability to solve problems in the absence of accurate mathematical models [13]. Surface roughness modeling in turning is considered complex process, so using the conventional techniques to model the surface roughness in turning results in significant discrepancies between simulation results and experimental data. Thus, this complex and highly time-variable process fits within the realm of neuro-fuzzy techniques. The application of a neuro-fuzzy inference system is used for prediction and overcomes the limitations of a fuzzy inference system such as the dependency on the expert for fuzzy rule generation and design of the non- adaptive fuzzy set.

3.1. Structure of The Adaptive Neuro-Fuzzy Inference System

Adaptive neuro-fuzzy inference system is a fuzzy inference system implemented in the framework of an adaptive neural network. By using a hybrid learning procedure, ANFIS can construct an input-output mapping based on both human-knowledge as fuzzy if-then rules and approximate membership functions from the stipulated input-output data pairs for neural network training. This procedure of developing a FIS using the framework of adaptive neural networks is called an adaptive neuro fuzzy inference system (ANFIS). There are two methods that ANFIS learning employs for updating membership function parameters: 1) backpropagation for all parameters (a steepest descent method), and 2) a hybrid method consisting of backpropagation for the parameters associated with the input membership and least squares estimation for the parameters associated with the output membership functions. As a result, the training error decreases, at least locally, throughout the learning process. Therefore, the more the initial membership functions resemble the optimal ones, the easier it will be for the model parameter training to converge. Human expertise about the target system to be modeled may aid in setting up these initial membership function parameters in the FIS structure [14-15]. The general ANFIS architecture is shown in Fig 1.

Five network layers are used by ANFIS to perform the following fuzzy inference steps. (i) Input fuzzification, (ii) Fuzzy set database construction, (iii) Fuzzy rule base construction, (iv) Decision making, and (v) Output defuzzification.

For instance assume that the FIS has two inputs x_1 and x_2 and one output y . For the first order Sugeno fuzzy model, a typical rule set with two fuzzy if-then rules can be expressed as:

$$\text{Rule 1: IF } (x_1 \text{ is } A_1) \text{ AND } (x_2 \text{ is } B_1) \text{ THEN } f_1 = p_1 x_1 + q_1 x_2 + r_1 \quad (1)$$

$$\text{Rule 2: IF } ((x_1 \text{ is } A_2) \text{ AND } (x_2 \text{ is } B_2)) \text{ THEN } f_2 = p_2 x_1 + q_2 x_2 + r_2 \quad (2)$$

Where A_1 , A_2 and B_1 , B_2 are the membership functions for the input x_1 and x_2 , respectively, p_1 , q_1 , r_1 and p_2 , q_2 , r_2 are the parameters of the output function. The functioning of the ANFIS is described as:

Layer 1: Calculate Membership Value for Premise Parameter

Every node in this layer produces membership grades of an input parameter. The node output

$$O_{1,i} = \mu_{A_i}(x_1) \text{ for } i=1,2, \text{ or} \quad (3)$$

$$O_{1,i} = \mu_{B_{i-2}}(x_2) \text{ for } i=3,4 \quad (4)$$

Where x_1 (or x_2) is the input to the node i ; A_i (or B_{i-2}) is a linguistic fuzzy set associated with this node. $O_{1,i}$ is the membership functions (MFs) grade of a fuzzy set and it specifies the degree to which the given input x_1 (or x_2) satisfies the quantifier. MFs can be any functions that are Gaussian, generalized bell shaped, triangular and trapezoidal shaped functions. A generalized bell shaped function can be selected within this MFs and it is described as:

$$\mu_{A_i}(x_1) = \frac{1}{1 + \left| \frac{x_1 - c_i}{a_i} \right|^{2b_i}} \quad (5)$$

Where a_i , b_i , c_i is the parameter set which changes the shapes of the membership function degree with maximum value equal to 1 and minimum value equal to 0.

Layer 2: Firing Strength of Rule

Every node in this layer, labeled Π , whose output is the product of all incoming signals:

$$O_{2,i} = w_i = \mu_{A_i}(x_1) \mu_{B_i}(x_2) \text{ for } i = 1, 2 \quad (6)$$

Layer 3: Normalize Firing Strength

The i^{th} node of this layer, labeled N , calculates the normalized firing strength as,

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1,2 \quad (7)$$

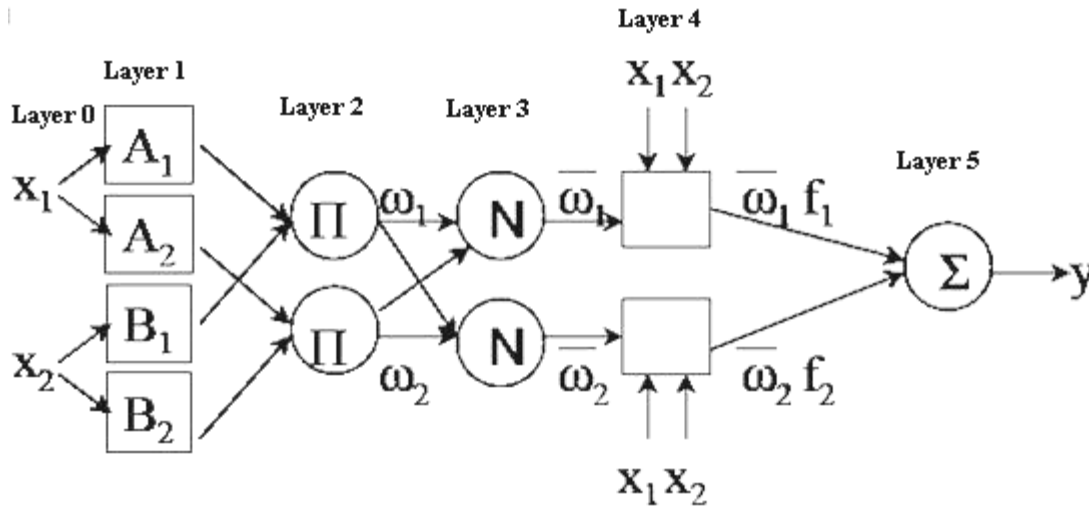


Figure 1. ANFIS architecture.

Table 1. Chemical composition of Aluminum Alloy 6082

Composition	weight (%)
Copper	0.1 (max)
Magnesium	0.4-1.2
Silicon	0.6-1.3
Iron	0.6
Manganese	0.4-1.0
Chromium	up to 0.25
Others	0.3
Aluminum	balance

Table 2. Physical properties of Aluminum alloy 6082.

Property	Value
Density	2.70 g/cm ³
Melting point	555°C
Modulus of Elasticity	70 G Pa
Electrical Resistivity	0.038x10 ⁻⁶ Ω .m
Thermal Conductivity	180 W/m K
Thermal Expansion	24x10 ⁻⁶ /K

Layer 4: Consequent Parameters

Every node i in this layer is an adaptive node with a node function,

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x_1 + q_i x_2 + r_i) \quad (8)$$

Where $\bar{w}_i$ is the normalized weighting factor of the i^{th} rule, f_i is the output of the i^{th} rule and p_i, q_i, r_i is consequent parameter set of this node.

Layer 5: Overall Output

The single node in this layer is a fixed node labeled Σ , which computes the overall output as the summation of all incoming signals:

$$\text{Overall output} = O_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (9)$$

ANFIS requires a training data set of desired input/output pair $(x_1, x_2, \dots, x_m, y)$ depicting the target system to be modeled. ANFIS adaptively maps the inputs $(x_1, x_2, \dots, x_m)$ to the outputs (y) through MFs, the rule base and the related parameters emulating the given training data set. It starts with initial MFs, in terms of type and number, and the rule base that can be designed intuitively. ANFIS applies a hybrid learning method for updating the FIS parameters. It utilizes the gradient descent approach to

fine-tune the premise parameters that define MFs. It applies the least-squares method to identify the consequent parameters that define the coefficients of each output equation in the Sugeno-type fuzzy rule base. The training process continues till the desired number of training steps (epochs) or the desired root mean square error (RMSE) between the desired and the generated output is achieved. In addition to the training data, the validation data are also optionally used for checking the generalization capability of FIS.

4. Experimental Details

The experiments were conducted according to full factorial design. The cutting parameters selected for the present investigation is cutting speed (V), feed (f) and depth (d) of cut. Since the considered variables are multi-level variables and their outcome effects are not linearly related. It has been decided to use three level tests for each factor. The machining parameters used and their levels are given in Table 3. The machining parameters, actual setting values and average surface roughness values are presented in Table 4. All the experiments were conducted on CNC Turning Lathe with the following specifications: Swing Over the Bed: 150mm, Swing Over Cross Slide: 50mm, Distance Between Centers: 300mm, Spindle Power

Table 3. Machining Parameters and their Levels.

Control parameters	Unit	Symbol	Levels		
			Level 1	Level 2	Level 3
Cutting Speed	m/min	v	95	105	115
Feed rate	mm / rev	f	0.02	0.04	0.06
Depth of cut	mm	d	0.5	0.75	1.0

1 HP, Spindle Speed (step less):0-3000rpm, Spindle Bore: 21mm, Spindle Taper: MT3, Tailstock Taper: MT2, the Tool Holder used for Turning operation was a WIDAX tool holder SDJCR 1212 11F3 and the tool material used for the study was Carbide Cutting Tool.

The average surface roughness (R_a) which is mostly used in industrial environments is taken up for the present study. The roughness was measured number of times and averaged. The average surface roughness is the integral absolute value of the height of the roughness profile over the evaluation length and was represented by the following equation.

$$R_a = \frac{1}{L} \int_0^L |Y(x)| dx \quad (10)$$

Where L is the length taken for observation and Y is the ordinate of the profile curve. The surface roughness was measured by using Surtronic 3⁺ stylus type instrument manufactured by Taylor Hobson with the following specifications. Traverse Speed: 1mm/sec, Cut-off values 0.25mm, 0.80mm and 2.50mm, Display LCD matrix, Battery Alkaline 600 measurements of 4 mm measurement length. The surfaces are cleaned and positioned using a V-block before each measurement. The actual setting values for the design matrix [16] and experimental results are shown in Table 4.

5. Results and Discussion

The ANFIS model has been developed as a function of machining parameters using twenty seven train data presented in Table 4. The fuzzy logic toolbox of MATLAB 7.0 was used to train the ANFIS and obtain the results. Different ANFIS parameters were tested as training parameters in order to achieve the perfect training and the maximum prediction accuracy. Fig 2 shows the fuzzy inference system (FIS) of ANFIS. The three inputs and one output and their final fuzzy membership functions

are shown in Fig 2. A total of 78 network nodes and 27 fuzzy rules were used to build the fuzzy inference system. A triangular membership functions were used to train ANFIS because it achieved the lowest training error of (0.1666) at 10 epochs, as shown in the training curve of Fig 3. A perfect training is clear from Fig 3. Three triangular membership functions were used for inputs (V, f and d). Fig 4 shows the comparison between the experimental and predicted values by the ANFIS and RSM model for training data. The predicted values by ANFIS and RSM model for training data are presented in Table 4. The average percentage deviation for training data set in the prediction of Surface roughness using ANFIS and RSM model is found to be 9.75%, 15.57% respectively.

5.1. Validation Runs

The models developed by ANFIS and RSM are validated using the validation data presented in Table 5. The predicted results were presented in Table 5. The predicted surface roughness values with the actual experimental values of surface roughness were plotted and shown in Fig 5. The average percentage deviation in the prediction of Surface roughness using ANFIS and RSM is found to be 3.29% and 15.86% respectively.

6. Conclusions

An adaptive neuro-fuzzy system and RSM is applied to predict the surface roughness during the turning process. The machining parameters were used as inputs to the ANFIS and RSM to predict surface roughness. The following conclusions can be drawn from this study:

- The ANFIS model could predict the surface roughness for training data with an average percentage deviation of 9.75% when a triangular member ship function is applied or 90.25% accuracy, while RSM model could predict the surface roughness for training data with an average percentage deviation of 15.57% or 84.43% accuracy from training data set.
- The ANFIS model could predict the surface roughness for testing or validation data set with an average percentage deviation of 3.29% when a triangular member ship function is applied or 96.71% accuracy, while RSM model could predict the surface roughness for training data with an average percentage deviation of 15.86% or 84.14% from validation data set. The accuracy of the developed model can be improved by including more number of parameters.

Table 4. Experimental Conditions, results (Experimental and Predicted).

v	f	d	Experimental Ra	Predicted Ra (RSM)	%Deviation (RSM)	Predicted Ra (ANFIS)	%Deviation (ANFIS)
95	0.02	0.50	1.706	2.38069	39.5481	2.0651	21.040
95	0.02	0.75	2.286	2.74749	20.1877	2.6736	16.950
95	0.02	1.00	5.393	5.14752	4.5518	6.5284	21.050
95	0.04	0.50	2.720	2.56557	5.6776	3.2926	21.050
95	0.04	0.75	2.366	2.33230	1.4243	2.8010	18.380
95	0.04	1.00	4.460	4.13224	7.3489	5.3989	21.050
95	0.06	0.50	2.546	2.09469	17.7262	3.0820	21.050
95	0.06	0.75	1.640	1.26132	23.0902	1.9742	20.370
95	0.06	1.00	2.006	2.46119	22.6914	2.4283	21.050
105	0.02	0.50	3.613	2.19788	39.1675	3.9571	9.520
105	0.02	0.75	2.540	2.76319	8.7870	2.6973	6.190
105	0.02	1.00	5.773	5.36171	7.1244	5.7739	0.015
105	0.04	0.50	2.113	2.60985	23.5140	2.3143	9.520
105	0.04	0.75	2.080	2.57507	23.8014	2.2046	5.990
105	0.04	1.00	4.890	4.57352	6.4720	5.3565	9.530
105	0.06	0.50	1.753	2.36605	34.9715	1.9200	9.520
105	0.06	0.75	2.213	1.73119	21.7718	2.4206	9.380
105	0.06	1.00	2.333	3.12955	34.1427	2.5552	9.520
115	0.02	0.50	2.200	2.47596	12.5436	2.2232	1.050
115	0.02	0.75	3.080	3.23977	5.1873	3.0902	0.330
115	0.02	1.00	5.760	6.03680	4.8056	5.9086	2.570
115	0.04	0.50	3.413	3.11502	8.7307	3.4490	1.050
115	0.04	0.75	3.660	3.27874	10.4169	3.6871	0.740
115	0.04	1.00	4.956	5.47569	10.4861	5.0818	2.530
115	0.06	0.50	2.840	3.09830	9.0951	2.8699	1.050
115	0.06	0.75	2.726	2.66194	2.3500	2.7357	0.350
115	0.06	1.00	5.006	4.25880	14.9261	5.1351	2.570
Average % Deviation: 15.5756%							9.750%

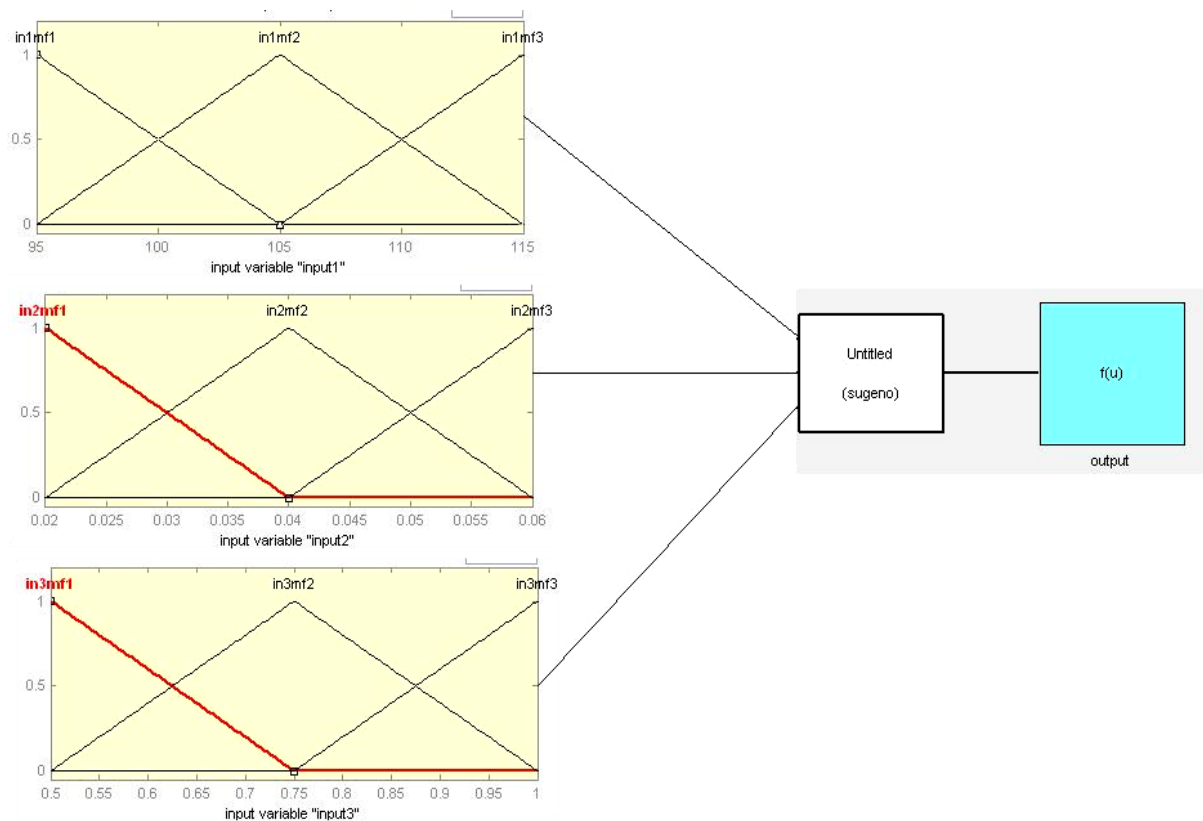


Figure 2. Fuzzy inference system for surface roughness prediction

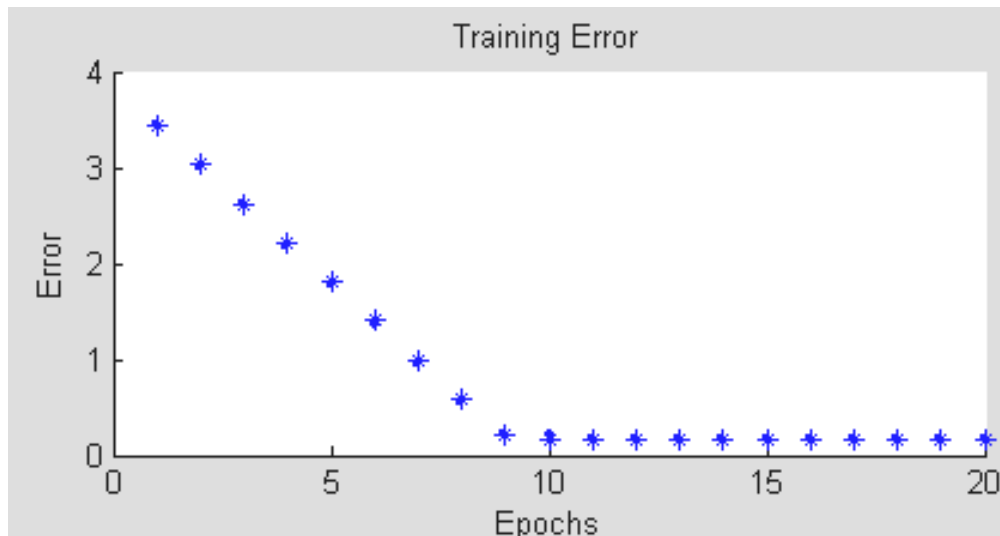


Figure 3. ANFIS Training Curve.

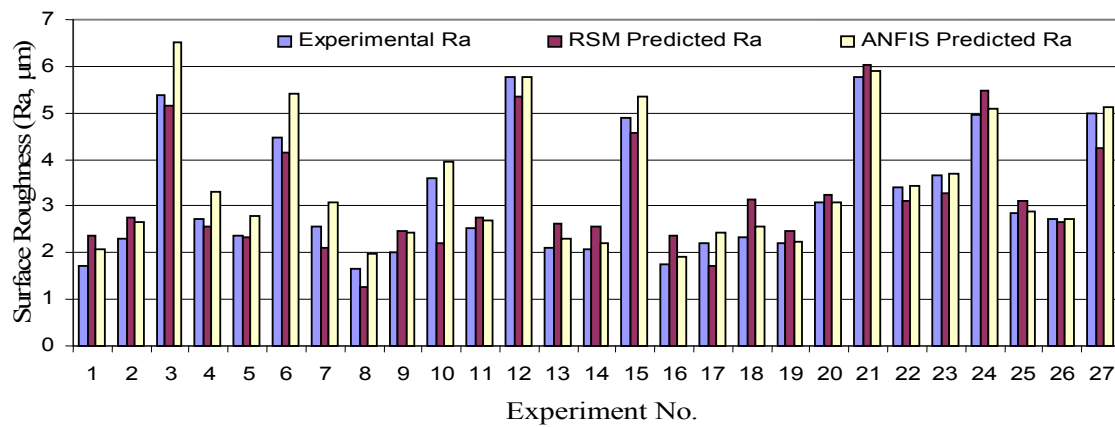


Figure 4. comparison between experimental and predicted values for training data.

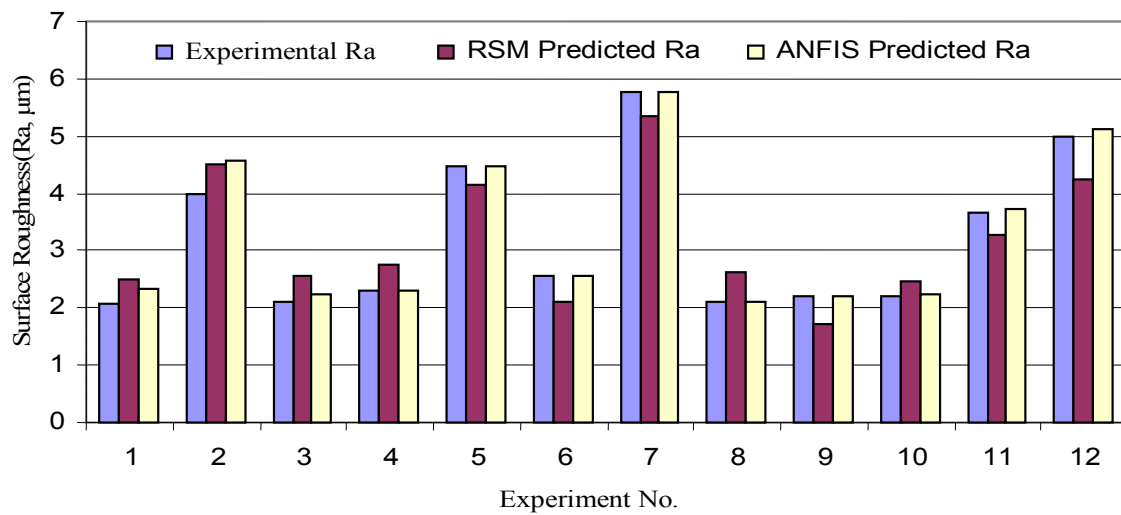


Figure 5. ANFIS Validation diagram.

Acknowledgement

The Authors are thankful to the Hon'ble Chairman Dr. M.Santhi Ramudu, Managing Director Mr. M. Sivaram, Principal Dr. T. Jaya Chandra Prasad and H.O.D of M.E, Rajeev Gandhi memorial College of Engg& Technology, Nandyal for providing the facilities to carry out the research work.

References

- [1] M. S. Lou, J.C. Chen, C. M. Li, "Surface roughness prediction for CNC end milling". *Journal of Industrial Technology*, Vol.15, No.1, Nov 1998 – Jan 1999, 2-6.
- [2] P. V. S. Suresh, P. V. Rao, S. G. Desmukh, "A genetic algorithm approach for optimization of the surface roughness prediction model". *International Journal of Mach Tools & Manufacture*, Vol. 42, 2002, 675-680.
- [3] C. A. Van Luttervelt, T. H. C. Childs, I. S. Jawahir, F. Klocke, P. K. Venunod, "Present situation and future trends in modeling of machining operations". *Process Report of the CIRP working group on 'Modeling of Machining Operations'*, *Annals of the CIRP*, Vol. 47, No.2, 1998, 587-626.
- [4] K. Taraman, B. Lambert, "A surface roughness model for a turning operation". *International Journal of production Research*, Vol. 12, No.6, 1974, 691-703.
- [5] M. Hasegawa, A. Seireg, R.A. Lindberg, "Surface roughness model for turning". *Tribology International* December, 1976, 285-289.
- [6] R.M. Sundaram, B.K. Lambert, "Mathematical models to predict surface finish in fine turning of steel". *International Journal of Production Research*, Vol. 19, Part-I, 1981, 547-556.
- [7] R.M.D.E. Dimla, P.M. Lister, N.J Leighton, "Neural network solutions to the tool condition monitoring problem in metal cutting-A review critical review of methods". *International Journal of Machine Tools and Manufacture*, Vol. 39, 1997, 1219-1241.
- [8] A Mital, M. Mehta, "Surface roughness prediction models for fine turning". *International Journal of Production Research*, Vol. 26, 1988, 1861-1876.
- [9] S. G. Ahmed, "Development of a prediction model for surface roughness in finish turning of aluminum". *Journal of Sudan Engineering society*, Vol. 52, No.45, 2006, 1-5.
- [10] D. Singh, P. Venkateswara Rao, "A surface roughness prediction model for hard turning process". *International Journal of Advanced Manufacturing*, Vol. 32, No.11-12, 2007, 1115-1124.
- [11] L. Zhanjie, Y. Bing, T. Meili, "Prediction of surface roughness of difficult-to-cut material by HSM based on RBF neural network". *6th International conference on Instrumentation, measurement, circuits and systems*, Hangzhou, China, 2007.
- [12] C. Lu, J. Costes, "Surface profile prediction and analysis applied to turning process". *International Journal of Machining and Machinability of Materials*, Vol. 4, No. 2-3, 2008, 158-180.
- [13] S. M. Samhour, B.W. surgenor, "surface roughness in grinding: on-line prediction with adaptive neuro-fuzzy inference system". *Transactions of NAMRI/SME*, Vol. 33, 2005, 57-64.
- [14] J.S.R. Jang, "ANFIS: adaptive-network- based fuzzy inference system". *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 23, 1993, 665-685.
- [15] J.S.R. Jang, C.T. Sun, "Neuro-fuzzy modeling and control". *Proceedings of the IEEE*, Vol. 83, No.3, 1995, 378-406.
- [16] Montgomery D. C. *Design and analysis of Experiments*. 2nd ed. New York: John Wiley and sons; 1984.

Wavelet Decomposition for the Detection and Diagnosis of Faults in Rolling Element Bearings

J. Chebil ^{a,*}, G. Noel ^b, M. Mesbah ^c, M. Deriche ^d

^a ECE Dep., Faculty of Engineering, International Islamic Univ. Malaysia, Jln Gombak, KL, Malaysia.

^b Ecole Nationale Supérieure des Télécommunications. Paris, France.

^c Signal Processing and Consultancy Group, University of Queensland, Brisbane, Queensland, Australia.

^d Dep. of Electrical Eng., King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

Abstract

Condition monitoring and fault diagnosis of equipment and processes are of great concern in industries. Early fault detection in machineries can save millions of dollars in emergency maintenance costs. This paper presents a wavelet-based analysis technique for the diagnosis of faults in rotating machinery from its mechanical vibrations. The choice between the discrete wavelet transform and the discrete wavelet packet transform is discussed, along with the choice of the mother wavelet and some of the common extracted features. It was found that the peak locations in spectrum of the vibration signal could also be efficiently used in the detection of a fault in ball bearings. For the identification of fault location and its size, best results were obtained with the root mean square extracted from the terminal nodes of a wavelet tree of Symlet basis fed to Bayesian classifier.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Discrete Wavelets Transform; Discrete Wavelet Packet Transform; Ball Bearing Fault Detection.

1. Introduction

Development of real-time fault detection and identification technologies will allow a migration from expensive scheduled based maintenance to the more efficient, less costly alternative of condition-based maintenance. One of the principal tools for diagnosing early faults has been vibration analysis [1-2]. Considerable research has been carried out previously for the development of various algorithms for bearing fault detection and diagnosis. These algorithms can be classified into time domain, frequency domain, time-frequency domain, higher order spectral analysis, neural-network and model based techniques [3-7].

Various time domain statistical parameters have been used as trend parameters to detect the presence of incipient bearing damage. Kurtosis and skew values of vibration signals are used in [8] for detection of bearing faults at early stages in their development. The paper in [9] presents a study on the application of sound pressure and vibration signals to detect the presence of defects in a rolling element bearing using a statistical analysis method. The most important shortcoming of the statistical analysis approach is its inability to detect bearing defects at later stages. In the frequency domain approach the major frequency components of vibration signals and their amplitudes are used for trending purposes. The frequency

characteristics of the vibration for a defective bearing subject to various load conditions are investigated in [10]. Envelope analysis, originally known as the high frequency resonance technique, is the most commonly used frequency analysis technique for the detection and diagnosis of bearing faults. The technique is studied in detail in [11]. One of the problems with envelope analysis and the other frequency domain approaches is that, they require the bearing defect frequencies be known or pre-estimated. The other shortcoming is the increasing difficulty in analyzing the vibration spectrum when the signal to noise ratio is low and the vibration spectrum has a large number of frequency components due to the complexity of the system [3]. Bi-coherence spectra are used in [12] to derive features that relate to the condition of a bearing. Neural networks are also applied to bearing fault detection and diagnosis [13-14]. Time-frequency domain techniques use both time and frequency domain information allowing for the investigation of transient features. A number of time-frequency domain techniques have been proposed including Short Time Fourier Transform (STFT), the Wigner- Ville Distribution (WVD), and the Wavelet Transform (WT) [1], [4-6], [14-15]. This paper presents results of wavelet analysis in the detection and diagnosis of ball bearing faults. A brief description of the typical bearing faults is given along with an overview of wavelet analysis in the next subsections.

The structure of this paper is as follows. In Section 2, a description of the typical faults of the bearing is presented while the basic concepts in wavelet analysis are explained

* Corresponding author. jalel@iiu.edu.my.

in Section 3. The data used for monitoring and diagnosis is described in Section 4, whereas the detailed fault diagnosis procedure based on wavelet analysis is discussed in Section 5. The last section concludes the paper.

2. Bearing Condition Monitoring

Bearing condition monitoring has received considerable attention for many years due to the fact that the majority of the problems in rotating machines are caused by faulty bearings. A schematic diagram of rolling element bearing is shown in Figure 1a. The typical failure mode of rolling element bearings is a localized defect, which occurs when a piece of material on the contact surface is dislodged during operation. The dislodgement is mostly caused by fatigue cracking under cyclic contact

stress. In general, a ball bearing has three main components that can typically experience damage: the rolling elements, the inner race and the outer race [16].

During bearing operation, wide band impulses are generated when rollers pass over the defect at a frequency determined by shaft speed, bearing geometry, and defect location. Some of the vibrational modes of the bearing and its supporting structure will be excited by the periodic impulses, and a distinct bearing signature will be generated. The leading edge of each impulse typically comprises a very sharp rise that corresponds to the impact between a roller and the defect. The ringing then decays with an approximately exponential envelope as the energy is dissipated by internal damping [7, 17] as shown in Figure 1b.

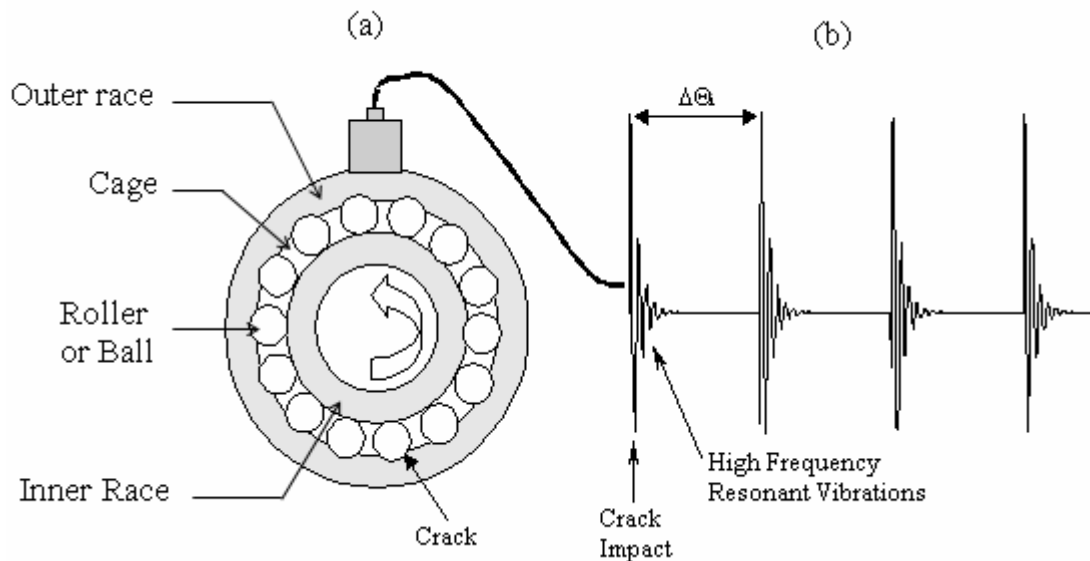


Figure 1. (a) Schematic diagram of rolling element bearings, and (b) the typical time waveform due to a crack on the outer race of a rolling element bearing. [17]

Applying Fourier transform to this type of signals results in a peak at the impact frequency along with harmonics due to the spike-resonance nature of the signal. However, the bearing fault component is often difficult to be distinguished due to the high levels of noise and other fault sources in the vicinity of the bearing fault frequencies. In addition, the frequency domain approaches are incapable of detecting nonstationary signals [18]. These problems can be overcome by using the wavelet analysis, which provides multi-resolution in time-frequency distribution for easier detection of abnormal vibration signals. Next section presents a brief summary about wavelet technique.

3. Wavelet Analysis

The wavelet transform has emerged as an efficient tool to deal with non-stationary signals such as vibrational signal waveforms [19-20]. It offers simultaneous interpretation of the signal in both time and frequency domain which allows local, transient or intermittent components to be exposed. Such components are often obscured due to averaging inherent within spectral only methods such as the Fourier transform. Wavelet transform

can be continuous or discrete. The continuous wavelet transform reveals more details about a signal but its computational time is enormous. For most applications, however, the goal of signal processing is to represent the signal efficiently with fewer parameters and less computation time. The discrete wavelet transform (DWT) can satisfy these requirements.

The DWT employs a dyadic grid and orthonormal wavelet basis functions and exhibits zero redundancy. The DWT computes the wavelet coefficients at discrete intervals (integer power of two) of time and scales [20]. The computed DWT coefficients can be used to form a set of features that unambiguously characterize different types of signals. The dilation function of the DWT can be represented as a tree of low and high pass filters, with each step transforming the low pass filter into further lower and higher frequency components as shown in Figure 2. The original signal is successively decomposed into components of lower resolution, while the high frequency components are not analysed any further. The low-frequency components of the signal are called approximations, while the high-frequency components are called details. For example, if F_s is the sampling frequency, then the approximation of an N level DWT decomposition corresponds to the frequency band

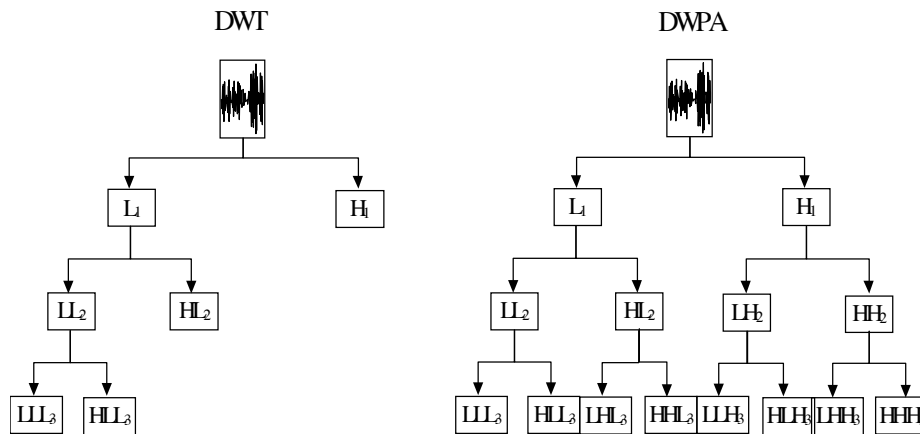


Figure 2. Filter bank representation of the DWT and DWPA decompositions [17].

$[0, \frac{F_s}{2^{N+1}}]$, whereas the detail covers the frequency range

$$[\frac{F_s}{2^{N+1}}; \frac{F_s}{2^N}].$$

The power and the flexibility of the DWT can be enhanced by using the discrete wavelet packet transform, (DWPT). Unlike the DWT, which only decomposes the low frequency components (approximations), DWPT utilises both the low frequency components (approximations), and the high frequency components (details) [20-21]. From this family of bases, a method for choosing the optimum scheme for a particular signal can be developed. This process requires a lot of a-priori information such as the choice of a mother-wavelet, the level of decomposition, and the features to be extracted. In addition, an algorithm has to be found for the selection of the best basis.

4. Vibration Data Acquisition

The experiments presented in this paper used the vibration data obtained from the Case Western Reserve University Bearing Data Centre [22]. The data were collected from an accelerometer mounted on the housing of an induction motor system coupled to a load that can be varied within the operating range of the motor. The data collection was done at two locations, one at the drive-end bearing and the other at the fan-end bearing. Data was gathered for four different conditions: (i) normal (N); (ii) inner race fault (IRF); (iii) outer race fault (ORF); (iv) ball fault (BF). Faults were introduced into the drive end bearing by using electro-discharge machining. For inner race and ball fault cases, the size of the fault is 0.007, 0.014 or 0.021 inches. For outer race fault case, the size of the fault is either 0.007 or 0.021 inches. The data is sampled at a rate of 12 kHz and the duration of each vibration signal was 10 seconds. All the experiments were repeated for four different load conditions: 0, 1, 2 and 3 horse power (HP). Therefore, experimental data consisted of 8 vibration signals for normal condition and 24 vibration signals for the inner race and ball fault conditions. For the outer race faulty case there were 23 vibration signals.

5. Experimental Results

The first step in a diagnosis of a ball bearing condition is to detect the presence of a fault, then identify its location and its size. Each step is discussed separately in the next sub-sections.

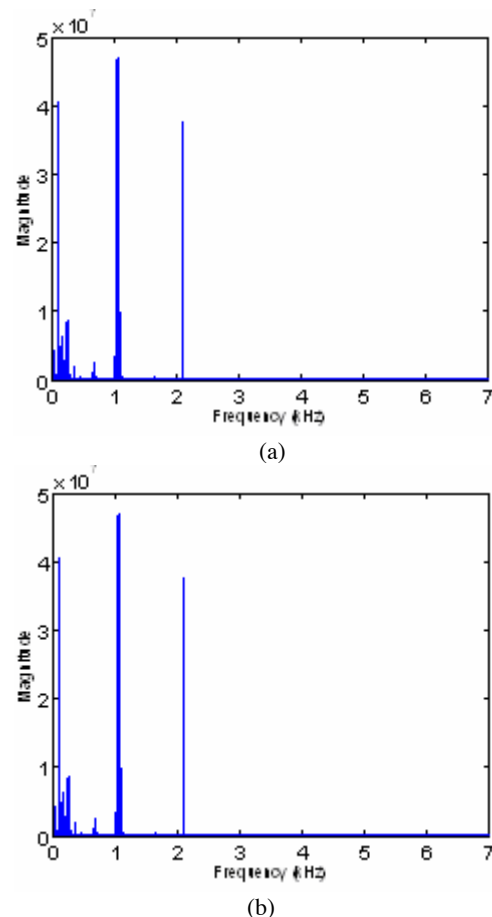


Figure 3. Typical spectrum of vibration data for (a) normal bearing, and (b) faulty bearing (IRF).

5.1. Fault Detection

Fault detection is an important and critical step. Any method developed for this purpose should be highly accurate. In this section, two approaches are investigated.

The first approach is based on the mechanical phenomenon of resonance. For a normal behaviour, the mechanical system is built to avoid resonance (Figure 3.a). But when there is a fault in the bearing, the resonant frequencies are likely to appear in the spectrum as peaks in high frequencies (Figure 3.b). The location of the frequencies peaks can be used to distinguish between normal and abnormal behaviours. In the first approach, the locations of the three first dominant frequency peaks obtained from the signal spectrums are used as features to discriminate between healthy and faulty behaviours.

Table 1. Frequency sub-bands for third level DWT decomposition.

Frequency Bands	1	2	3	4
Frequency range (Hz)	[0, 750]	[750,1500]	[1500, 3000]	[3000,6000]
	A3	D3	D2	D1

The second approach is based on the fact that cracks are translated into transient and high frequency phenomena in the vibration signal. Consequently, the abnormal behaviour can be detected by analysing the percentage of energy contained in high frequencies. In this study, DWT decomposition is applied to the vibration data for normal and abnormal bearing using mother wavelet Daubechies (Db1). Extensive experiments have showed that third level decomposition is sufficient for the problem of fault detection. The original signal is decomposed into four components: third level approximation A3, third level detail D3, second and first level details D2 and D1. The frequency sub-bands corresponding to each component of the signal are shown in Table1. It was found that when this approach is applied to the available data, the percentage of energy contained in the higher frequencies is large if the bearing is faulty and small if the bearing is healthy. Figure 4 shows sample results of the average energy contained in the frequency sub-bands 1, 2, 3 and 4 for vibration data of normal and abnormal bearing. The results demonstrate that the average energy in the first band is always the highest if the bearing is normal while it will be highest in other frequency bands if the bearing has defects. From these experiments, we could efficiently distinguish between normal and abnormal ball bearing behaviours by comparing the average energy of each sub-band.

In order to simulate noisy environments and compare the efficiency of the two methods, a white Gaussian noise with various power levels is added to the data. The detection results were perfect as long as the SNR is greater than -5 dB for the first method and greater than -20 dB for the second one. Therefore, for a noisy environment, the detection of a fault in a ball bearing is far more efficient

with a system based on frequency peaks location obtained from the Fourier analysis rather than based on energy per band obtained from the DWT decomposition.

5.2. Fault Localisation

For many applications, the detection of the fault might not be sufficient; it is also essential to determine its location. For the ball bearing, the fault can generally be located at three places: the inner race, the ball and the outer race. Thus, the identification of the fault location can be looked at as a classification problem where each class represent one fault location. To have high classification accuracy, adequate and reliable features should be extracted from the data. In this section, two wavelets based techniques are explored and applied to the vibration data. The signals are decomposed using third level DWPT or DWT decomposition, and features such as root mean square (RMS), variance and norm are extracted from the terminal nodes. A Bayesian classifier is used to segregate between different classes where each class represents one type of fault location. A Bayesian classifier is dealing with a simple probabilistic classifier based on applying Bayes' theorem with strong independence assumptions combined with a decision rule. One common rule is to pick the hypothesis that is most probable using the well-known maximum a posteriori or MAP decision rule [23]. Using this method, a number of experiments were carried out with the aim of comparing the performance of the DWPT and the DWT and finding the best mother wavelet that produces the best extracted features. Noise with three power levels was added to the data with SNR equal to 120, 20 and 0 dB. The noisy data is added to test the strength of the proposed methods.

In order to compare the performance of the two types of decompositions when used with the Bayesian classifier, the classification accuracy has been evaluated for various extracted features and for different wavelet bases. Tables 2 and 3 present the results for the fan-end bearing and for the drive-end bearing respectively. The tables show that DWPT is more efficient than DWT in identifying the fault location for the two above-mentioned cases. However, DWT-based technique can still be used if a lower level of computation is required.

Table 2. Classification accuracy (%) using the extracted features from the DWPT and DWT on the fan-end bearing data.

Feature	Wavelet	IRF*		BF*		ORF*	
		DWPT	DWT	DWPT	DWT	DWPT	DWT
RMS	Db4	100	88	100	82	100	100
RMS	Db6	100	85	100	91	100	100
RMS	Sym4	100	79	100	88	100	100
RMS	Sym6	100	85	100	94	100	100
Norm	Db4	73	61	100	94	100	97
Norm	Db6	85	55	100	100	100	91
Norm	Sym4	64	67	100	97	100	88

*IRF: inner race fault, BF: ball fault, ORF: outer race fault

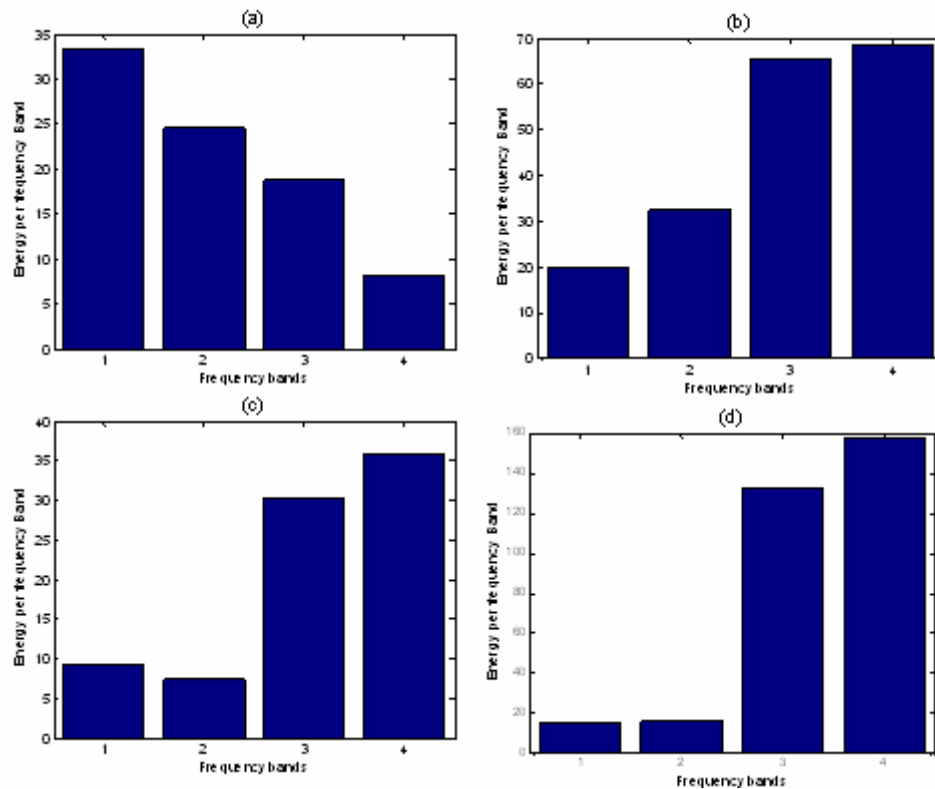


Figure 4. Average energy contained in the frequency sub-bands 1, 2, 3 and 4 for vibration data for rolling element bearing with (a) no fault, (b) IRF, (c) BF, and (d) ORFz

Table 3. Classification accuracy (%) using the extracted features from the DWPT and DWT on the drive-end bearing data.

Feature	Wavelet	IRF		BF		ORF	
		DWPT	DWT	DWPT	DWT	DWPT	DWT
RMS	Db4	100	97	67	6	100	100
RMS	Sym4	100	97	78	30	100	100
RMS	Sym6	100	100	70	6	100	100
Norm	Db4	100	100	100	89	100	94
Norm	Sym6	100	95	100	91	97	100

The next step is to determine the best mother wavelet that produces the best DWPT-based features. Table 4 shows the classification accuracy results obtained using the extracted features from the DWPT of the data for various mother wavelets: Daubechies family (Db) and Symlets family (Sym). The root mean square (RMS) and the norm are found to be the best features extracted from the DWPT. The choice of the mother wavelet is not as critical as its size and could be selected from Db4, Db5, Db6, Sym4, Sym5, and Sym6. When the combined RMS and Sym6 were applied to the database, a perfect classification was achieved.

5.3. Fault Size

The results of the previous experiments do not solve the issue of the fault size. The problem is now to determine

the best information that can help in the classification of the size of the faults. First, we assume that the diagnosis is going to adopt a multi-level classification as shown in Figure 5. The Bayesian method, as explained in the previous section, is used here on signals belonging to a specific type of fault location. The signals are decomposed by a third level DWT decomposition to obtain four terminal nodes for the drive-end bearing and to second level decomposition for the fan-end bearing. The study of the results for the two bearings shows that RMS with a mother wavelet among Db4, Db5, Db6, Sym4, Sym5, and Sym6 is a suitable choice.

Table 4. Classification accuracy (%) after a DWPT for the drive-end bearing (DE) and a fan-end bearing (FE) data.

Feature	Wavelet	IRF		BF		ORF	
		DE	FE	DE	FE	DE	FE
Variance	Db4	100	88	0	91	100	100
Variance	Sym6	100	97	0	94	100	100
RMS	Db4	100	100	69	100	100	100
RMS	Db6	100	100	69	100	100	100
RMS	Sym6	100	100	75	100	100	100
Norm	Db4	100	73	100	100	100	100
Norm	Sym6	100	67	100	100	97	100

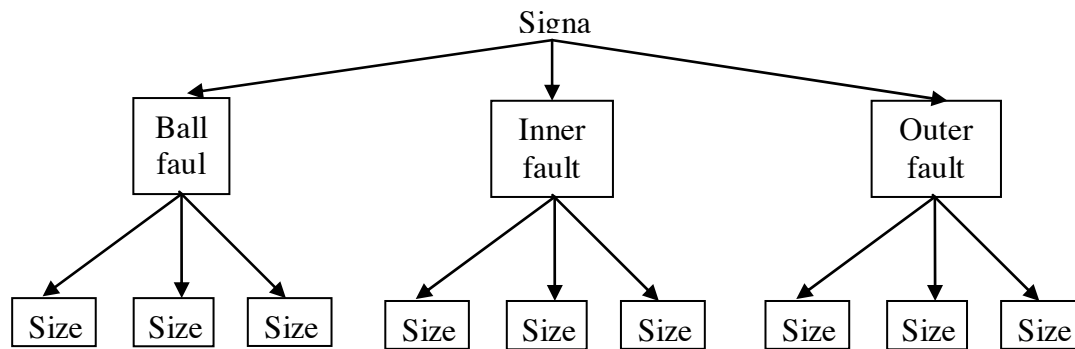


Figure 5. Multilevel classification where Size 1= 0.007 inch, Size 2= 0.014 inch and Size 3= 0.021 inch.

6. Conclusion

This paper shows that wavelet-based analysis techniques can be efficiently used in condition monitoring and fault diagnosis of bearings. In the first part of the paper, it was found that the peak locations in spectrum of the vibration signal could be efficiently used in the detection of a fault in ball bearings. For the identification of fault location and its size, the RMS extracted from the terminal nodes of a wavelet tree can be reliably used as discriminating feature. It was found that the choice of the mother wavelet Sym6 combined with the use of the RMS feature produce excellent classification results.

References

- [1] S. Wadhwani, S. P. Gupta and V. Kumar, "Wavelet based vibration monitoring for detection of faults in ball bearings of rotating machines". *Journal Inst. Eng. (India) –EL*, Vol. 86, 2005, 77-81.
- [2] S. Pati, J. Mathew, P.K. RajendraKumar. "Bearing signature analysis as a medium for fault detection: A Review". *Transactions of ASME, Journal of Tribology*, Vol. 130, No. 1, 2008, 014001-1 to 7.
- [3] H. Ocak, K.A. Loparo, "A new bearing fault detection and diagnosis scheme based by on hidden Markov". *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP 2001*, Salt Lake City-UT, USA, 2001.
- [4] E. K. Lada, J. C. Lu, J. R. Wilson, "A wavelet-based procedure for process fault detection". *IEEE Transactions on Semiconductor Manufacturing*, Vol. 15, No. 1, 2002, 79-90.
- [5] M. H. Kahaei, M. Torbatian, J. Poshtan, "Detection of bearing faults using Haar wavelets". *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E89-A(3), 2006, 757-763.
- [6] X. Lou, K. A. Loparo, F. M. Discenzo, J. Yoo, A. Twarowski, "A wavelet-based technique for bearing diagnostics". *International Conference on Acoustics, Noise and Vibration*, Montreal, Canada, August 2000.
- [7] J. Altmann, J. Mathew, "High frequency transient analysis for the detection and diagnosis of faults in low speed rolling element bearings". *Proceedings of Asia Pacific Vibration Conference '97*, Vol. 2, Kyungju, Korea, 1997.
- [8] H. R. Martin, F. Honarvar, "Application of statistical moments to bearing failure detection". *Journal of Applied Acoustics*, Vol. 44, 1995, 67-77.
- [9] R. B. W. Heng, M. J. M. Nor, "Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition", *Journal of Applied Acoustics*, Vol. 53, No. 1-3, 1998, 211-226.
- [10] Y. T. Su, S. J. Lin, "On initial fault detection of a tapered roller bearing: frequency domain analysis". *Journal of Sound and Vibration*, Vol. 155, No. 1, 1992, 75-84.
- [11] P. D. McFadden, J. D. Smith, "Vibration monitoring of rolling element bearings by high freq. resonance technique: A Review". *Tribology International*, Vol. 77, 1984, 3-10.
- [12] C. James Li, J. Ma, B. Hwang, "Bearing localized defect detection by bicoherence analysis of vibrations". *Journal of Engineering for Industry*, Vol. 117, November 1995, 625-629.
- [13] M. Subrahmanyam, C. Sujatha, "Using neural networks for the diagnosis of localized defects in ball bearings". *Tribology International*, Vol. 30, No. 10, 1997, 739-752.
- [14] C. James Li, Jun Ma, "Wavelet decomposition of vibrations for detection of bearing-localized defects". *Independent Nondestructive Testing and Evaluation (NDT&E) International*, Vol. 30, No. 3, 1997, 143-149.
- [15] Gary Y. Yen, Kuo-Chung Lin, "Wavelet packet feature extraction for vibration monitoring". *Proceedings of the IEEE International Conference on Control Applications*, August 1999, 1573- 1578.
- [16] S. H. Upadhyay, S. P. Harsha, S. C. Jain, "Nonlinear vibration signature analysis of high speed rotor due to defects of rolling element". *Advances in Theoretical and Applied Mechanics*, Vol. 1, No. 7, 2008, 301-314.
- [17] J. Altmann, J. Mathew, "DWPA best basis demodulation for the detection and diagnosis of faults in rolling element bearings". *Proceedings of the 1st Australasian Conference on Systems Integrity and Maintenance*, Surfers Paradise, Australia, October 3, 1997, 331-339.
- [18] P.W. Tse, Y. H. Peng, Richard Yam, "Wavelet analysis and envelope detection for rolling element bearing fault diagnosis: Their effectiveness and flexibilities". *Journal of Vibration and Acoustics*, Vol. 123, No. 3, July 2001, 303-310.
- [19] V. Giurgiutiu, A. Cuc, P. Goodman, "Review of vibration-based helicopters health and usage monitoring methods". *55th Meeting of the Society for Machinery Failure*

- Prevention Technology, Virginia Beach, VA, USA, April 2001.
- [20] Mertins A. Signal analysis: wavelets, filter banks, time-frequency transforms and applications. John Wiley & Sons Ltd; 1999.
- [21] J. Altmann, J. Mathew, "Multiple band-pass autoregressive demodulation for rolling-element bearing fault diagnosis. Mechanical Systems and Signal Processing, Vol. 15, No. 5, 2001, 963-977.
- [22] Case Western Reserve University Bearing Data Center. <http://www.eecs.case.edu/laboratory/bearing/download.htm>, 2009.
- [23] Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Naive_Bayes_classifier, 2009.

Acoustic Diagnosis Technique for Machine Condition Monitoring

M. A. Nawafleh ^{a,*}, N.Al-Kloub ^b

^aDepartment of Civil Engineering, Al-Hussen Bin Talal University, Ma'an, Jordan

^bDepartment of Mechanical Engineering, Al-Balqa' Applied University, Amman, Jordan

Abstract

This work presents a real time analyzer which allows extremely and accurately finding the noisiest of acoustic fluctuations in machines due to imposing spectra on the screen against each other. Technical and acoustic diagnostics of machines were predicted using the real time analyzer. The schemes and diagrams presented in this paper allow effectively using the studied real time analyzer for many other purposes.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Analyzer; Real Time; Spectrum; Noise Meter; Filters.

1. Introduction

The development of modern machines is impossible without the constant control of its operation. Moreover, it is necessary to conduct an analysis of processes occurring in units of machines and also understanding the mechanisms and the performances of the previous generation of similar machines [1-3]. Most of the operations in the modern machines include acoustics and technical diagnostics which require the using of analyzers in real time [4 - 6].

Let's consider what opportunities open using a real time analyzer. For example, using of analog-digital correction and most applicable spectral analysis in 1/3 Octave band of frequencies. This spectrum meter can perform the automatic spectral analysis

in a range from 25 Hz up to 20 kHz with the help 30 Octave filters, and also if necessary under the characteristics A, B, C, D of noise meter [7-10]. The given device provides the analysis of a spectrum of signals and averaging of signals in time. The device can also estimate the maximal amount and record the received signals. Also, visual supervision on the indicator of a spectrum of signals and digital selective estimation in one of channels can be stipulated by the device. At a dynamic range of

60 dB, the resolving power of the device makes 0.2 dB. Exclusive accuracy of gauging thus is provided. In this work, the reliability and accuracy of the application of analyzer used for acoustic diagnostics was studied in more details and investigated.

2. Analyzers Application for Acoustic Diagnostics

The acoustic analyzer is operating by receiving a signal from the converter (sensor placed on the examined object

or in the given point) which comes on broadband amplifiers (the preamplifier usually is on object together with the converter). Then the signal in each channel passes through the filter, square-law detector, amplifier and threshold circuit.

Circuits of averaging of memory and commutation further follow, then the signals come on the indicator on which for each analysis they are submitted as horizontal strips. The second channel begins from correcting filters (A- D) and gives integrated values which are mentioned before. In figure 1[2], the fields of the tolerance for one-third octave filters of the analyzer are presented.

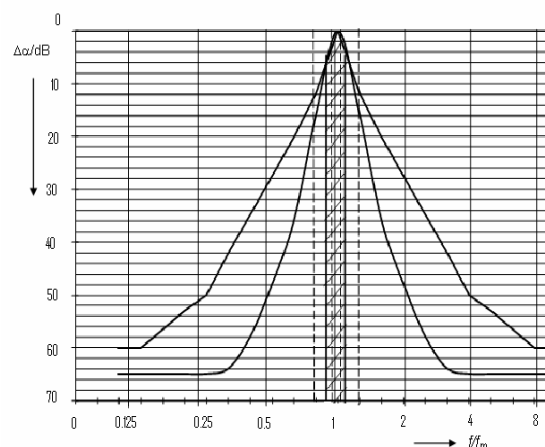


Figure. 1 Field of the tolerance for the relative characteristic of attenuation $\Delta\alpha$ for one-third octave filters of rated frequency f/f_m and the relative characteristic of attenuation for channels of filters 25 Hz... 20 kHz of the analyzer of a spectrum [2].

* Corresponding author. m_nawafleh@hotmail.com.

In table 1, the correcting values of parameter of attenuation f/f_M are shown. For the channels of filters, the field of the tolerance for the relative characteristic of attenuation $\Delta\alpha$ for one-third octave filters is valid. The image of the real characteristic of the attenuation of one-third octave filters for channels of 25 Hz ...20 kHz was imposed on a field of the tolerance [11]. The curves are well coincide and as a whole are symmetric. The relative attenuation of 60 dB is already achieved at < 0.4 and $> 2.5 f/f_M$.

Table 1. Parameters of attenuation.

f/f_M	$\Delta\alpha$
≤ 0.1250	$60 \text{ dB} \leq \Delta\alpha < \infty$
0.2500	$50 \text{ dB} \leq \Delta\alpha < \infty$
0.7937	$13 \text{ dB} \leq \Delta\alpha < \infty$
0.8909	$-0.5 \text{ dB} \leq \Delta\alpha \leq 6 \text{ dB}$
0.9439	$-0.5 \text{ dB} \leq \Delta\alpha \leq 1 \text{ dB}$
1.0000	$-0.5 \text{ dB} \leq \Delta\alpha = 0 \text{ dB}$
1.0595	$-0.5 \text{ dB} \leq \Delta\alpha \leq 1 \text{ dB}$
1.1225	$-0.5 \text{ dB} \leq \Delta\alpha \leq 6 \text{ dB}$
1.2599	$13 \text{ dB} \leq \Delta\alpha < \infty$
4.0000	$50 \text{ dB} \leq \Delta\alpha < \infty$
≥ 8.0000	$60 \text{ dB} \leq \Delta\alpha < \infty$

Below 0.25 and above 4.0 f/f_M there is a restriction of attenuation. In this case attenuation makes 62 ...66 dB depending on the channel of the filter. In the pass band no significant pulsation is observed.

Accordingly, the calibration on transfer factor of a microphone is applied, when its value is known, but the precondition is that this value remains constant. For test control and inspection measurements, the periodic control of microphones is necessary [12]. The calibration is based on basic sound pressure,

$$P_0 = 2 \cdot 10^{-5} \text{ N/m}^2 = 2 \cdot 10^{-5} \text{ Pa} \approx 0 \text{ dB}$$

For microphones, the working sensitivity a_k and working factor of transfer B_k can be calculated using the following ratio:

$$a_k = 20 \lg \left(\frac{B_k}{B_0} \right) \text{ dB} \quad (1)$$

Predetermined factor of transfer B_0 is accepted 10 mV/Pa. For capacitors measuring microphones, the values of the working factor of transfer is B_k and the working sensitivity a_k are:

for 1-inch microphone $B_k \approx 50 \text{ mV/Pa}$, $a_k \approx 14 \text{ dB}$
for 1/2-inch microphone $B_k \approx \text{mV/Pa}$, $a_k \approx 0 \text{ dB}$
for 1/4-inch microphone $B_k \approx 2 \text{ mV/Pa}$, $a_k \approx -14 \text{ dB}$

For calibration of the inlet to the analyzer of a spectrum moves a calibrating voltage 100 mV and at the expense of change of amplifying and basic level of digital indication adjust the level of sound pressure, at which the microphone would produce a voltage of 100 mV. This

calibrating level L_∇ makes for a microphone with working sensitivity a_k :

$$L_\nabla = 144 \text{ dB} - a_k \quad (2)$$

For microphones with $a_k = 14 \text{ dB}$ levels of a voltage and sound pressure thus are simultaneously correctly shown.

Let P_∇ - sound pressure, at which the microphone produces a voltage of 100 mV, then, the calibrating level can be written as:

$$L_\nabla = 20 \lg \frac{P_\nabla}{P_0} = 20 \lg p_\nabla / \text{Pa} - \lg p_0 / \text{Pa} \text{ dB} \quad (3)$$

The working sensitivity is determined as the 20-multiple logarithm of the relation between voltages which is raised at sound pressure p, to the agreed factor of transfer

$$10 \text{ mV/Pa. } a_k = 20 \lg \frac{u}{\frac{p}{10 \text{ mV}}} \cdot \frac{Pa}{Pa}$$

$$\text{This implies: } \frac{a_k}{20} = \left(\lg \frac{u}{10 \text{ mV}} - \lg \frac{p}{Pa} \right)$$

Further follows

$$\lg \frac{p_0}{Pa} = \lg / 2 \cdot 10^{-5} / = -4,7, \text{ hence } L_\nabla = 20/1 - \frac{a_k}{20 \text{ dB}} + 4,7$$

$$\text{dB, } L_\nabla = 114 \text{ dB} - a_k$$

From the given below monogram (fig. 2), it is possible to take the calibrating level L_∇ , if the working factor of transfer B_k or working sensitivity a_k is known. Only for microphone cells the factor of transfer and amplifying are given at absence of load B_L by mV/Pa and dB accordingly,

$$a_L = 20 \lg \left(\frac{B_L}{B_0} \right) \quad (4)$$

For microphone preamplifiers the rate of attenuation a_V is shown, it is possible to find working factor of transfer: $a_k = a_L - a_V$.

The rate of attenuation a_V depends on the connected capacity of a source from given microphone cell. Rates a_V for average capacities of cells of various types (for example, $a_V = 68$ for 1-inch cell) are therefore specified.

At the end, the analyzer allows, with the use of adjusting digital chain, effectively to accept a signal in a wide dynamic range of levels or voltage acting on the inlet of the amplifier, fig. 3.

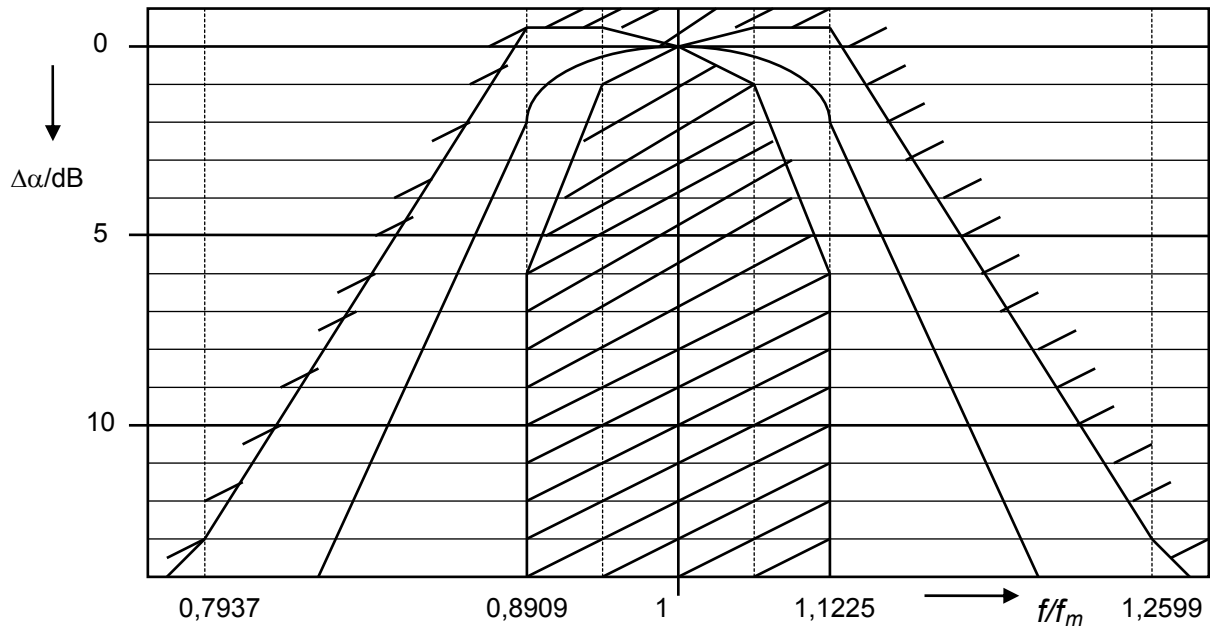
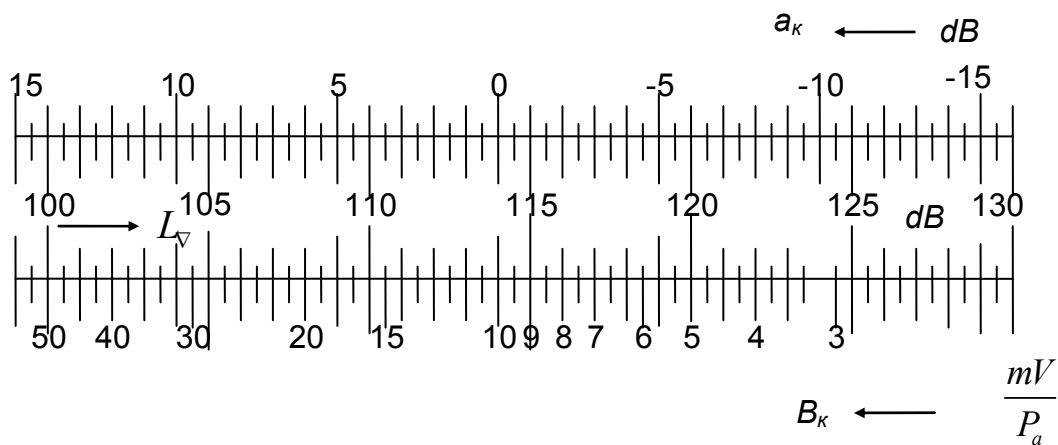


Figure 2. Fragment of figure with a site of the characteristic within a pass ban.

Figure 3. Monogram for finding L_v depending on B_k or a_k .

Thus, by using the developed measuring scheme it is possible, with practically complete technical reliability, to get spectra of levels of sound pressure in 1/3-octave bands of frequencies (octaves), for example, before and after covering of metal plates used for manufacturing of thin-walled protecting structures of the equipment.

Actually, the philosophy of measurement with the use of the analyzer and the measuring scheme consists of serial

putting of damping coverings on the calibrated metal plates and measurement of acoustic efficiency of damping of a specific radiating surface.

To get the authentic results, the exact adjustment of a sound signal is important and its adequate presentation in an electrical signal with the subsequent giving to a scale of the meter (Fig. 4).

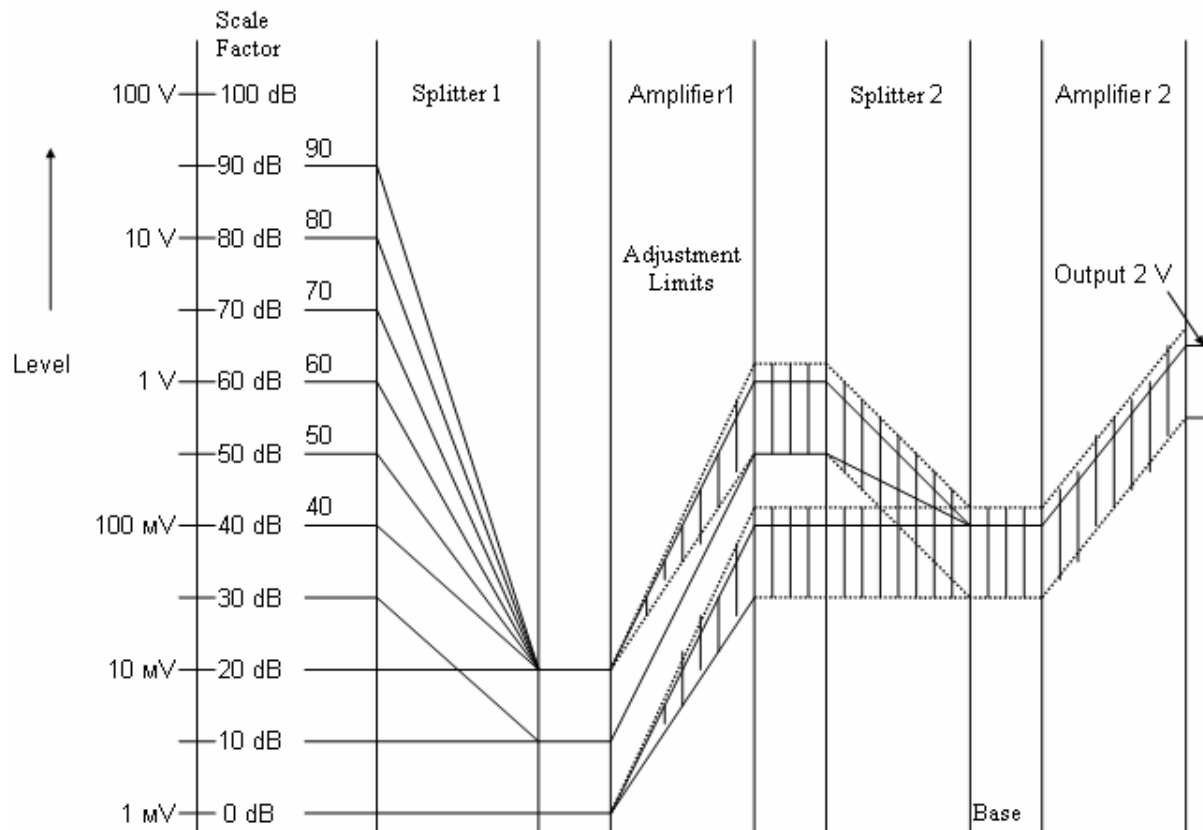


Figure 4. Diagram of a level of the Inlet amplifier.

3. Conclusion

The real time analyzer allows extremely authentically finding the noisiest source of acoustic fluctuations in the machine due to imposing spectra on the screen against each other. Separately there is a problem of technical and acoustic diagnostics of machines. The schemes and diagrams, submitted in this paper, allow effectively enough using the analyzer for many other purposes. The analyzer accepts a signal of a wide dynamic range of levels or voltage acting on the inlet of the amplifier. The rate of attenuation depends on the connected capacity of a source from a given microphone cell.

References

- [1] M.A. Alnawafleh, O.N. Nizhibitsky, "Effective methods analysis for machines noise control". *Journal of Applied Sciences*, Vol. 4, No. 2, 2004, 197-200.
- [2] T. Yoshioka, "Development of diagnosis system for machines using compound detection of acoustic emission and vibration". *Japanese Journal of Tribology*, Vol.51, No. 4, 2006, 503-513.
- [3] M. Kozochkin, N. Kochinev, F. Sabirov, "Diagnostics and monitoring of complex production processes using measurement of vibration-acoustic signals". *Measurement Techniques*, Vol. 49, No. 7, 2006, 672-678.
- [4] A. Lucifredi, P Silvestri, "Virtual experimental modal analysis: an application of simulation models to diagnostics". *The Fifth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM 2008 / MFPT 2008)*, Scotland, UK, 2008.
- [5] P. Santin, G. Sedmak, "A real-time spectrum analyser with on-line definition of the confidence levels". *Astrophysics and Space Science*, Vol. 48, No. 1, 1977, 57-63.
- [6] R.C.Wu, C.T. Chiang, J.I. Tasi, "The real-time optimal spectrum analysis system based on personal computer". *Applied Simulation and Modeling (ASM)*, Greece, 2006.
- [7] Khla M, Muller XA. *Hand book in applied acoustics*. Leningrad, shipbuilding; 1980.
- [8] N.D. Stewart. *Spinning Noise Textile Industries*. V.141, 1977.
- [9] F.Q. Wu, G. Meng, "Feature extraction based on the 3D spectrum analysis of acoustic signals to identify rotor malfunction". *International Journal of Advanced Manufacturing Technology*, Vol. 28, No. 11-12, 2006, 1146-1151.
- [10] Nizhibitsky O.N. *Hand book of machines vibration*. 6 Vol. Moscow, 1980.
- [11] Kluken E E, Kolesnikov A E. *Acoustical measurements*. Leningrad, shipbuilding; 1982,
- [12] Keluva A E. *Hand book of noise occupational control*. Shipbuilding, Moscow; 1797.

Vibration Analysis of Plates With Spot Welded Stiffeners

S. M. Nacy^{*}, N. K. Alsaheb, F. F. Mustafa

Al-Khawarizmi College of Engineering, University of Baghdad, Jaderyia, Baghdad, Iraq

Abstract

In this work both theoretical and experimental investigation was carried out to study the effect of residual stresses on the vibrational characteristics of plates with spot welded stiffeners at different boundary conditions. Expressions of the exact frequency equation were derived. Finite element modeling (FEM) was adopted to predict the tendon force produced due to spot welding and to find the natural frequencies at different modes. Different experimental models were tested to backup the results obtained theoretically. It was found that both theoretical and experimental results are in good agreement.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Stiffened Plates; Spot Welding, Residual Stresses; Vibration Analysis; Finite Element Analysis.

Nomenclature

a, b	Plate side length (mm)
D	Flexural rigidity of an isotropic plate (N.mm)
D _x , D _y	Flexural rigidity of an orthotropic plate in x and y directions
D _{xy}	Torsional rigidity of an isotropic plate
G _{xy}	Shear modulus of orthotropic material
h	Plate thickness (mm),
N _x	Edge forces per unit distance (N/mm)
T	Kinetic energy of the element (W)
t	Time (sec)
U	Strain energy stored in complete plate (W)
U _b	Strain energy stored due to bending (W)
U _t	Strain energy stored due to twisting (W)
U _r	Strain energy stored due to concentrated force (W)
w	Displacement components in z directions
x, y	Cartesian coordinates
C-F-S-C	Clamped-Free-Simply-Clamped
ρ	Mass density (Kg/mm ³)
ω	Angular frequency (rad/S)
ω _f , ω _s	Angular frequency without and with residual stresses non dimensional frequency. λ

1. Introduction

The wide use of stiffened structural elements in engineering began mainly with the application of steel plates for hulls of ships, steel bridges and aircraft structures. Structures consisting of thin stiffened plates have now also found wide applications in modern industry.

The stiffening usually have a small part of the total weight of the structure, substantially influence their strength and performance under different load conditions. Recently, [1- 3] studied the free vibration analysis of stiffened plates and shells, it was found that the stiffeners shape and distribution have great effect on the natural frequencies and mode shapes of the plate.

On the other hand, resistance spot welding is a process used for joining faying surfaces. Major advantages of resistance spot welding are high speed and suitability for automation. Many researches have been published regarding joining strength and residual stresses of spot welds [4 - 6].

The main objectives of this investigation is to study the effect of spot distribution and residual stresses, induced from spot welding, on the vibrational characteristics of plates with spot welded stiffeners at different boundary conditions. Expressions of the exact frequency equation were derived. Finite element modeling was adopted to predict the tendon force produced due to spot welding, and to find the natural frequencies and mode shapes. Different experimental models were tested to backup the results obtained theoretically.

1.1. Frequency Equation

Due to the existence of stiffeners, the stiffened plate is treated as an orthotropic material. The fundamental equation for small deflection theory of bending of thin plates is used to give details of the theoretical analysis of residual stresses that result from welding and its effect on the natural frequencies and mode shapes. The governing differential equation of deflection for an orthotropic plate, subjected to a force (N_x per unit length) acting on the edges of the plate, can be written as,

$$D_x \frac{\partial^4 w}{\partial x^4} + 2H \frac{\partial^4 w}{\partial x^2 \partial y^2} + D_y \frac{\partial^4 w}{\partial y^4} = N_x \frac{\partial^2 w}{\partial x^2} \quad (1)$$

Where, $H = D_{xy} + 2 G_{xy}$,

^{*} Corresponding author. somernacy@gmail.com.

and D_x , D_y , D_{xy} , G_{xy} represent the flexural and torsional rigidities of an orthotropic plate, respectively.

The analysis of deformation and stress in an elastic body can be accomplished by employing the energy method. Conjunction of energy method with Rayleigh's principle is used to obtain an approximation to the modal frequencies of transverse vibration of thin plate subjected to longitudinal mid plane tension N_x per unit width as shown in figure 1. For a small rectangular element of plate, the strain energy stored is the sum of the work done by the bending moment δU_b , twisting moment δU_t , and the mid plane force δU_r . If the maximum kinetic energy of the element is δT , then the angular frequency ω of the plate may be deduced from the energy equation, where the integrations are taken over the whole plate.

$$D_x \frac{\partial^4 w(x, y, t)}{\partial x^4} + 2H \frac{\partial^4 w(x, y, t)}{\partial x^2 \partial y^2} + D_y \frac{\partial^4 w(x, y, t)}{\partial y^4} + \rho h \frac{\partial^2 w(x, y, t)}{\partial t^2} = 0 \quad (4)$$

The displacement function $w(x, y, t)$ is approximated by means of the expansion,

$$w(x, y, t) \approx w(x, y) \sin \alpha t = \sin \alpha t \sum_{ij} C_{ij} X_i(x) Y_j(y) \quad (5)$$

Table 1 shows expressions of the displacement function for different boundary conditions. Substituting each expression separately in eq.(3), integrating over the given domains, applying boundary conditions and rearranging terms to obtain the final form of the frequency equation as listed below,

$$a^4 \frac{\rho h}{D_x} \omega_f^2 = [C_1 + C_2 \frac{a^2}{b^2} (\frac{D_{xy} + 2G_{xy}}{D_x}) + C_3 \frac{a^4}{b^4} \frac{D_y}{D_x}] \quad (6)$$

and

$$\omega_s^2 = \omega_f^2 + A \frac{N_x}{\rho h a^2} \quad (7)$$

where C_1 , C_2 , C_3 and A are constants depending on the boundary conditions as tabulated in table 2. Sample of calculation of these constants is clarified in Appendix.

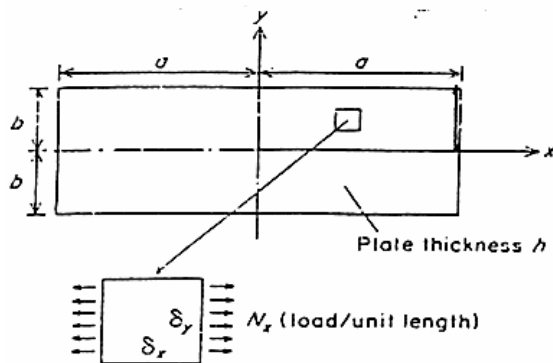


Figure 1. Geometry and mid-plan loading of a rectangular plate.

2. Finite Element Modeling

The FEM analysis was carried out in two steps (coupled-field analysis). A non-linear transient thermal analysis was conducted first to obtain the global

$$\int dU_b + dU_t + dU_r = \int dT \quad (2)$$

Eq. (2) can be written as follows,

$$\int_{-a}^a \int_{-b}^b \left[D_x \left(\frac{\partial^2 w}{\partial x^2} \right)^2 + 2D_{xy} \left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) + 4G_{xy} \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 + D_y \left(\frac{\partial^2 w}{\partial y^2} \right)^2 \right] dx dy + \frac{1}{2} N_x \left(\int_{-a}^a \int_{-b}^b \left(\frac{\partial w}{\partial x} \right)^2 dx dy - \int_{-a}^a 2b \left(\frac{\partial w}{\partial x} \right) dx \right) = \frac{1}{2} \rho h \omega^2 \int_{-a}^a \int_{-b}^b w^2 dx dy \quad (3)$$

Free, transverse vibrations of the structural system under study are governed by the differential system,

temperature history generated during the welding process. A stress analysis was then developed with the nodal temperatures obtained from the thermal analysis, which are applied as "body force" in the subsequent stress analysis. Then by using the result from the stress analysis with pre-stress on, dynamic structural analysis was achieved. ANSYS software was implemented to achieve such a task. The accuracy of the FEM depends on the density of the mesh used in the analysis. Therefore, it was necessary to have a more refined mesh closer to the weld nugget, while in regions located away from weld-nugget a more coarse mesh was used, as shown in figure 2. After solving the non-linear transient heat transfer and residual stress model, modal analyses were achieved to determine the natural frequencies and mode shapes of a structure.

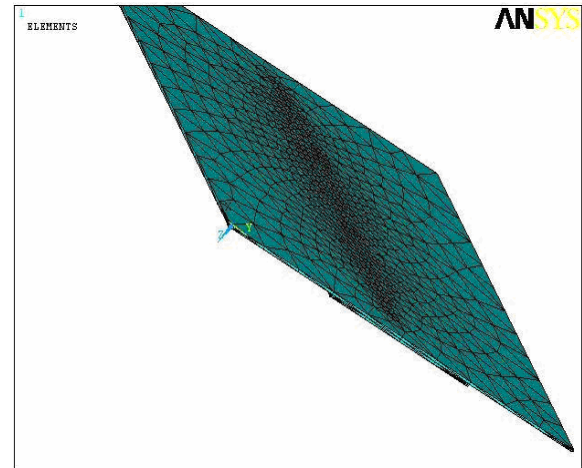


Figure 2. Finite element mesh.

3. Experimentation

Experimental tests were designed in order to measure the natural frequency and mode shape of the plates with spot welded stiffeners. A photograph of the instruments used is shown in figure 3, a simplified block diagram for these instruments is presented in figure 4. The sheet

Table 1 Displacement functions.

Boundary Condition	Displacement Function
C-C-C-C	$w = C[(\frac{2x}{a})^2 - 1]^2 [(\frac{2y}{b})^2 - 1]^2 \sin \omega t$
C-C-C-S	$w = C[(\frac{2x}{a})^2 - 1]^2 [1 + \frac{y}{b} - 6(\frac{y}{b})^2 - 4(\frac{y}{b})^3 + 8(\frac{y}{b})^4] \sin \omega t$
C-S-C-S	$w = C[(\frac{2x}{a})^2 - 1]^2 [1 - \frac{6}{5}(\frac{2y}{b})^2 + \frac{1}{5}(\frac{2y}{b})^4] \sin \omega t$
S-S-S-S	$w = C[1 - \frac{6}{5}(\frac{2x}{a})^2 + \frac{1}{5}(\frac{2x}{a})^4] [1 - \frac{6}{5}(\frac{2y}{b})^2 + \frac{1}{5}(\frac{2y}{b})^4] \sin \omega t$
S-S-S-C	$w = C[1 - \frac{6}{5}(\frac{2x}{a})^2 + \frac{1}{5}(\frac{2x}{a})^4] [1 + \frac{y}{b} - 6(\frac{y}{b})^2 - 4(\frac{y}{b})^3 + 8(\frac{y}{b})^4] \sin \omega t$
C-C-S-S	$w = C[1 + \frac{x}{a} - 6(\frac{x}{a})^2 - 4(\frac{x}{a})^3 + 8(\frac{x}{a})^4] [1 + \frac{y}{b} - 6(\frac{y}{b})^2 - 4(\frac{y}{b})^3 + 8(\frac{y}{b})^4] \sin \omega t$
C-C-C-F	$w = C[(\frac{2x}{a})^2 - 1]^2 [\frac{56}{17}(\frac{y}{b}) + \frac{24}{17}(\frac{y}{b})^2 - \frac{32}{17}(\frac{y}{b})^3 + \frac{16}{17}(\frac{y}{b})^4] \sin \omega t$
C-C-S-F	$w = C[1 + \frac{x}{a} - 6(\frac{x}{a})^2 - 4(\frac{x}{a})^3 + 8(\frac{x}{a})^4] [\frac{56}{17}(\frac{y}{b}) + \frac{24}{17}(\frac{y}{b})^2 - \frac{32}{17}(\frac{y}{b})^3 + \frac{16}{17}(\frac{y}{b})^4] \sin \omega t$
S-C-S-F	$w = C[1 - \frac{6}{5}(\frac{2x}{a})^2 + \frac{1}{5}(\frac{2x}{a})^4] [\frac{56}{17}(\frac{y}{b}) + \frac{24}{17}(\frac{y}{b})^2 - \frac{32}{17}(\frac{y}{b})^3 + \frac{16}{17}(\frac{y}{b})^4] \sin \omega t$

Table 2. Values of the constants C_1 , C_2 , C_3 , A

Boundary Condition	C_1	C_2	C_3	A
C-C-C-C	504	288	504	-17.531
C-C-C-S	504	272.83	238.73	-12.86
C-S-C-S	504	236.89	97.54	-11.44
S-S-S-S	97.54	194.87	97.54	-9.72
S-S-S-C	97.54	224.43	238.73	-10.586
C-C-S-S	238.73	258.48	238.73	-11.368
C-C-C-F	504	69.312	32.358	12
C-C-S-F	238.73	64.77	32.358	11.368
S-C-S-F	97.54	56.24	32.358	9.87

material employed in this investigation for all tests was an austenitic stainless steel AISI 304 sheet with nominal thickness of 0.6mm.

The welding process was implemented using the spot welding machine type (P1-Rectifier Press Spot Welding Machine Schlatter), with a force transducer, a displacement transducer, a Rogowski coil for measuring

the welding current and a temporizer for controlling the welding cycle (squeezing, welding and holding times). The welding machine has a maximum force of 1,885 daN, a supply pressure of 6 bar and a nominal welding power of 200 KVA. The electrodes have a truncated conical shape with a flat circular contacting area of 5 mm diameter, made of pure copper, which has high



Figure 3. Measuring Instruments.

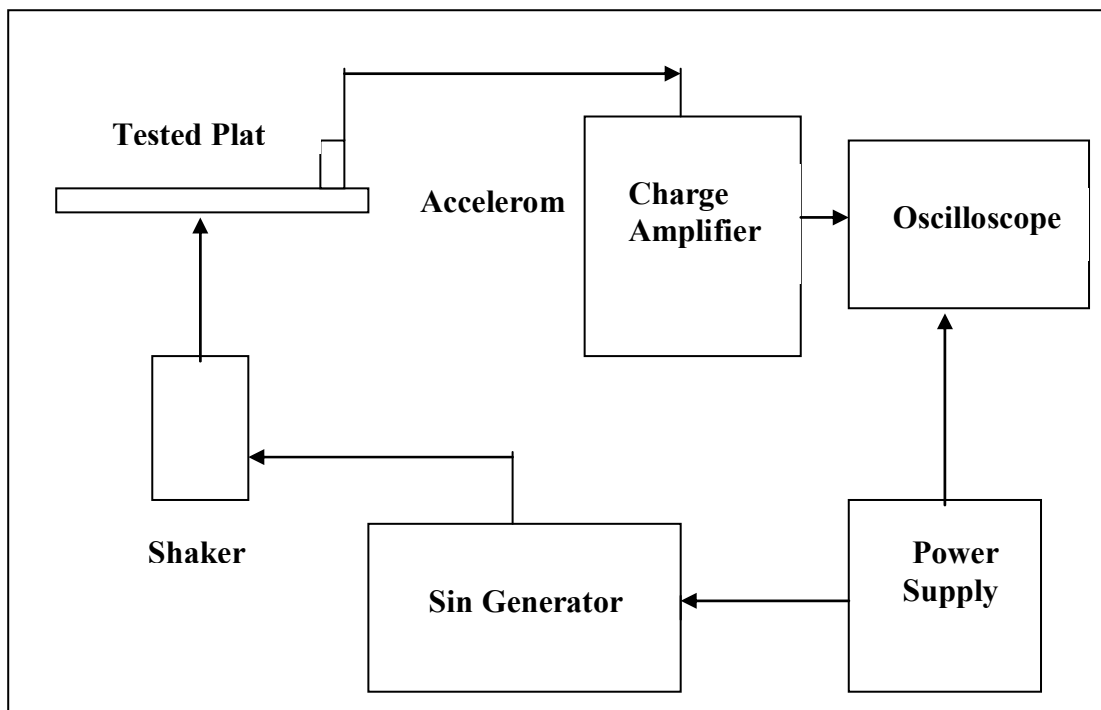


Figure 4. Block Diagram of the Measuring Instruments.

thermal and electrical conductivity, and were chosen in accordance with the ISO 5182 standard.

A rectangular stainless steel plate with dimensions of (120mm, 100mm, 0.6 mm) stiffened by another stainless steel plate of (120mm, 40mm, 0.6mm) in the longitudinal direction, is considered in this study. Five groups of plates prepared as the stiffened plate specimens, with different spot welding position are illustrated in figure 5.

The measurement of natural frequencies has been done with and without including the effect of residual stresses for each specimen to find the shift in this frequency due to spot welding.

The frequency response for each stiffened plate was investigated by slowly increasing the driving frequency of the vibrator by means of the sine generator. The natural frequency was distinguished by observing the

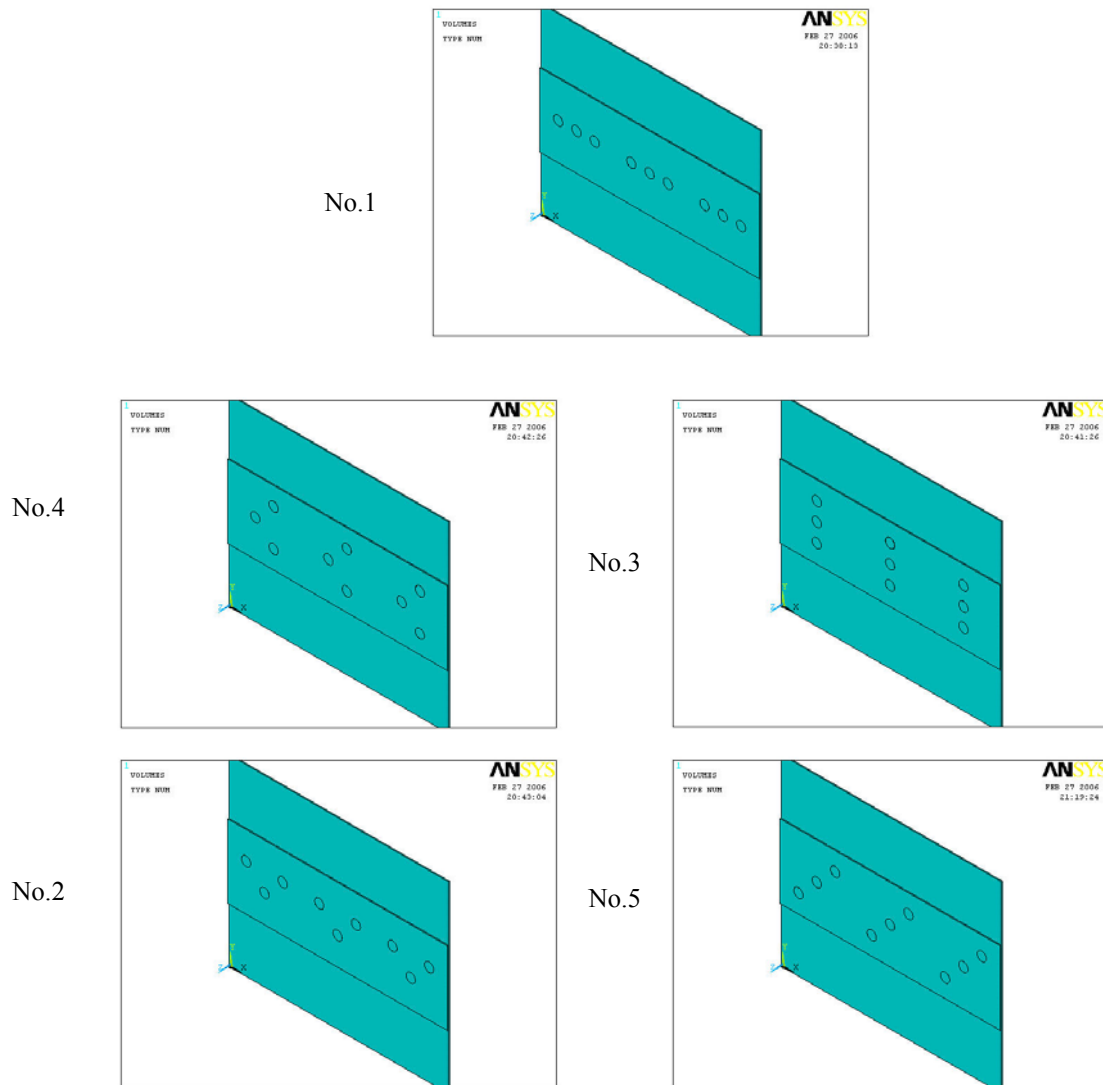


Figure 5. Different stiffened plate models.

sharp increase in amplified of the pickup output, which was amplified and displayed on the oscilloscope.

Heat treatment was employed to spot weld joint to investigate their effect on the natural frequency and mode shape of stiffened plate.

The stress relieving heat treatment was employed at a temperature of (750 C), for 3 minutes duration, then the weldments was cooled by air.

4. Results and Discussions

In order to clarify the results obtained from this investigation, three main effects were taken into consideration, namely, spot welds distribution, residual stresses and boundary conditions. Results are presented in tables 3 and 4. The first two natural frequencies for the five models considered in this study with a clamped-free-free-free boundary condition are tabulated in table 3. Although differences are small, but it is clear that the highest natural frequency is gained by model 5. Both numerical and experimental results insure that the inclusion of residual stresses tends to raise the natural frequency, for these stresses are mainly tensile in nature,

Table 3. Natural frequencies (Hz) for C-F-F-F boundary condition.

Model No.	Mode No.	FEA Results		Experimental Results	
		With Residual	Without Residual	With Residual	Without Residual
1	1 st	73.934	69.949	69.2	65.5
	2 nd	152.06	145.62	144.2	137.8
2	1 st	72.895	68.84	67.9	63.8
	2 nd	165.07	157.98	154.1	144.8
3	1 st	74.329	70.363	69.8	65.6
	2 nd	164.44	156.03	153	146.1
4	1 st	73.241	69.166	70.1	67.2
	2 nd	157.26	147.08	145.8	139
5	1 st	74.3	71.046	69.5	65.7
	2 nd	165.87	157.06	157.8	147.8

thus increasing the lateral stiffness of the plates, leading to an increase in natural frequency. The amount

of increase in natural frequency due to residual stresses is small because these stresses are acting on small areas (spots) as compared to the area of the plate.

Nodal line and mode shapes for model 1 with clamped-free-free-free boundary conditions are shown in figure 6. Mode shapes obtained numerically and experimentally are almost identical. The maximum difference in natural frequency happened to be in the first mode with a value of 6.841%, which is almost acceptable.

In order to verify the derived expressions for natural frequencies with and without the inclusion of residual stresses, values of the fundamental natural frequency for model 1 at different boundary conditions were calculated and listed in table 4. Natural frequencies obtained analytically and numerically are in good agreement, the maximum difference happened to be in the range of 1.7%. This indicates good proof for the validity of the derived frequency expressions.

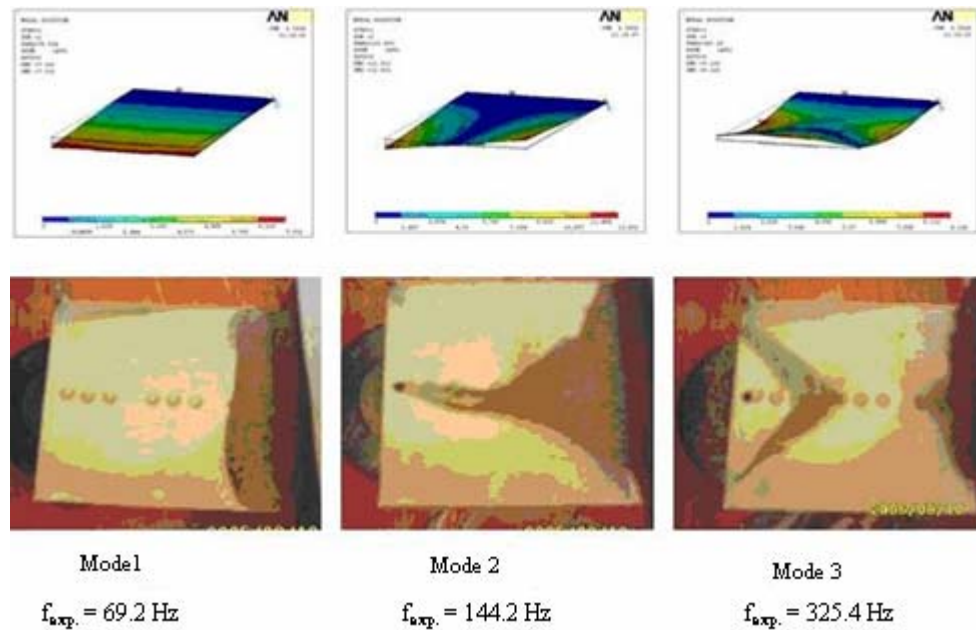


Figure 6. Nodal line and mode shapes of model 1 with C-F-F-F boundary condition.

Table 4. Fundamental natural frequencies (Hz) for different boundary conditions.

No.	B-C Description	Analytical Results		FEA Results		
		With Residual	Without Residual	With Residual	Without Residual	
1	C - C - C - C	775.153	773.963	770.797	768.3	
2	C-S-C-C	582.432	581.269	582.61	577.67	
3	C-S-C-S	443.812	442.454	437.91	436.39	
4	S-S-S-S	362.257	360.84	358.46	357.39	
5	S-C-S-S	522.304	521.23	520.63	518.65	
6	C-S-S-C	545.639	544.54	542.79	539.21	
7	C-F-C-C	372.397	370.698	366.06	363.06	
8	C-F-S-C	310.972	309.043	307.06	304.65	
9	S-F-S-C	267.60	265.659	264.35	263.12	

5. Conclusions

According to the results obtained, it can be seen that both theoretical and experimental results showed good agreement. It was found that residual stresses produced in each nugget have significant effect on the natural frequency of the plate, where natural frequency increases when residual stresses are included. This effect varies depending on the boundary conditions of the plate and on the distribution of the weld spots.

References

- [1] M. Guo, I.E. Harik, W.X. Ren, "Free vibration analysis of stiffened laminated plates using layered finite element method". *Structural Engineering and Mechanics*, Vol. 14, 2002, 245-262.
- [2] Jung JY, Kim JH. Vibration control of stiffened plates with integrated piezoelectrics. School of Mechanical and Aerospace Engineering. Coll. of Engineering, Seoul, Korea; 2002.
- [3] S.M. Nacy, M.Q. Abdullah, M.M. Ali, "Free vibration analysis of stiffened conical shell". *Journal of Engineering, Coll. of Engineering*, Vol. 8, 2002, 263-275.
- [4] Berglund D, Runnemalm H. Comparison of deformation pattern and residual stresses in finite element models of a TIG-Welded stainless steel plate. Volvo Aero Corporation, Manufacturing Process Development, Sweden; 2002.
- [5] D.H. Bae, I.S. Sohn, J.K. Hong, "Assessing the effects of residual stresses on the fatigue strength of spot welds". *Welding Journal*, 2003, 18-23.
- [6] Xin L. Finite element analysis of residual stress generation during spot welding and its affect on fatigue behavior of welded j. Ph.D. Thesis, University of Missouri-Columbia, December; 2005.

Appendix

This appendix shows a sample of calculation for C_1, C_2, C_3 , and A for the case of C-C-C-C. Consider an approximate solution with one term

$$w = C \left[\left(\frac{2x}{a} \right)^2 - 1 \right]^2 \left[\left(\frac{2y}{b} \right)^2 - 1 \right]^2 \sin \omega t$$

Satisfying the boundary conditions:

$$w = w_{x=\pm(a/2)} = 0$$

$$w = w_{y=\pm(b/2)} = 0$$

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x_{x=\pm(a/2)}} = 0$$

$$\frac{\partial w}{\partial y} = \frac{\partial w}{\partial y_{y=\pm(b/2)}} = 0$$

Substitution $w(x,y)$ in eq (3) one obtains the following frequency equation.

$$a^4 \frac{\rho h}{D_x} \omega_s^2 = [504 + 288 \frac{a^2}{b^2} (\frac{D_{xy} + 2G_{xy}}{D_x}) + 504 \frac{a^4}{b^4} \frac{D_y}{D_x}] +$$

$$17.531 a^2 \frac{N_x}{D_x} = \lambda^2$$

Optimal Forward Kinematics Modeling of Stewart Manipulator Using Genetic Algorithms

A. Omran ^{a,*}, M. Bayoumi ^b, A Kassem ^b, G. El-Bayoumi ^b

^a Department of Aerospace Engineering, Old Dominion University, Norfolk, VA, USA

^b Department of Aerospace Engineering, Cairo University, Giza, Egypt

Abstract

In this paper, a new technique for modeling the forward kinematics of a Stewart manipulator is presented based on a predicted squared error cost function. Initially, the inverse kinematics model was introduced in detail for seeding the regression points. Genetic algorithms were then employed to search for optimal model structure from a set of candidate regression predictors. The proposed technique provides a significant reduction in the expenses of memory and computation compared to the traditional Taylor series expansion model. Such reduction increases the suitability of the model for a task space control application of the Stewart manipulator.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Stewart Forward Kinematics; Predicted Squared Error; Genetic Algorithm.

1. Introduction

Stewart manipulators have been used in numerous industrial applications. Their potential benefits include a high force to weight ratio, accuracy, and rigidity. The Stewart manipulator consists of a moving plate connected to a base plate by six legs. Each leg has an upper part sliding inside a lower part. This gives the upper plate six degrees of freedom: three translational motions (surge, sway, and heave) and three rotational motions (pitch, roll, and yaw). The acceleration forces of this moveable plate can emulate the physical feeling of piloting an aircraft in forward, backward, or turning motions. Such capability makes Stewart manipulators the best candidates for flight simulators. This perspective attracts many researchers to design a controller that can sustain a more realistic motion with minimum error. Two schemes are commonly used to control the Stewart manipulator: joint space control and task space control. The joint space scheme is developed using the information of joint displacement. Each leg of the manipulator is controlled as a single-input single-output (SISO) system as shown in Fig. 1-a. The joint space control involves the inverse kinematics model, which has a closed form. Although the joint space control scheme has been given great attention [1-5], it does not achieve high tracking performance because of the nonlinear coupling in the system.

Task space control, on the other hand, achieves high performance tracking by counting the system coupling, and it has been widely investigated [6-10]. In this scheme, the framework is multi-input multi-output (MIMO) as

shown in Fig. 1-b. Thus, the forward kinematics is involved instead of inverse kinematics. This scheme is exacerbated by the fact that the Stewart manipulator forward kinematics has no closed form solution. For example, Dietmaier showed that there are 40 such possible solutions [11].

Many studies have attempted different approaches to simplify the Stewart manipulator forward kinematics problem. Ilian *et al.* [8] presented a new closed-form solution of the problem, yet it required three extra sensors. Pratik *et al.* [7] and Sadjadian *et al.* [12-13] used the neural network approach. The accuracy of this approach is very sensitive to the structure of the neural network. For example, Sadjadian *et al.* [13] showed that different structures of neural networks lead to different accuracy in modeling the forward kinematics of Stewart manipulator. Liu *et al.* [14-15] proposed a numerical algorithm based on fundamental geometric operation with three nonlinear simultaneous algebraic equations, which is impractical for control process. Sadjadian and Taghirad [12] derived a fourth order Taylor series expansion using least square estimation (LSE). This technique has an over parameterization by 774 regression coefficients. Sheng *et al.* [16] presented a hybrid immune algorithm, genetic algorithm, and fuzzy system method for solving the forward kinematic problem of Stewart manipulator. The produced model is very complicated and has not yet been applied to any simulation or experimental test case. Tarokh [17] proposed an algorithm for solving the forward kinematics of Stewart manipulator. The algorithm consists of two phases: an on-line phase and an off-line phase. A database of the relationship between links and poses is decomposed into cells.

* Corresponding author. aomra001@odu.edu.

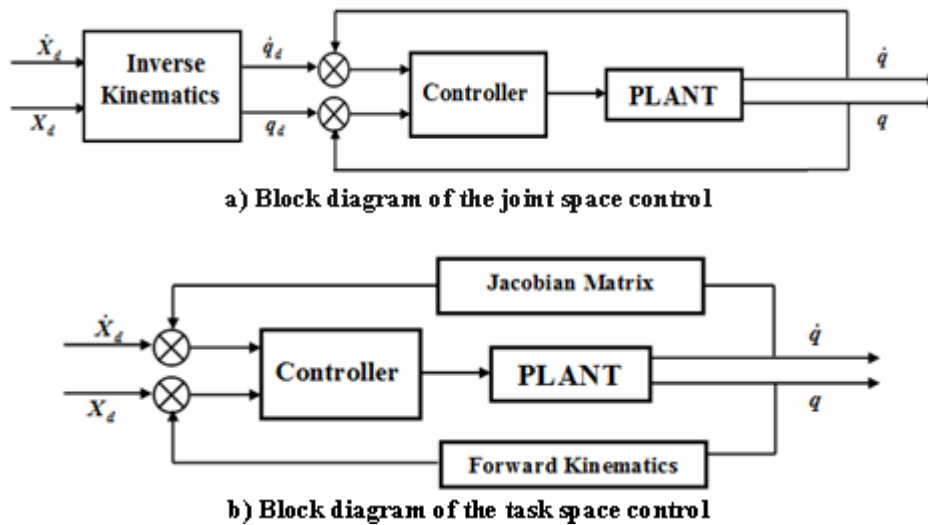


Figure 1. Comparison between the block diagram of the task space control and the block diagram of the joint space control: X refers to task space displacements and q refers to joint space displacements.

During the online phase, given link lengths, the appropriate cell is identified, and the poses are computed. Chen *et al.* [18] proposed a nonlinear observer based method to solve the forward kinematics. This observer uses the inverse kinematics as a baseline to update the model.

The current research offers an efficient technique for modeling the Stewart manipulator forward kinematics. This technique uses the predicted squared error (PSE) cost function to find an optimal structure for the forward kinematics model by reducing the number of terms in the model without affecting the required accuracy for control applications. Six optimization problems are considered to define the task space coordinates (surge, sway, heave, roll, pitch, and yaw) as functions of joint space coordinates (length of each leg). The candidate modeling predictors were selected from a third order polynomial's terms in six-dimensions (joint space). Then, the genetic algorithms (GA) was used to search for an optimal structure to minimize PSE cost function using the so-called "Switching Factor". The results of this research can be used to save the computation and memory costs. This paper is organized as follows: in section two, the description of inverse kinematics enables one to determine the link lengths in terms of desired/specified upper platform position and angular orientations. Section three includes a brief discussion of PSE regression model. Section four presents the use of genetic algorithm to search for optimal model structure. Section five offers the results of the simulation used to examine the proposed technique. Finally, section six is the conclusion.

2. Inverse Kinematics Model

There are two frames describing the motion of the moving plate: an inertia frame (X, Y, Z) located at the center of the base plate and a body frame (x_B, y_B, z_B) located at the center of the moving plate with the z_B -axis pointing outward. The angle between the local x_B -axis of the moving plate and the line of the joint J_{ui} is denoted by β_i as shown in Fig. (2).

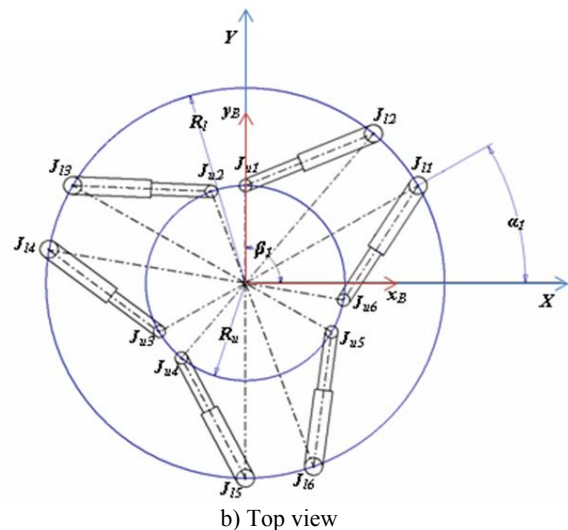
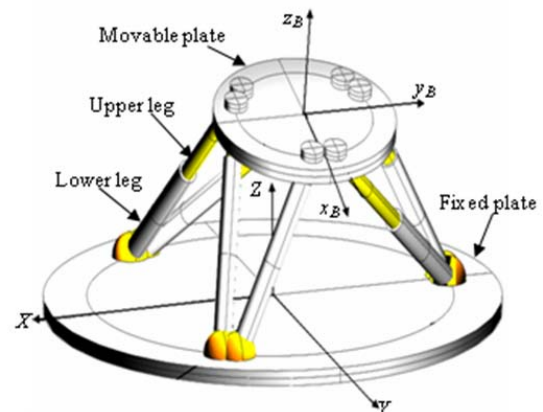


Figure 2. Stewart manipulator.

The position of the joint J_{ui} in the plate body frame is

$$P_{J_{ui}}^B = \begin{bmatrix} X_{ui}^B & Y_{ui}^B & Z_{ui}^B \end{bmatrix}^T = [R_u \cos(\beta_i) \quad R_u \sin(\beta_i) \quad 0]^T \quad i = 1, 2, \dots, 6 \quad (1)$$

In the same manner, an angle α_i is defined between the inertia X -axis and the line of the joint J_{li} . The position of the joint J_{li} in the inertia frame is defined as:

$$P_{J_{li}}^I = \begin{bmatrix} X_{li}^I & Y_{li}^I & Z_{li}^I \end{bmatrix}^T = [R_l \cos(\alpha_i) \quad R_l \sin(\alpha_i) \quad 0]^T \quad i = 1, 2, \dots, 6 \quad (2)$$

The upper plate has a capability for 6-DOF motion (three rotational motions and three translational motions). The rotational motions of the plate are defined by Euler angles in sequence 1-2-3. Thus the transformation from the body frame (x_B, y_B, z_B) to the inertia frame (X, Y, Z) is given by the Matrix R_{plate} :

$$R_{plate} = R_z^T(\psi) R_y^T(\theta) R_x^T(\varphi) = \begin{bmatrix} C_\psi & -S_\psi & 0 \\ S_\psi & C_\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} C_\theta & 0 & S_\theta \\ 0 & 1 & 0 \\ -S_\theta & 0 & C_\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & C_\varphi & -S_\varphi \\ 0 & S_\varphi & C_\varphi \end{bmatrix} \quad (3)$$

$$= \begin{bmatrix} C_\theta C_\psi & S_\varphi S_\theta C_\psi - C_\varphi S_\psi & C_\varphi S_\theta C_\psi + S_\varphi S_\psi \\ C_\theta S_\psi & S_\varphi S_\theta S_\psi + C_\varphi C_\psi & C_\varphi S_\theta S_\psi - S_\varphi C_\psi \\ -S_\theta & S_\varphi C_\theta & C_\varphi C_\theta \end{bmatrix}$$

where C refers to angle cosine and S refers to angle sine. The angles ψ , θ , and φ are Euler angles. In addition to the rotation, one should consider the translation vector T_{plate}^I as:

$$T_{plate}^I = \begin{bmatrix} x(t) \\ y(t) \\ z(t) + h \end{bmatrix} \quad (4)$$

where h is the initial height of the upper plate's center. The trajectory of the upper plate's center is defined by $x(t)$, $y(t)$, and $z(t)$. The position of the joint J_{ui} in inertial frame (X, Y, Z) is then calculated as:

$$P_{J_{ui}}^I = \begin{bmatrix} X_{ui}^I & Y_{ui}^I & Z_{ui}^I \end{bmatrix}^T = R_{plate} P_{J_{ui}}^B + T_{plate}^I \quad (5)$$

Each leg has three degrees of freedom: two rotational and one translational motion. Thus a spherical joint is employed to connect the upper part of each leg by the movable plate while the lower part is connected to the base plate by a universal joint as shown in Fig. (3). The length vector of the i^{th} leg L_i^I can then be computed from Eq. (2) and (5) as:

$$L_i^I = P_{J_{ui}}^I - P_{J_{li}}^I \quad i = 1, 2, \dots, 6 \quad (6)$$

By substituting Eq. (3) and Eq. (4) in Eq. (5), and considering the square value of vector L_i^I in Eq. (6), the relationship between the joint space variables and task space variables can be summarized as (Equation. 7):

$$L_i^2 = (C_\theta C_\psi R_u \cos(\beta_i) + (S_\varphi S_\theta C_\psi - C_\varphi S_\psi) R_u \sin(\beta_i) + x(t) - R_l \cos(\alpha_i))^2 + \\ (C_\theta S_\psi R_u \cos(\beta_i) + (S_\varphi S_\theta S_\psi + C_\varphi C_\psi) R_u \sin(\beta_i) + y(t) - R_l \sin(\alpha_i))^2 + \\ (-S_\theta R_u \cos(\beta_i) + S_\varphi C_\theta R_u \sin(\beta_i) + z(t) + h)^2 \quad (7)$$

where $i = 1, 2, \dots, 6$. Based on Eq. (7), the inverse kinematics has a closed form. On the other hand, it is "difficult" to develop any closed form for the forward kinematics.

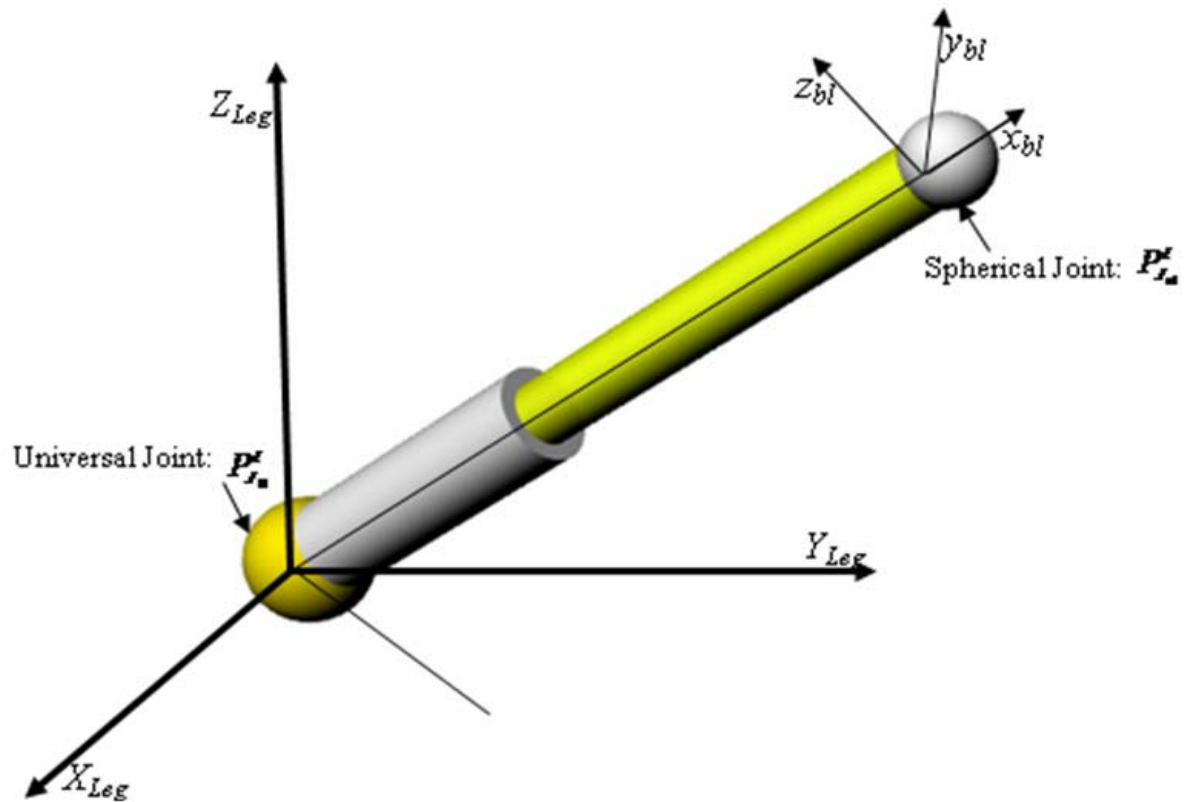


Figure 3. Leg mechanism of Stewart manipulator.

3. Regression Model and Predicted Squared Error

If an N -dimensional vector of independent responses $Y = [y_1 \ y_2 \ \dots \ y_N]^T$ is modeled in terms of linear combinations of K modeling predictors $Q = [q_1 \ q_2 \ \dots \ q_K]$, then:

$$Y = \gamma_1 q_1 + \gamma_2 q_2 + \dots + \gamma_K q_K + \varepsilon = Q\gamma + \varepsilon \quad (8)$$

where $\gamma = [\gamma_1 \ \gamma_2 \ \dots \ \gamma_K]$ is an K -dimensional constant model vector to be determined, and ε denotes the N -dimensional modeling residual vector. Each q_i is an N -dimensional vector function depending on the independent variables. In conventional least square estimation (LSE), the optimization goal is to minimize the sum of square errors between the estimated and measured values by defining a scalar function J .

$$J = \sum_{i=1}^N \varepsilon_i^2 = \varepsilon^T \varepsilon = (Y - Q\gamma)^T (Y - Q\gamma) \quad (9)$$

By minimizing J , the regression coefficients are calculated by:

$$\gamma = (Q^T Q)^{-1} Q^T Y \quad (10)$$

Although LSE minimizes the error between the measured and estimated data, this method cannot ensure an optimal regression structure. Thus, considering many regression predictors sometimes leads to over parameterization and noisy model. Many methodologies have been proposed to search for the optimal regression structure; such as, backward elimination, stepwise, and others. One of the most efficient techniques to find the minimum number of predictors with minimum error is the predicted squared error (PSE). This technique has an optimization problem because the structure of the model should be chosen to minimize F_{PSE} defined by [19]:

$$F_{PSE} = \frac{2J}{N} + \frac{K}{N} \sigma_{\max}^2 \quad (11)$$

The constant $\sigma_{\max}^2$ is the prior upper-bound estimate of the square error between future data and the model. The F_{PSE} function depends on the mean squared fit error J/N and a term proportional to the number of terms in the model. Thus $\sigma_{\max}^2$ is considered as the residual variance estimate for a constant model structure equal to the mean of the measured response values by:

$$\sigma_{\max}^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 \quad (12)$$

where $\bar{y}$ is the mean value of vector Y .

4. Optimization Procedure for Modeling the Forward Kinematics

Genetic algorithm (GA) is now considered as one of the most popular optimization and search techniques. The first obvious application for the algorithm was traced back to 1962 when Holland introduced the algorithm in his work studying adaptive systems [20]. The algorithm then received an enormous exploration by Goldberg [21]. The main advantages of GA are its global optimization performance and the ease of distributing its calculations among several processors or computers as it operates on the population of solutions that can be evaluated concurrently. It is a very simple method, generally applicable, not inclined to local optimization problems that arise in a multimodal search space, and no needs for special mathematical treatment. Moreover the algorithm is more applicable for the discontinuous problem, as in the case of the present study, unlike the conventional gradient-based searching algorithms.

Basically, genetic algorithm works based on the mechanism of natural selection and evolutionary genetics. The algorithm starts by coding the variables to binary strings (chromosomes). Every chromosome has n genes. The gene is a binary bit by value zero or one. Three main operations control the procedure of the GA: reproduction, crossover, and mutation. Reproduction is processing to select the parent from a generation. The process is based

on survival of the fittest (highest performance index). In this way, the reproduction process guides the search for the best individuals (high performance index). After the individuals are selected, the crossover process is then used to swap between two chromosomes by specific probabilistic decision. The crossover process generates offspring carrying mixed information from swapped parents (chromosomes). Mutation is the mechanism to prevent the algorithm from local optimal points by adding some degree of randomness. The process is performed by alternation of the gene from zero to one or from one to zero with the mutation point determined uniformly at random. The mutation rate should be considered carefully since the higher mutation rate means more number of generations are required for algorithm convergence and a low mutation rate may lead to a convergence for a local minimum. The algorithm maintains a constant size of generation by selecting the fittest chromosomes from parents and offsprings. The algorithm iteratively operates to converge for schema matches by some tolerance. Further description of genetic algorithms can be found in Goldberg [21-22].

In the current research, the optimization procedure is developed by using a switching factor SF . This factor $SF = [sf_1 \ sf_2 \ \dots \ sf_K]$ is a K -dimensional vector. Each element (sf_k) of this vector is a binary bit to disable or enable the vector q_i as revealed by:

$$SF = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 1 \end{bmatrix}$$

$$Q = \begin{bmatrix} q_{1,1} & q_{2,1} & q_{3,1} & \dots & q_{i-1,1} & q_{i,1} & q_{i+1,1} & \dots & q_{K-1,1} & q_{K,1} \\ q_{1,2} & q_{2,2} & q_{3,2} & \dots & q_{i-1,2} & q_{i,2} & q_{i+1,2} & \dots & q_{K-1,2} & q_{K,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ q_{1,N} & q_{2,N} & q_{3,N} & \dots & q_{i-1,N} & q_{i,N} & q_{i+1,N} & \dots & q_{K-1,N} & q_{K,N} \end{bmatrix} \quad (13)$$

If the value of this bit is "1", the equivalent predictor vector q_i is then counted in the structure of the regression model; if the value is "0", q_i is omitted. Searching for the optimal SF or optimal structure model is a discontinued optimization problem. The main reason to use the GA in the current problem is that the most conventional optimization techniques depend on slope or curvature of the response surface which is not available in the current problem. The procedure can be summarized as:

1. Generate a binary vector SF .
2. Construct the Q matrix from Eq. (13).
3. Compute the regression constant model vector γ using LSE in Eq. (10).
4. Compute the F_{PSE} cost function in Eq. (11).

An example is listed here to clear the idea behind the switching factor:

If a set of N data $X = [x_1 \ x_2 \ \dots \ x_N]^T$ and $y = [x_1 \ x_2 \ \dots \ x_N]^T$ is collected.

Assume, there are five functions considered to be candidates for modeling $y = f(x)$, these functions are compile in a vector Q as

$$Q = \begin{bmatrix} 1 & x^2 & \cos(x) & \log(x) & \frac{1}{x^2 + 1} \end{bmatrix} \quad (14)$$

In this example, there are $32 = 2^5$ possibilities to build the regression model. The proposed switching factor SF is used to code these possibilities. Thus, the switching factor is a 5-dimensional vector. Each element is either "1" or "0". If the $SF = [1 \ 0 \ 1 \ 0 \ 0]$, then the first and third candidates are only passed to the LSE or the selected regression structure is $y = \gamma_1 + \gamma_2 \cos(x)$. Based on this regression structure the coefficients γ_1 and γ_2 are computed and the PSE cost function is evaluated as a function of number of terms $K = 2$ and the error between the estimated and the measured vector y . Then, GA to qualify the performance of each SF , where 32 possibilities are available to construct the structure of the regression model for minimum F_{PSE} , repeats this procedure. Note: the model is based on the assumption that manipulator has a rigid links or no uncertainty source.

5. Simulation Results and Discussion

The proposed scheme is applied to the Stewart platform with parameters given in Table 1. A population of 15000 samples was seeded from the inverse kinematics model. The regression predictors were generated from a third order polynomial in the sense that any analytical function can be expanded in polynomial form. Each term of this polynomial is a six-dimensional function given by:

$$q_i = \prod_{j=1}^6 L_i^{j_i} \quad \text{and} \quad \sum_{j=1}^6 j_i \leq 3 \quad i = 1, 2, \dots, 6 \quad (15)$$

where $j_i = 1, 2$, and 3 is a power index. The number of predictors generated by Eq. (15) is 84. Six optimization problems are then addressed to model the translational $\{x(t), y(t), z(t)\}$ and rotational $\{\varphi(t), \theta(t), \psi(t)\}$ motions in terms of joint space variables $\{L_i(t), i = 1, 2, \dots, 6\}$. Genetic algorithms then propagate the search for optimal SF for each motion. Since 2^{84} possible structures are assigned to be candidates for the optimization search, the length of each chromosome is coded by 84 bits. For the six optimization problems, the mutation rate is 20%. Each generation has a fixed population size 20 or no generation overlap. The optimization code was run many times with different starting guess (initial solution) and it converges to the same solution every time. The algorithm is highly convergent for all the optimization problems. The average number of generations for convergence is 3500. The resultant optimal regression structure is listed in Appendix A.

Table 1. Manipulator parameters used in the simulation.

Variable	Description	Value	Unit
L_u	Length of upper leg	0.95	m
L_l	Length of lower leg	0.95	m
R_u	Radius of the movable plate	0.4	m
R_l	Radius of base plate	1	m
$\bar{\alpha}$	Joint angles of base plate	$[-50, 50, 70, 170, -170, -70]$	deg
$\bar{\beta}$	Joint angles of upper plate	$[-2, 2, 118, 122, -122, -118]$	deg

Table 2 shows the computational expenses of the model listed in Appendix A. The computational expense of each expression is computed as the number of terms and the number of multiplication operations. For example, the yaw expression has 30 terms: 21 cubic/tri-linear terms, 8 quadratic/bi-linear terms, and one linear term. The total number of the multiplication in this expression is 80 ($21 \times 3 + 8 \times 2 + 1$) operations. For the full Taylor expansion [12], the total number of terms is 84: 56 cubic/tri-linear terms, 21 quadratic/bi-linear terms, 6 linear terms, and 1 constant term. The total number of multiplication is then 216 operations. Such comparison proves that the proposed technique has the capability to deliver an optimal forward kinematics model for Stewart manipulator. Herein, the

optimality is defined in terms of saving memory and computation expenses. The yaw motion example indicates that there is 65% saving in term of terms and 63% saving in number of multiplications. This example is the worst in the overall model (see Table. 2).

Table 2. Computational expenses of the proposed model in the form of number of terms and number of multiplications for each expression

	Surge	Sway	Heave	Roll	Pitch	Yaw
Number of terms	24	24	25	26	22	30
Number of multiplications	62	66	67	65	48	80

Two tracks are selected to test the model adequacy. The first track is a smooth movement from an initial point $\{x = 10 \text{ cm}, y = -10 \text{ cm}, z = -8 \text{ cm}, \varphi = 10 \text{ deg}, \theta = -10 \text{ deg}, \text{ and } \psi = -10 \text{ deg}\}$ to a final point $\{x = -10 \text{ cm}, y = 10 \text{ cm}, z = 8 \text{ cm}, \varphi = -10 \text{ deg}, \theta = 10 \text{ deg}, \text{ and } \psi = 10 \text{ deg}\}$. The time horizon of this track is 20 sec. This track is assigned to have zero velocities and accelerations for the initial and final points. After the track is generated, the inverse kinematics is then used to find the equivalent joint displacements $\{L_i(t), i = 1, 2, \dots, 6\}$ as shown in Fig. (4). The second track is a sinusoidal movement. The task displacement amplitudes of these sinusoidal waves are $A_x = 15 \text{ cm}, A_y = 15 \text{ cm}, A_z = 10 \text{ cm}, A_\varphi = 10 \text{ deg}, A_\theta = 10 \text{ deg}, \text{ and } A_\psi = 5 \text{ deg}$. The resultant joint displacements of the second track are shown in Fig. (5). The two tracks have been selected to cover a wide range of displacements and velocities inside the workspace of the manipulator. The task space displacements $\{x(t), y(t), z(t), \varphi(t), \theta(t), \text{ and } \psi(t)\}$ for the first and second tracks are now reproduced by the model listed in Appendix A. Figures (6) and (7) show a comparison between the original track and the simulation results using the proposed model.

Figure (8) shows the variation of the percentage error for the two tracks with time. The results indicate that the percentage error is within $\pm 4\%$. The pitch motion has the worst error signal in the first track while the roll motion has the worst error signal in the second track. However, both of them are still inside the acceptable level of the error (less than 5%). The influence of the amplitude and the frequency variations on the error signal for the second track is shown in figure (9). The simulation results show that the roll motion has the highest maximum error for all amplitudes and frequencies. The percentage of the maximum error (ME) for position and velocity roll motion is used as an indicator for the error. There is insignificant variation of ME % with frequency. Such result is expected since the model is a kinematic one. the model delivers the task space displacements as functions of the joint space displacements without any derivatives (see Appendix A). Increasing the amplitude increases the percentage of the maximum error with a quadratic shape.

Table (3) summarizes the statistics of errors from the approximate model in the two test cases. MSE is the mean squared error. ME is the maximum error during the track. The accuracy of the model is tested for both position and velocity since the accuracy in the velocity should be considered for a controller designing purpose (PD controller). The maximum position error in the first track

for sway motion is 0.28 cm and for the pitch motion is 0.41 deg. On the other hand, the maximum velocity error that appeared in sway motion is 0.23 cm/s and in the pitch motion is 0.34 deg/sec. This is sufficiently adequate for building and designing a task space control for the Stewart manipulator. Moreover, the model has a simple computational manipulation.

The authors previously developed the workspace for the same manipulator parameters in reference [4]. The workspace indicates that the heave motion limits are 50 cm

downward and 65 cm upward, if all the rest of motions have zero values. A new trajectory operating close to these boundaries (singularity-nearby) is generated in Fig. 10-a. This test case investigates the singularity-nearby performance. Figure 10-b shows that the accuracy level is within the accepted range, however as the manipulator heads to the limit values as the error increases. All these observations emphasize the applicability of the proposed model to be imbedded in the control loop more efficiently than other approaches.

Table 3. Statistical Comparison between the first and the second tracks: MSE refers to the mean squared error during the track and ME refers to the maximum error during the track

	Position MSE		Velocity MSE		Position ME		Velocity ME	
	Track 1	Track 2	Track 1	Track 2	Track 1	Track 2	Track 1	Track 2
Surge	1.03 cm^2	0.81 cm^2	$0.04 \text{ cm}^2/\text{s}^2$	$0.05 \text{ cm}^2/\text{s}^2$	0.13 cm	0.17 cm	0.09 cm/s	0.08 cm/s
Sway	1.12 cm^2	1.45 cm^2	$0.15 \text{ cm}^2/\text{s}^2$	$0.52 \text{ cm}^2/\text{s}^2$	0.28 cm	0.27 cm	0.04 cm/s	0.23 cm/s
Heave	1.25 cm^2	0.86 cm^2	$0.22 \text{ cm}^2/\text{s}^2$	$0.56 \text{ cm}^2/\text{s}^2$	0.11 cm	0.04 cm	0.03 cm/s	0.20 cm/s
Roll	0.21 deg^2	1.09 deg^2	$0.19 \text{ deg}^2/\text{s}^2$	$0.78 \text{ deg}^2/\text{s}^2$	0.11 deg	0.39 deg	0.11 deg/s	0.28 deg/s
Pitch	2.06 deg^2	2.00 deg^2	$0.96 \text{ deg}^2/\text{s}^2$	$1.89 \text{ deg}^2/\text{s}^2$	0.41 deg	0.33 deg	0.12 deg/s	0.34 deg/s
Yaw	1.05 deg^2	0.46 deg^2	$0.45 \text{ deg}^2/\text{s}^2$	$0.32 \text{ deg}^2/\text{s}^2$	0.30 deg	0.14 deg	0.06 deg/s	0.13 deg/s

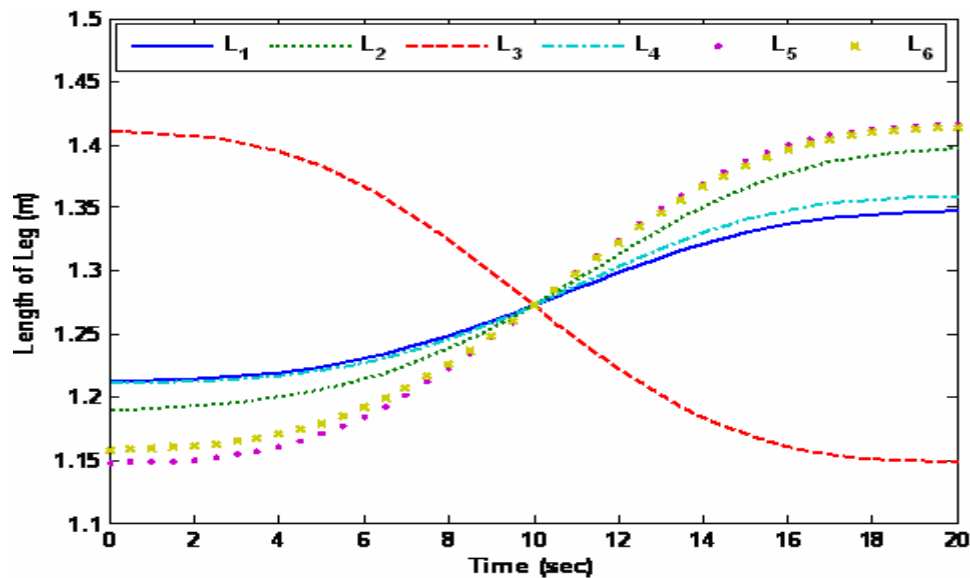


Figure 4. The time history of the joint displacements for the first track.

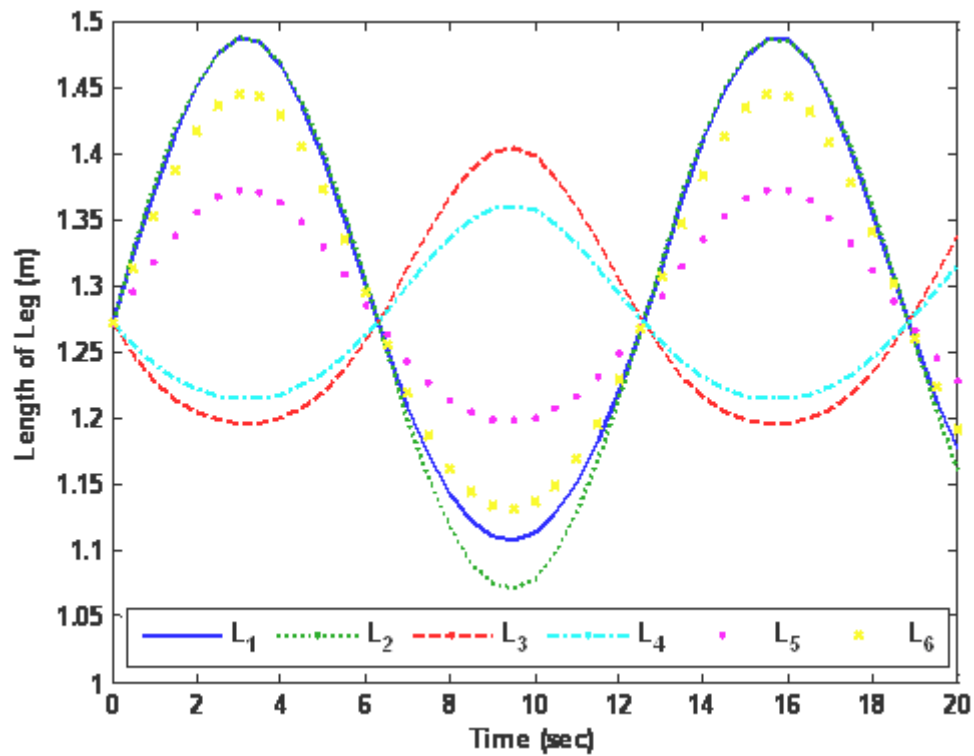
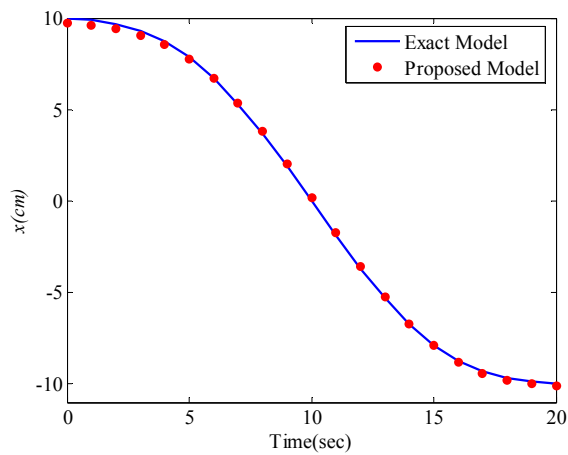
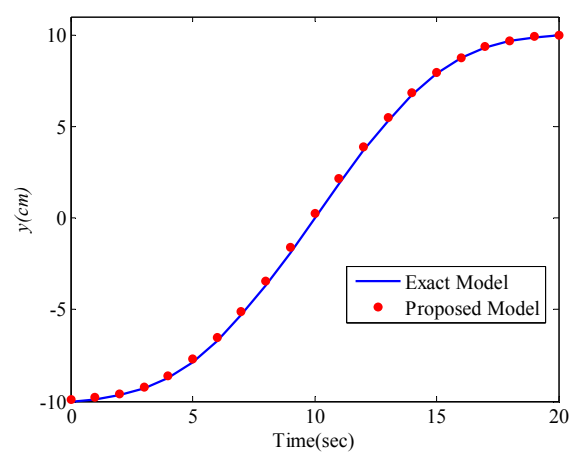


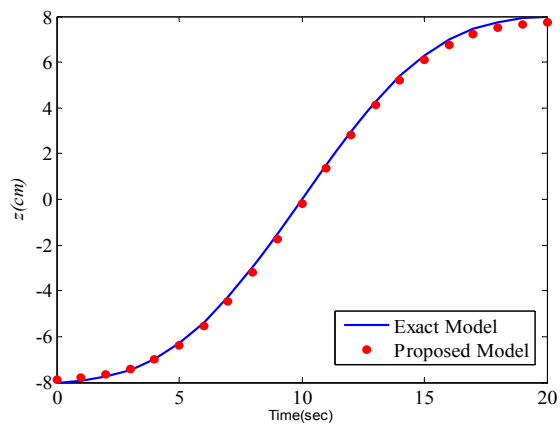
Figure 5. The time history of the joint space displacements for the second track.



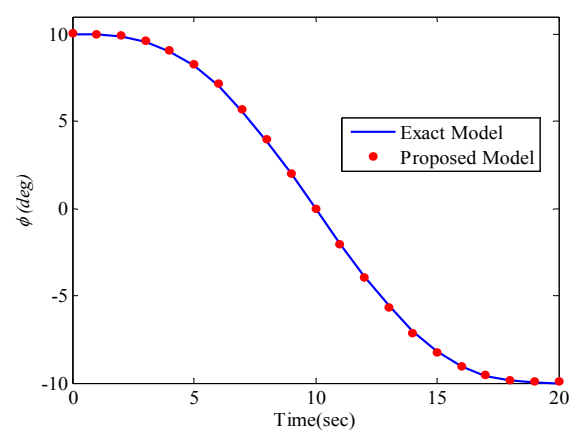
a) Surge motion of the first track



b) Sway motion of the first track

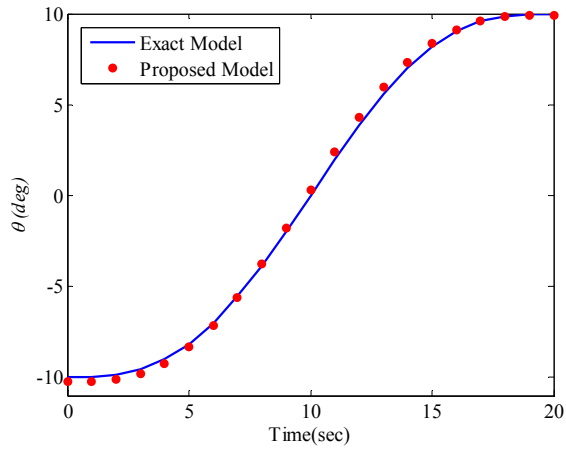


c) Heave motion of the first track

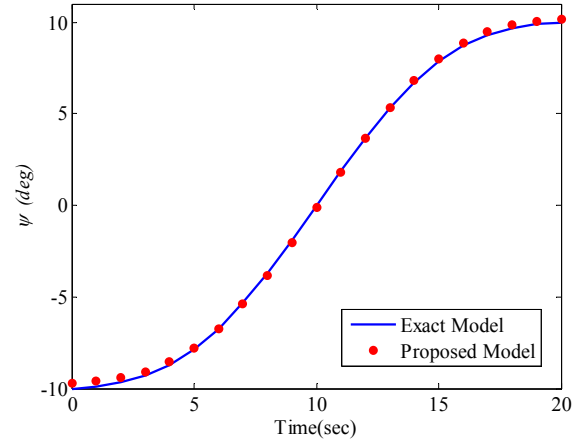


d) Roll motion of the first track

Figure 6. continues next page.....

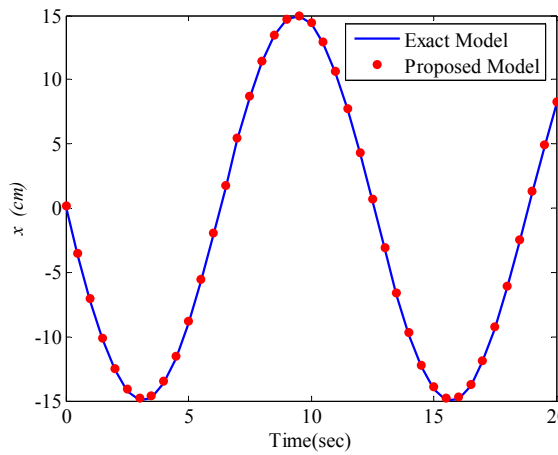


e) Pitch motion of the first track

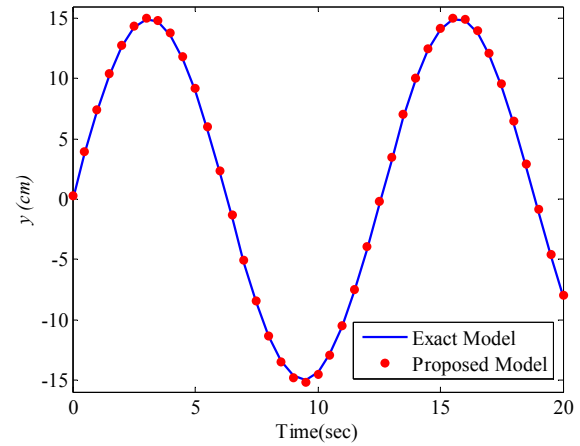


f) Yaw motion of the first track

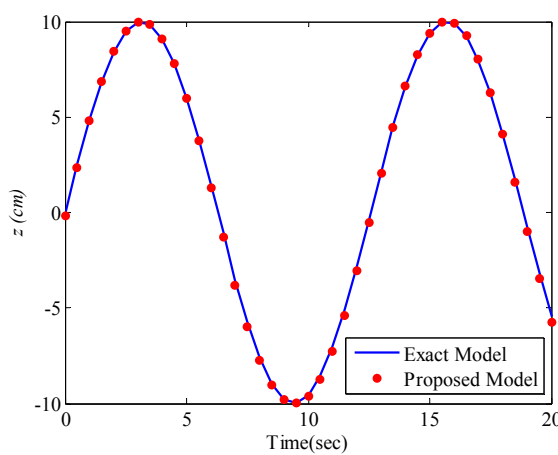
Figure 6. Comparison between the exact model and the results from the proposed approximate model for the first track: smooth motion between two different points.



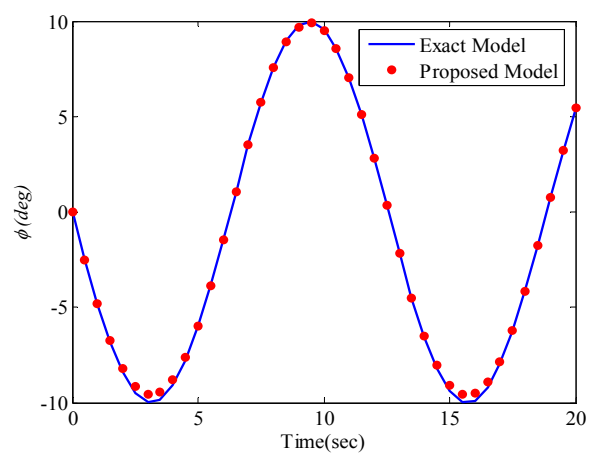
a) Surge motion of the second track



b) Sway motion of the second track

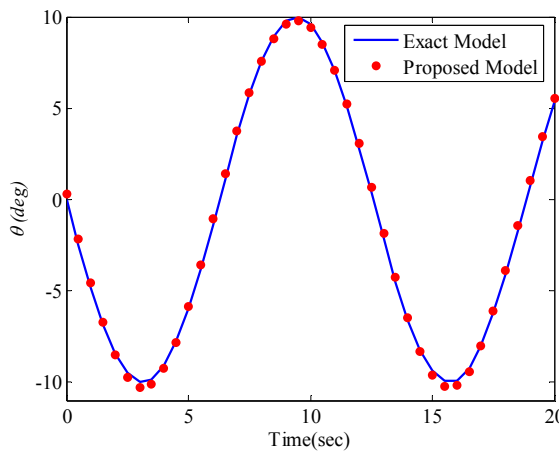


c) Heave motion of the second track

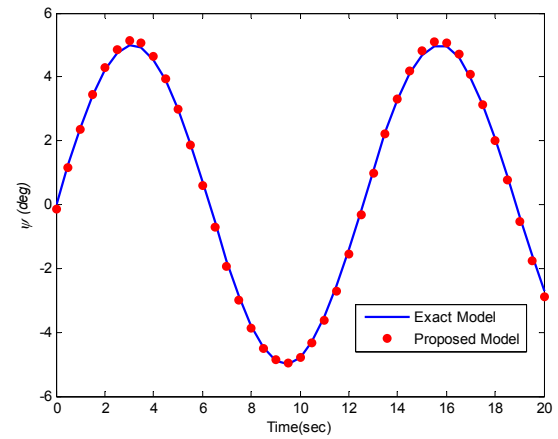


d) Roll motion of the second track

Figure 7. continues next page...

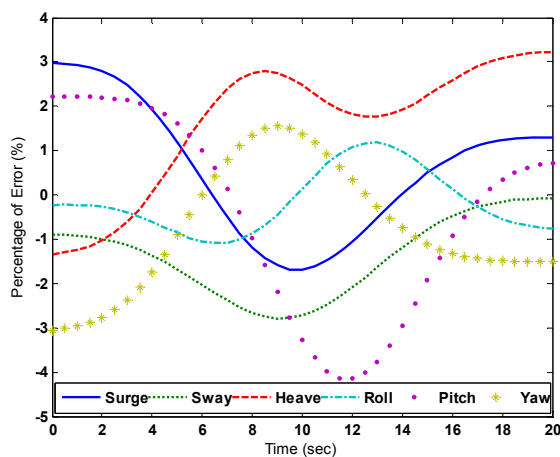


e) Pitch motion of the second track

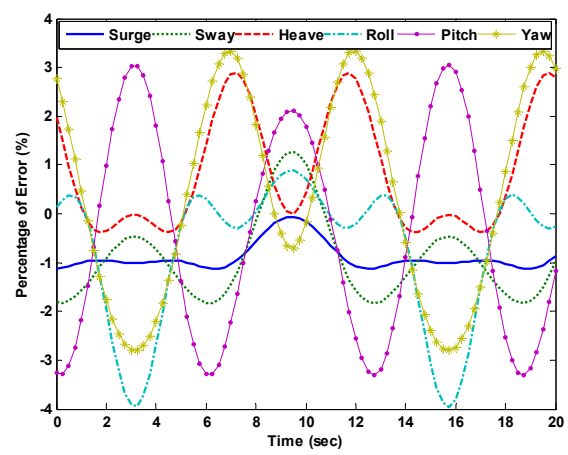


f) Yaw motion of the second track

Figure 7. Comparison between the exact model and the results from the proposed approximate model for the second track: sinusoidal motion with frequency 0.5 rad/s.

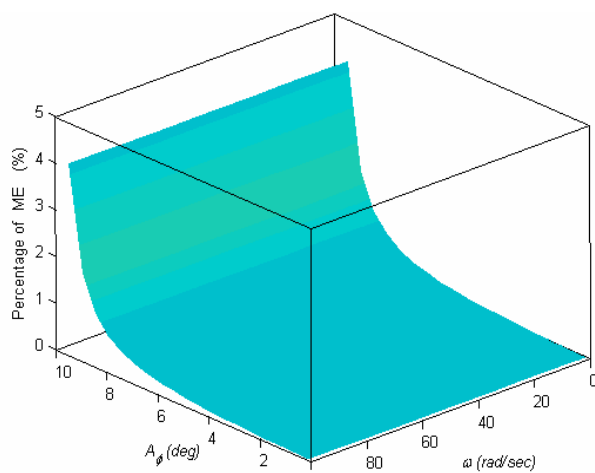


a) First track

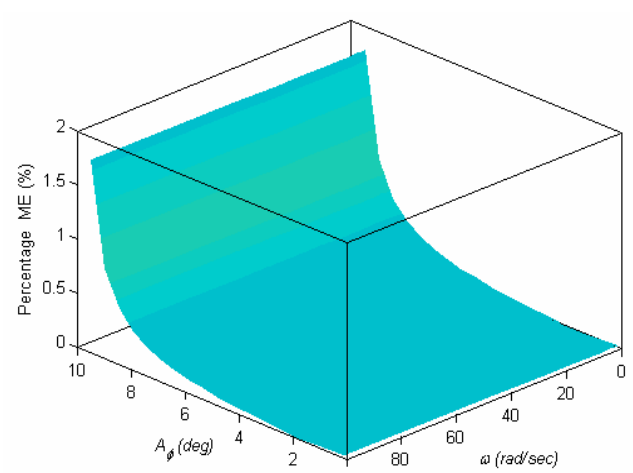


b) Second track

Figure 8. The variation of the error between the exact and the proposed approximate model.



a) Position Error



b) Velocity Error

Figure 9. The variation of the percentage maximum error (ME) for the roll motion and velocity with amplitude and frequency.

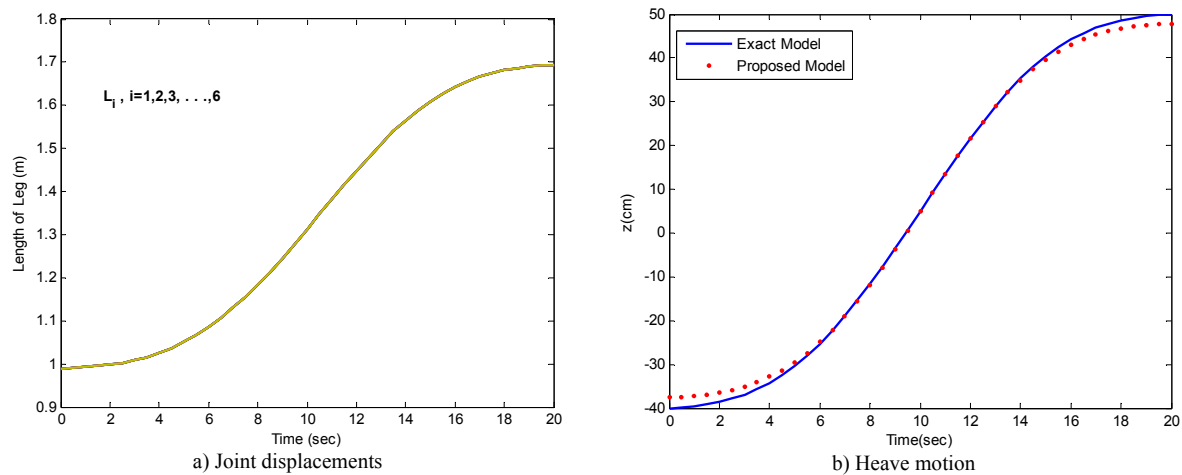


Figure 10. The performance of the proposed model in the singularity-nearby vicinity.

6. Conclusion

In this paper, we propose a new procedure to model the Stewart manipulator forward kinematics. The modeling procedure is based on predicted squared error cost function. This cost function is a trade-off between a minimum error and a minimum number of terms in the regression model. To demonstrate the technique feasibility, a simulation by numerical example has been developed. 2^{84} candidate regression structures were selected to model the forward kinematics. The structures' predictors were generated from a third order polynomial in a six-dimensional space. Using a genetic algorithm as a global optimization technique has successfully solved the six optimization problems. In this way, a simple structure of forward kinematics for the Stewart manipulator was developed. The model has the capability to reduce the computational expense by at least 65% from the traditional Taylor expansion previously investigated. Such reduction increases the suitability for the model in the real time control. Thus, most of previous in literature models seek the computational expenses, which restrict their practical ability. The model was tested by two generic tracks providing high precision results with maximum percentage error less than 5%. All these observations emphasize that the proposed model is quite applicable for a task space control.

References

- [1] I. Nag, W. Chong, "High speed tracking control of stewart platform manipulator via enhanced sliding mode control". IEEE International Conference on Robotics & Automation, Leuven, Belgium, 1998, 2716-2721.
- [2] C. Fang, C. Hung-Hsiang, L. Chin-Teng, "Fuzzy control of a six-degree motion platform with stability analysis". IEEE SMC Conference, Vol.11, 1999, 325-330.
- [3] Y. Su, Y. Duan, C. Zheng, Y. Zhang, G. Chen, J. Mi, "Disturbance-rejection high-precision motion control of a Stewart platform," IEEE transactions on control systems technology, Vol. 12, 2004, 364-374.
- [4] A. Omran, G. El-Bayoumi, M. Bayoumi, A. Kassem, "Genetic algorithm-based optimal control for a 6-DOF non redundant stewart manipulator". International Journal of Mechanical Systems Science and Engineering, Vol. 2, No. 2, 2008, 73-79.
- [5] A. Omran, A. Kassem, G. El-Bayoumi, M. Bayoumi, "Mission-based optimal control of stewart manipulator". Aircraft Engineering and Aerospace Technology Journal, Vol. 81, No. 3, 2009, 226-233.
- [6] P. Nanua, W. Kenneth, M. Vasudeva, "Direct kinematics solution of a stewart platform". IEEE Trans. Robot. Automat. Vol. 6, 1990, 438-444.
- [7] J. Pratik, Y. Sarah, "A hybrid strategy to solve the forward kinematics problem in parallel manipulators". IEEE Trans. Robot. Automat. Vol. 21, 2005, 18-25.
- [8] B. Ilian, R. Jeha, K. Sung-Gaun, L. Sun-Kyu, "A closed-form solution to the direct kinematics of nearly general parallel manipulators with optimally located three linear extra sensors". IEEE Trans. Robot. Automat, Vol. 17, 2001, 148-156.
- [9] D. Li, S. Salcudean, "Modeling, simulation, and control of a hydraulic stewart platform". IEEE Int. Conf on Robotics and Automation, Albuquerque, 1997, 3360-3366.
- [10] S. Fu, Y. Yao, "Comments on "A Newton-Euler formulation for the inverse dynamics of the stewart platform manipulator," Mech. Mach. Theory, Vol. 8, 2006, 1-3.
- [11] P. Dietmaier, "The Stewart-Gough platform of general geometry can have 40 real postures". In Advances in Robot Kinematics: Analysis and Control, J. Lenarcic and M. L. Hust, Eds. Norwell, MA: Kluwer, 1998, 7-16.
- [12] H. Sadjadian, H. Taghirad, "Comparison of different methods for computing the forward kinematics of a redundant parallel manipulator". Journal of Intelligent and Robotic Systems, Vol. 44, No. 3, 2005, 225-246.
- [13] H. Sadjadian, H. Taghirad, A. Fatehi, "Neural networks approaches for computing the forward kinematics of a redundant parallel manipulator". International Journal of Computational Intelligence, Vol. 2, No. 1, 2006, 40-47.
- [14] K. Liu, M. Fitzgerald, F. Lewis, "Kinematic analysis of a Stewart platform manipulator". IEEE Trans. Industrial Electronics, Vol. 40, No. 2, 1993, 282-293.
- [15] K. Liu, F. Lewis, F., M. Fitzgerald, "Solution of nonlinear kinematics of a parallel-link constrained Stewart platform manipulator". Circuits, Systems, and Signal Proc., Special Issue on "Implicit and Robust Systems", Vol. 13, No. 2 1994, 167-183.

- [16] L. Sheng , L. Wan-long, D. Yan-Chun, F. Liang, "Forward kinematics of the Stewart platform using hybrid immune genetic algorithm". Proceedings of the IEEE: International Conference on Mechatronics and Automation, 2006, 2230-2235.
- [17] M. Tarokh, "Real time forward kinematics solution for general Stewart manipulator". IEEE International Conference on Robotics and Automation, 2007, 901-906.
- [18] S. Chen, C. Huang, L. Fu, "Applying a nonlinear observer to solve forward kinematics of a Stewart platform". 17th IEEE International Conference on Control Applications, 2008, 1183-1188.
- [19] Barron R. Predicted squared error: A criterion for automatic model selection. Self-Organization Methods in Modeling arrow, S.J., ED., Marcel Dekker, Inc., New York; 1984.
- [20] Holland J. Adaptation in natural and artificial systems. The University of Michigan Press; 1975.
- [21] Goldberg D, David E. The design of innovation: lessons from and for competent genetic algorithms. Boston, MA: Kluwer Academic Publishers; 2002.
- [22] J. John, "Optimization of control parameters for genetic algorithms". IEEE Trans on System, Man, and Cybernetics, Vol. 16, No. 1, 1986, 566-574.

APPENDIX A: Resultant Model

$$\begin{aligned}
 x(L_1, L_2, L_3, L_4, L_5, L_6) = & \\
 & 0.229L_3L_6^2 - 0.049L_4L_6^2 - 0.167L_2L_6^2 - 0.341L_1L_6^2 - 0.048L_1L_4 - 0.233L_1^2L_3 + 0.303L_3L_6 - 0.281L_3^2L_6 \\
 & + 0.357L_1L_3^2 - 0.317L_1L_2 - 0.074L_2L_3 - 0.032L_1L_6 - 0.003L_3L_4^2 + 0.046L_1^2L_5 - 0.014L_2L_4L_5 - 0.003L_2^3 \\
 & + 0.534L_1L_2L_6 - 0.003L_2L_6 - 0.049L_1L_2L_5 + 0.063L_6 + 0.089L_6^3 - 0.0719L_5 + 0.044L_3^2 - 0.047L_1L_3L_4
 \end{aligned}$$

$$\begin{aligned}
 y(L_1, L_2, L_3, L_4, L_5, L_6) = & \\
 & 0.345L_2L_4 + 0.006L_1L_2^2 + 0.125L_3^2L_4 + 0.459L_1L_6 + 0.282L_1^2L_3 + 0.143L_5L_6^2 + 0.381L_5L_6 + 0.309L_1L_2L_5 \\
 & - 0.685L_1L_4 + 0.051L_1L_4^2 - 0.316L_6 + 0.076L_5^3 - 0.005L_4L_5L_6 - 0.256L_2L_3L_6 - 0.454L_1L_2L_3 + 0.362L_1L_2L_6 \\
 & - 0.187L_2L_4L_6 + 0.175L_3L_3L_5 - 0.545L_1L_5L_6 + 0.194L_1L_2L_4 - 0.049L_1^2 - 0.194L_1^2L_2 - 0.104L_4^3 - 0.124L_3^3
 \end{aligned}$$

$$\begin{aligned}
 z(L_1, L_2, L_3, L_4, L_5, L_6) = & \\
 & -1.210 + 0.175L_2L_4^2 - 0.151L_2^2L_4 - 0.405L_4^2L_6 - 0.065L_3L_6^2 - 0.407L_2L_3^2 + 0.072L_3^2L_4 + 0.282L_3L_6 + 0.002L_4^3 \\
 & + 0.413L_4L_5 - 0.229L_3^2L_6 + 0.076L_1L_6^2 - 0.294L_1L_5^2 + 0.049L_4L_6^2 - 0.222L_2L_6^2 + 0.505L_5^2L_6 + 0.252L_2L_3L_6 \\
 & + 0.290L_2L_6 - 0.436L_5L_6^2 + 0.739L_1L_2L_5 - 0.352L_1^2L_2 - 0.106L_1L_4L_5 + 0.039L_1L_3L_6 + 0.212L_1^2 + 0.025L_4^2
 \end{aligned}$$

$$\begin{aligned}
 \varphi(L_1, L_2, L_3, L_4, L_5, L_6) = & \\
 & -0.187L_4^2L_6 - 0.125L_1^2L_5 + 0.741L_1L_4L_6^2 + 0.582L_1^2L_6 - 0.204L_2^2L_3 - 0.048L_2L_6^2 + 0.435L_3L_5 + 1.46L_3L_4 \\
 & - 0.293L_3L_6^2 + 0.040L_1L_6^2 - 0.374L_1L_3^2 - 1.893L_1L_6 - 0.013L_2L_5^2 - 0.088L_1^2L_4 + 0.170L_4L_5^2 - 0.267L_4L_5L_6 \\
 & - 0.289L_3L_5L_6 - 0.541L_3L_4L_5 + 0.278L_2L_3L_5 - 0.563L_1L_4L_6 + 0.434L_1L_4L_5 + 0.863L_1L_3L_6 + 1.433L_2 \\
 & - 0.216L_1^2 - 0.027L_6^3 - 1.298L_4
 \end{aligned}$$

$$\begin{aligned}
 \theta(L_1, L_2, L_3, L_4, L_5, L_6) = & \\
 & 0.119L_2L_6^2 + 0.201L_2L_5 - 0.093L_4L_6^2 - 0.068L_1L_5^2 - 0.006L_5L_6 + 0.328L_3^2L_5 + 0.027L_4L_5^2 - 0.843L_1L_4^2 \\
 & - 0.155L_1L_2^2 - 0.268L_3L_4 - 0.201L_1L_3 + 1.445L_1L_4 - 0.982L_2L_3 + 0.135L_5^2L_6 - 1.063L_6 + 2.091L_5 \\
 & + 0.564L_1L_3L_4 - 0.538L_5^2 - 1.131L_1 - 0.088L_3^2 + 0.616L_4^2 - 0.083L_3L_5
 \end{aligned}$$

$$\begin{aligned}
 \psi(L_1, L_2, L_3, L_4, L_5, L_6) = & \\
 & -0.398L_2L_3^2 + 0.249L_2L_4 + 0.199L_1^2L_3 + 0.042L_2L_4^2 - 0.065L_1^2L_5 - 1.252L_1L_2^2 + 0.928L_2L_6 + 0.267L_1^2 \\
 & + 0.160L_4^3 - 0.598L_1L_3 - 0.102L_2L_5 + 0.513L_1L_6^2 + 1.400L_1^2L_2 + 0.212L_1L_2L_4 - 0.767L_1^2L_6 - 0.454L_1L_4^2 \\
 & + 0.545L_2^3 + 0.162L_4^2 - 0.089L_1^2L_4 - 0.506L_1L_2 - 0.406L_3L_6 + 0.525L_2L_4L_5 - 0.336L_2^2L_6 - 0.150L_4L_5^2 \\
 & + 0.113L_5^2L_6 - 0.556L + 0.350L_3L_1^2 - 0.505L_2L_4L_6 + 0.264L_2L_3L_6 - 0.439L_1L_5L_6
 \end{aligned}$$

An Intelligent Machine Condition Monitoring System Using Time-Based Analysis: Neuro-Fuzzy Versus Neural Network

M. Samhouri ^{a,*}, A. Al-Ghandoor ^a, S. A. Ali ^a, I. Hinti ^b, W. Massad ^a

^a Department of Industrial Engineering, College of Engineering, Hashemite University, P.O. Box 150459, 13115 Zarka, Jordan

^b Department of Mechanical Engineering, College of Engineering, Hashemite University, P.O. Box 150459, 13115 Zarka, Jordan

Abstract

Monitoring and predicting machine components' faults play an important role in maintenance actions. Developing an intelligent system is a good way to overcome the problems of maintenance management. In fact, several methods of fault diagnostics have been developed and applied effectively to identify the machine faults at an early stage using different quantities (Measures or Readings) such as current, voltage, speed, temperature, and vibrations. In this paper, an intelligent machine condition monitoring and diagnostic system is introduced with experimental verification. An adaptive neuro-fuzzy inference system (ANFIS) and a neural network system (NN) are used to monitor and predict the fault types of a critical mechanical element in the Potash industry (namely; a Carnallite surge tank pump). The system uses a piezoelectric accelerometer to generate a signal related to machine condition and fault type. Combinations of the vibration time signal features (i.e., root mean square, variance, skewness, kurtosis, and normalized sixth central moment) are used as inputs to both ANFIS and neural nets, which in turn output a value for predicted fault type. Experimental validation runs were conducted to compare the actual fault types with the predicted ones. The comparison shows that the adoption of the time root mean square and variance features achieved the minimum fault prediction errors for both ANFIS and neural nets. In addition, trapezoidal membership function in ANFIS achieved a fault prediction accuracy of 95%, whereas, a cascade forward back-propagation neural network achieved a better fault prediction accuracy of 99%.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Condition Monitoring; Time Analysis; Neuro-Fuzzy.

1. Introduction

In recent years, with the rapid development of condition monitoring and forecasting, information processing, fault detecting, and artificial intelligence technology, it has been possible and feasible to monitor and forecast equipment condition and assess its health online. It is well recognized that optimized maintenance practices within an industrial setting require the correct blend of maintenance strategies. Condition-based (reliability centered, predictive, proactive) maintenance is an important part of this blend for many compelling reasons [1].

Recently, there has been a significant amount of research effort directed towards developing and implementing useful automated machinery fault detection and diagnostic tools. Most of these tools have been based on various pattern-recognition schemes, knowledge-based systems or artificial neural networks systems. The main thrust of the work has been towards developing systems

that are not only objective in their treatment of data and presentation of results, but also flexible, thereby being applicable in a wide range of situations. A new method using fuzzy logic techniques to improve the performance of the classical inductive learning approach was presented by [2]. In [2], a hard cut point was proposed to discretize the continuous-valued attributes by using soft discretization to enable the systems have less sensitivity to noise. In [3], they used the concept of the fuzzy fractal dimension to measure the complexity of a time series of observed data from the plant. A method for analyzing and forecasting field failure data for repairable systems was proposed by [4]. This novel method constructed a predictive model by combining the seasonal autoregressive integrated-moving average (SARIMA) method and neural network model. In [5], they introduced a new combined method based on wavelet transformation, fuzzy logic, and neuro-networks for fault diagnosis of a triplex.

* Corresponding author. Samhouri@hu.edu.jo.

Most real life physical systems are nonlinear, ill-defined and uncertain which makes them difficult to model by conventional mathematical means. Furthermore, most industrial processes are based on the assumption that the process is a linear system. Fuzzy logic and neural networks have the potential to deliver successful solutions to problems that have previously proved difficult or impossible to handle by conventional linear methods [6-9]. Fuzzy logic and fuzzy inference systems have been shown to be effective techniques for the identification and prediction of complex, nonlinear, and vague systems. Fuzzy logic is particularly attractive due to its ability to solve problems in the absence of mathematical models.

Condition-based maintenance (CBM) is a method used to reduce the uncertainty of maintenance activities, and is carried out according to the need indicated by the equipment condition [10]. CBM assumes that existing indicative prognostic parameters can be detected and used to quantify possible failure of equipment before it actually occurs. Prognosis parameters provide the indication of potential problems and incipient faults which would cause the component or equipment to deviate from the acceptable performance level [11]. A number of computational tools have been developed for condition-based maintenance such as knowledge base [12-13], analytic hierarchy process [14-15], Petri nets [16], neural networks [17-18], and fuzzy logic and networks [19-20].

This study was motivated by the problem of improper condition-based maintenance strategy and a need of a process industry (Potash production) for a fast enough machine monitoring system to be employed as real-time fault detection system at their plants. In this paper, the maintenance records analysis is used to provide critical information from past experience to improve current maintenance process in the Potash industry. A neuro-fuzzy technique (i.e., ANFIS) and a neural network of time-based analysis are used to build an intelligent condition monitoring system to predict the type of fault or failure for one of the critical production units. Different combinations of five statistical parameters computed from the vibration time signal of a critical pump were fed as inputs into both ANFIS and neural network to output a value for the predicted fault. The procedure is illustrated using the vibration data of a carnallite surge tank pump with normal and faulty pump.

This paper is organized as follows: Section 1 is an introduction. Intelligent condition monitoring and fault diagnosis system will be given in the second section. Section 3 is about vibration data analysis and feature selection. Structure of adaptive neuro-fuzzy inference system and neural network will be discussed in section 4. ANFIS and neural-networks time-based fault diagnosis system is the subject of section 5. The last section is to conclude.

2. Intelligent Condition Monitoring and Fault Diagnosis System.

Condition monitoring is becoming popular in industry because of its efficient role in detecting potential failures. The use of condition monitoring techniques will generally improve plant production availability, and reduce downtime cost, but in some cases it also tends to

over-maintain the plant in question. If a hidden defect is already present, with the help of condition monitoring, the defect may be identified and corrective actions may be taken. It is noted however that for a cost-effective maintenance, advanced prediction of such a defect and its development is very important since ordering spare parts and possible production shutdowns for maintenance may be costly and require careful planning well before the failure actually occurs.

Condition monitoring traditionally means acquiring data from various classes of plant which gives an indication of the condition of machine. Condition monitoring is an essential element of predictive maintenance. An ideal condition monitoring system would accept measured data as input and will produce the operational status, a possible mode of failure and time to failure as output.

Many machinery fault diagnostic techniques use automatic signal classification in order to increase accuracy and reduce errors caused by subjective human judgment. Detection of machine faults like mass imbalance, rotor rub, shaft misalignment, gear failure, and bearing defects is possible by comparing the vibration signals of a machine operating with and without faulty conditions. These signals can also be used to detect the incipient failures of the machine components through online monitoring system, reducing the possibility of catastrophic damage.

In intelligent maintenance management systems, IMMS, the three "isolated islands" of the automation system (i.e., monitoring and forecasting, diagnosis and prognosis, and maintenance decision making) are integrated into an organic system, and maintenance improved by sharing information among these systems. Intelligent methods try to decode the intelligence supplied from the system. Artificial intelligence techniques such as expert systems, neural networks, genetic algorithms, and fuzzy logic, have been widely applied in mechanical equipment monitoring and diagnosis with different aspects and degrees. It also noticed that different techniques have their unique advantages and disadvantages, and usually cannot replace each other.

The main problem of the Potash industry (Arab Potash Company, APC) is the visual inspection of frequency analysis performed at the preventive maintenance department, and more importantly the low speed of the automatic frequency-based monitoring system which is no longer suitable for real-time applications. Although they have a very progressive maintenance software system used to assent the conjuncture of their machines, as an attempt to control the break downs and the health conditions of the machines, the visual inspection by maintenance staff results in big discrepancies between the predicted and actual faults, thus causing time delays, inconsistencies, increased inaccurate maintenance activities, and of course increased loss of money. The carnallite surge tank pump is one of the important components that undergoes repetitive failures which cause breakdowns for the process, Figures 1 shows one of the carnallite surge tank pumps.

To solve this problem, an adaptive neuro-fuzzy inference system (ANFIS) as well as a neural network was used in this paper to replace the human operator and the automatic frequency-based system for predicting the faults

types from the original time signal. The critical component (namely; the carnallite surge tank pump) was selected to apply our intelligent fault diagnosis system because this component has faced many unsuccessfully-predicted breakdowns and failures, which resulted in unnecessary money loss. The carnallite pump, shown in Figure 1, is considered a critical component because its breakdown could easily cause a production shut down or delay. A

special software (i.e., EMONITOR Odyssey Delux) is used at this Potash industry to display and manipulate the vibration time signals coming from a piezoelectric accelerometer mounted on the pump component. As given in Figure 2, the frequency pattern of the original time signal is clear, but visual and automatic inspection and prediction usually results in an inaccurate diagnosis of the fault type.



Figure 1. The carnallite surge tank pump.

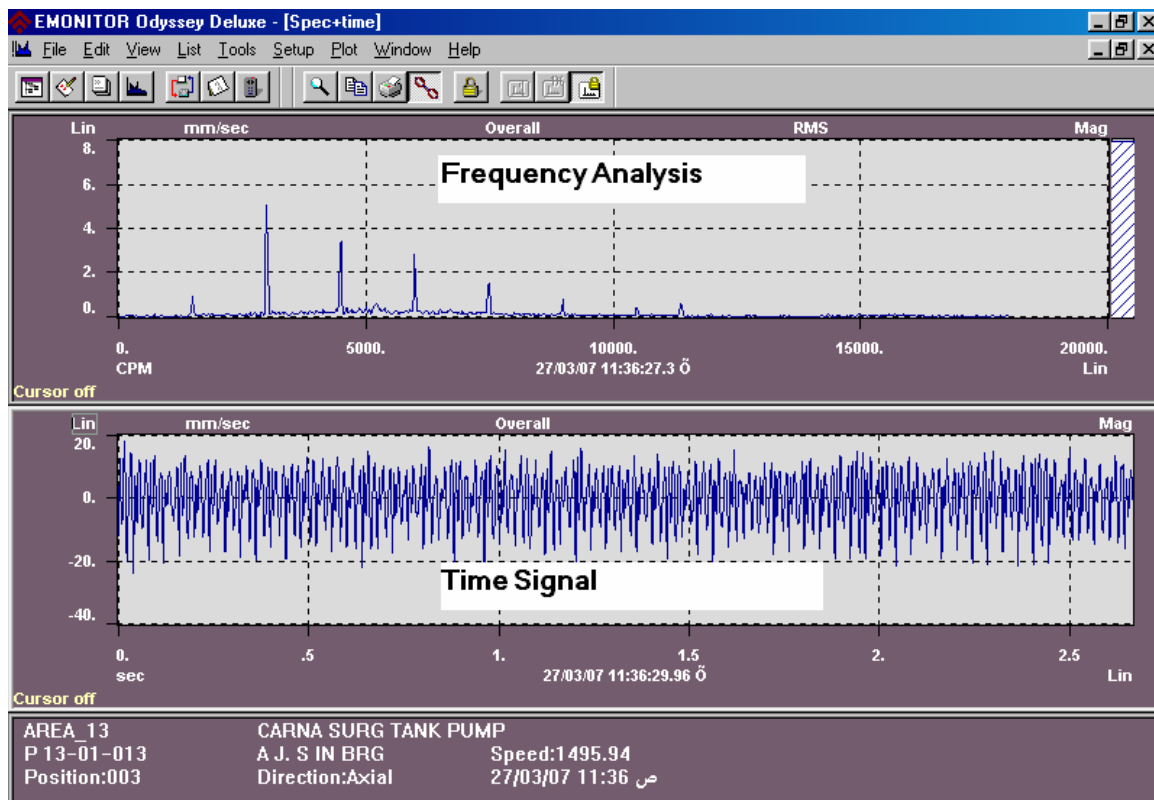


Figure 2. A sample of the vibration time signal in the axial direction along with its frequency-domain analysis using Odyssey software.

3. Vibration Data Analysis and Feature Selection

The traditional way of observing signals is to view them in what is called the time domain. The time domain is a record of what happened to a parameter compared to time. Typically, the signal would be displayed on an oscilloscope or a computer screen as given in Figure 2. In the analysis of time series signals, certain restrictions are imposed by the length of the data window (T), being analyzed and by the sampling rate (fs), used when digitizing continuous data [21]. A sample of time series segment of length $T = 2.8$ sec is shown in Figure 2. This is a standard time duration used at the Potash production plant to pick up useful vibration signals for frequency analysis.

Machines faults diagnosis and prediction requires generating representative and useful information about the vibration features by means of a sensor. Our approach to predict the fault type is to mount a piezoelectric accelerometer on the machine's component under study in order to give a time-series signal which is supposed to contain useful information about the machine's faults, failures and health conditions. The data used in building intelligent maintenance system has undergone several processing and analysis steps which will be described briefly in this section.

The first step is the vibration measurement by using sensor as shown in Figure 3; this sensor is a piezoelectric accelerometer. Accelerometers are absolute vibration transducers which produce a signal proportional to the vibration acceleration. The piezoelectric accelerometer is most attractive in view of its rigidity, wide frequency range, flat response and dynamic range, this sensor has the ability to measure the vibration in the three dimensions (namely; axial, horizontal, and vertical).

As an example, it is important to note that carrying out measurements on bearings readings should be taken in both radial and axial planes. Using both planes is an important method for distinguishing between various mechanical faults.

The piezoelectric accelerometer is connected with data base collection device which in turn is connected with a computer that has the analysis software (i.e., Emonitor odyssey deluxe) as shown in Figure 3, which in turn applies preliminary signal automatically on the vibration time signals (e.g., low-pass filtering, and windowing) and transfers the vibration from time domain to frequency domain. Figure 3 shows the Graphical User Interface for data transfer of the data base from the data collection device to computer. The output of these signal handling and processing when displayed on Emonitor software screen, is shown as the last step of Figure 3.

Upon the final output of signal processing steps is generated as illustrated in Figure 3, and before analyzing data using neuro-fuzzy or neural networks, the time data of 701 observations were divided into 3 bins each, with 234 non-overlapping samples in each bin. Each of these bins has been processed using MATLAB 7.0 to extract the following statistical five features:

1. Root mean square (rms),
2. Variance (σ^2),
3. Skewness (normalized third central moment, γ_3),
4. Kurtosis (normalized fourth central moment, γ_4) and

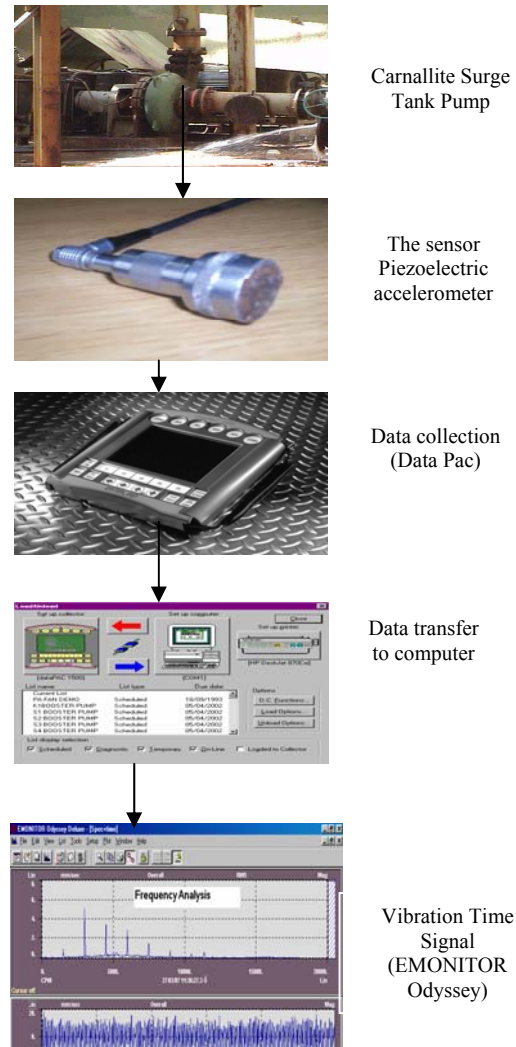


Figure 3. Experimental setup used to generate vibration time signals (i.e., training data)

5. Normalized sixth central moment (γ_6).

The above-mentioned statistical parameters of the time-domain signal could be evaluated as follows:

$$rms = \left(\sum \frac{(y_{i_{avg}})^2}{n} \right)^{1/2} \dots \dots \dots (1)$$

$$\sigma^2 = E(y_{i_{avg}})^2 \dots \dots \dots (2)$$

$$\gamma_3 = \frac{E(y_{i_{avg}})^3}{\sigma^3} \dots \dots \dots (3)$$

$$\gamma_4 = \frac{E(y_{i_{avg}})^4}{\sigma^4} \dots \dots \dots (4)$$

$$\gamma_6 = \frac{E(y_{i_{avg}})^6}{\sigma^6} \dots \dots \dots (5)$$

Where E is the expected value, y_i is the time signal amplitude, $y_{i_{avg}} = y_i - \mu$, and the mean is estimated as: $\mu = E\{y_i\}$. These five features resulted from previous step were used as inputs and the faults codes of Table 1 as output into ANFIS or neural network toolboxes of Matlab 7.0. Table 2 gives a whole set of data for component item 13 (i.e., carnallite surge tank pump), axial direction, depending on two of the time signal features (i.e., rms and σ^2), along with the fault type codes as given in Table 1. These data points of Table 2 were divided into training set

Table 2. The training data set for component 13, axial direction, with rms and σ^2 as inputs.

rms1	σ^2_1	rms2	σ^2_2	rms3	σ^2_3	Actual Fault code
3.7307	13.981	3.8662	15.012	3.9502	15.671	300
1.2485	1.5655	1.223	1.5022	1.2897	1.6706	200
1.3565	1.848	1.3335	1.786	1.3234	1.759	200
0.9928	0.9899	1.1611	1.3539	1.0622	1.133	200
1.3402	1.8039	1.3107	1.7254	1.171	1.3771	200
1.6552	2.7515	1.6109	2.2908	1.6855	2.8531	200
1.215	1.4827	1.2306	1.521	1.2481	1.5643	200
13.076	171.72	12.666	161.12	22.64	514.77	300
2.8199	7.9862	2.7284	7.4761	2.7346	7.5099	200
1.7922	3.2259	1.5147	2.3042	1.3804	1.9136	200
1.3467	1.8215	1.5302	2.3514	1.6215	2.6406	200
2.443	5.9941	2.819	7.9807	2.8086	7.9223	200
1.8628	3.4851	1.7569	3.1	1.6668	2.79	200
0.9843	0.9731	1.0165	1.0377	0.984	0.9723	200
3.1338	9.8633	3.1908	10.225	3.0204	9.1623	300
1.9599	3.8577	1.8879	3.5796	2.08	4.3449	200
0.907	0.8261	0.8984	0.8107	0.9357	0.8793	200
2.0023	4.0265	1.9251	3.7218	1.8842	3.5653	200
1.0872	1.1872	1.0067	1.0177	1.0271	1.0595	200
2.5425	6.4922	2.5336	6.4465	2.4124	5.8446	200
2.2107	4.9087	2.0619	4.2697	2.2094	4.9023	200
1.2672	1.6126	1.2585	1.5906	1.2272	1.5125	200
2.4477	6.0169	2.3541	5.5657	2.1948	4.8379	200
5.6621	32.197	5.8764	34.68	5.6977	32.603	100
6.637	44.24	7.0581	50.023	6.7727	46.067	100
1.206	1.4608	1.1843	1.4083	1.1547	1.3391	200
2.6178	6.8826	2.6075	6.8281	2.3234	5.4214	200
2.3611	5.599	2.687	7.251	2.5142	6.3485	200
1.7136	2.9492	1.8835	3.5627	1.7775	3.1732	200
1.6569	2.7572	1.4991	2.2569	1.3466	1.8212	200
0.8988	0.8114	0.8949	0.8042	0.8966	0.8073	200
0.0989	0.0098	0.1002	0.0101	0.0924	0.0086	200
2.5588	6.5757	2.4251	5.9065	2.3858	5.7163	200
1.2306	1.5209	1.1383	1.3013	1.0534	1.1144	200
0.8541	0.7326	0.8338	0.6982	0.9651	0.9354	200
3.0268	9.2009	2.9115	8.5136	3.042	9.2938	300
1.7095	2.935	1.8268	3.3518	1.7048	2.9189	200
5.6362	31.904	5.5425	30.852	5.6648	32.227	100
3.8295	14.729	4.0158	16.197	3.9914	16.001	300
1.5041	2.272	1.457	2.1401	1.6345	2.6796	200
1.1725	1.3806	1.114	1.2462	1.1229	1.2662	200
1.0424	1.0912	1.067	1.1433	1.0358	1.0776	200
2.1943	4.8359	2.4021	5.7949	2.3776	5.6775	200
3.8507	14.892	4.0665	16.608	3.731	13.98	300

Table 2 continues next page

rms1	σ_{21}	rms2	σ_{22}	rms3	σ_{23}	Actual Fault code
1.3712	1.8883	1.6033	0.2146	1.2241	1.5049	200
1.5978	2.5639	1.5616	2.4491	1.6925	2.8769	200
2.2375	5.0282	11.056	0.5714	2.7876	7.8044	200
2.9763	8.8968	4.1149	17.059	4.0959	16.827	200
1.6633	2.7785	1.6581	2.7613	1.509	2.2869	200
1.1405	1.3063	1.0985	1.2119	1.1265	1.2745	200
1.4592	2.1385	1.2842	1.6564	1.1575	1.3455	200
2.8833	8.3492	2.8825	8.3449	3.0087	9.0914	200
0.4385	0.1931	0.5096	0.2609	0.538	0.2907	200
5.1991	27.148	4.8232	23.364	4.7715	22.864	100
2.3357	5.479	2.9529	8.7569	3.3841	11.501	200
11.441	131.46	11.178	125.5	9.1815	84.662	200
2.1325	4.567	2.049	4.2164	1.8905	3.589	200
2.7235	7.4494	3.0313	9.2283	2.8	7.8737	200
0.9063	0.8249	1.0504	1.1081	1.1852	1.4108	200
1.2134	1.4788	0.664	0.4428	0.7107	0.5073	200
2.6068	6.8247	2.4486	6.0204	2.4909	6.2323	200
1.336	1.7925	1.2764	1.6361	1.291	1.6739	200
1.2019	1.4509	1.2263	1.5102	1.1169	1.2528	200
1.9368	3.7674	1.9661	3.8821	2.0943	4.4048	200
2.0871	4.375	2.3831	5.7027	2.254	5.1035	200
1.1115	1.2408	1.1403	1.306	1.183	1.4055	200
0.8634	0.7487	0.9165	0.8436	0.9324	0.8732	200
2.41	5.833	2.4545	6.0507	2.5355	6.4566	200
2.2684	5.1684	2.5308	6.433	2.7589	7.6451	200
5.9603	35.678	5.6239	31.765	5.8444	34.305	100
3.5534	12.681	4.432	19.727	3.6933	13.699	300
2.6698	7.1583	2.3951	5.7612	2.7892	7.8134	200
1.3009	1.6997	1.5103	2.2908	1.531	2.3541	200

(73 points) and testing set (19 points) in order to train and test ANFIS and neural networks as will be discussed in sections 5 and 6.

4. Neural Networks and Adaptive Neuro-Fuzzy Inference System

4.1. Neural Networks

Neural networks are universal function approximators. They are "model-free estimators" [22]. The first mathematical model of a neuron was proposed by [32] in 1943. It was a binary device using binary inputs, binary output, and a fixed activation threshold. In general, an artificial neural network, ANN (or simply neural network, NN) is a computational model defined by the following four parameters:

- Type of neurons (also called nodes).
- Connectionist architecture the organization of the connections between neurons.
- Learning algorithm.
- Recall algorithm.

Figure 4 shows artificial neural network architecture. The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. The activity of each layer could be summarized as follows: The activity of the input units represents the raw information that is fed into the network. The activity of each hidden unit is determined by the activities of the input units and the behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units. A one-layer network with R input elements and neurons are illustrated in Figure 4.

4.2. Adaptive Neuro-Fuzzy Inference System (ANFIS)

Adaptive neuro-fuzzy inference system (ANFIS) is a fuzzy inference system implemented in the framework of an adaptive neural network. By using a hybrid learning procedure, ANFIS can construct an input-output mapping based on both human-knowledge as fuzzy If-Then rules and stipulated input-output data pairs for neural networks

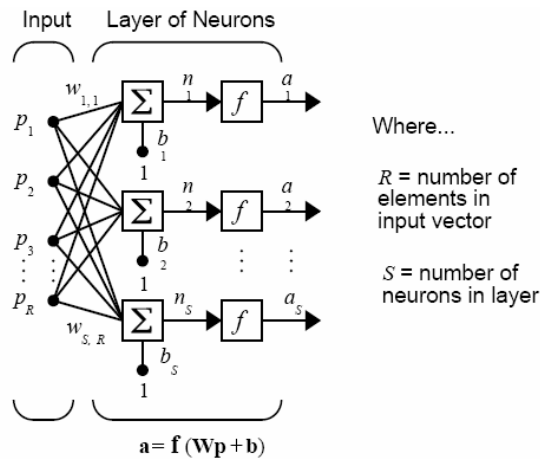


Figure 4. Artificial neural network [33].

training. ANFIS architecture is shown in Figure 5, where x and y are the inputs, f is the output, A_i and A_n^2 are the input membership functions, w_i and w_n^2 are the rules firing strengths. Five network layers are used by ANFIS to perform the fuzzy inference process.

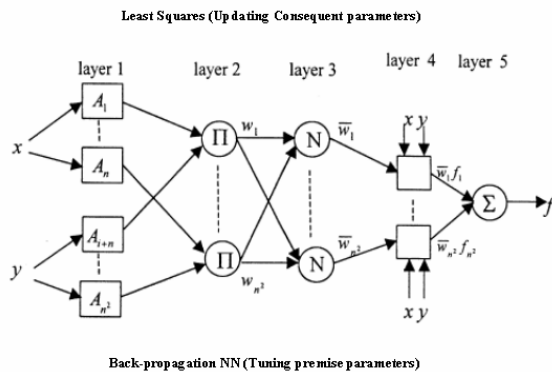


Figure 5. ANFIS architecture [33]

ANFIS is more powerful than the simple fuzzy logic algorithm and neural networks, since it provides a method for fuzzy modeling to learn information about the data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data [23].

The architecture of ANFIS, illustrated in Figure 5, has five layers to accomplish the tuning process of the fuzzy modeling system. The five layers are:

1. Layer 1: Every node in this layer is an adaptive node with a node function (i.e., membership function). Parameters of membership functions are referred to as premise or antecedent parameters.
2. Layer 2: Every node in this layer is a fixed node, which multiplies the incoming signals and sends the product out. Each node represents the firing strength of a fuzzy rule.
3. Layer 3: Every node in this layer is a fixed node which calculates the ratio of the one firing strength to the sum of all rules' firing strengths. The outputs of this layer are called normalized firing strengths.
4. Layer 4: Every node in this layer is an adaptive node with a node function (i.e., linear combination of input variables). Parameters in this layer are referred to as consequent parameters.

5. Layer 5: The single node in this layer is a fixed node that computes the overall output as the summation of all incoming signals

In the next section, fault-diagnosis systems based on ANFIS and neural networks will be presented.

5. Intelligent Fault-Diagnosis Systems

5.1. Anfis-Based Fault-Diagnosis System

ANFIS prediction of machine's fault types starts by obtaining the data set (input-output data pairs) and dividing it into training and testing or validating data sets. The training data set is used to find the initial premise parameters for the fuzzy membership functions by equally spacing each membership function. The testing data used to validate the system.

Using ANFIS editor of Matlab 7.0, statistical input data (i.e., rms , σ^2 , γ_3 , γ_4 , and γ_6) were used to train and test the system. In fact, these five statistical features in each bin will complicate the structure of ANFIS. Therefore, combinations of two features were tested to build the system. Tables 3 and 4 shows the input features ranges and a sample of features used for training, respectively.

Table 3. Inputs features with their ranges which were used to generate the training data for the time-based ANFIS fault-diagnosis system

Feature	Meaning	Range
rms	Root mean square	0.09 - 46
σ^2	Variance	0.008 - 960
γ_3	Skewness	-0.5 - 0.6
γ_4	Kurtosis	1.5 - 6.0
γ_6	Normalized sixth central moment	3E-6 - 5E+10
FC	Fault Code	100 - 600

Table 4. Sample of ANFIS training data.

Component number	13_11_a_01	13_11_a_02	13_11_a_03	13_11_a_04	13_11_a_05	13_11_a_06
RMS	7.5613	3.7307	1.8738	13.1524	1.2485	1.3565
VAR	57.4192	13.9806	3.5261	173.7325	1.5655	1.848
RMS	7.6958	3.8662	1.7561	12.6339	1.223	1.3335
VAR	59.48	15.0118	3.0971	160.3005	1.5022	1.786
RMS	7.5859	3.9502	1.7603	22.7711	1.2897	1.3234
VAR	57.7924	15.6711	3.1119	520.7485	1.6706	1.759
Fault Code	100	300	200	300	200	200

The total number of data was 92 points. 73 points were used for training and 19 points (i.e., 20% of the total data points in order to make the process valid statistically), which are different and independent of the training data, were used for testing. In addition, each of the original data points (i.e., training and testing) is an average of 4 readings (replicates) in order to insure the statistical validity of this work. Using a given input/output data set, the toolbox function ANFIS constructs a fuzzy inference system (FIS) whose membership function parameters are tuned (adjusted) using either a backpropagation algorithm alone or a mixture of backpropagation and least squares (hybrid method). In this study, the membership function parameters were updated using the hybrid method.

ANFIS takes the experimental data of the vibration features in each bin (rms1(RMS1), σ_{21} (VAR1), rms2(RMS2), σ_{22} (VAR2), rms3(RMS3), σ_{23} (VAR3)) as input training data of the system. Different ANFIS parameters were tested as training parameters in order to achieve the perfect training and the maximum prediction accuracy.

The training data set has been used to set the initial parameter of the (ANFIS) model. This model has been trained with different parameter in order to get the minimum training and testing error. 83 data points out of 92 total points were adopted for training the system, the remaining 19 points were devoted to test and validate the system. Some fault codes did not appear in the training/testing data because these faults did not occur on the machine's components during the study period or in the machine's history, and consequently were excluded from the training/testing data.

A total of 216 fuzzy rules were used to build the fuzzy inference system. A Gaussian membership function (MF) was adopted to train ANFIS because it achieved minimum training error at epoch 170, as shown in the training curve of Figure 6. Figure 7 shows that the system is very well-trained to predict the machine's fault type.

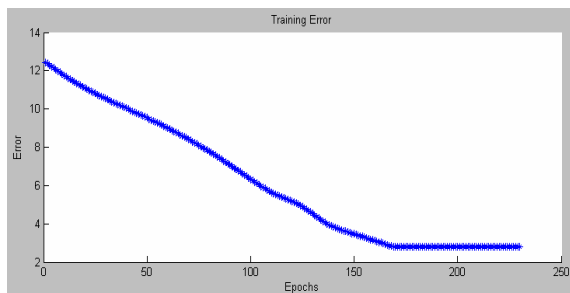


Figure 6. ANFIS training curve.

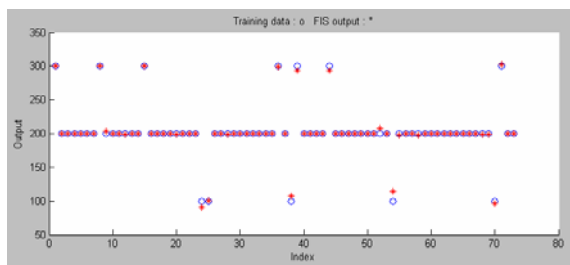


Figure 7. Actual and Predicted fault type values.

A perfect training is clear in this figure. Three Gaussian membership functions (MF) were used for root mean square inputs (RMS1, RMS2, RMS3) while another two Gaussian membership functions were adopted for variance inputs (VAR1, VAR2, VAR3). The final ANFIS-tuned (MF) for all input features are illustrated in Figures 8. The training root mean square error (RMSE) was dropped from 44.7159 when using root mean square (rms) and centralized six momentums (γ_6) were used as inputs to the ANFIS model to 9.608 when using root mean square (rms) and variance (σ^2) as input features to the model. The training remains constant after 170 epochs which means no improvement occurs after this epoch. The tuning trials of input features selection for the ANFIS system are highlighted in Table 5. The final trained fuzzy inference system (FIS) for predicting the fault types is illustrated in Figure 9.

5.2. Neural-Networks-Based Fault-Diagnosis System

A neural network system (i.e., nntool in Matlab 7.0) can be considered as a parameterized nonlinear map. However, in this study, the neural network parameters (i.e., Root mean square and Variance) have been selected as inputs, with failure code as the output.

First, the training data was used to find the appropriate network between the input and the target (desired output) to realize the actual output. The error between each pair was computed and the overall training error was determined.

A multi-layer forward neural network (ANN) is used for the computation. This network is also a cascade-forward backpropagation network. The characteristic features of time domain signals of the system with normal and faulty conditions have been used as inputs to this ANN structure, including input, hidden and output layers. The input layer contains for selected features from the time domain. The output layer contains of nodes indicating the fault type code. The final neural network used in this study consists of the input layer, one hidden layer and the output layer. The input layer has nodes representing the features extracted from the measured vibration signals. The number of neurons in the first hidden layer was 20. The number of output nodes is only one. The ANN was trained and implemented using the MATLAB neural network toolbox using backpropagation with Levenberg–Marquardt algorithm. For training, maximum iteration number (epoch) of 6000 was used. The initial weights and biases of the network were generated automatically by the program.

All the five statistical input features (i.e., rms, σ^2 , γ_3 , γ_4 , and γ_6) were used for training and testing the neural network. The combination of root mean square and variance lead to the best result in testing and training. The training error was reduced from 29.1715 for kurtosis and six momentum feature to .04205 for root mean square and variance. The tuning trials of input features selection for the neural-networks system are highlighted in Table 6. Figure 10 shows the training and testing curves of the root mean square and variance model with performance equal to 0.92 that indicates the behavior of the network.

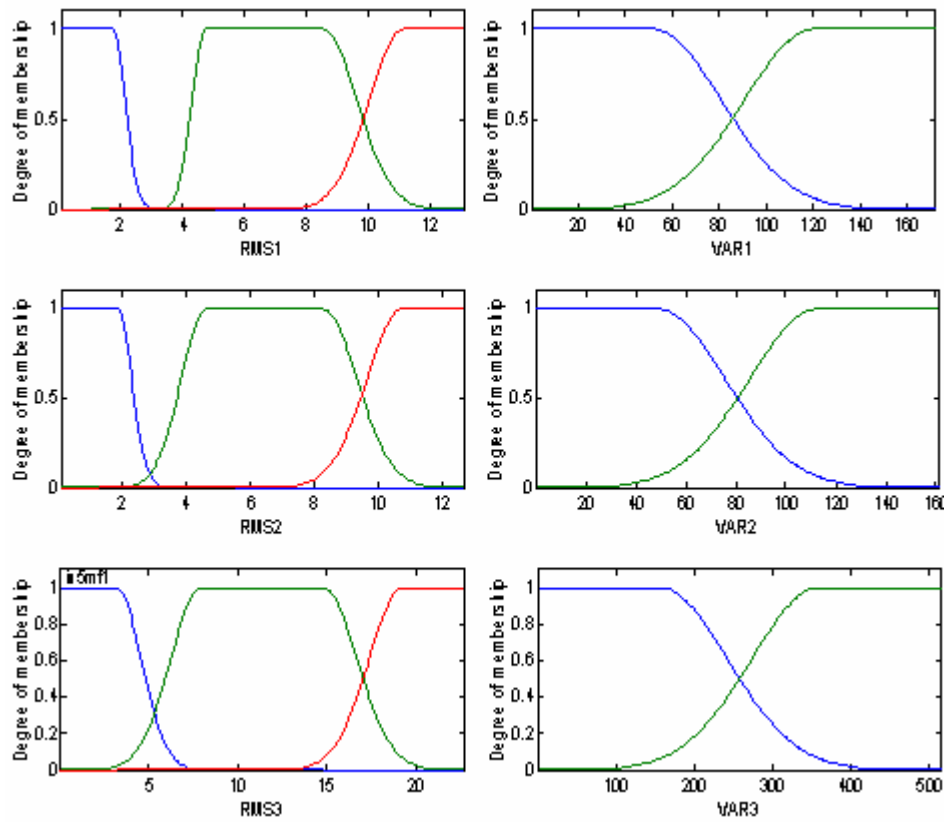
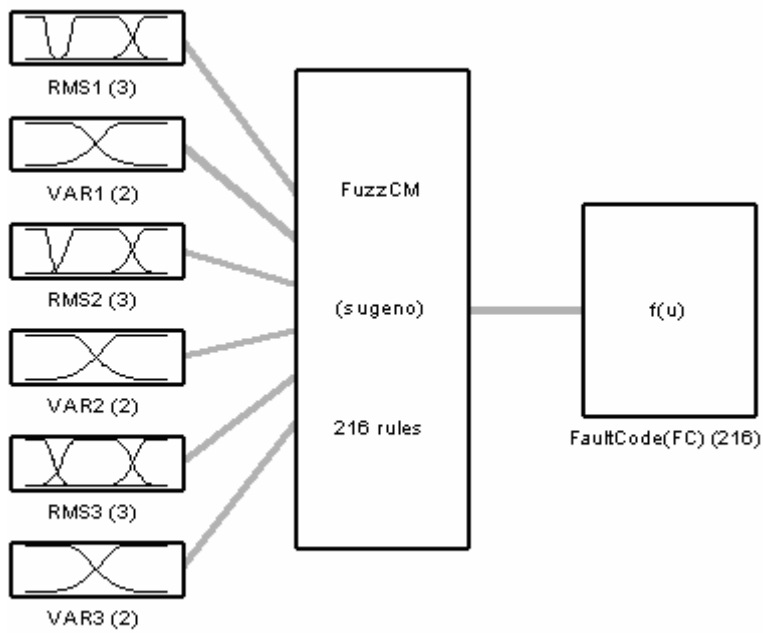


Figure 8. Final membership functions (MF) for input features.

Table 5. Tuning trials of input features (ANFIS)

MF type (Trapezoidal) , number of MFs =3 4 3 4 3 4 , Epochs =5					
	rms	σ^2	γ_3	γ_4	γ_6
rms		Train error=9.608 Test error=15.329	Train error=.0902 Test error =34.84	Train error=2.044 Test error =99.719	Train error=18.737 Test error =72.28
σ^2			Trainers=37.169 Test error =73.328	Train error=30.61 Test error =31.62	Train error=33.236 Test error =31.108
γ_3				Trainers=21.146 Test error=684.99	Train error=35.084 Test error =53.781
γ_4					Train error=44.7159 Test error =44.27
γ_6					



System FuzzCM 6 inputs , 1 outputs , 216 rules

Figure 9. The final fuzzy inference system (FIS) for faults prediction.

Table 6. Tuning trials of input features selection (Neural Networks).

	rms	σ^2	γ_3	γ_4	γ_6
rms		Train error=.04205 Test error=1.056	Train error=14.695 Test error=42.414	Train error=21.752 Test error=41.522	Train error=24.99 Test error=45.457
σ^2			Train error=21.698 Test error=14.112	Train error=1.2286 Test error=1.5624	Train error=20.398 Test error=21.041
γ_3				Train error=27.137 Test error=57.250	Train error=24.745 Test error=44.79
γ_4					Train error=29.171 Test error=34.783
γ_6					

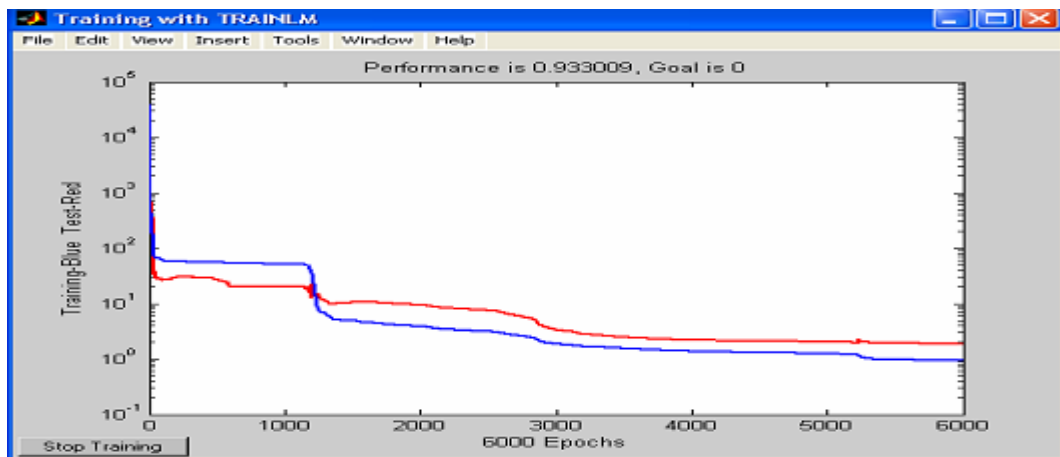


Figure 10. Training and testing of the neural network with inputs (rms and σ^2).

5.3. Models Validation

The ANFIS and neural networks prediction models for machine's faults were validated by selecting a certain number of data points (i.e., 19 points), different from the other 73 points used for ANFIS and neural nets training. Each validation data point (i.e., rms and σ^2) in the three bins, as given in Table 7, was fed into the system, and then the predicted fault type code (i.e., FC) were computed with

the actual values of FC. The average percent errors in the ANFIS fault prediction is 3%, achieving a satisfactory accuracy of prediction of 97% as illustrated in Table 7, and the percent error in the Neural nets fault diagnosis is 0.8%, achieving a much better accuracy (i.e., 99.2% as given in Table 8) than the ANFIS prediction system. Table 8 illustrates that the neural networks-predicted values are a close match of the actual ones.

Table 7. Validation table for the ANFIS fault prediction system

(rms, σ^2), ANFIS, Time-domain								
rms1	σ^2 1	rms2	σ^2 2	rms3	σ^2 3	Actual Fault code	Predicted Fault code	Error%
13.15	173.73	12.63	160.30	22.77	520.75	300	300	0
6.58	43.51	6.72	45.32	6.56	43.23	100	107	7
1.95	3.84	1.92	3.69	1.86	3.46	200	200	0
1.98	3.93	1.91	3.68	1.87	3.52	200	200	0
1.02	1.05	1.08	1.17	1.03	1.07	200	200	0
3.97	15.85	3.83	14.75	3.93	15.54	300	287	4.3
1.80	3.26	1.78	3.17	1.95	3.81	200	200	0
6.50	42.46	6.72	45.32	6.64	44.33	100	112	12
1.45	2.10	1.42	2.03	1.31	1.73	200	200	0
1.37	1.88	1.43	2.06	1.41	1.99	200	200	0
7.65	57.42	7.69	59.48	7.59	57.79	100	105	5
5.87	34.66	4.90	24.12	5.55	30.93	100	92	8
4.09	16.86	3.90	15.31	4.48	20.13	300	270	10
13.15	173.73	12.63	160.30	22.77	520.75	300	300	0
3.34	11.19	3.54	12.59	4.08	16.74	300	317	5.7
3.92	15.42	3.94	15.56	4.15	17.28	300	287	4.3
1.87	3.53	1.76	3.09	1.76	3.11	200	200	0
1.41	2.00	1.43	2.04	1.44	2.08	200	200	0
2.69	7.30	2.06	4.26	2.06	4.26	200	201	0.5
Average Percent Error						3 %		

Table 8. Validation table for the neural-networks fault prediction system

(rms, σ^2), Neural Nets, Time-domain								
rms1	σ^2 1	rms2	σ^2 2	rms3	σ^2 3	Actual Fault code	Predicted Fault code	Error%
13.15	173.73	12.63	160.30	22.77	520.75	300	300.29	0.10
6.58	43.51	6.72	45.32	6.56	43.23	100	102.65	2.65
1.95	3.84	1.92	3.69	1.86	3.46	200	199.47	0.27
1.98	3.93	1.91	3.68	1.87	3.52	200	199.38	0.31
1.02	1.05	1.08	1.17	1.03	1.07	200	200.61	0.31
3.97	15.85	3.83	14.75	3.93	15.54	300	299.56	0.15
1.80	3.26	1.78	3.17	1.95	3.81	200	202.20	1.10
6.50	42.46	6.72	45.32	6.64	44.33	100	97.70	2.30
1.45	2.10	1.42	2.03	1.31	1.73	200	199.31	0.35
1.37	1.88	1.43	2.06	1.41	1.99	200	199.70	0.15
7.65	57.42	7.69	59.48	7.59	57.79	100	103.30	3.30
5.87	34.66	4.90	24.12	5.55	30.93	100	102.50	2.50
4.09	16.86	3.90	15.31	4.48	20.13	300	300.28	0.09
13.15	173.73	12.63	160.30	22.77	520.75	300	299.10	0.30
3.34	11.19	3.54	12.59	4.08	16.74	300	300.00	0.00
3.92	15.42	3.94	15.56	4.15	17.28	300	299.00	0.33
1.87	3.53	1.76	3.09	1.76	3.11	200	200.00	0.00
1.41	2.00	1.43	2.04	1.44	2.08	200	199.20	0.40
2.69	7.30	2.06	4.26	2.06	4.26	200	199.80	0.10
Average Percent Error						0.8 %		

6. Conclusions

An adaptive neuro-fuzzy system and a neural network are applied to predict the fault types of a mechanical system (carnallite surge tank pump). The time domain features (rms and σ^2) were used as inputs to ANFIS as well as neural nets to predict machine's fault type.

The following conclusions can be drawn from this study:

- The average percent error predicted by ANFIS with the trapezoidal membership function in the axial direction is only 5 %, achieving an accuracy of 95% using time-domain features.
- The average percent error predicted by neural-network with cascade network type in the axial direction is only 0.7 %, achieving an accuracy of 99.3 % using time-domain features.
- Analysis of time domain in both ANFIS and NN shows that, the most significant group of vibration signals and the characteristic features were root mean square and variance.
- Artificial neural networks (ANN) have potential applications in automated detection and diagnosis of machine conditions. Many of the ANNs for machine condition monitoring used the preprocessed frequency-domain features of the measured vibration signals
- ANFIS technique in parallel with time-based analysis can be used to predict and diagnose the machine's faults and failures. It is believed that this approach can be applied to identify other maintenance-related parameters.
- Neural networks, fuzzy logic and Neuro-fuzzy systems have an inherent shortcoming that they need to be retrained for different process parameters.
- A future work may be focused on constructing a real-time condition monitoring system by implementing an adaptive self-learning intelligent system for predicting and diagnosing machine's faults online.

References

- [1] C. Mechefske, "Hydro electric Generating Unit MCM: A Comprehensive System". The Journal of Non-Destructive Testing and Condition Monitoring, Vol. 36, 1994, 238-245.
- [2] Y. Peng, "Intelligent Condition Monitoring Using Fuzzy Inductive Learning". Journal of Intelligent Manufacturing, Vol. 15, 2004, 373-380.
- [3] O. Castillo, P. Melin, "A Hybrid Fuzzy Fractal Approach for Time Series Analysis and Plant Monitoring". International Journal of Intelligent Systems, Vol. 17, 2002, 751-765.
- [4] L. Tong, Y. Liang, "Forecasting Field Failure Data for Repairable Systems Using Neural Networks and SARIMA Model". International Journal of Quality and Reliability Management, Vol. 22, No. 4, 2005, 410-420.
- [5] F. Kong, R. Chen, "A Combined Method for Triplex Pump Fault Diagnosis Based on Wavelet Transform, Fuzzy Logic, and Neuro-Networks". Mechanical Systems and Signal Processing, Vol. 18, 2004, 161-168.
- [6] L. Zadeh, "Outline of a New Approach to Analysis of Complex Systems and Decision Process". IEEE Transactions on Systems, Man, and Cybernetics, 1973, 28-44.
- [7] P. Wasserman, T. Schwartz, "Neural Networks, Part 2: What are they and why is everybody so interested now?". IEEE Expert, Vol. 8, 1998, 10-15.
- [8] T. Fukuda, T. Shibata, "Theory and Applications for Industrial Control Systems". IEEE Transactions on Industrial Electronics, Vol. 39, No. 6, 1992, 472-489.
- [9] C. Klimasauskar, "Neural Networks: An Engineering Perspective". IEEE Communications Magazines, Vol. 15, 1992, 50-53.
- [10] British Standard Institution, BS3811, "British Standard Glossary of Maintenance Management Terms in Terotechnology".
- [11] R. Yam, P. Tse, L. Li, P. Tu, "Intelligent Predictive Decision Support System for Condition-Based Maintenance". International Journal of Advanced Manufacturing Technology, Vol. 17, 2001, 383-391.
- [12] A. Borden, "Designing and Maintaining Decision-making Processes". AGARD Conference Proceedings, Paris, France, 1993.
- [13] M. Liberatore, A. Stylianou, "Using Knowledge-based Systems for Strategic Market Assessment". Information and Management, Vol. 27, No. 4, 1994, 221-232.
- [14] T. Saaty, Multi-criteria Decision Making The Analytic Hierarchy Process. RWS Publication, 1990.
- [15] M. Davies, "A Multi-criteria Decision Model Application for Managing Group Decisions". Journal of the Operational Research Society, Vol. 45, No. 1, 1994, 47-58.
- [16] C. Kocourek, "A Petri-net based Design Decision Support System". Proceedings of the IASTED International Conference, Applied Modeling and Simulation, Vancouver, BC, Canada, 1993, 108-114.
- [17] A. Hurson, S. Pakzad, B. Lin, "Automated Knowledge Acquisition in a Neural Network Based Decision Support System for Incomplete Database System". Microcomputers in Civil Engineering, Vol. 9, No. 2, 1994.
- [18] T. Quah, C. Tan, T. Srinivasan, "A Fuzzy-neural-network Decision Support Systems for Bond Trading". Proceedings of International Symposium on Artificial Intelligence, The Artificial Intelligence Technology Transfer Conference, Monterrey, Mexico, 1993.
- [19] A. Ishikawa, "The New Fuzzy Delphi Methods: Economization of GDS (group decision support)". Proceedings of the 26th Hawaii International Conference on System Sciences, Wailea, HI, USA, 1993.
- [20] C. Schrunder, J. Gallertly, J. Bicheno, "A Fuzzy Knowledge-based Decision Support Tool for Production Operations Management". Expert Systems, Vol. 11, No. 1, 1994, 3-11.
- [21] M. Rogers, K. Hrovat, T. Reckart, M. Moskwitz, K. Mcpherson, "Accelerometer data analysis and Presentation Techniques". Technical Report No. 138ptN97, NASA, Lewis Research Center, Cleveland, Ohio, 1997
- [22] B. Kosko, Neural Networks and Fuzzy systems: A Dynamical Systems Approach to Machine Intelligence. Prentice-Hall, Englewood Cliffs, NJ, 1992.
- [23] J. Jang, 1993 "ANFIS: adaptive-network-based fuzzy inference system". IEEE Transactions on Systems, Man and Cybernetics, Vol. 23, 1993, 665-685.

Effect of Exhaust Gas Recirculation (EGR) on the Performance and Emission Characteristics of Diesel Engine with Sunflower Oil Methyl Ester

K. Rajan ^{a,*}, K. R. Senthilkumar ^b

^a Department of Mechanical Engineering, Dr.M.G.R Educational and Research Institute, India

^b Professor, Sri Sastha Institute of Engineering and Technology, India

Abstract

Transesterified fuels (biodiesel) from vegetable oils are alternative fuels for diesel engines. They are renewable and offer potential reduction in CO and HC emissions due to higher O₂ contents in vegetable oil. Many research studies have reported that exhaust from biodiesel fuel has higher NO_x emissions while HC and PM emissions are significantly lower than operated with diesel fuel. The aim of the present investigation is to reduce NO_x emissions. Exhaust gas recirculation (EGR) is one of the most effective technique for reducing NO_x emissions in compression ignition engines. A twin cylinder four stroke water cooled direct injection (DI) diesel engine was used for conducting test with (Sunflower methyl ester:SFME) biodiesel blends with diesel fuel combined with EGR technique. The results showed that for a 7.5kW power output, B20 SFME with 15% EGR rate produce 25% less NO_x emissions compared to diesel fuel for the same level smoke emissions.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: Emission; Exhaustgas Recirculation; Sunflower Oil, Compression Ignition; Biodiesel.

1. Introduction

Compression ignition engine are preferred prime movers due to excellent drivability and higher thermal efficiency. Despite their advantages they produce higher levels of NO_x and smoke emissions which will more harmful to human health. Hence stringent emission norms have been imposed. In order to meet the emission norms and also the fast depletion of petroleum oil reserves lead to the research for alternative fuels for diesel engines. Biodiesel from vegetable oils are alternative to diesel fuel for diesel engines. The use of biodiesel in diesel engines does not require any engine modification. Biodiesel gives considerably lower emissions of PM, carbon monoxide (CO) and hydrocarbon (HC) without any fuel consumption or engine performance penalties. Many researchers have found that with biodiesel fueled engine produces higher NO_x emissions compared to diesel [9-13]. To achieve reductions in NO_x emissions, exhaust gas

recirculation (EGR) can be used with biodiesel in the diesel engines. EGR is an effective technique of reducing NO_x emissions from the diesel engine exhaust [1-4]. Controlling the NO_x emissions primarily requires reduction of in-cylinder temperatures [2, 3]. However, the application of EGR results in higher fuel consumption and emission penalties, also EGR increases HC, CO, and PM emissions along with slightly higher specific fuel consumption [13]. EGR rates are sufficient for high load,

also as the load increases, diesel engines tend to generate more smoke because of reduced oxygen. Therefore, EGR, although effective to reduce NO_x, further increases the smoke and PM emissions [5]. Abd-Alla et al [1] performed experiments on a 9.0 kW rated power dual fuel (gaseous fuel- methane with diesel as pilot fuel) mode direct injection diesel engine to study the effect of inlet air temperature by the way of mixing of hot EGR and addition of diluents gas such as CO₂ and N₂. They reported that the addition of CO₂ gas in the intake charge resulted in moderate reduction of NO_x emission but Unburnt hydrocarbon emission (UBHC) was increased. By increasing the intake charge temperature resulting in increase of NO_x emission with decrease in UBHC, the brake thermal efficiency and power output increased due to reduced ignition delay. Also they suggested that the performance was improved at low load condition when the intake air temperature was increased. Deepak Agarwal et al [2] conducted a test on a single cylinder DI diesel engine and measured the performance and emission characteristics with rice bran methyl ester (RBME) and its blends as fuel with EGR system. They optimized and reported that 20% biodiesel blends with 15% EGR produce the less NO_x, CO and HC emissions and also improved thermal efficiency and reduced BSFC.

Hountalaous et al [3] using 3D-multi dimensional model to examine the effect of EGR temperature on a turbocharged DI diesel engine with three different engine speeds, and they reported that high EGR temperature affects the engine brake thermal efficiency, peak combustion pressure, air fuel ratio and also soot emissions, and the combined effect of increased temperature and

* Corresponding author. kr_iceer@yahoo.co.in

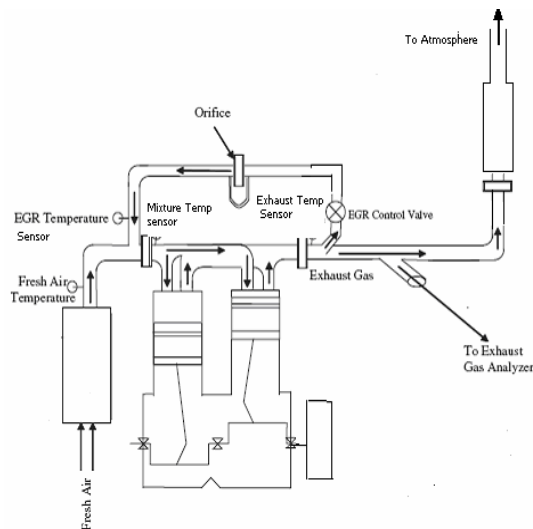


Figure 1. Experimental setup using EGR

Table 2. Engine specifications.

Make	Kirloskar
Engine type	Twin cylinder, Vertical direct Injection, Water cooled.
Bore/stroke	80/110mm
Rated power	7.4 kW
Compression ratio	16.5:1
Rated speed	1500 rpm

can be used to reduce NO_x and smoke intensity simultaneously. A series of exhaustive engine test were carried out which levels, an optimum of 15% EGR can be used with B20 and B40 SFME of biodiesel. The performance and emission data were analyzed for thermal efficiency, BSFC, exhaust gas temperature, HC, CO, NO_x and smoke emissions.

3.1. Performance Analysis

Engine tests were carried out using diesel at 1500 rpm and different EGR rates in order to study the effect of EGR on the smoke density and NO_x concentration in the exhaust emissions. Higher amount of smoke in the exhaust is observed when the engine is operated with EGR compared to without EGR. Smoke emissions increases with increasing engine load and EGR rates. EGR reduces availability of oxygen for combustion of fuel, which results in relatively incomplete combustion and increased formation of PM and reducing NO_x emissions from diesel engine. Using biodiesel in diesel engine, smoke is decreased with increase in NO_x. Thus, biodiesel with EGR can be used to reduce NO_x and smoke intensity simultaneously. A series of exhaustive engine test were carried out which levels, an optimum of 15% EGR can be used with B20 and B40 SFME of biodiesel. The performance and emission data were analyzed for thermal efficiency, BSFC, exhaust gas temperature, HC, CO, NO_x and smoke emissions.

3.1.1. Brake Thermal Efficiency

Figure 2 shows the variations of brake thermal efficiency of diesel and sunflower methyl ester (SFME) with and without EGR. It is observed that from

the figure the brake thermal efficiencies are increased with increase in load with EGR at lower load due to re-burning of hydrocarbons that enter in to the combustion chamber with the recirculated exhaust gases and at full load operation the brake thermal efficiency not affected by exhaust gases. The brake thermal efficiencies are improved with increasing concentration of bio diesel and its diesel blends due to the higher oxygen present in the bio diesel. B20 SFME with 15 % EGR shows 4 % increase in brake thermal efficiency at lower load (0-75%) compared to diesel without EGR.

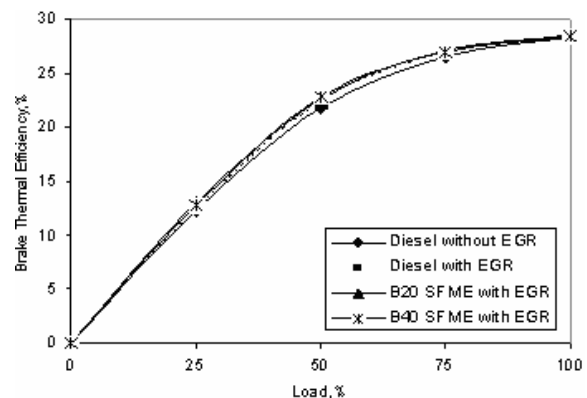


Figure 2. Variation of Brake thermal efficiency with Load.

3.1.2. Brake Specific Fuel Consumption (BSFC)

Figure 3 shows the variations of BSFC for diesel and biodiesel blends and with and without EGR. The brake specific fuel consumptions are lower for diesel at lower loads operated with EGR when compared to without EGR. However, at higher engine loads, BSFC with EGR is almost similar to that of without EGR for diesel fuel. The brake specific fuel consumptions are increased with increasing concentration of biodiesel blends when the engine is operated on biodiesel blends with EGR due to lower calorific value and high viscosity of the sunflower methyl ester when compared to diesel with and without EGR. The brake specific fuel consumptions are increased about 10 % for B20 and 15 % for B40 sunflower methyl ester blends at full load operation with EGR.

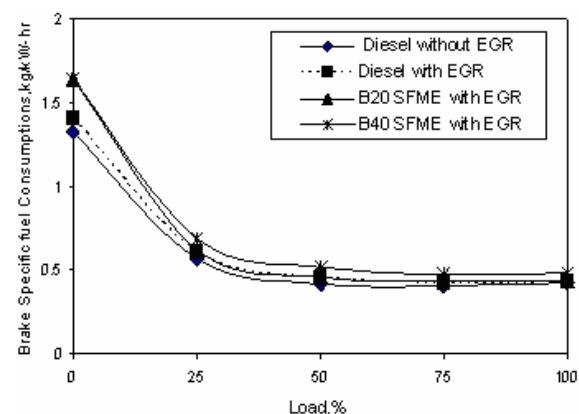


Figure 3. Variations of Brake Specific fuel Consumptions with Load.

3.1.3. Exhaust Gas Temperature

Figure 4 shows the variations of exhaust gas temperature with diesel and blends of biodiesel with EGR. It can be observed that with increase in load, exhaust

gas temperature also increases. The exhaust gas temperature was found to be lower for EGR-operated engine with diesel due to lower availability of oxygen for combustion and higher specific heat of intake exhaust gas air mixture. The temperature of the exhaust gases for B20 and B40 sunflower methyl ester were observed lower than the diesel without EGR.

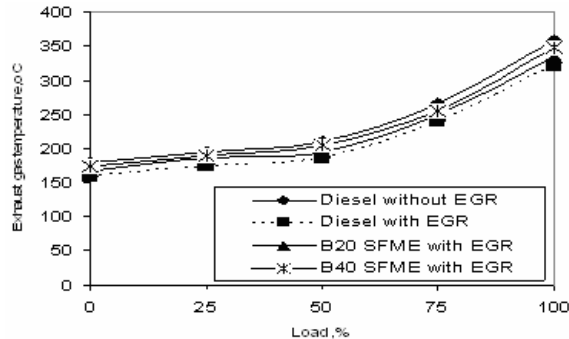


Figure 4. Variation of Exhaust gas Temperature with Load.

3.2. Emission Characteristics

3.2.1. Unburnt Hydro Carbon Emissions (UBHC)

Figure 5 shows the variations of UBHC emissions of diesel and sunflower methyl ester blends with and without EGR. The UBHC increases with increase in load and EGR rate, because of lower oxygen content available for combustion, that is lower excess oxygen concentration results rich mixture which results incomplete combustion and results higher hydro carbon emission. It is also observed from the graph the 20% and 40% biodiesel blend with 15% EGR gives 5% and 15% lower UBHC emissions with full load compared to diesel without EGR.

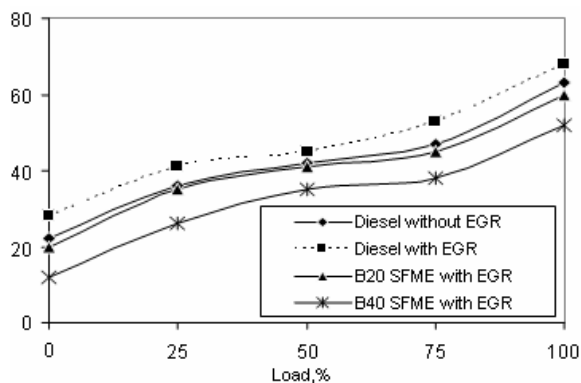


Figure 5. Variations of Hydrocarbon Emission with Load

3.2.2. Carbon Monoxide Emissions (CO)

Figure 6 shows the variations of CO emissions of diesel and sunflower methyl ester with EGR and without EGR. The CO increases with increase in load and EGR rate. However, CO emissions of SFME were comparatively lower. Higher values of CO were observed at full load for both diesel and biodiesel fuels with EGR. For biodiesel, the excess oxygen content is believed to have partially compensated for the oxygen deficient operation under EGR. Dissociation CO_2 to CO at peak loads where high combustion temperatures and comparatively fuel rich operation exists, can also contribute to higher CO emissions. It is observed that from the graph CO

emissions are 10% and 20% lower for 20% and 40% biodiesel blends respectively with full load compared with diesel when the engine is operated with EGR.

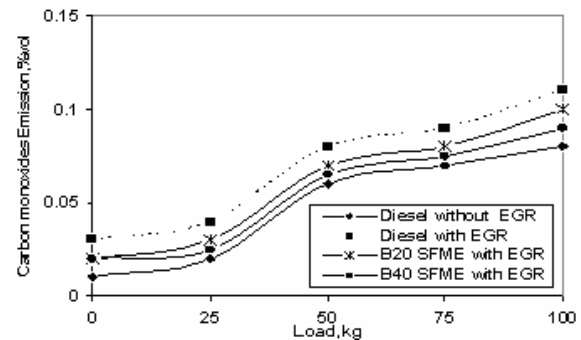


Figure 6. Variations of CO Emissions with Load.

3.2.3. Nitrogen Oxides Emission (NO_x)

Figure 7 shows the variations of NO_x emissions of diesel and sunflower methyl ester with and without EGR. The degree of reduction in NO_x at higher at higher loads. The reasons for reduction in NO_x emissions using EGR in diesel engines are reduced oxygen concentration and decreased the flame temperatures in the combustion chamber. However, NO_x emissions in case of biodiesel blends without EGR are higher than diesel due to higher temperatures prevalent in the combustion chamber. It is also observed from the graph the 20% and 40% biodiesel blends have 25% and 14% lower NO_x emissions respectively with full load when compared to diesel fuel without EGR.

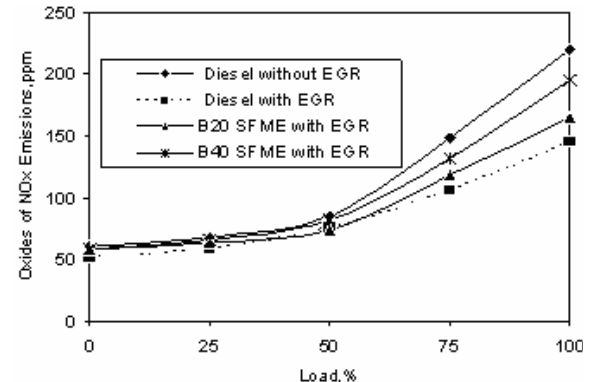


Figure 7. Variations of NO_x Emissions with Load.

4. Conclusion

In this report the performance and emission characteristics of diesel fuel and diesel-SFME blends with exhaust gas recirculation were investigated. The results obtained of this study are summarized as follows.

1. Methyl ester of sunflower oil was prepared with lye catalyst NaOH and methanol.
2. Compared with conventional diesel fuel, the exhaust NO_x was reduced about 25% at 20% biodiesel blends with 15% EGR due to less oxygen available in the recirculated exhaust gases which lowers the flame temperature in the combustion chamber.
3. SFME blend with 15% EGR, which improves the 4% of brake thermal efficiency and 10% increase in BSFC due to lower calorific value of the biodiesel.

4. The total unburnt HC and CO emissions were decreased by 5 % and 10 % for 20 % biodiesel blends respectively compared to diesel fuel with EGR and smoke emissions were observed as increases, due to incomplete combustion.
5. Engine operation with biodiesel while employing EGR were able to reduce 25 % NO_x, and reduction in brake thermal efficiency and increase in smoke, CO and UBHC were observed compared to diesel.

References

- [1] G. H. Abd-Alla, H. A. Soliman, O. A. Badr, M.F. Rabbo, "Effects of diluents and intake air temperature in exhaust gas recirculation of an indirect injection dual fuel engine". *Energy Conversion Management*, Vol.42, 2001, 1033-1045.
- [2] D. Agarwal, S. Sinha, A. K. Agarwal, "Experimental investigation of control of NO_x emissions in biodiesel – fueled compression ignition engine". *Renewable Energy*, Vol. 31, 2006, 2356-2369.
- [3] D. T. Hountalas, G. C. Mavropoulos, K. B. Binder, "Effect of exhaust gas recirculation (EGR) temperature for various EGR rates on heavy duty DI diesel performance and emissions". *Energy*, Vol. 33, 2008, 272-283.
- [4] K. Santoh, L. Zhang, H. Hatanaka, T. Takatsuki, K. Yokoto, "Relationship between NO_x and SM emissions from DI diesel engine with EGR". *Society of Automotive engineers of Japan*, Vol.18, 1997, 369-375.
- [5] M. Zheng, G. T. Reader, J. G. Hawley, "Diesel engines exhaust gas recirculation—a review on advanced and novel concepts". *Energy Conversion Management*. Vol.45, 2004, 883–900.
- [6] N. Nabi, S. Akhter, Z. Shahadat, "Improvement of engine emissions with conventional diesel fuel and diesel –biodiesel blends". *Bio Resource Technology*, Vol. 97, 2006, 372-378.
- [7] V. Pradeep, R. P. Sharma, "Use of HOT EGR for NO_x control in a compression ignition engine fuelled with bio-diesel from *Jatropha* oil". *Renewable Energy*, Vol. 32, 2007, 1136-1154.
- [8] N. Saravanan, G. Nagarajan, K.M. Kalaiselvan, C. Dhanasekaran, "An experimental investigation on hydrogen as a dual fuel for diesel engine system with exhaust gas recirculation technique". *Renewable Energy*, Vol. 33, 2008, 422-427.
- [9] S.Y.E. Mohamed, "Effect of exhaust gas recirculation on some combustion characteristics of dual fuel engine". *Energy Conversion Management*, Vol.44, 2003,707-721.
- [10] Tsolakis.A, Megaritis.A, "Exhaust gas reforming of rapeseed methyl ester for reduced exhaust emissions of CI engines", *Biomass & Bioenergy*, Vol.27 ,2007,493- 505.
- [11] A. Tsolakis, A. Megaritis, M.L.Theinnoi, "Engine performance and emissions of a diesel engine operating on diesel-RME (rapeseed methyl ester) blends with EGR (exhaust gas recirculation)". *Energy*, Vol. 32, 2007, 2072-2080.
- [12] Y. Yoshimoto, H. Tamaki, "Reduction of NO_x and smoke emissions in a diesel engine fueled by biodiesel emulsion combined with EGR". *SAE Transaction*, 2001-01-0649, 467-475.
- [13] Y. Yoshimoto, M. Onodera, H. Tamaki, "Reduction of Nox and smoke and BSFC in a Diesel Engine Fueled by BiodieselEmulsion with Used Frying oil". *SAE Transaction*,1999-01-3598, 1913-1929.

The Use of Rational Design in the Development of an Improved Plug Tool for the Rotary Tube Widening Process

S. D. Al-Shobaki *, A. K. A. Al-Dahwi , R. H. Fouad

Department of Industrial Engineering, Hashemite University, Jordan

Abstract

To improve the performance of any metal forming process, a minimization in the process load parameters is needed. The used die geometry significantly affects the process performance. Several die designs were developed to increase efficiency and reduce defects. This paper investigates the effect of the widened plug profile on the tube widening process, and considers the use of the Constancy of the Ratios of successive generalized Homogeneous Strain (CRHS) increments concept for different die profiles. Five widening process parameters are considered: the axial plug load, the widening torque, the variation in temperature, the widened tube thickness, and the widened tube quality. The investigation shows that an improvement in the widening process parameters was achieved using (CRHS) plug design leading to a 25% increase in the maximum widening ratio compared to other plug designs. This was achieved at relatively low tube ends temperature and an improved quality of widened tube ends.

© 2009 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: : Tool Design; Tube Widening Process; CRHS Concept; Widening Load Parameters; Temperature.

1. Introduction

Many deformation processes have been used as manufacturing processes to widen the tubes ends [1] in order to facilitate the requirements for overlapping joining of tubes and pipes. The sealing condition in the joining portions of the pipes networks may play an important reason in their failure [[2], and calls for improving the quality of the tube end to minimize that problem.

Different deformation process techniques have been adopted to widen the tube ends with certain widening ratio. A plug maybe inserted as a tool to widen the tube end. This has given a limitation in the maximum achievable widening ratio (the ratio of the final internal tube diameter to the initial internal tube diameter). The process is critical as the quality of the widened tube end depends on several process parameters, such as the plug geometry and its axial load and torque [3].

The use of hydrostatic pressure and a press with ball-shaped tool were also considered in different studies for tube widening [4]. However, these techniques required the use of complex tooling and the amount of load and consumed power were high.

Ball shaped and truncated conical plugs have also been used in conjunction with the rotary forming process, in which the formed elements are rotated during forming [5] [6]. It was established that the rotary process gives high quality products with good geometry shapes. This led to an increase in the use of rotary forging and tube spinning in producing flanges and elongated tubes [7] [8].

The main aim of using different widening tools for tube ends in rotary forming processes was to investigate their effect on the process parameters. The ball shaped tool, for instance, has been used to determine the maximum widening ratio, which can be achieved with the increase in ball diameter and tube thickness. Using these techniques, a maximum widening ratio of 2.0 was achieved [5]. The process axial load, specific widening pressure and its defects have also been given. When the ball rests on bore, a force is required to overcome friction and a sharp increase in load is noticed.

It can be concluded from the previous research that the widening tool shape with the tube rotary has a large effect on the process parameters and calls for selecting a rational tool design method to improve them. The concept of the Constancy of the Ratios of successive generalized Homogeneous Strain increments (CRHS) has been adopted in many metal- forming processes, such as: tube elongation [9], tube piercing [10], and the extrusion of tubes and cans [11–12]. It has been concluded that the designed tools using the CRHS concept give high products quality while minimizing the load required in performing the process. This is referred to as the reduction of the redundant shear strain and consequently of the work redundancy.

The ever-increasing demand for high quality product in metal-forming processes necessitated the establishment of working zone improved geometry by suggesting rational methods of tool design. Nevertheless, the majority of these methods can only deal with specific processes. The CRHS concept showed a substantial reduction in the magnitude of in-homogeneous deformation [13]. The incidence of redundancy shearing strains was estimated and the

* Corresponding author. sshobaki@hu.edu.jo.

correlation between force parameters and the work redundancy of the system was made.

Furthermore, and to show the superiority of the curved dies over the conventional conical dies for a wide range of reductions and frictional conditions, Blazynski *et al.* [11] reviewed many investigations which were based on slip-line field, upper-bound and viscoplasticity analytical methods, to show the superiority of the curved dies over the conventional conical dies for a wide range of reductions and frictional conditions.

The CRHS concept is adopted in this study for the plug, due to its successful application in other metal forming processes and because it produces a curved profile with different rates of deformation.

2. Theoretical Background

The axial plug load and the torque are the main two force parameters that must be considered in any seamless tube manufacturing, using a rotary forming process [10]. Therefore and in order to improve the performance of any of these processes, these two parameters must be minimized as much as possible.

As the plug tool is used to widen the tube end in the rotary tube process, the forward movement of that plug inside the tube leads to increasing the axial plug load to overcome the friction, and dislocate the tube material size which is required to achieve the designed widening ratio of the tube end. This is also what is expected for the process torque parameter. From that, the geometry of the widening plug profile plays an important role in the magnitude of the axial plug load and the torque as the variation of the plug profile affects the contact area between the tube and the plug and the total dislocation distant of the tube material size under widening.

The CRHS concept, which has been applied to determine the widening plug profile, relies on the consideration of the homogeneous strain level (ϵ_H) in (n) equispaced transverse sections of the pass and is defined by the following equation [9]:

$$Z_n / Z_{(n-1)} = (Z_1)^{S(n-1)} \quad (1)$$

Where Z is defined by the physical dimensions of the work piece. Thus $\ln(Z) = \epsilon_H$. S is the rate of deformation unrelated to time. The value of S determines the mode of deformation, and so, $S = 1$ corresponds to a uniform rate of flow (UCRHS), $S > 1$ corresponds to an accelerated rate (ACRHS), and $S < 1$ corresponds to a decelerated rate (DCRHS). Thus there are three sets of plug profiles that can be designed for each value of tube widening ratio used in this research.

Copper metal has been used as a model material to determine the effect of the widening plug profile on the process parameters. Thus is due to reasons of economy and the limitation of laboratory facilities. It is important to note that the basic requirement demanded of a model material is that its mechanical behaviour should resemble, as closely as possible, the prototype material required to be simulated [14] [15]. This technique has solved many problems in estimating the actual work parameters in metal-forming processes.

3. Experimental Setup

To determine the effect of the widening plug profile on the performance of the process, Copper tubes have been used with different wall thickness of 2, 2.5, 3, and 3.5 mm and with a constant internal diameter of 12 mm. Regarding the widening plugs, different plug diameters of 18, 22, 24, 28, 30 and 32 mm have been used to give different widening ratios; 1.50, 1.83, 2.00, 2.33, 2.50 and 2.67. It has to be noted here that the plug profile shape is governed by the constant (S), which determines the rate of deformation for each diameter.

The designed plug profiles of three different rates of deformation (S); 0.9, 1.0 and 1.1 were based on the CRHS design concept. For practicality, a plug advance (The length of the plug profile, X_p) of 12.5 mm has been used [9] [10]. The respective changes in plug profiles and their dimensions are shown in Figure 1 and Figure 2. Plugs specifications are given in Table 1.

To run the experimental part of the research, an automated drill machine with a dynamometer device (Model No. BKM2000 – TeLC Co.) was used. Technical modifications were made in order to facilitate the measurement of the axial load and torque during the widening process of the tube ends with different plug profiles. At the same time, a temperature measurement sensor (TeLC Co.) was used to measure the temperature increase during the operation of widening the tubes end. The process was carried out at a feed rate of 0.1 mm/rev. and a spindle speed of 475 rpm. Oil was used as a lubricant. The experimental process set up is shown in Figure 3.

4. Experimental Results and Analysis

To show the effect of the widening plug profile type on the process performance, several experiments were carried out for the different plug dimensions shown in Table 1 and the process parameters were measured to check the widened portion of the tube and the quality of deformation. The results will be considered in the next five sections.

4.1. Axial Plug Load

The values of this process parameter were obtained using the load measurement Dynamometer. Figure 4 shows a plot of these results for different widening ratios, tube thicknesses and plug profiles. It can be seen that the axial plug load increases with the increase in the widening ratio. This is due to the increase in the contact area between the widening plug profile and the tube material. Table 1 has shown that the surface area of the plug profile is higher as the widening ratio and plug diameter increase. Again, the axial plug load increases with the increase of the tube thickness, since the displacement of the material volume increases in unit time.

Figure 4 also shows that the ACRHS plugs give the lowest values of axial plug load, while the DCRHS plugs profile result in the highest. The reason of that can be explained with reference to Figure 5, which shows the mechanism of the tube metal flow during the widening process using different plug profiles. It can be easily noticed that the lowest metal flow displacement in the deformation zone occurs using the ACRHS plug profile.

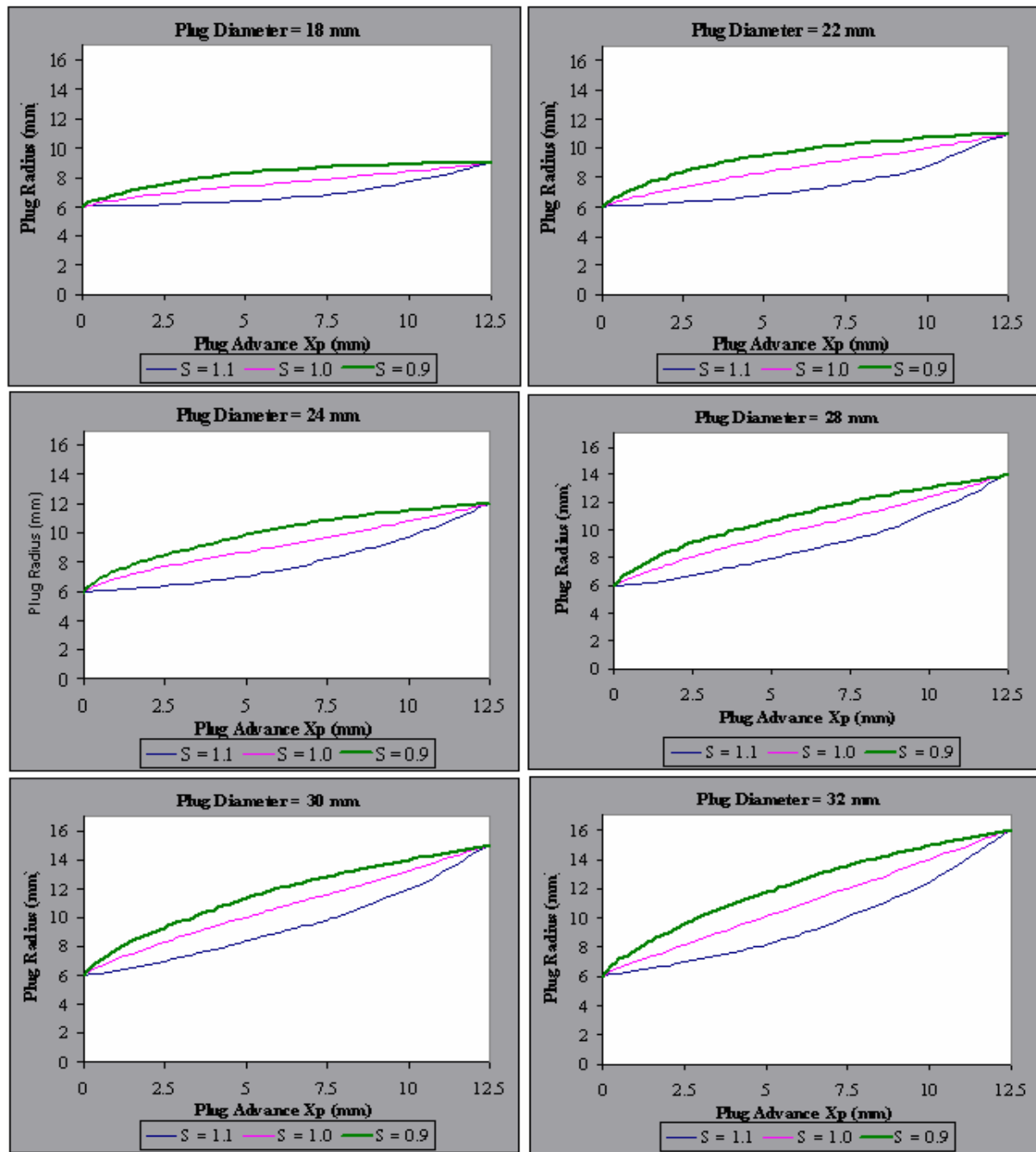


Figure 1. CRHS – Concept design of widening plug profile (S is the rate of deformation).

Opposite to that, the metal flow displacement is highest with the DCRHS plug profile.

The experiments have shown that using the ACRHS plug profile and for the widening ratios of 2.50 and 2.67, the shape of the widened tube end has become a round flattened flange. This is due to the magnitude of the inclination angle of the ACRHS plug profile, as it is increased, at the rear portion, with the increase of the plug diameter. As a result, this pushes the tube metal out of the cylindrical part of the widening plug, as illustrated in Figure 6.

4.2. Widening Torque

The variation of the widening process torque with the different widening ratios, tube thicknesses and plug profiles is given in Figure 7. As seen, the general behaviour of the widening torque increases with the increase of the widening ratio, which is similar to the increases in the axial plug load. This is due to the increase in both of the contact area between the widening plug profile and the tube material, as the plug diameter increases, and to the increase of the plug profile length, as shown in Table 1. Also, the widening torque increases with the increasing of the tube thickness and this is due to

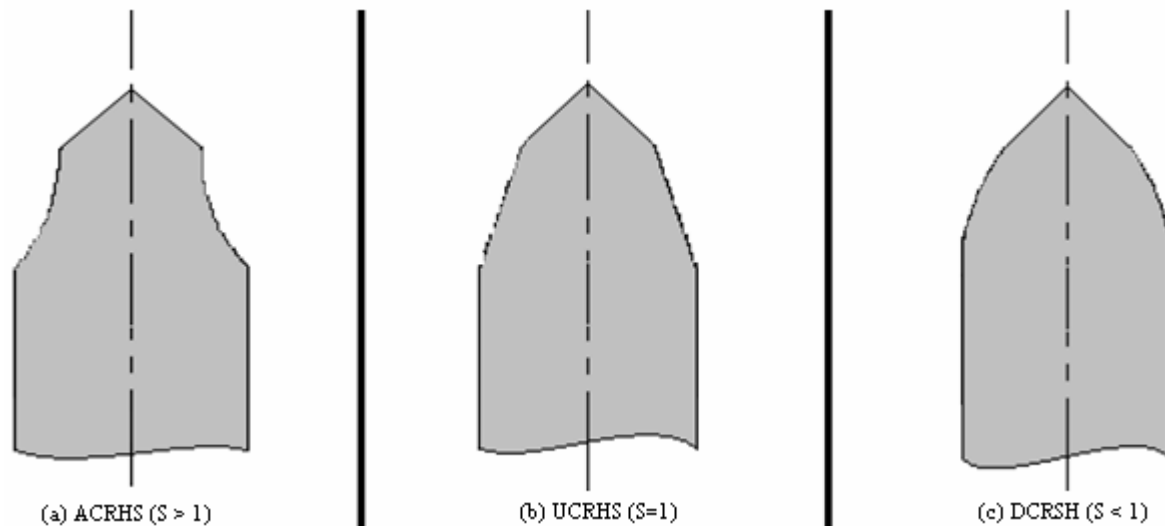


Figure 2. CRHS – Concept design of widening plug profile (S is the rate of deformation).

Table 1 Plug specifications.

Plug no.	Plug type	Deformation rate (s)	Plug diameter (mm)	Profile length (mm)	Plug profile surface area (mm) ²
1	DCRHS	0.9	18.0	15.442	97.028
2			22.0	16.161	101.545
3			24.0	16.597	104.283
4			28.0	17.321	108.831
5			30.0	17.749	111.523
6			32.0	18.178	114.217
7	ACRHS	1.1	18.0	15.090	94.813
8			22.0	15.799	99.274
9			24.0	16.227	101.958
10			28.0	16.940	106.439
11			30.0	17.367	109.117
12			32.0	17.793	111.798
13	UCRHS	1.0	18.0	14.992	94.195
14			22.0	15.698	98.631
15			24.0	16.119	101.276
16			28.0	16.827	105.728
17			30.0	17.250	108.385
18			32.0	17.675	111.057

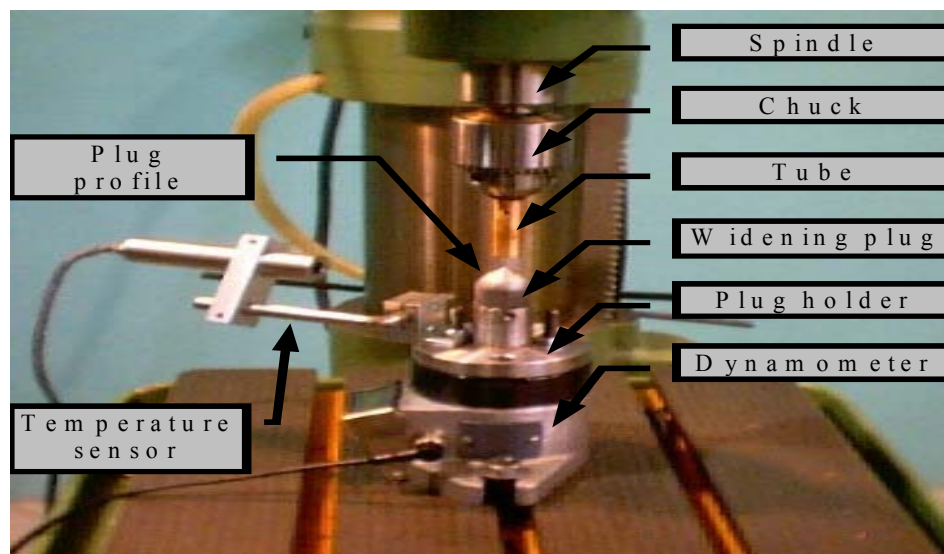


Figure 3. Experimental process set up.

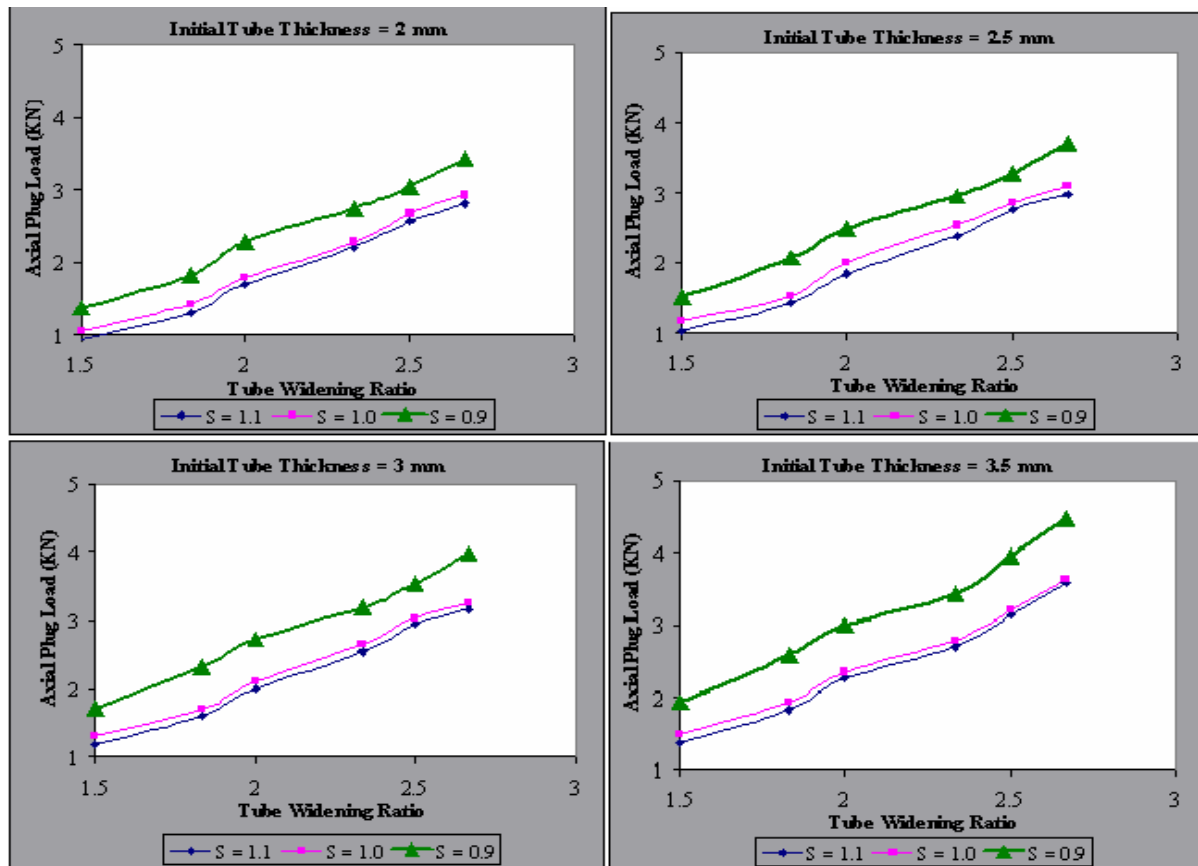


Figure 4. Relation between axial plug load and widening ratio at different tube thicknesses.

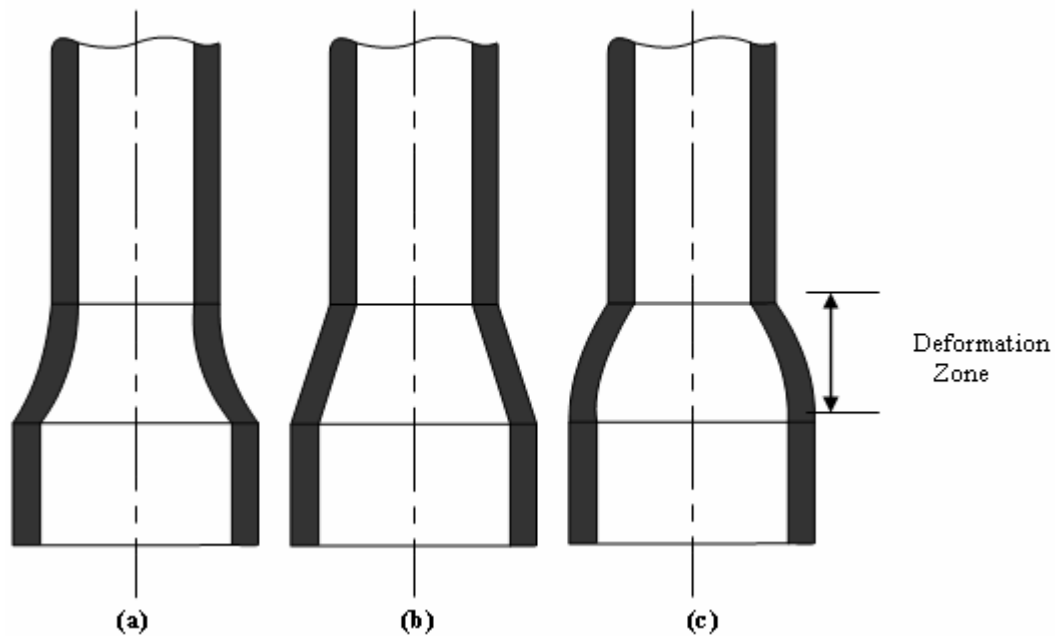


Figure 5. Mechanism of tube metal flow for different plug profiles, (a) ACRHS (b) UCRHS (c) DCRSH.

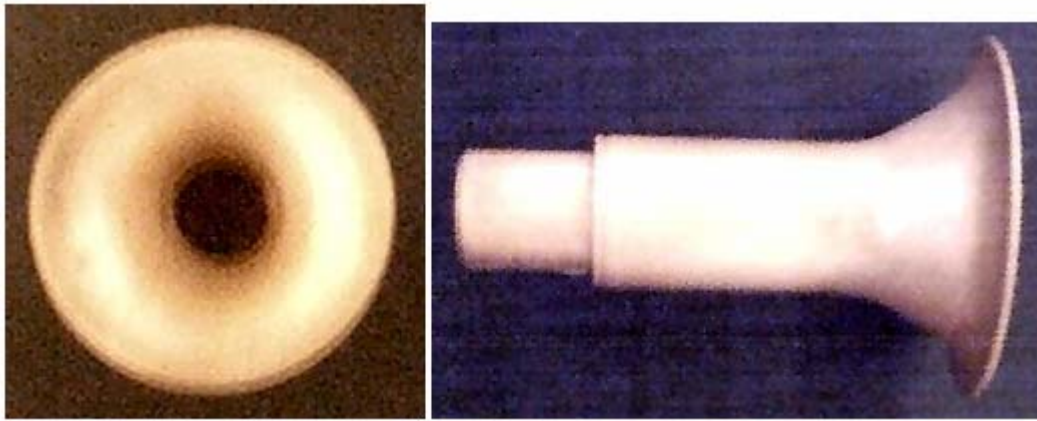


Figure 6. Round flatted flange at the tube ends: Top and side views.

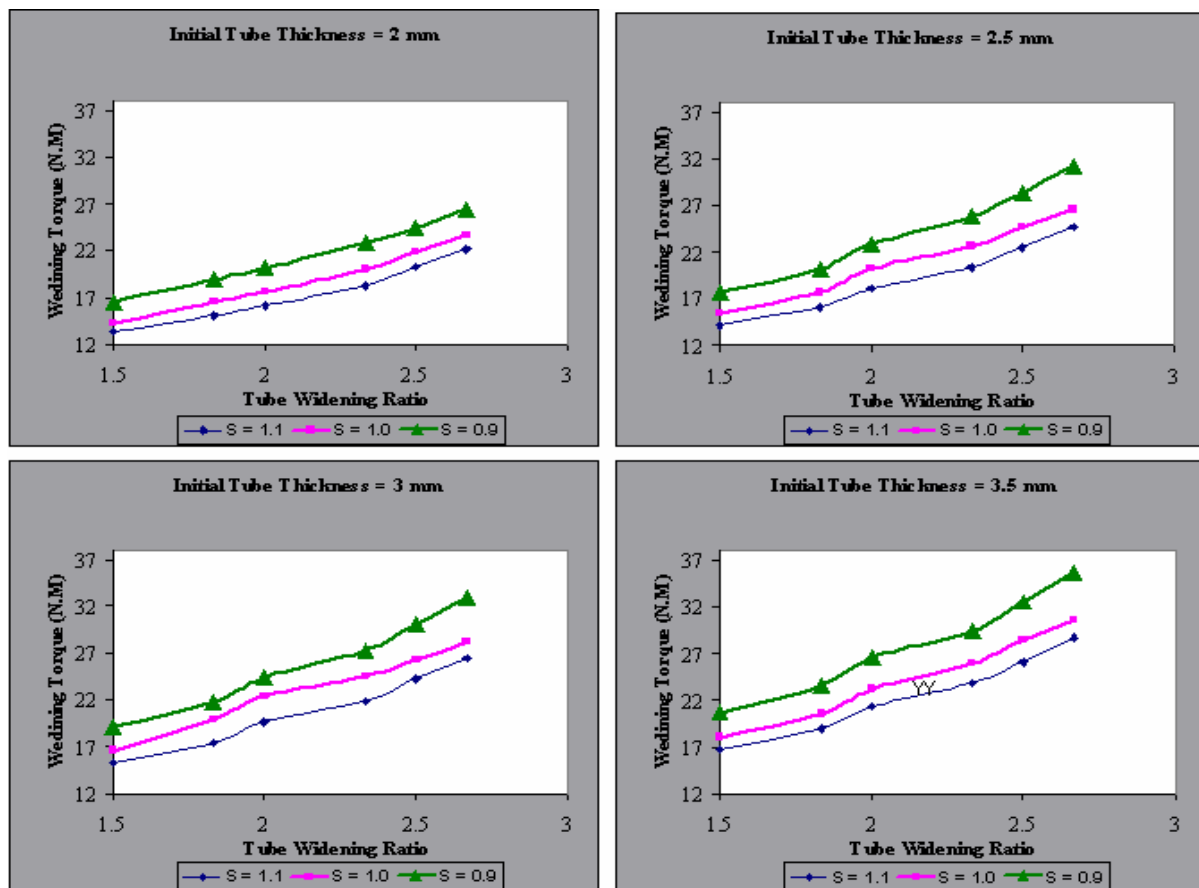


Figure 7. Relation between widening process torque and widening ratio

the same reason of increasing in material volume displaced in unit time.

It can be indicated again from Figure 6 that ACRHS plugs give the lowest level of widening torque and the DCRHS plugs give the highest due to the same reasons mentioned in the analysis of the axial plug load.

4.3. Variation of Temperature

Analysis of the axial plug load and the widening torque showed that ACRHS plugs give the lowest level of axial plug load and torque. The reason for this was that the widened tubes end experienced less metal flow

displacement in the deformation zone. To verify this explanation, the temperature that results from the widening process was measured using a temperature sensor, as shown in the experimental setup in Figure 3.

Figure 8 shows the variation of temperature during the widening process relative to the widening ratios, tube thicknesses and plug profiles. As expected, the ACRHS plugs profiles gives the least amount of temperature increase compared to the other profiles.

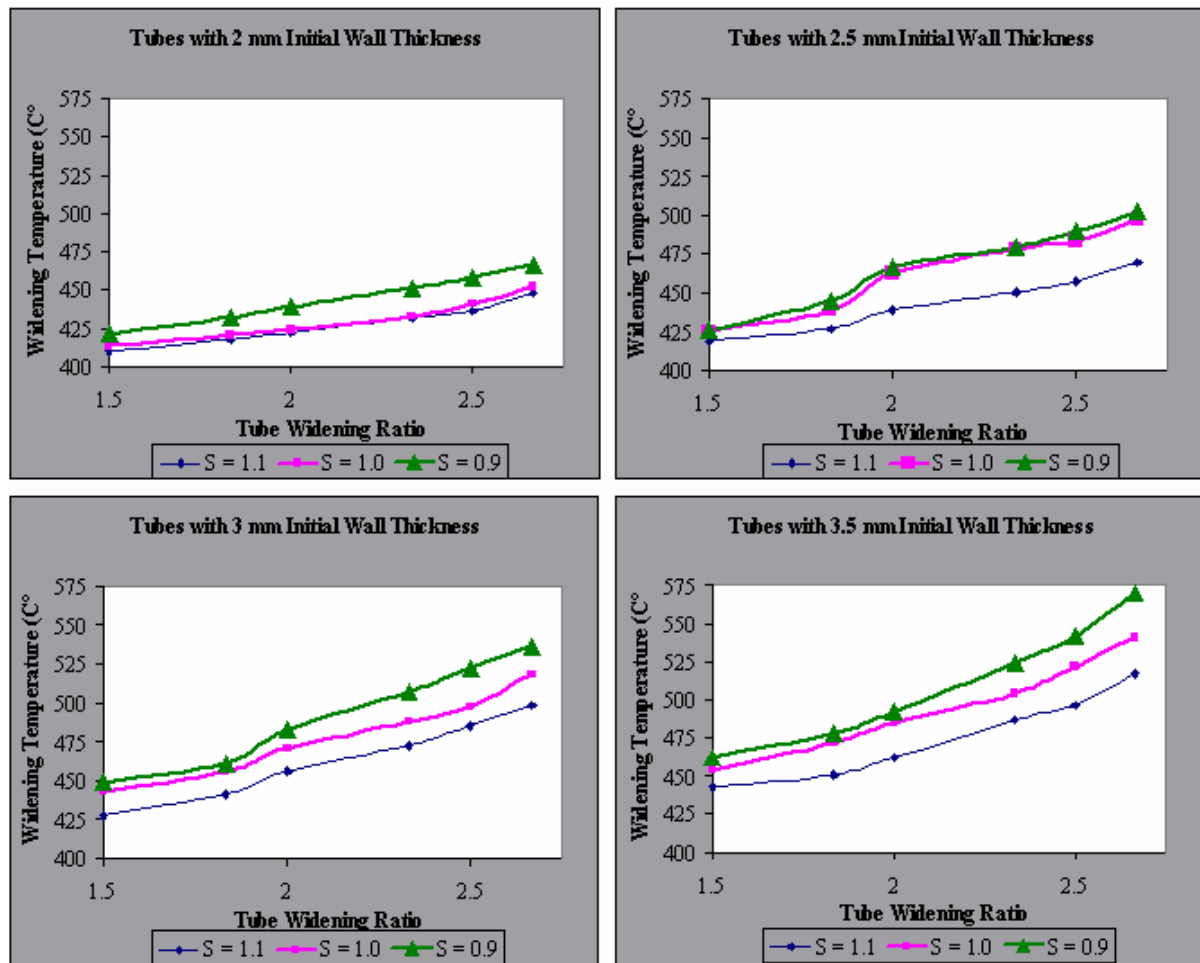


Figure 8. Temperature variation with the widening ratio for different tub thicknesses.

It has been known that one of the many factors that lead to the temperature increase in the deformation processes is the volume of the internal distortion that occurs in the deformed material. It is important to note that with large material displacement, which may be performed in such deformation process, a high quantity of the internal distortion will be deformed. This is the reason why the ACRHS plug profiles give a small increase in temperature during the process due to the small material displacement achieved with those plug profiles.

4.4. Widened Tube Thickness

Figure 9 shows the variation in the final tube thickness relative to the widening ratios, tube thicknesses and plug profiles. Note that the increase of the widening ratios leads to a high reduction in tube thicknesses. This is due to the high tube material expansion i.e. high material stretching or high incidents of strain and as a result of that high reduction in the tube thicknesses.

As expected, the ACRHS plugs profile give less reduction in the tubes thicknesses. This is again due to the less tube material displacement or material expansion in the deformation zone and as shown in Figure 5. This can be clarified from the figure, where for the highest widening ratios, the differences in tube thickness variations relative to the type of the plug profiles are well defined.

4.5. Widened Tube Quality

One of the most important requirements from any use of deformation process is to gain high quality products. This can be verified if the product has, for example, high dimensional accuracy, precise shape, and minimum internal and external defects. These three quality considerations may be simulated in the case of the widened tube ends using three parameters; the roundness of the tube end, the uniformity of the tube end thickness and the formation of tube tearing. Accounting for these quality parameters can minimise the problem and failures in pipes networks.

It has been found that the use of the concept of the Constancy of the Ratios of successive generalized Homogeneous Strain increments (CRHS) for the widening plug profiles give a good effect on these parameters. This can be seen in Figure 10, where all the plug profiles show a perfect round tube ends in addition to a uniform thickness for all tube thicknesses and widening tube ratios. An exception to this case is the case of the ACRHS plug profiles, where a round flattened flange resulted when widening ratios of 2.50 and 2.67 were attempted. Furthermore, no tearing defects occurred during the process with the used widening ratios, even with those widened tubes that have been deformed to a flattened flange.

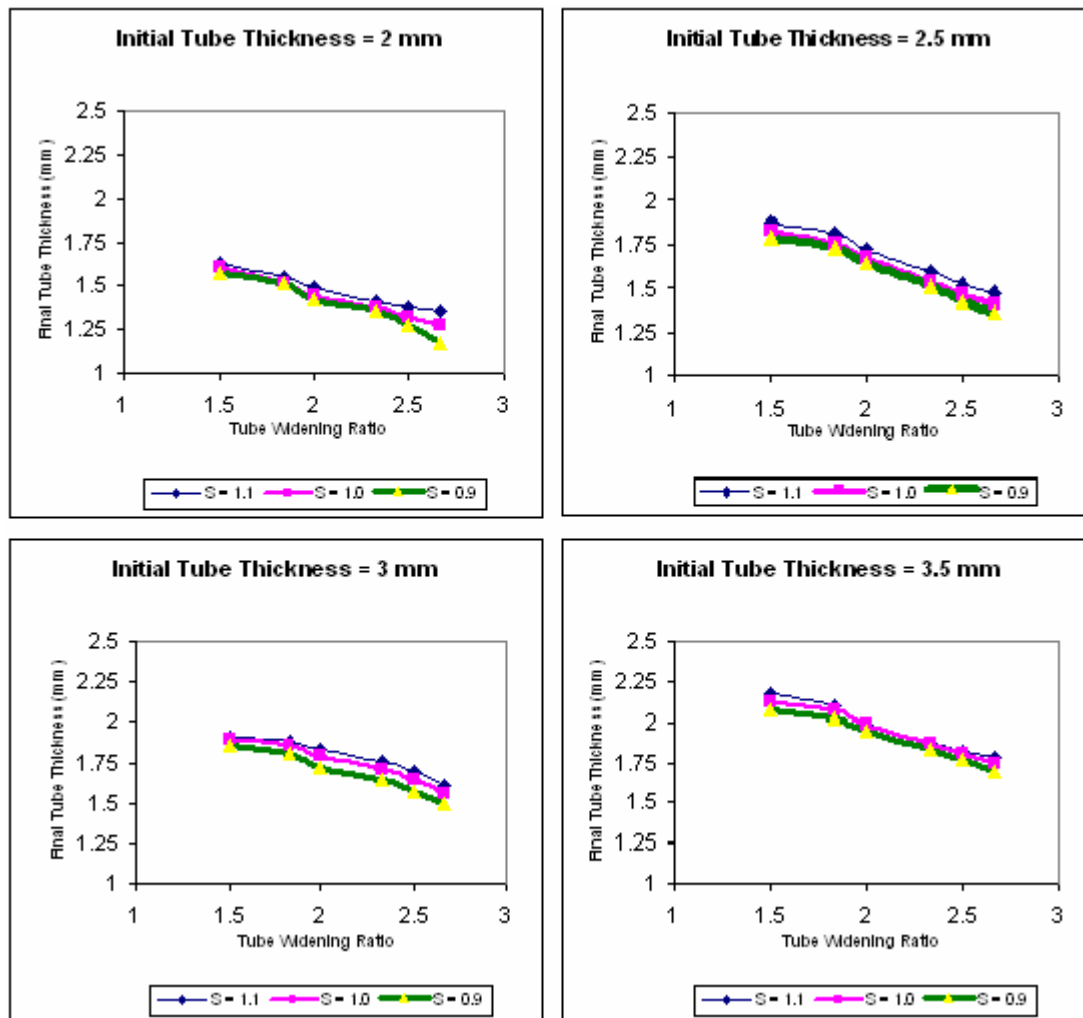


Figure 9. Relation between final tube thickness and widening ratio.



Figure 10. Widened tube ends profile: Top and side view.

5. Conclusions

The experimental results for the widening tubes end process have shown that the process parameters are completely affected by the geometry of the widened plug profile. They have also shown that the use of the concept of Constancy of Ratios of successive generalized Homogeneous Strain increments (CRHS) has significantly improved the performance of the process. It gives a maximum widening ratio of 2.67 under all the conditions selected for the process, which represents a 25% increase in the widening ratio compared to other used techniques. Adding to that, the produced widened tube ends are in high dimensional precision and uniform thicknesses without any tube wall tearing.

References

- [1] Kalpakjian S. Manufacturing processes for engineering materials. 4th ed. Addison- Wesley Publishing Company; 1984.
- [2] Colangelo VJ, Heiser FA. Analysis of metallurgical failures. New York: John Wiley & Sons; 1974.
- [3] Blazynski TZ. Metal forming- tool profiles and flow. 1st ed. London: Macmillan Press; 1976.
- [4] M. Ahmed, M.S. Hashmi, "Aspects of tooling and machine design for the hydrostatic bulge forming of tubular components". Advances in Material Processing Technologies, Vol. 95, 1995, 1455- 1461.
- [5] M. Abdel-Rahman, F.A. Mohamed, S. Zein El-Abden, "Rotary Widening of aluminum tube ends on the lathe using a ball- shaped tool". Proceeding of the 1st Int. Industrial Eng. Conference, Amman, Jordan, 2001, 145- 154.
- [6] O.I. Ahmed, M.N. El-Sheikh, A.A Ibrahim, "An investigation into flaring of tube ends using truncated conical plugs". Bull Fac Eng. Assuit University, Vol. 17, No. 2, 1989, 47- 52.
- [7] S. Choi, K.H. Na, H.J. Kim, "Upper bound analysis of the rotary forging of a cylindrical billet". J. Mat. Process Tech., Vol. 67, 1997, 78- 82.
- [8] Y. Can, A.N. Bramley, "An analysis of power requirements in tube spinning". Advances in Manufacturing Technology, Proc. 12th Conference on Manufacturing Research, University of Bath, A. Bramley ed., 1996, 54- 60.
- [9] Blazynski TZ. Design of tools for deformation processes. London: Elsevier Applied Science Publishers; 1986.
- [10] A.K.A. Al-Dahwi, T.Z. Blazynski, "Inhomogeneity of flow, force parameters and pass geometry in rotary cone-roll tube piercing". Mat.-wiss.u. Werkstofftech, Vol. 23, 1992, 29- 38.
- [11] T.Z. Blazynski, R.S. Lee, "Pass design and redundant strains in forward tube extrusion". J. Mech. Working Tech., Vol. 9, 1984, 313- 324.
- [12] A.K.A. Al-Dahwi, A. Shahab, "Rational die design and the correlation of the strain redundancy- the relative die pressure in forward extrusion of Cans". Proceeding of the 1st Int. Industrial Eng. Conference, Jordan, 2001, 155- 165.
- [13] A.K.A. Al-Dahwi, A. Shahab, "Pass profile and the assessment of its effect in forward extrusion of Cans". J. of Materials Processing and Manufacturing Science, Vol. 10, No. 2, 2001, 117- 124.
- [14] T. Altan, H.J. Henning, A.M. Sabroff, "The use of model materials in predicting forming loads in metal-working". J. Eng. Ind., Trans. ASME (B), Vol. 92, 1970, 444-452.
- [15] R.S. Lee, T.Z. Blazynski, "Mechanical properties of a composite wax model material simulating plastic flow of metals". J. Mech. Work. Technology, Vol. 9, 1984, 301-312.



الجامعة الهاشمية



المملكة الأردنية الهاشمية

المجلة الأردنية
للمهندسة الميكانيكية والصناعية

JJMIE

مجلة علمية عالمية محكمة

<http://jjmie.hu.edu.jo/>

ISSN 1995-6665

المجلة الأردنية للهندسة الميكانيكية والصناعية

مجلة علمية عالمية محكمة

المجلة الأردنية للهندسة الميكانيكية والصناعية: مجلة علمية عالمية محكمة أسستها اللجنة العليا للبحث العلمي في وزارة التعليم العالي والبحث العلمي، الأردن، وتصدر عن عمادة البحث العلمي والدراسات العليا، الجامعة الهاشمية، الزرقاء، الأردن .

هيئة التحرير

رئيس التحرير:

الأستاذ الدكتور موسى محسن
قسم الهندسة الميكانيكية، الجامعة الهاشمية، الزرقاء، الأردن .

الأعضاء:

الأستاذ الدكتور عدنان الكيلاني الجامعة الأردنية	الأستاذ الدكتور بلال العكش الجامعة الهاشمية
الأستاذ الدكتور أيمن المعاينة جامعة مؤتة	الأستاذ الدكتور علي بدران الجامعة الأردنية
الأستاذ الدكتور محمد النمر جامعة العلوم والتكنولوجيا الأردنية	الأستاذ الدكتور نسيم سواقد جامعة مؤتة

مساعد رئيس هيئة التحرير:

الدكتور أحمد الغندور
الجامعة الهاشمية

فريق الدعم:

تنفيذ وإخراج

م. أسامة الشريط

المحرر اللغوي

الدكتور وائل زريق

ترسل البحوث إلى العنوان التالي :

رئيس تحرير المجلة الأردنية للهندسة الميكانيكية والصناعية
عمادة البحث العلمي والدراسات العليا
الجامعة الهاشمية
الزرقاء - الأردن

هاتف : 3903333 00962 5 فرعي 4147

Email: jjmie@hu.edu.jo

Website: www.jjmie.hu.edu.jo