ISSN Print : 0377 - 2969
ISSN Online : 2306 - 1448
Vol. 53(2), June 2016

PROCEEDINGS
OF THE PAKISTAN ACADEMY OF SCIENCES:
A. Physical and Computational Sciences

PAKISTAN ACADEMY OF SCIENCES
ISLAMABAD, PAKISTAN

Proceedings of the Pakistan Academy of Sciences, published since 1964, is quarterly journal of the Academy. It publishes original research papers and reviews in basic and applied sciences. All papers are peer reviewed. Authors are not required to be Fellows or Members of the Academy, or citizens of Pakistan.

Annual Subscription: **Pakistan:** Institutions, Rupees 2000/- ; Individuals, Rupees 1000/-
Other Countries: US$ 100.00 (includes air-lifted overseas delivery)

HEC Recognized, Category X; PM&DC Recognized

Published by **Pakistan Academy of Sciences**, 3 Constitution Avenue, G-5/2, Islamabad, Pakistan
Tel: 92-5 1-9207140 & 9215478; **Fax:** 92-51-9206770; **Website:** **www.paspk.org**

Printed at **PanGraphics (Pvt) Ltd.,** No. 1, I & T Centre, G-7/l, Islamabad, Pakistan
Tel: 92-51-2202272, 2202449 Fax: 92-51-2202450 E-mail: pangraph@isb.comsats.net.pk

PROCEEDINGS
OF THE PAKISTAN ACADEMY OF SCIENCES:
A. Physical and Computational Sciences

CONTENTS

Volume 53, No. 2, June 2016 **Page**

Research Articles

Submission of Manuscripts: Manuscripts may be submitted online as E-mail attachment. Authors must consult the ***Instructions for Authors*** at the end of this issue or at the Website: www.paspk.org

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 111–130 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Deformable Registration Methods for Medical Images: A Review Based on Performance Comparison

Fakhre Alam[1*], Sami Ur Rahman[1], Adnan Khalil[1], Shah Khusro[2], and Muhammad Sajjad[3]

[1]Department of Computer Science and Information Technology,
University of Malakand, Malakand, Pakistan
[2]Department of Computer Science, University of Peshawar, Peshawar, Pakistan
[3]Department of Computer Science, Islamia College University Peshawar, Peshawar, Pakistan

Abstract: Deformable registration methods are widely used for the accurate registration of objects with large-scale deformation. In this paper, we present a detail review on performance analysis of deformable registration methods. We comprehensively review each registration method and describe its features, advantages, issues and challenges. Deformable registration methods are further quantitatively compared and evaluated based on a set of criteria, which estimate the performance of each method. The performance of registration methods is estimated using root mean square error (RMS), mutual information (MI), computational time complexity and memory requirement. It is found in our analysis that every registration method has its own strength to register deformable objects. However, due to large-scale variations in deformable objects most of the registration methods are not still a perfect choice in clinical applications. Therefore, advanced and powerful registration methods are needed to develop in future, which can precisely, efficiently, and automatically register medical images with large-scale deformations.

Keywords: Medical image processing, image registration, deformable registration, imaging modalities

1. INTRODUCTION

In the last two to three decades, tremendous growth has been observed in the area of medical image analysis due to the development of automated, efficient, accurate, and non-invasive devices. Medical imaging [1-4] is used in several applications in the clinical setting such as diagnostic setting, planning and procedures. The role of medical imaging is now not only limited to simple visualization and inspection of anatomic structures but it is also used as a tool for surgical and radiotherapy planning, inter and intra-operative navigation and for tracking the progress of disease [5].

The rapid advancement in this challenging area has driven the need for sophisticated pre-processing techniques. Image registration [6-18] is the most important step in medical image processing in which one to one geometric correspondence between source and target images is established. The source image is fixed image, which remain unchanged during registration process while the target image is the moveable image super-imposed on the source image. In image registration, both images represent the same organ/tissue obtained from either a same modality but with different time frames/ angles or by different imaging modalities. The basic purpose of establishing geometric correspondence between source and target image is to obtained useful and complementary information. This is done by quantitatively comparing their features using sets of parameters which includes feature detection, feature matching, transform model estimation and image re-sampling and transformation [19-21].

Recent developments and wide spread use in imaging modalities [9, 10, 22-24] provides an easy to use platform for the radiologists and surgeons to obtain useful information from human anatomy. However, each modality show different types of information i.e. either anatomical (showing mainly morphology) or functional (showing mainly information on the metabolism of the

Received, January 2016; Accepted, June 2016
*Corresponding author: Fakhre Alam; Email: fakrealam@uom.edu.pk

fundamental anatomy). Modalities such as X-rays, magnetic resonance imaging (MRI) , computed tomography (CT) and ultrasound (US) [25] are used to obtain anatomical information while single-photon emission computed tomography (SPECT) and positron emission tomography (PET) are used to extract functional information. However, both MRI (DWI, DCE-MRI, and MRSI) and US (Elastography, contrast enhanced US) can also be used to obtain functional information. The proper integration of useful information from two separate images taken with different types of modalities is often required in the clinical tracks of events. This is done by the registration process which brings the separate images obtained from different modalities into special alignment and integrates useful information from them.

Deformable registration is widely used in computer assisted surgery and radiotherapy for the accurate voxel by voxel mapping of medical images with large-scale local and global deformations. Deformable registration also enhances the planning, execution and evaluation of surgical procedures [26]. In deformable registration, a special association between source and target image is established during transformation. The correspondence between transformations signals are usually performed locally in a non-linear and dense fashion [27].

Deformable registration is one of the best choices for the analysis of medical images obtained either by the same or by different imaging modalities having high degree of functional and anatomical variability. Deformable registration algorithms either operate on images features such as lines, counters and points/ landmarks or on their gray levels i.e. directly on pixel or voxel data [28]. Algorithms belong to deformable registration can successfully determine the local differences in the anatomy and accordingly resolve them.

One of the main challenges today for deformable registration methods are how to properly validate them on clinical data. The lack of adaptation in clinical workflow is due to their limited availability and high computational requirements [29]. Several other challenges includes recovering a local transformation that align two signals that have a non-linear relationship, proper alignment of tissue having sudden change in volume, registering poor and non-diagnostic quality images, and designing image similarity for multi-model scans.

The importance and popularity of deformable registration have led to several survey papers, which are listed in Table 1, together with the publication years and topics. In general, each paper covers only a subset of the topics in deformable registration. For Example, the work of Sotiras et al. [27] is one of the comprehensive review on

Table 1. Surveys on deformable registration.

Year	Reference	Topic
1996	[5]	Deformable Models in Medical Image Analysis: A Survey
2008	[97]	Objective assessment of deformable image registration in radiotherapy: A multi-institution study
2010	[90]	Implementation and evaluation of various demons deformable image registration algorithms on a GPU
2011	[98]	Deformable Medical Image Registration: Setting the State of the Art with Discrete Methods
2013	[99]	Evaluation of various deformable image registration algorithms for thoracic images
	[27]	Deformable Medical Image Registration: A Survey
	[10]	Survey of Medical Image Registration
2015	[30]	Evaluation of various Deformable Image Registrations for Point and Volume Variations

deformable medical image registration in which they attempt to give an overview on the recent advances in the field and evaluate each component of deformable registration. Although the work of Sotiras et al provide in-depth systematic review but they not cover the quantitative evaluation of deformable registration methods. Similarly the review of Mani et al. gives a short overview and the pros and cons of several types registration methods including deformable registration [10]. The study of Han et al. [30] evaluated accuracy of various DIR algorithms using variations of the deformation point and volume. Mclnerney and Terzopoulos [5] present a detail review on the development and application of deformable models to problems of fundamental importance in medical image analysis, such as segmentation and registration.

We have performed the experiments on the datasets of two 2D rat lung images, the one after inspiration of air into lungs (source image) and the second after exhalation (target image). The size of each image is 128 * 126 and the physical spaces are one millimeter along x axis and one millimeter along y axis. The datasets and parameters are obtained from [31].

The main scope of this paper is focused on the quantitative evaluation on deformable registration methods. Furthermore, recent developments and challenges are also analyzed. The evaluation parameters are outlined in Fig 1 which includes mutual information (MI), root mean square error (RMS), computation time and memory space occupied during execution. We have found the accuracy and efficiency of each deformable registration method with the help of these parameters.

The main contribution of this review are as follows.

- Deformable registration methods which includes finite element model (FEM) based, BSplines, level set motion, BSplines multi-grid, warping with kernel Splines, warping with BSplines, asymmetric demon and symmetric demon deformable registration methods are discussed in a clearly organized manner, and their performance are shown.
- To examine the state of the art, each registration method involved in medical image processing are discussed in detail. The merits and limitations of each method is summarized. Our main focus is on the detail estimation of accuracy and efficiency because the performance of deformable registration is not been surveyed previously using these parameters.
- We discuss the future work on deformable image registration methods
- This work attempts to provide a theoretical foundation and compact platform for researchers by evaluating the important aspects of deformable image registration methods.
- It will also help clinicians by providing relevant and quantitative information on diagnostic, surgical and treatment planning which will eventually improve their knowledge on this challenging area of research.
- The above mentioned contributions clearly distinguish our survey from the existing surveys on deformable registration methods. To our knowledge, our survey is the broadest.

The remainder of this paper is organized as follows: Section 2 briefly reviews the work related to image registration. Section 3 categorize and presents the methods for deformable registration in detail. Section 4 analyzes the performance of each method. Section 5 summarizes this paper.

2. MEDICAL IMAGE REGISTRATION

Image segmentation and registration are the two main areas of medical image processing. Image segmentation divides an image into different segments of interest while image registration establishes a one to one correspondence between two different images of the same organ. Literature study dictates that image segmentation and registration are the most challenging areas of medical image processing and a lot of research work is available on it [5, 8-10, 19, 20, 22, 32-39]. Fig 2 show the process of registration of human head 3D images acquired using CT and MRI scanners. In the registration process, source CT image which is the most suitable to represent anatomical information i.e. bones is mapped on MR

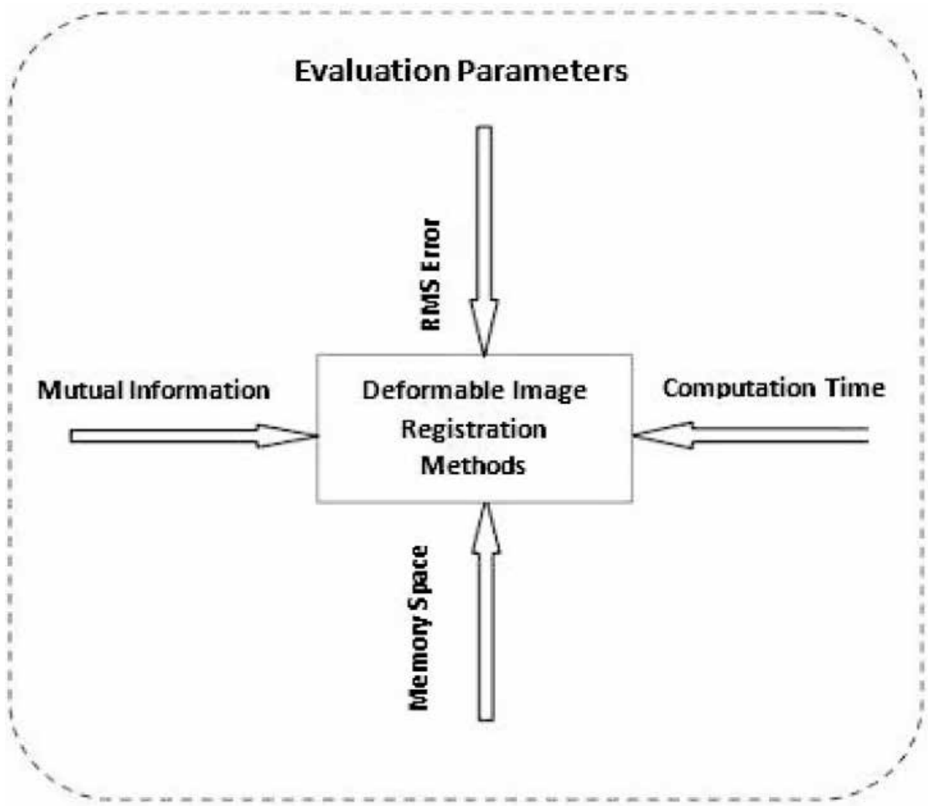


Fig. 1. Evaluation parameters for deformable registration methods.

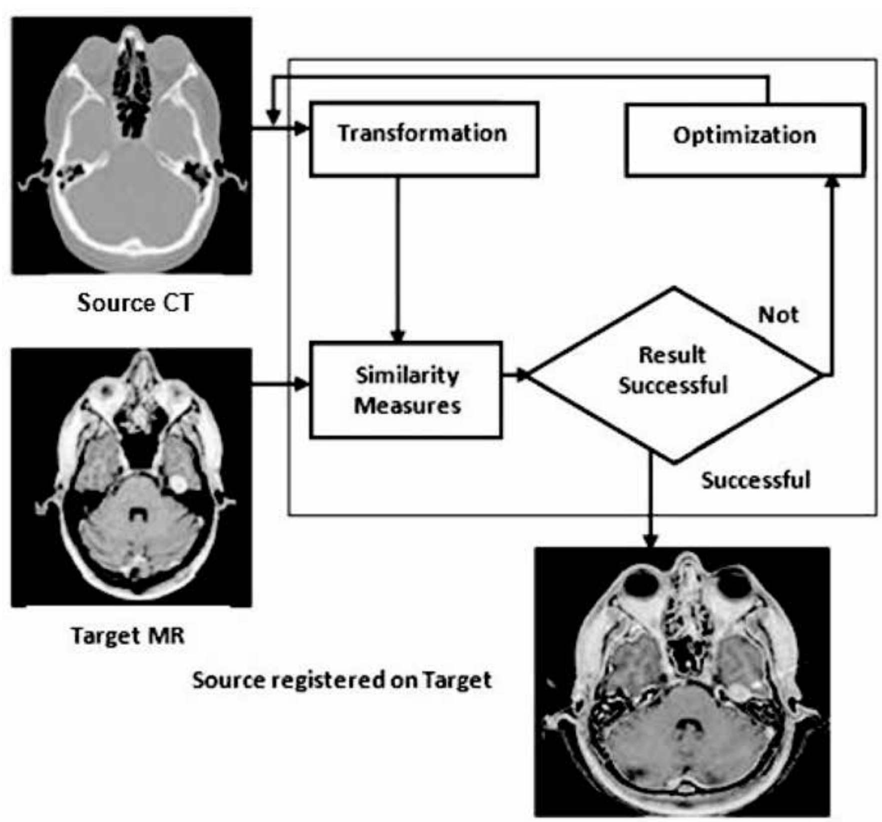


Fig. 2. CT and MR registration of brain images.

image which shows functional information such as soft tissues and tumor [40, 41]. It is shown in Fig 2 that the registered image provide more useful information not only about the tumor but also about the soft tissues and bony structures.

It is also shown in Fig. 2 that the process of registration is based on three important steps i.e. transformation, similarity measures and optimization techniques. The steps involved in registration processes are interrelated and iterative. In the transformation step the points or landmarks in source image space is mapped to the corresponding target image space [42]. A transformation function maps source image to target image considering image dimensionality, accuracy, and computational speed [43, 44]. Similarity measure, is another important step in image registration because it can precisely estimate the correspondence between source and target images. Similarity measure estimate the degree of matching between source and target images and it is based on pixel intensities and patterns, cross-correlation, anatomical structures and mutual information [20, 43, 45-47]. Angle of view, time interval and sensor used are the essential parameters for estimating the similarity measure of input images. It is important to choose best similarity measures while dealing medical images taken from human organs with constant variation and movements during the course of time.

As mentioned earlier that registration is an iterative process, at every iteration similarity is checked between source and target image. If the similarity is not according to the requirement of successful registration then the process is optimized [42, 43, 48-50] to further find the best alignment between source and target image as shown in Fig 2. In this procedure, similarity parameters obtained from the earlier steps are updated (either increased or decreased) tell the optimum values. Several types of optimization methods with pros and cons are available for the registration of medical images. The popular among them includes quasi-Newton optimization, evolutionary strategy, genetic algorithm, stochastic approximation, iterative closest point, powell's method, downhill simplex method, steepest gradient descent and the conjugate gradient method [43, 44, 51-55].

3. DEFORMABLE REGISTRATION METHODS

Deformable registration is a fundamental technique for the analysis of mono and multi-modal images of deformable organs such as heart, lungs, breast and kidney. Deformable organs naturally show consistent deviations due to breathing and movement. Therefore, the precise identification and localization of potential tumor tissue is difficult and challenging. Furthermore, several other issues such as the proper identification of both anatomical and functional contents and automatic voxel-by-voxel transformation are also successfully done using the advance methods of deformable registration. Fig 3 [26] shows the registration of multi-modal CT and PET images of unresectable pancreatic cancer. In the Fig , the anatomical contents e.g. bones and hard tissues are apparent in the CT image (A) while functional contents such as metabolism are clear in the PET image (B). However, the pancreatic tumor is not clearly visible in both A and B. The registered image (C) at the bottom contains the properties of both CT and PET and thus the tumor is more visible as indicated by the arrow.

Beside multi-modality images, the analysis of pre and post interventional images obtained from mono modality is also important for treatment success. Such types of mono modal images obtained at different time are also precisely analyzed with deformable registration methods. However, all the methods use different types of operations to perform registration on the set of 2D and 3D images, which might be good for some situation but not necessarily suitable for others [56].

Deformable registration methods use several types of complicated models to estimate the internal behavior of deformable tissues. Among them, some of the popular models are finite element model (FEM), elastic model, viscous fluid model and radial basis function [57, 58]. FEM decomposes images into several desirable regions containing soft tissues, create volume meshes for particular tissue and allocate a precise tissue property for volume meshes. FEM is a powerful computational tool applicable to several types of deformable registration algorithms. Elastic model is also a useful physical model for the proper transformation

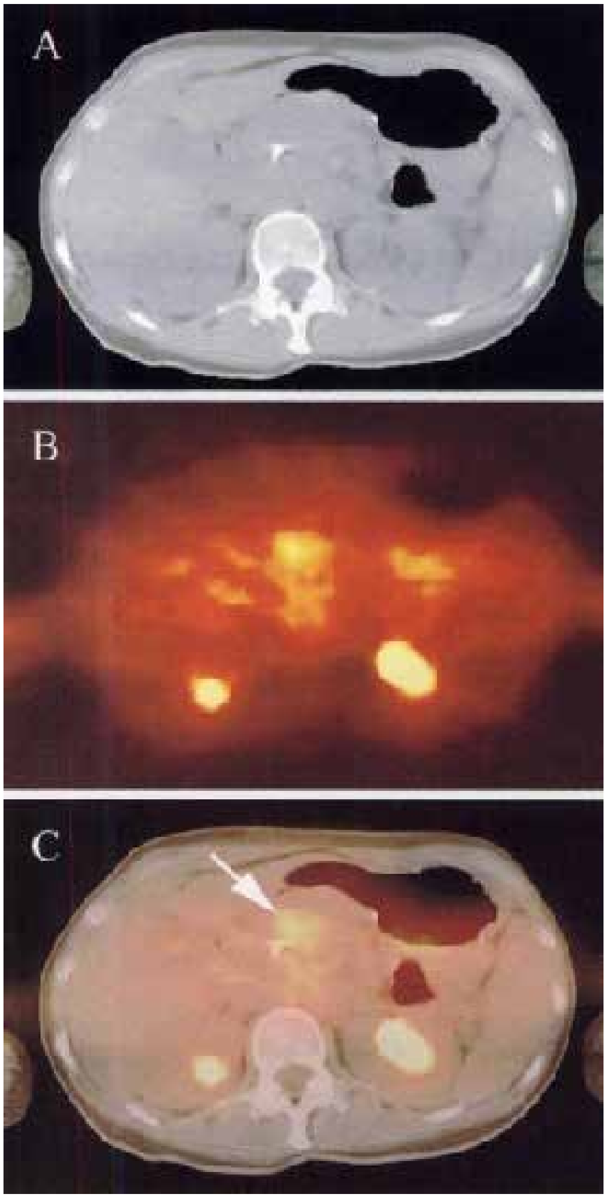


Fig. 3. Deformable registration of multi-modal CT and PET images of the abdomen.

of detailed local deformations during registration process [27, 59]. This model considers images as deforming elastic objects and estimates the small differences between them using external body forces. Deformable registration using elastic model can give robust and fast results and can also efficiently describe multimodal deformation. Contrast to elastic model which is best suited for small deformation in the tissues, viscous fluid model is the right choice for image transformation when the number of deformation is more in the tissues [60]. However, high computational time complexity of viscous fluid model is the main reason behind its less widespread use and popularity. Radial basis function is another model for the estimation of local differences and point correspondence in deformable registration. Radial basis function is mostly used for the interpolation of images with local distortion and differences [61]. This function minimizes dissimilarity between source and target image and provide a smooth resultant image with more clear information.

Registration methods belong to deformable models are more complex as compared to registration methods based on rigid models. Therefore, to further improve the performance and reduce computational complexity, researchers are trying to develop more effective deformable registration methods from time to time. Therefore, there is a need to further investigate deformable registration methods with high performance and that might be suitable in almost every type of scenario. Recently several types of deformable registration methods have been developed as shown in Fig 4. These registration methods are further discussed in the subsections below and their quantitative evaluation and comparison is described in section 4.

3.1 FEM-Based Image Registration

Several types of physical and biomechanical models are available for the analysis of deformable registration. Finite element method (FEM) is the most popular model among them which successfully compute the biomechanical properties of human tissues and special positions of anatomy [62]. FEM effectively estimates large local deformations in the various tissues during registration process. This model treats human organs images as elastic bodies and applies non-uniform meshes to the important features in them which improve the accuracy of registration [62-64]. In order to generate registered image and solve linear system of equations, FEM uses linear elasticity and static analysis assumptions. Efficient modeling of material properties in object and providing better global solution to the entire image domain are also the main features of FEM. However, the accuracy of registration should be considered while estimating material properties of object because using FEM in deformable registration the accuracy rely on material properties.

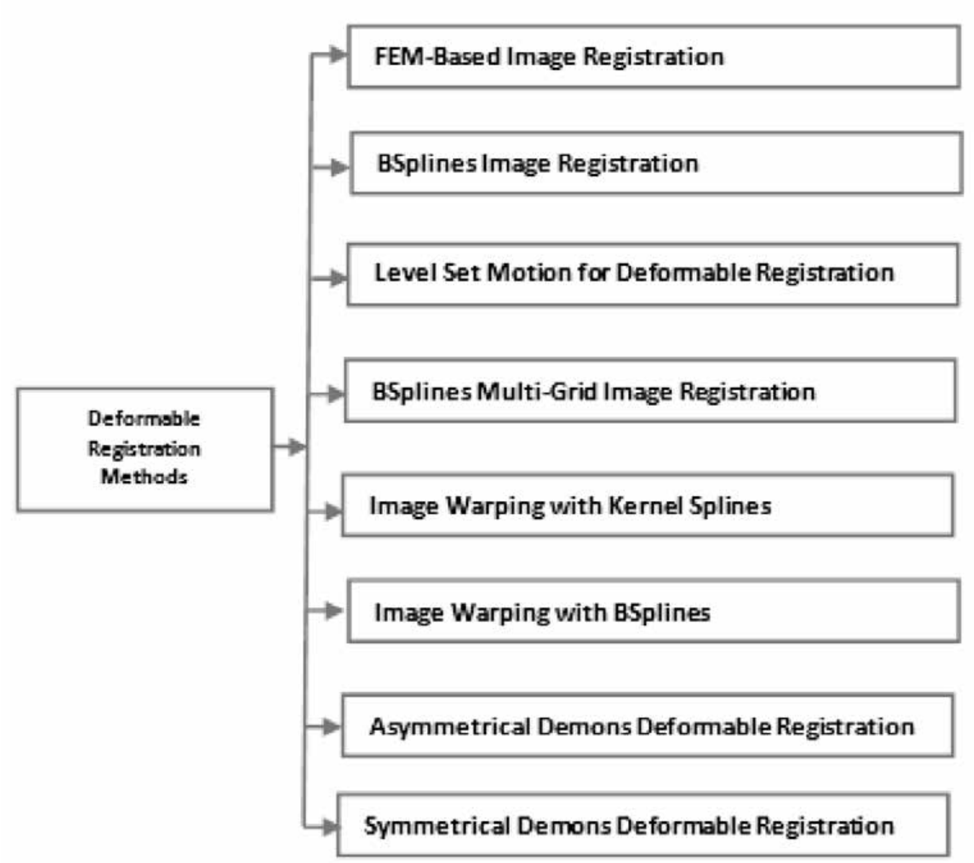


Fig. 4. Deformable registration methods.

The basic idea behind finite element method is to discretize image domain in groups and apply mechanical or physical forces on them. It is due to these forces that the relations between different types of tissues are formed which help to estimate deformations and obtained tissue properties from the estimated deformations. The important applications of FEM are in brain modeling and simulation during image guided surgery, physical integration of deformable registration methods, analyzing mechanical deformations during biopsy examinations and testing, reconstruction and improvement of elastic properties in deformable tissues such MR breast, lung and prostate images. Configuration of boundary conditions is another important feature of FEM in deformable registration [65]. However, due to complexity in the geometries of human anatomical structures FEM faces several difficulties while estimating boundary conditions during image-guided radiation therapy. Therefore, the computational complexity is more that affect the speed and efficiency of registration method. FEM based deformable registration methods are mostly used in a situation where the high accuracy is desired. Registration results are also better in case of low contrast tissues because FEM provides several mechanisms based on elasticity and biomechanics to align low contrast tissues with high one.

3.2 B-Splines Image Registration

Sampling frequency changes, proper estimation of un-known pixel values from known pixel values, transformation of a digital image to analog image and increase in the resolution are the necessary steps in image processing. These important steps in image processing are performed by the techniques of interpolation. Basis Splines (B-Splines) are widely used piecewise continuous interpolation functions for the analysis of digital images. B-Splines functions are based on polynomials and are spread over an interval of unit width in which the number of pieces is preoperational to the order of splines [66-68]. These functions efficiently detect and compute image shapes, curves and contours having large variations and transform them into continuous images. Registration using B-Splines model also generate several types of non-linear elastic deformations. B-Splines execute images locally and with multi-scale processing due to their excellent localizations and multi-resolution properties.

Deformable image registration provides flexible mapping of local key points between source and target images due to high degree of freedom in transformation. The transformation become more smoother when used with B-Splines functions because of their fast interpolation schemes , generation of a vector field on interested volume and provision of compact support [69, 70]. In deformable transformation using B-Splines interpolation, several coefficient values are defined and distributed on a grid at continuous interval due to which wide variety of deformation become possible. However, the use of more coefficient values increases the running time of transformation and effect the efficiency of registration.

3.3 Level Set Motion for Deformable Registration

Level set motion is another important and widely used framework to represent deformable objects. Level set uses global regularization of its shapes while dealing with 3D interfaces and automatically detect the boundaries of interested regions during image segmentation and registration [71]. Level set methods give promising results while dealing with deformable models because it represents the models as 3D images. In the 3D images, the intensity at each voxel is considered as distance measurements to object surface. In the distance measurements, object internal values are negative and the outer values are positive [72-74]. In deformable registration using level set motion algorithm, the histogram of source and target images are matched with user specified number of quantile values. In other words, during registration, intensity values are same at both images while representing the same homologous points on an object with pixels.

Medical image registration and segmentation with level set motion give successful results and are more popular than other types of techniques such as mesh. Robustness in noisy conditions, changing and extracting curved objects with complex topology and multidimensional implementation/ computation of motion of the interface with simple and compact mathematical notation [73, 74]. However, level set motion framework has several limitations and need more improvements. Mesh framework such as FEM has built-in ability to track vertices but level set motion framework is no capability to do so. More features are lost in level set motion framework because resembling is performed at each time step with low pass filter. Moreover, high computation time, less stability to multi-resolution images and organs anatomy with high variations are some other limitations of deformable registration using level set motion which requires further improvement.

3.4 B-Splines Multi-grid Image Registration

Modifying grid of control points and increasing similarity measures are the main steps in registration using B-Splines approach to deform an image. In order to obtain more intensity values and improve the robustness of registration, a multi-resolution approach is followed. In this approach, a grid containing multiple resolution levels i.e. low and high are placed on the images [75-77]. Registration process based on B-Splines multi-grid interpolation is performed in two steps. In the first step, images with low resolution levels are estimated and transformed. In the second step, the transformation is again performed on the images with high resolution values propagating estimated parameters to them which are obtained in the first step.

Registration methods based on B-Splines multi-grid/ multi-resolution approache perform better then single-grid/ resolution approach. The speed and efficiency of registration method is enhanced by mapping the spatial resolution of the underlying image model to the step size of the registration method. Performing maximum iterations on coarsest resolution is also the unique property of B-Splines multi-grid registration. Avoiding local minima due to smoothing effect of pyramid and well handling of local and global errors and invert-ability in deformable objects are some other features of B-Splines multi-grid registration. On the other side, B-Splines multi-grid registration gives weak results while operating on images with high distortion and rotational deformations [78-80]. Moreover, this method is also unsuitable in mesh refinement procedures with large scale adaptive procedures due to its high setup cost.

3.5 Image Warping with Kernel Splines

In image warping, a mapping function is applied on digital image which make distortion in it and as a result a new position is created for each landmark point in the image. The basic purpose is to create a possible similarity between set of images. Several types of filtering operations such as smoothing, edge detection, elimination of noise/blur and intensity enhancement is performed on digital images through the processes of convolution [81]. Convolution is performed through several types (i.e. size 2*2, 3*3 and 4*4) of matrices also called kernel. Kernel consists of mathematical values which are applied to pixel values resulting modified image with different properties is formed. The resultant image greatly depends on the values used and on the size of kernel. Kernel spline is popular and widely used method for image warping and it uses a 3D mapping function to find out information about landmark points or pixel intensities in both source and target objects [82, 83]. This is done by localizing a mapping area of pixel intensities in which the size of area under consideration is find out by increasing the distance into twice between the source and target point.

Kernel spline interpolation functions can effectively model deformation field in landmark based image registration. A smooth global transformation is performed with kernel spline by computing local controlled deformation of landmark points [84, 85]. Image deformation model is created by joining spline kernel over the area of image under consideration in the form of rectangular grid.

3.6 Image Warping with BSplines

Beside computer vision and multimedia application, image warping has also got importance in the fields of medical image processing such as image morphing and deformation. Several powerful warping methods are developed in the recent years for the precise analysis and matching of medical images. BSplines warping is the popular and widely used among them because it precisely deforms medical images with its significant property of local control and global mapping. BSplines mapping functions perform free form deformation on the control points and establish a one to one correspondence between them to generate a warp image. Images having local distortion, irregularly spaced samples in them and have nonlinear distortions are successfully warped with BSplines interpolation [86-88]. Medical images with large local deformation such as coronary arteries/ cardiac images are also successfully warped with BSplines interpolation functions. However, accurate image warping with BSplines needs more computational time which makes it less desirable in a situation when the time constraints are important.

Image warping with BSplines is performed by defining local domain which contains many control points [88]. Subsequently, the local domain which is divided into several blocks is shifted to a new region. The warping is performed again and again on the shifted domain till the recovery of all pixels in the image. Images with inconsistent contrast between them are therefore precisely warp with BSplines. Furthermore, BSplines also provide one-to-one mapping property due to which there is no chance of distorted image to fold back upon itself.

3.7 Asymmetrical Demons Deformable Registration

Demons methods also called gray-scale automatic deformable algorithms are efficient and robust registration methods for medical images [28, 77, 89-91]. These algorithms are mostly used for intensity based image registration and perform operation on prominent and distinctive features of images. Demons algorithms greatly depend on the intensity or color change (image gradients) in the source and/or target images. Therefore, changing the gradients of input images strongly affects the accuracy of registration. Several types of demons algorithms are available with high capability to describe critical structure and trace potential differences/ similarities in source and target images during radiotherapy and image guided surgery.

Uniform one-to-one mapping of corresponding point landmarks in both directions is very important for the consistent and smooth registration of medical images [27]. However, majority of the available registration methods are asymmetrical and unidirectional because registering source image to target image does not show the same associations as registering target to source. Furthermore, most of the transformation functions such as linear and non-linear elastic functions, thin plate functions, cubic kernel and quadratic regularization functions are also perform asymmetrically during registration [92]. The difference in coordinate frames between source and target images due to the difference in the deformation field also gives better registration result when the transformation is asymmetric. To represent a valid probabilistic model on the basis of atlas based registration, the asymmetric approach provides more simplicity than symmetric registration. In image-guided surgery, image to atlas based registration using asymmetric transformation is mostly used for the precise analysis of tumor in human organs.

In asymmetric approach, registration relay on the choice of target domain because the estimation and mapping of point lands-marks is unidirectional. Therefore, unidirectional transformation of deformable images provides more efficiency than bi-directional transformation, particularly in the case of mono-modal image registration. However, the performance of bi-directional (symmetric) registration is much high on the basis of running time and estimation of total errors in the registration of multi-modal deformable images due to the availability of regional symmetric similarity measures [93].

3.8 Symmetrical Demons Deformable Registration

Pair-wise and bi-directional mapping of homologous point land-marks between source and target image are widely used methods for the accurate and consistent image registration. The pair-wise and bi-directional mapping of point landmarks in one coordinate to another coordinate is also called symmetric registration [94]. In symmetric registration both images are treated in the same manner and the mapping of similarity measures is performed from source image direction to target image direction and vice versa. Such type of two ways mapping provides more consistency and accuracy during image registration because the possibility of detecting different transformation components will be more. Furthermore, symmetrical registration also show robustness against local minima because in the optimization process it uses the gradients of both images

Contrast to symmetric image registration, asymmetric image registration maps point land marks between two images in a non-uniform and unidirectional way due to which the choice of target image usually influences the results of deformable image registration. Symmetric registration also eliminates the chance of inverse consistent errors which usually occurs due to the inconsistent voxel-by-voxel wise association between source and target images in both directions [95]. The inverse consistent function computes transformation information from source and target images and converts them into a common intermediate image. This function also ensures that the transformation from source image side is the inverse of the transformation from the target image side. On the down side, the two way transformation require more computational time and is difficult to implement especially in case of image registration with iterative optimization. Furthermore, symmetrical registration cannot perform well in the registration of image to template (atlas) based registration [96] because in this type of registration the mapping is mostly performed in a single direction for the template to exhibit a suitable probabilistic model.

4. PERFORMANCE EVALUATION

The performance evaluation of different deformable registration methods is always a difficult task faced by the researchers. The main reason is the unavailability of related pixel information between source and target images. We have evaluated the performance of deformable registration methods by implementing them in C++ based on the Insight Segmentation and Registration Toolkit (ITK) [72]. The computer system used for testing the performance of registration methods is Core i5 with 4GB RAM. The performance i.e. accuracy and efficiency was evaluated by testing the eight variants of deformable registration methods. We have implemented the registration methods on the datasets of two 2D lung images of living rat obtained from ITK software package [72]. For experimental analysis, images were taken at different times, the one after inspiration of air into lungs (source image) and the second after exhalation (target image). In the quantitative analysis, mutual information, root mean square error (RMS), computation time and occupied memory were used as evaluation metrics. Since deformable images hold high local variation, therefore, for each method we have performed three types of registration for rat lungs images: registering normal images, registering the same images by inducting 0.001% and 0.002% Gaussian noise. Table 2 shows the overall quantitative analysis of each method on different levels. The basic aim was to estimate the effect of change or noise on the accuracy and efficiency of each registration method.

Accuracy evaluation in deformable registration is a challenging task due to variations in every voxel points. In our experiment, this is performed by estimating mutual information and root mean square errors between source and target images. On the other hand, efficiency of each method is calculated based on computation time and memory space occupied by each method.

To find out the accuracy of each registration method, the values of mutual information and RMS errors estimated in Table 2 at noise levels (0%, 0.001% and 0.002%) were further listed in Table 3 and Table 4. On the bases of data obtained in Table 3 and Table 4, accuracy of deformable registration methods are graphically shown in Fig 5 and Fig 6 respectively. Mutual information (MI) and RMS are the two essential parameters to estimate the accuracy of registration method. These two parameters are widely adopted by a large number of researchers in the medical image processing community. MI estimates the similarity measures between source and target images through pixel-by-pixel correspondence while RMS error is a standard statistical metric for error prediction. The registration method is more accurate if the values of MI are maximum and RMS values are minimum.

It is shown in Fig 5 that in most registration methods, the introduction of noise has little effect on the values of mutual information. However, in the experiments, we obtained high mutual information values for FEM based registration method. Similarly, we estimate good results while testing level set method for mutual information. Therefore, FEM based and level set registration methods provide more accuracy as compare to others deformable registration methods. On the other hand, the accuracy of warping with kernel splines is much low in our experiment due to small mutual information values obtained from the registration of source and target images at different noise levels.

We have also tested the accuracy of each registration method based on RMS error at different noise levels. Accuracy based on RMS error is shown in the Table 4 which is further plotted in Fig 6. In this experiment, we obtained minimum RMS error values for FEM based registration and high values for warping with kernel splines. Therefore, the accuracy of FEM based registration based on MI and RMS error is much high than other types of deformable registration methods. However, the accuracy of warping with kernel splines is low compared to other methods due to low MI and high RMS error values obtained in our experiments.

In order to determine the efficiency of deformable registration methods, they are tested on the basis of computation time and occupied memory space. Registration method is more efficient if it takes minimum time and less memory space during execution. The computation time estimated for all registration methods in our case at different levels of Gaussian noise are listed in Table 5 and graphically shown in Fig. 7. It is shown in the Fig that the warping with kernel splines is the most efficient method at 0% and 0.001% Gaussian noise because it takes less time than others methods. Similarly, asymmetric demon registration also provides consistent efficiency at different noise levels. On the other hand, the efficiency of FEM based method is much low due to excessive time it takes during registration of source and target images. We have also estimated the efficiency of each registration method on the bases of occupied memory space. Values obtained for memory space occupied by each method at different noise levels are listed in

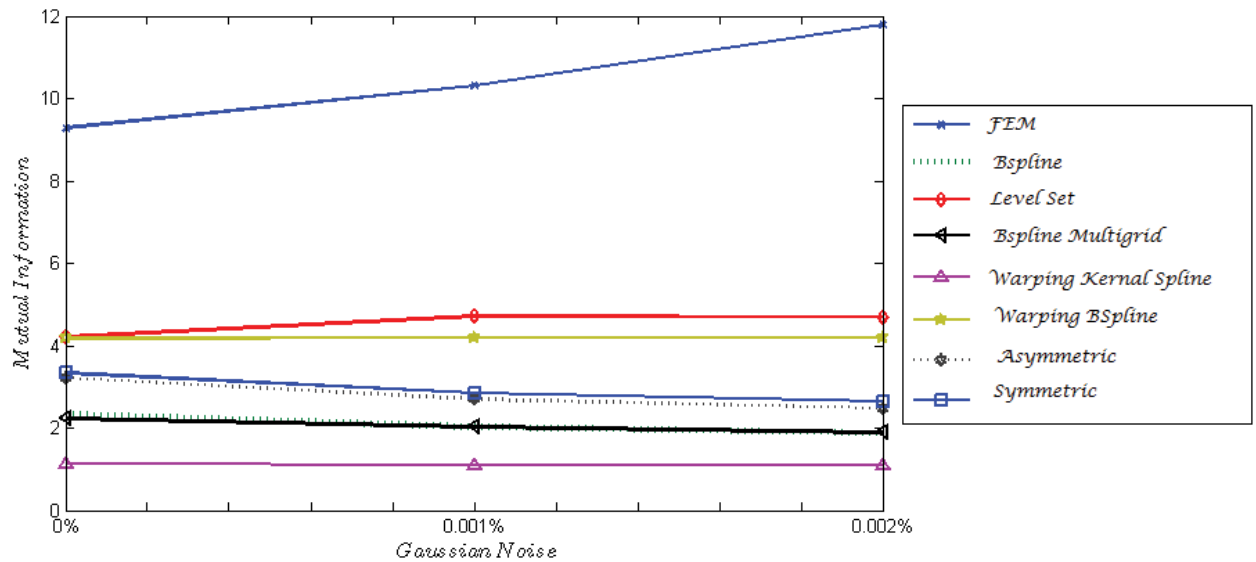


Fig. 5. Graphical representation of MI at different levels of Gaussian noise.

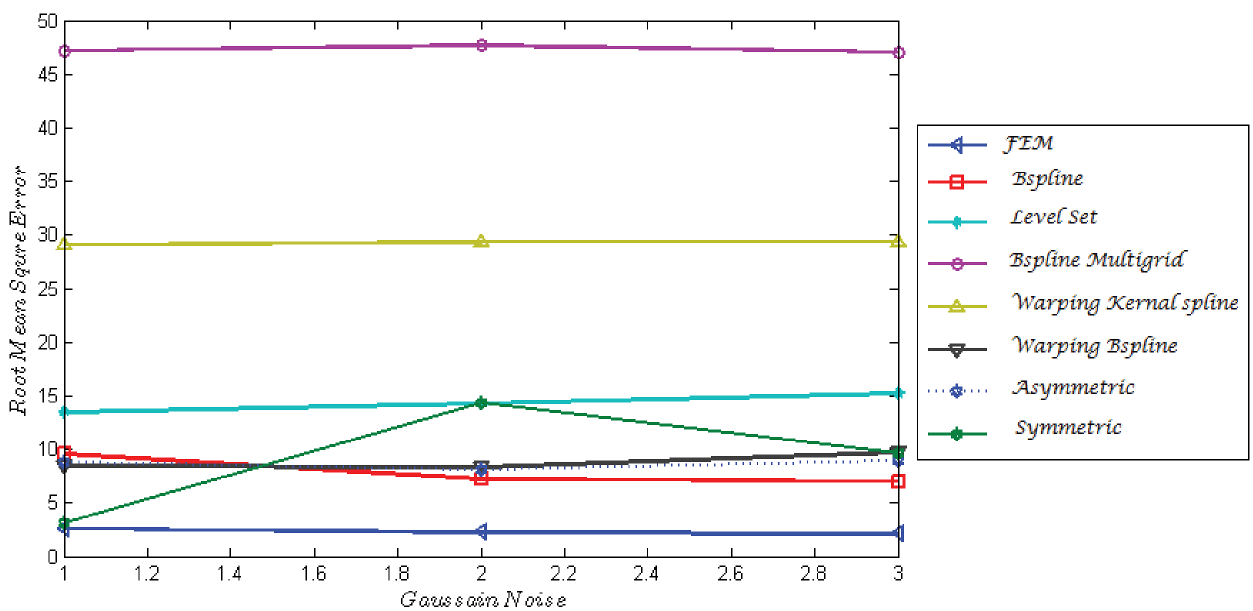


Fig. 6. Graphical representation of RMS errors at different levels of Gaussian noise.

Table 6 and their graphical representation is shown in Fig 8. In this case, the efficient methods are BSplines multi-grid and FEM based registration due to the minimum memory space occupied at different noise levels. However, the memory spaces occupied by BSlines registration, warping with kernel and BSplines methods are much high due to which they are less efficient.

After the detail evaluation and testing of deformable registration methods we came up with certain conclusions. In our experiments, the performance of FEM based registration is high than other types of registration methods. This is due to the sharing of excessive amount of mutual information, less number of RMS errors in registration and taking less memory space during execution. We also estimate high computational time during execution for FEM registration, which negatively affect image registration. Minimizing the computational time for FEM based registration will make it a perfect

Table 2. Experimental evaluation of deformable registration methods.

Parameters	FEM based registration			BSplines registration			Level set registration			BSplines multi-grid registration			Warping with kernel splines			Warping with BSplines			Asymmetric demons registration			Symmetric demons registration		
	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)	Noise (0%)	Noise (0.001%)	Noise (0.002%)
Mutual Information	9.276	10.305	11.785	2.3286	2.021	1.87	4.212	4.7103	4.7043	2.221	2.028	1.886	1.142	1.105	1.109	4.170	4.189	4.189	3.219	2.694	2.473	3.332	2.855	2.640
Root Mean Square Error (mm)	2.590	2.270	2.125	3.087	14.320	15.981	9.602	7.221	7.044	13.46	14.28	15.23	47.2	47.67	47.074	29.124	29.337	29.337	8.442	8.321	9.674	8.785	8.081	8.911
Computation Time (Seconds)	1.024	1.065	1.018	0.453	0.532	0.535	0.397	0.435	0.596	0.532	0.453	0.543	0.315	0.463	0.478	0.643	0.745	0.764	0.386	0.547	0.452	0.642	0.7032	0.796
Memory (kb)	788	1048	1096	5656	4360	4432	3320	2065	1896	136	136	136	3922	4037	4054	3912	3991	4193	3206	1960	1960	3160	1760	1920

choice for the successful registration of deformable images. It is also observed that the performance of warping with kernel splines in terms of accuracy is low due to high number of RMS errors arises during registration and sharing of low mutual information. However, the efficiency in terms of computational time is high for warping with kernel splines method. Likewise, BSplines multi-grid registration provides more efficiency in terms of computational time and occupied memory space. Therefore, the improvement in terms of accuracy can also make this method a perfect choice in clinical applications.

Deformable registration methods are widely used for the accurate registration of objects with large deformation. Therefore, several registration methods are available which automatically register medical images. After a thorough analysis of each method, we find out that every method provide their own strength and flexibility for the precise and efficient registration of medical images. However, due to complex and difficult to calculate deformation field, most of the registration methods cannot perform perfectly in clinical applications. Therefore, generic and powerful registration methods are required to be developed in the future, which can precisely, efficiently and automatically register medical images with large-scale deformations.

5. CONCLUSIONS

Deformable image registration is a challenging task in medical image analysis due to different imaging conditions, variability in anatomical structures and elasticity of the body and organs. In this article, we have experimentally evaluated the existing deformable registration methods on the images of rat lungs to estimate their performance. Although several automatic deformable registration methods are available applicable for single modality, linear and small local deformation, universal and generic methods are still a problem in clinical applications. To perfectly register medical images obtained through multimodality with uncertain and complex features of deformable objects further research is needed. However, in our analysis, we came up with a conclusion that FEM based registration method obtained excellent performance in terms of

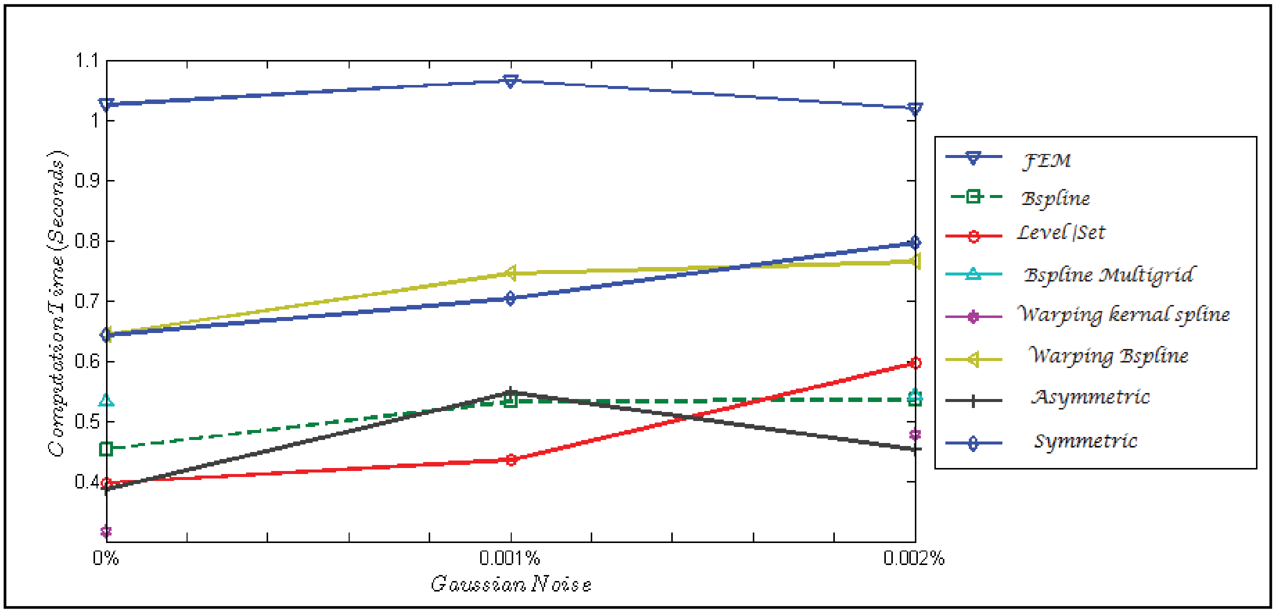


Fig. 7. Graphical representation of computation time at different levels of Gaussian noise.

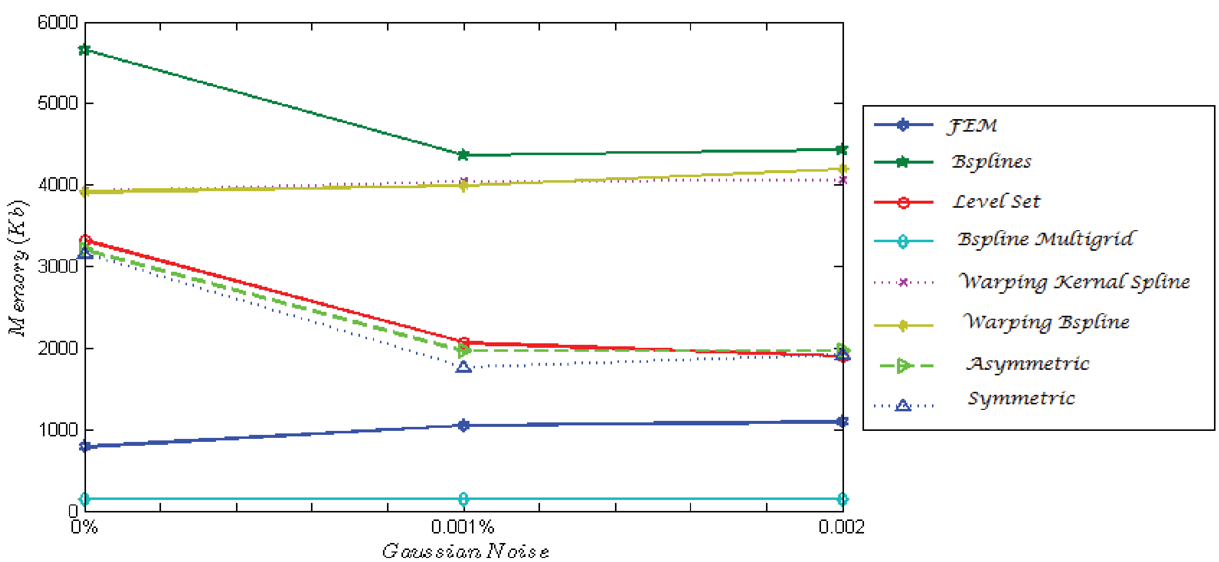


Fig. 8. Graphical representation of memory space occupied by deformable registration methods at different levels of Gaussian noise.

Table 3. MI of deformable registration methods at different noise levels.

	FEM based registration	BSplines registration	Level set registration	BSplines multi-grid registration	Warping with kernel splines	Warping with BSplines	Asymmetric demons registration	Symmetric demons registration
Noise (0%)	9.276	2.328	4.212	2.221	1.142	4.170	3.219	3.332
Noise (0.001%)	10.305	2.0211	4.7103	2.028	1.105	4.189	2.694	2.855
Noise (0.002%)	11.785	1.872	4.7043	1.886	1.109	4.189	2.473	2.640

Table 4. RMS error of deformable registration methods at different noise levels.

	FEM based registration	BSplines registration	Level set registration	BSplines multi-grid registration	Warping with kernel splines	Warping with BSplines	Asymmetric demons registration	Symmetric demons registration
Noise (0%)	2.590	3.087	9.602	13.46	47.23	29.124	8.442	8.785
Noise (0.001%)	2.270	14.320	7.221	14.289	47.678	29.337	8.321	8.081
Noise (0.002%)	2.125	9.602	7.044	15.235	47.074	29.337	9.674	8.911

Table 5. Computation time for deformable registration methods at different noise levels.

	FEM based registration	BSplines registration	Level set registration	BSplines multi-grid registration	Warping with kernel splines	Warping with BSplines	Asymmetric demons registration	Symmetric demons registration
Noise (0%)	2.590	3.087	9.602	13.46	47.23	29.124	8.442	8.785
Noise (0.001%)	2.270	14.320	7.221	14.289	47.678	29.337	8.321	8.081
Noise (0.002%)	2.125	9.602	7.044	15.235	47.074	29.337	9.674	8.911

Table 5. Computation time for deformable registration methods at different noise levels.

	FEM based registration	BSplines registration	Level set registration	BSplines multi-grid registration	Warping with kernel splines	Warping with BSplines	Asymmetric demons registration	Symmetric demons registration
Noise (0%)	1.0245	0.453	0.397	0.532	0.315	0.643	0.386	0.642
Noise (0.001%)	1.065	0.532	0.435	0. 453	0. 463	0.745	0.547	0.703
Noise (0.002%)	1.0186	0.535	0.596	0.543	0.478	0.764	0.452	0.796

Table 6. Memory Space Occupied by Deformable Registration Methods at Different Noise Levels.

	FEM based registration	BSplines registration	Level set registration	BSplines multi-grid registration	Warping with kernel splines	Warping with BSplines	Asymmetric demons registration	Symmetric demons registration
Noise (0%)	788	5656	3320	136	3922	3912	3206	3160
Noise (0.001%)	1048	4360	2065	136	4037	3991	1960	1760
Noise (0.002%)	1096	4432	1896	136	4054	4193	1960	1920

accuracy and memory space. Our experiments also confirm that in terms of efficiency, the performance of BSplines multi-grid and warping with kernel splines is excellent compared to other methods.

The future work is to further improve the performance of deformable registration methods we obtained in our experiment and to develop an advance registration method applicable for the registration of several types of deformable objects. In the next work, we will also evaluate the performance of deformable registration methods on larger datasets containing 3D and 4D deformable image and for a broad range of medical image registration methods according to the criteria we adopted in this work.

6. REFERENCES

1. Bushberg, J. T., J.A. Sielert, E.M. Leidholdt, jr., & J.M. Boone. *The Essential Physics of Medical Imaging*. Lippincott Williams & Wilkins, Philadelphia, USA (2002).
2. Flower, M.A. *Webb's Physics of Medical Imaging,* 2nd ed. Taylor & Francis, Boca Raton, US (2012).
3. Metter, R.L.V., J. Beutel, & H. L., Kundel. *Handbook of Medical Imaging, Volume 1. Physics and Psychophysics,* vol. 1. Society of Photo-Optical Instrumentation Engineers, Bellingham, Washington, USA (2000).
4. Angenent, S., E., Pichon, & A. Tannenbaum, A., Mathematical methods in medical image processing, *Bulletin (New Series) American Mathematical Society,* Vol. 43, p. 365-396, doi: http://dx.doi.org/10.1090/S0273-0979-06-01104-9 (2006).
5. McInerney, T. & D. Terzopoulos, *Deformable Models in Medical Image Analysis: A Survey, Medical Image Analysis,* Vol. 1, p. 91–108. doi:10.1016/S1361-8415(96)80007-7 (1996).
6. Brown, L.G. *A Survey OfImage Registration Techniques,* ACM Computing Surveys (CSUR), vol. 24, p. 325-376, www.geo.uzh.ch/microsite/rsl-documents/research/SARlab/.../Bro92.pdf (1991).
7. Fahmi Khalifa, G.M.B., Georgy Gimel'farb, Jasjit S. Suri & Ayman S. El-Baz. State-of-the-art medical image registration methodologies: A survey, In: *Multi Modality State-of-the-art Medical Image Segmentation and Registration Methodologies,* R. A. Ayman S. El-Baz, Andrew F. Laine, & Jasjit S. Suri (Ed.), Springer, New York, p. 235-280, DOI: 10.1007/978-1-4419-8195-0_9 (2011).
8. Hajnal, J., D., Hawkes, & D. Hill. *Medical Image Registration*. CRC Press LLC, http://www.ncbi.nlm.nih.gov/pubmed/11277237 (2001).
9. Maintz, J.B A. & M.A. Viergever, A survey of medical image registration, *Medical Image Analysis*, vol. 2, pp. 1-36, DOI: 3D10.1.1.110.3014 (1998).
10. Mani, V.R.S. & D.S. Rivazhagan. Survey of medical image registration. *Journal of Biomedical Engineering and Technology* 1: 8-25, DOI: http://pubs.sciepub.com/jbet/1/2/1/ (2013).
11. Nielsen, L.K., *Elastic Registration of Medical MR Images.* Thesis in Computational Science, Department of Mathematics and Neuroinformatics and Image Analysis Group, Department of Physiology, University of Bergen, folk.uib.no/nmaxt/thesis/nielsen.pdf (2003).
12. Oliveira, F. & J. Tavares. Medical image registration: a review, *In: Medical Imaging: Technology and Applications*, K. I. Troy Farncombe, (Ed), CRC Press, p. 619-660, http://www.ncbi.nlm.nih.gov/pubmed/22435355 (2013).
13. Salvi, J., C. Matabosch, D. Fofi, & J. Forest. A review of recent range image registration methods with accuracy evaluation. *Image Vision Comput.,* vol. 25, p. 578-596, DOI: 10.1016/ j.imavis.2006.05.012 (2007).
14. Saxena, S. & R., Kumar. A survey of recent and classical image registration methods. *International Journal of Signal Processing*, *Image Processing and Pattern Recognition,* 7: 167-176, www.sersc.org/journals/IJSIP/vol7_no4/16.pdf (2014).
15. Schiller, J. *Basics of Medical Image Segmentation and Registration.* Technische Universität München, http://vision.in.tum.de/publications (2005).
16. Van den Elsen, P.A., E.J.D. Pol, & M. A. Viergever. Medical image matching-a review with classification, *Engineering in Medicine and Biology Magazine*, IEEE 12: 26-39, DOI: 10.1109/51.195938 (1993).
17. Wyawahare, M., P., Patil, & H., Abhyankar. Image registration techniques: an overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition,* vol. 2, No. 3, http://www.sersc.org/journals/IJSIP/vol2_no3/2.pdf (2009).
18. Zitova, B. & J. Flusser. Image registration methods: A survey. *Image and Vision Computing* 21: 977–1000, DOI:10.1016/S0262-8856(03)00137-9 (2003).
19. Aiming, L. *A Survey of Medical Image Registration*, Project Report: Advanced Video and Image Processing ECE 738, Available: homepages.cae.wisc.edu/~ece738/projs03/lu.doc (2000).
20. Goshtasby, A.A. *Image Registration: Principles, Tools and Methods*. Springer, DOI: 10.1007/978-1-4471-2458-0 (2012).
21. Pitiot, A. (2005, 27-8-2014). *Medical Image Registration Taxonomy.* Available: http://www.ariser.info/training/imgproc.php (2005).
22. Bhadoria, S., P., Aggarwal, C.G., Dethe, & R., Vig, Comparison of segmentation tools for multiple modalities in edical imaging. *Journal of Advances in Information Technology* vol. 3, DOI:10.4304/jait.3.4.197-205 (2012).
23. Doi, K. Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology. *Physics in Medicine and Biology* 51: 5-27, DOI:10.1088/0031-9155/51/13/R02 (2006).
24. Maes, F., D., Vandermeulen, & P. Suetens. Medical image registration using mutual information. *Proceedings of the IEEE,* 91: 1699-1722, DOI: 10.1109/JPROC.2003.817864 (2003).
25. Schramek, G.G.R., D. Stoevesandt, A. Reising, J.T. Kielstein, M. Hiss, & H. Kielstein. Imaging in anatomy: a comparison of imaging techniques in embalmed human cadavers. *BMC Medical Education,* vol. 13, p. 143, DOI: 10.1186/1472-6920-13-143 (2013).
26. Schwarz, L. A. *Non-rigid Registration Using Free-form Deformations.* Ph.D. dissertation, Department

of Computer Science, Technical University of Munich, http://campar.in.tum.de/twiki/pub/Main/LorenSchwarz/thesis-070510.pdf (2007).

27. Aristeidis Sotiras, C.D., N. Paragios. Deformable medical image registration: A survey. *IEEE Trans Med Imaging*. 32: 1153–1190, DOI: 10.1109/TMI.2013.2265603 (2013).
28. Wang, H., L. Dong, J. O'Daniel, R. Mohan, A.S. Garden, K.K. Ang, *et al.* Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. *Physics in Medicine and Biology,* 50: 2887–2905, DOI: 10.1088/0031-9155/50/12/011 (2005).
29. Jolesz, F. *Intraoperative Imaging and Image-Guided Therapy*: Springer, New York, DOI: DOI 10.1007/978-1-4614-7657-3 (2014).
30. Han, S.C., S. H. Choi, S. Park, S. S. Lee, H. Jung, M.-S. Kim, *et al.* Evaluation of various Deformable Image Registrations for Point and Volume Variations. *arXiv preprint arXiv:1503.03622*. DOI: 10.3938/jkps.67.218 (2015).
31. *InsightSoftwareConsortium/ITK*, cited 2015 (28-12-15), Available: https://github.com/InsightSoftwareConsortium/ITK/tree/master/Examples/Data (2015).
32. Hu, Grossberg, & Mageras. Survey of recent volumetric medical image segmentation techniques. In: *Biomedical Engineering,* Carlos Alexandre Barros de Mello (Ed.), p. 315-337, DOI: 10.5772/7865 (2009).
33. Pham, D. L., C., Xu, & J.L. Prince. Current methods in medical image segmentation. *Annu Rev Biomed Eng,* 2: 315-37, DOI: 10.1146/annurev.bioeng.2.1.315 (2000).
34. Erdt, M., S. Steger, & G. Sakas. Regmentation: A New View of Image Segmentation and Registration. *Journal of Radiation Oncology Informatics* 4: 1-23, DOI: 10.5166/jroi-4-1-19 (2012).
35. Mishra, A., A. Rai, & A. Yadav. Medical image processing: A challenging analysis. *International Journal of Bio-Science and Bio-Technology* 6: 187-194, http://dx.doi.org/10.14257/ijbsbt.2014.6.2.19 (2014).
36. Phamy, D.L., C. Xu, & J. L. Prince. A survey of current methods in mdical image segmentation. *Annual Review of Biomedical Engineering* 2: 315-338, DOI: 10.1146/annurev.bioeng.2.1.315 (1998).
37. Yoo, T.S. *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis*. AK Peters Ltd, DOI: 10.1201/b10657 (2004).
38. Zhang, H., J. E. Fritts, & S.A. Goldman. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding.* 110: 260-280. DOI: 10.1016/j.cviu.2007.08.003, (2008).
39. Zuva, T., O. Olugbara, S. Ojo, & S. Ngwira. Image segmentation, available techniques, developments and open issues. *Canadian Journal on Image Processing and Computer Vision* 2: 20-29, https://www.scienceopen.com/document/vid/de286b73-697c-4035-bb39-87255f1b494b (2011).
40. Wang, X., L. Li, C. Hu, J. Qiu, Z. Xu, & Y. Feng. A comparative study of three CT and MRI registration algorithms in nasopharyngeal carcinoma. *Journal of Applied Clinical Medical Physics,* vol. 10, DOI: http://dx.doi.org/10.1120/jacmp.v10i2.2906 (2009).
41. Tate, P. M., D. M. A. Chapman, & R. Polytechnic. The assessment of magnetic imagery for computer assisted spinal surgery: *Int'l Archives of Photogrammetry and Remote Sensing,* 33: 09–816, www.isprs.org/proceedings/XXXIII/congress/.../809_XXXIII-part5.pdf (2000).
42. Crum, W.R., T. Hartkens, & D.L.G. Hill. Non-rigid image registration: theory and practice: *The British Journal of Radiology* 77: 140–153, DOI: 10.1259/bjr/25329214 (2004).
43. Gomes, P.N.S. *Image Registration and its Relevance in Plantar Pressure Images*. Thesis Integrated Master in Bioengineering, Faculty of Engineering, The University of Porto, https://web.fe.up.pt/~tavares/downloads/publications/relatorios/Pedro_Gomes_Monografia.pdf (2013).
44. Perperidis, D. *Spatio-temporal Registration and Modelling of the heart using cardiovascular MR Imaging.* PhD thesis, Department of Computing, Imperial College London, University of London, DOI:10.1007/s11263-009-0212-6 (2005).
45. Markelj, P., D. Tomaževič, B. Likar, & F. Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis* 16: 642-661, DOI: 10.1016/j.media.2010.03.005 (2012).
46. Rueckert, D. & P. Aljabar, Nonrigid registration of medical images: Theory, methods, and applications. *Signal Processing Magazine, IEEE* 27: 113-119, DOI: 10.1109/MSP.2010.936850 (2010).
47. Shekhar, R., V. Walimbe, S. Raja, & V. Zagrodsky, Automated 3-dimensional elastic registration of whole-body PET and CT from separate or combined scanners. *Journal of Nuclear Medicine,* 46: 1488-1496., DOI:10.1016/j.radphyschem.2015.01.022 (2005).
48. Edward, C. *Medical Image Registration: A Review of Existing Methods and Preliminary Numerical Results*. MS Thesis, Rice University, http://hdl.handle.net/1911/17867 (2005).
49. Faliagka, E., G. Matsopoulos, A. Tsakalidis, J. Tsaknakis, & G. Tzimas, Registration and fusion techniques for medical images: Demonstration and evaluation. In: *Biomedical Engineering Systems and Technologies*. vol. 127. Springer Berlin Heidelberg,

p. 15-28, DOI: 10.1007/978-3-642-18472-7_2 (2011).

50. Tavares, J.M. Towards more effective medical image analysis based on computational image registration methods, In: *3rd International Workshop on Standardization of 3D Based Medical Technologies, Università Roma Tre, Rome, Italy*, *DOI:*10.1080/10 255842.2012.670855 (2013).
51. Dai, Y.-H. A family of hybrid conjugate gradient methods for unconstrained optimization. *Mathematics of Computation,* 72: 1317-1328, www.ams.org/mcom/2003-72-243/...03.../S0025-5718-03-01491-1.pdf (2003).
52. Duy, L.V.N.C. *Review and Enhancement Optimization methods in Image Registration.* MEng thesis, Department of Computing, Imperial College London, www.doc.ic.ac.uk/~vl05/master_thesis.pdf (2009).
53. Hansen, N. & A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* 9: 159-195, DOI: 10.1162/106365601750190398 (2001).
54. Dennis, J.E. & J.J. Moré. Quasi-Newton methods, motivation and theory. *SIAM Review* 19: 46-89, DOI:10.1137/1019005 (1977).
55. Nocedal, J. , Wright, S. J. *Numerical Optimization*. Springer, DOI: 10.1007/978-0-387-40065-5 (2006).
56. Ou, Y., S.P. Weinstein, E.F. Conant, S. Englander, X. Da, B. Gaonkar, *et al.* Deformable registration for quantifying longitudinal tumor changes during neoadjuvant chemotherapy. *Magnetic Resonance in Medicine* 73: 2343-2356, DOI:10.1002/mrm.25368 (2014).
57. Crouch, J.R., S.M. Pizer, E.L. Chaney, Y.-C. Hu, G.S. Mageras, & M. Zaider. Automated finite element analysis for deformable registration of prostate images. *IEEE Transactions on Medical Imaging,* vol. 26, DOI:10.1109/TMI.2007.898810 (2007).
58. Yu, H. *Automatic Rigid and Deformable Medical Image Registration*. PhD thesis, Mechanical Engineering, Worcester Polytechnic Institute. https://www.wpi.edu/Pubs/ETD/Available/etd.../Thesis_victor.pdf (2005).
59. Rohr, K. Elastic registration of multimodal medical images: A survey. *Auszug aus: Künstliche Intelligenz,* vol. 14, p. 11-17, http://www.fbki.kuenstliche-intelligenz.de/fileadmin/template/main/archiv/2000_3/rohr1.pdf (2000).
60. Maes, F., D. Vandermeulen, & P. Suetens. Medical image registration using mutual information. *Proceedings of the IEEE* 91: 1699 - 1722, DOI: 10.1109/JPROC.2003.817864 (2003).
61. Fornefett, M., K. Rohr, & H.S.F. Stieh. Elastic registration of medical images using radial basis functions with ompact support, In: *Proc. Computer Vison and Pattern Recognition* (CVPR'99), Fort Collins, CO, USA, DOI: 10.1109/CVPR.1999.786970 (1999).
62. Julia, S., C. Tanner, A. D. Castellano-Smith, A. Degenhard, M. O. Leach, D.R. Hose, *et al.* Validation of nonrigid image registration using finite-element methods: application to breast MR images. *IEEE Transactions on, Medical Imaging,* 22: 238-247, DOI: 10.1109/TMI.2002.808367 (2003).
63. Lee, H.-P., M. Lin, & M. Foskey. Physically-based validation of deformable medical image registration. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*. vol. 5242, D. Metaxas, L. Axel, G. Fichtinger, G. Székely, (Ed). Springer, Berlin, p. 830-838. DOI: 10.1007/978-3-540-85990-1_100 (2008).
64. Samavati, N. *Deformable Multi-Modality Image Registration Based on Finite Element Modeling and Moving Least Squares*. Master of Applied Science thesis, Electrical & Computer Engineering, McMaster University, Hamilton, Ontario, Canada, www.ece.mcmaster.ca/~sirouspour/documents/thesis/thesis_Navids.pdf (2009).
65. Zhong, H., J. Kim, H. Li, T. Nurushev, B. Movsas, & I.J. Chetty. A finite element method to correct deformable image registration errors in low-contrast regions. *Physics in Medicine and Biology* 57: 3499-3515, 05/11, DOI: 10.1088/0031-9155/57/11/3499 (2012).
66. Diou, A., Y. Voisin, L. F. C. L. Y. Voon, G. Moreels, & S, J. Clairemidi. Ome applications of B-spline functions in image processing, In: *Graphics/Vision Interface Conference (VI), Vancouver, Canada*, www.cipprs.org/papers/VI/VI1998/pp325-332-Diou-et-al-1998.pdf (1998).
67. Sarrut, D. & J. Vandemeulebroucke. B-LUT: Fast and low memory B-spline image interpolation. *Computer Methods and Programs in Biomedicine* 99: 172-178, DOI: 10.1016/j.cmpb.2009.11.013 (2010).
68. Unser, M. Splines: a perfect fit for signal and image processing. *Signal Processing Magazine, IEEE* 16: 22-38. DOI: 10.1109/79.799930 (1999).
69. Se Young, C. & J.A. Fessler. Regularized methods for topology-preserving smooth nonrigid image registration using B-spline basis, In: *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium,* p. 1099-1102. DOI: 10.1109/ISBI.2008.4541192 (2008).
70. Wu, Z., E. Rietzel, V. Boldea, D. Sarrut, & G.C. Sharp. Evaluation of deformable registration of patient lung 4DCT with subanatomical region segmentations. *Medical Physics* 35: 775-781, DOI: 10.1118/1.2828378 (2008).
71. Vandemeulebroucke, J., O. Bernard, J. Kybic, P. Clarysse, & D. Sarrut. Automatic motion mask

extraction for deformable registration of the lungs. In: *XVIth International Conference on the Use of Computers in Radiation Therapy, Amsterdam*, DOI: 10.1118/1.3679009 (June 2010).

72. Ibanez, L. & W. Schroeder. *The ITK Software Guide 2.4* Kitware Inc, www.itk.org/ItkSoftwareGuide (2003).
73. Lucas, B.C., M. Kazhdan, & R.H. Taylor. Spring level sets: A deformable model representation to provide interoperability between meshes and level sets. *IEEE Transactions on Visualization and Computer Graphics* 19: 852–865. DOI:10.1109/TVCG.2012.162 (2013).
74. Wilson, D. & S. Laxminarayan. *Handbook of Biomedical Image Analysis: Volume 3: Registration Models*: Springer, USA, DOI: 10.1007/b104805 (2007).
75. Celaya-Padilaa, J.M., J. Rodriguez-Rojas, V. Trevino, & J.G. Tamez-Pena. Local image registration a comparison for bilateral registration mammography. p. 892210-892210-9, DOI: 10.1117/12.2035516 (2013).
76. Metz, C.T., S. Klein, M. Schaap, van Walsum, & W.J. Niessen. Nonrigid registration of dynamic medical imaging data using nD + t B-splines and a groupwise optimization approach. *Medical Image Analysis,* 15: 238-249, DOI: 10.1016/j.media.2010.10.003 (2011).
77. Zhou, W. & Y. Xie, Interactive multigrid refinement for deformable image registration. *BioMed Research International,* vol. 2013, p. 9, http://dx.doi.org/10.1155/2013/532936 (2013).
78. Unser, M.A., A. Aldroubi, & C.R. Gerfen. Multiresolution image registration procedure using spline pyramids. vol: 2034, p. 160-170, DOI: 10.1117/12.162061 (1993).
79. Sampath, R.S. & G. Biros. Parallel elastic registration using a multigrid preconditioned Gauss-Newton-Krylov solver, grid continuation and octrees. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.297.211 (2009).
80. Klein, A., D.-J. Kroon, Y. Hoogeveen, L.J.S. Kool, W. K. J. Renemab, & C. H. Slump. Multimodal image registration by edge attraction and regularization using a B-spline grid. In: *Medical Imaging Lake Buena Vista,* Florida, USA, DOI: 10.1117/12.878267 (2011).
81. Kim, S. & R. Casper. *Applications of Convolution in Image Processing with MATLAB.* University of WashingtonAugust 20, www.math.washington.edu/~wcasper/math326/projects/sung_kim.pdf (2013).
82. Choi, Y. & S. Lee, Injectivity conditions of 2D and 3D uniform cubic B-spline functions. *Graphical Models* 62: 411–427, DOI:10.1006/gmod.2000.0531 (2000).
83. Whitbeck, M. & H. Guo. Multiple landmark warping using thin-plate splines. In: *IPCV'06, citeseerx.ist.psu.edu/viewdoc/summary*? doi=10.1.1.92.2533 (2006).
84. Dai, L. & R.N. Jazar. *Nonlinear Approaches in Engineering Applications: Applied Mechanics, Vibration Control, and Numerical Analysis (*illustrated ed). Springer, DOI: 10.1007/978-3-319-09462-5 (2014).
85. Mattes, D., D.R. Haynor, H. Vesselle, T K. Lewellen, & W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging,* 22: 120-128, http://www.ncbi.nlm.nih.gov/pubmed/12703765 (2003).
86. Nejati, M., S. Sadri, & R. Amirfattahi. Nonrigid image registration in digital subtraction angiography using multilevel B-spline. *BioMed Research International,* vol. 2013, p. 12, DOI: 10.1155/2013/236315 (2013).
87. Zhiyong, X. & G.E. Farin. Image registration using hierarchical B-splines. *IEEE Transactions on Visualization and Computer Graphics,* 10, 85-94, DOI: 10.1109/TVCG.2004.1260760 (2004).
88. Masataka, S., Y.-W. Chen, & H. Aoki. Automatic transformation of "KOGAO" (small face) based on Fast B-spline approximation. *Journal of Information Hiding and Multimedia Signal Processing* 2: 192-203, http:// bit.kuas.edu.tw/~jihmsp/2011/vol2/JIH-MSP-2011-03-008.pdf (2011).
89. Gu, X., X. Jia, B. Dong, Gautier, Q., & S. Jiang. A contour-guided demons deformable image registration algorithm for Adaptive radiotherapy. *International Journal of Radiation Oncology • Biology • Physics* 81: S803-S804, DOI: 10.1088/0031-9155/58/6/1889 (2013)
90. Gu, X., H. Pan, Y. Liang, R. Castillo, D. Yang, D. Choi, *et al.*, Implementation and evaluation of various demons deformable image registration algorithms on a GPU. *Physics in Medicine and Biology* 55: 207, DOI: 10.1088/0031-9155/55/1/012 (2010).
91. Wong, S.M., K.S. Ng, & T.S. Li. Computing deformable image registration by inverse multiquadric demons algorithm. *IACSIT International Journal of Engineering and Technology,* Vol. 6, No. 2, April (2014).
92. Cachier, P. & D. Rey. Symmetrization of the Non-rigid Registration Problem Using Inversion-Invariant Energies: Application to Multiple Sclerosis. In: *Proceedings of the Third International Conference on Medical Image Computing and Computer-Assisted Intervention,* DOI: 10.1007/978-3-540-40899-4_48 (2000).
93. Yang, X., Z. Li, & M. Pan. Asymmetrical interpolation methods and applications in medical

image registration. *Journal of Computational Information Systems* 8: 99–106, www.jofcis.com/publishedpapers/2012_8_1_99_106.pdf (2012).
94. Rogelj, P. & S. Kovacic. Symmetric image registration. p. 334-343, DOI: 10.1.1.58.950 (2003).
95. Heinrich, M.P., M. Jenkinson, M. Bhushan, T. Matin, F.V. Gleeson, S. M. Brady, *et al.* MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis,* 16: 1423-1435, DOI: 10.1016/j.media.2012.05.008 (2012).
96. Sabuncu, M. R., B.T.T. Yeo, K. Van Leemput, T. Vercauteren, & P. Golland. Asymmetric image-template registration, In: *Medical Image Computing and Computer-assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention,* vol. 12, p. 565-573, DOI: 10.1007/978-3-642-04268-3_70 (2009).
97. Kashani, R., M. Hub, J.M. Balter, M.L. Kessler, L. Dong, L. Zhang, *et al.* Objective assessment of deformable image registration in radiotherapy: A multi-institution study. *Medical Physics* 35: 5944-5953, DOI: 10.1118/1.3013563 (2008).
98. Glocker, B., A. Sotiras, N. Komodakis, & N. Paragios. Deformable medical image registration: setting the state of the art with discrete methods. *Annual Review of Biomedical Engineering* 13: 219-244, DOI: 10.1146/annurev-bioeng-071910-124649 (2011).
99. Kadoya, N., Y. Fujita, Y. Katsuta, S. Dobashi, K. Takeda, K. Kishi, *et al.* Evaluation of various deformable image registration algorithms for thoracic images. *Journal of Radiation Research* 5: 175-182*,* DOI: 10.1093/jrr/rrt093 (2013)

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 131–144 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Resource Optimization in Job-shop Scheduling using Ant-Colony-Optimization Metaheuristic

Shahid A. Khan[1], Muhammad A. Alam[2]*, and Muhammad Umer[1]

[1]Abasyn University, Islamabad, Pakistan
[2]Iqra University, Islamabad, Pakistan

Abstract: Present study elucidates the probity of ant colony optimization metaheuristic in minimizing the makespan by efficiently allocating jobs to workstations in general aviation maintenance. The metaheuristic technique is applied to real workplace problems in general aviation sector of Pakistan to resolve scheduling quandaries of XT-10 helicopters inspection in Burq Air Services (Pseudo names of organization and helicopter to keep anonymity). Secondary data for processing times of jobs at workstations was obtained from job cards and process sheets. Matlab codes were developed for reaching the optimal scheduling. Results indicated almost 25% improvement in efficiency, and proffered a customized yet efficient solution to scheduling problem in real aviation maintenance setup. The study posited that with the slight adjustment, the present model could be applied to other variants of job-shop, service industry, and similar areas of social sciences.

Keywords: Ant-colony-optimization, heuristic, job-shop, scheduling problem, makespan, aviation, helicopter

1. INTRODUCTION

One of the means to achieve competitive advantage is resource optimization, i.e., judicious and efficient utilization of available resources. Scheduling is one of the most effective tools for achieving resource optimization. Its main objective is to minimize makespan, i.e., completing all the jobs in shortest time span. It aims at finding most efficient mix of machines, jobs (tasks) and workers so that human resource and machines are utilized optimally. It entails allocating shared resources over time to mutually competing activities. One of the main purposes of scheduling is to increase efficiency. In other words, scheduling is the optimization of multiple jobs and limited resources for efficiency, and is an important area of research [12].

Allocation of jobs to workstations is a basic yet intricate scheduling problem for two reasons. Firstly, scheduling possibilities increase as factorial of number of jobs. For example, scheduling problem involving 4 jobs presents 4! or 24 scheduling possibilities, while a problem involving 25 jobs present 25! or 15,511,210,043,330,985, 984,000,000 scheduling possibilities. Secondly, various assumptions are to be met in order to solve scheduling problems. Nonetheless, in majority of cases all these assumptions do not correspond to the actual problems, resultantly the quality of model and solution is compromised. Conventionally, experienced managers are employed for job-shop scheduling using manual methods to deal with various uncertainties and intricacies. With increase in the number of scheduling possibilities, manual technique becomes increasingly complex. Hence, seeking the best possible way to complete the required work in the quickest possible manner becomes a challenge. To address this issue, initially various calculus and exact techniques were presented, e.g., Fibonacci method, enumerative techniques etc. But these techniques have various limitations i.e. calculus techniques have tendency to stick to local optima, while the enumerative techniques demand considerable computational effort and long times to reach at optimal solution [30]. Alternatively, various heuristic techniques were developed.

Received, September 2015; Accepted, April 2016
*Corresponding author: Muhammad A. Alam; Email: alamgemex@yahoo.com

Heuristic techniques provide good results in shorter times with nominal computational resources but the optimal solution is not guaranteed i.e. guaranteed optimal solution is traded off with less computational effort and time. Heuristics are basically approximation algorithms, which are employed in cases where exact solution is either not available or where exact methods take too long to reach a solution. Job scheduling problem and time constraints have been studied extensively and several heuristic approaches have been developed for its solution [16]. Metaheuristics and the proposed algorithms evaluated under different scenarios demonstrated that these strategies outperform other available methods [7, 16, 35]. These techniques are being increasingly employed for optimization problems. Each heuristic technique has its own advantages and limitations. Ant-Colony-Optimization (ACO) is one of the heuristic based techniques that have been used extensively to solve complex problems. It is claimed that this technique provides the best solution in a very short time with nominal computation resources [31, 33]. It uses artificial intelligence (AI) and is metaheuristic in nature, and can be applied to different optimization problems with few job specific modifications [12]. Prior researchers have extensively tested ACO metaheuristic to resolve variety of complex problems [e.g., 10, 13, 20, 23, 27] but it has never been applied to job-shop scheduling problem in aviation maintenance.

This study endeavors to address job-shop scheduling problem (JSSP) in aviation maintenance setup utilizing ACO technique, and formulates workable model to significantly minimize the makespan. In so doing, it seeks to answer two main queries: a) How to formulate problem and develop an algorithmic tool based on ACO metaheuristic that minimizes the makespan? b) How to apply ACO metaheuristic to a real practical problem? The answer to these queries addressed in this study will contribute to the existing body of knowledge on applied optimization and metaheuristic in aviation maintenance.

1.1 Theoretical Foundation

This study is conceptually inspired by prior work of Umer et al. [34] and Aftab et al. [4]. Theoretically, it approaches JSSP through ontological premise of determinism i.e. all events (e.g., job-shop maintenance scheduling) have causes, and one events (e.g., inspection) can be linked to another event through general laws [2]. Epistemologically, the trade-offs between theory development and applying it for resolving job-shop scheduling problems in aviation sector is dealt with by foundationalism, and espousing a positivist methodology [29]. A concise outline of basic features presented herein provided foundation for model development.

1.2 Ant-Colony-Optimization

Ant-Colony-Optimization (ACO) algorithm is a probabilistic technique employed for solving computational problems. Its application enables finding good or good enough solutions to optimization problems by determining paths through graphs. It has been developed based upon Evolutionary Algorithm, which utilizes guided random search techniques. ACO can be traced back to theory of 'stigmergy' presented by Pierre-Paul Grassé in the year 1959 [14]. Stigmergy means communication and coordination between agents through indirect means like pheromone level in ant's system. In 1980s research was conducted on the ants' colonies, their behavior and social aspects. In 1991, Dorigo in his PhD thesis first presented the concept of ant colony for solving optimization problems [11]. In mid 80s and late 90s other researchers like Stutzle, Hoos, and Binachietc worked on the development of ACO concept on variety of problems and applications. Subsequent developments in ACO framework have enabled its application to diverse problems [10, 13, 20, 23, 27]. Resultantly ACO has emerged as one of the preferred application tools and promising area of research [12]. JSSP using ACO metaheuristic was first attempted in 1994. However, with the passage of time ACO metaheuristic was applied to various other variants of job-shopissues [12].

Recent literature published on ACO metaheuristic envisages its growing popularity. Jing and Tomohiro [21] presented hybrid approach using two optimization techniques, ACO and Tabu search. It was applied toflexible job-shop scheduling problem (FJSSP) with multi objectives, the primary objective being makespan minimization. Many researchers claim that proposed hybrid approach provide far superior results than the other optimization algorithm [e.g., 10, 13, 20, 21, 23, 27, 32, 37]. Remarkable improvements were achieved with regard to time required to solve the problems while maintaining good accuracy closer to genetic algorithm and exhaustive search [4]. Huang et al. [19] came up with new method named 2PH-ACO (two pheromone Ant-Colony-Optimization), a variant of ACO metaheuristic. It was applied to FJSSP, and results obtained were better than the traditional ACO. More recently, the concept of Neural Augmented ACO (NaACO) was coined to address worker assignment problem besides JSSP by combining Artificial Neural Network (ANN) and ACO [33, 34]. Application of ACO on various types of JSSP and FJSSP validated its application [26].

1.3 Job-Shop Scheduling Problems and Makespan

Job-shop is characterized by low volume of production and high customization. A job-shop entails various operations to be performed in a definite sequence. The machine sequence of jobs (also called process plan) is fixed, however, the problematic is to find a particular sequence on machines "m" in such a way that all "n" number of jobs are completed most efficiently. This is called 'Job-shop Scheduling Problem' (JSSP). The main objective is most optimal utilization of resources aiming at efficiency and elimination of all sorts of wastes [25, 36]. It is done by allocating resources like machines and workers to jobs in such a way that idle time of machines and workers are minimized and jobs are completed in minimum possible time. In Job-shop, machines are arranged according to functions or processes, and this arrangement is termed as process layout [17]. Conventionally scheduling has been carried out manually for simple scenarios. However with the increase in scheduling possibilities, the scheduling become increasing complex where manual working does not remain a feasible option and require solution through advanced techniques like heuristics or calculus based techniques.

The reported history of JSSP spans over 5 decades. Fisher and Thomson introduced famous 10 x 10 problem (ten jobs and ten machines). This particular instance of 10-job 10-machine problem remained unresolved for over 25 years [8]. The most commonly used measure of efficiency in JSSP is makespan (time required to complete all the jobs). Makespan is dependent upon the sequence in which jobs are fed to machines. The other measures of efficiency include job lateness, job tardiness, job flow time etc. Job-shop scenario of 'n' jobs and 'm' machines presents $(n!)^m$ scheduling possibilities. If we consider single sequence of processing through all the machines than scheduling possibilities reduce to n! Thus scheduling of 3 jobs on 2 machines in a single sequence presents six scheduling possibilities (a simple manual method i.e. Johnson's rule).

With increase in number of scheduling possibilities, manual methods like Johnson's rule do not remain a workable option, and various exact techniques were presented. However, these techniques require a lot of computation effort and unrealistically long times to reach to an optimal solution. Another issue is that there are many problems for which exact solutions are not possible. Thus scheduling problems presents hardest optimization problems which are NP-complete [6]. In the similar vein, various heuristic based techniques were presented. These techniques provide good results in shorter times with nominal computation resources however best results are not guaranteed. Flexible Job-shop Scheduling Problem (FJSSP) is an extension of JSSP that incorporates flexibility of route, and provides decision point for assignment to more than one machine. Likewise, non-deterministic polynomial times i.e. 'NP complete problem' is a problem in which there is no efficient way to reach the solution directly and no fast solution to such problem is known. If any of the currently known algorithms is used, time required to solve the problem increases remarkably with increase in solution space. In other words best solution for these problems is not possible within polynomial bounded computation time [12]. Solutions of such problems are obtained through methods like heuristic or approximation algorithms.

1.4 Optimization

Optimization simply means achieving the best solution of a problem under given set of constraints. When best solution is not possible due to nature of problem, limited computation ability and longer times involved, then good or good-enough solutions are searched. Optimization problems are maximization or minimization problems. An objective function and set of constraints are defined for the problem. Solutions that satisfy the constraints form a set of feasible solutions from which best or good solution is selected. The easiest way is to find out all the possible solutions (exhaustive search) and select the best one. But, number of possible solutions becomes too large for slightly complex problem that exhaustive search becomes inappropriate. Resultantly, the approximation algorithms are the wise option. At European based airports, Ravizza et al. [28] found optimization approach more convincing for settling ground movement of aircrafts. Optimization problem in which the feasible solution consists of discrete members is termed as combinatorial optimization. On a broader level, JSSP can be classified as static combinatorial problem as variables like number of machines and number of jobs are discrete and fixed.

NP hard problems can be solved by calculus methods or through application of algorithms. The algorithms can be exact or approximate. Exact algorithms provide best solutions. However, their usage is constrained due to two factors. First is their inability to solve beyond certain number of dimension, and second is that even nominally complex problems require long times to reach solution. That is why their usage for optimization problem after certain level becomes inappropriate. On the other hand, approximate or partial algorithms provide solution to optimization problems in reasonable time frame. They are also known as heuristic [12, 30].

An optimization problem has many possible solutions called solution space. Each individual possible solution is called a candidate solution. In algorithm based optimization methods working principle the solution is searched in the solution space and results with desired accuracy are obtained. Method of search progression and criteria for results is defined pre hand. In the exhaustive search methods, all the possible ways are explored and investigated, thus best solutions are guaranteed. However, these methods become infeasible for sizably voluminous solution space. In comparison, the local search methods apply metaheuristic method that iteratively seeks neighboring solution candidates only starting from an initial one. The termination may be according to lapse of defined time or non-improvement in results after certain iterations. This method has limitation that at times solution may only be locally optimum, but not globally. Global optimum means that the solution obtained is the best one amongst the entire solution space, while local optimum means solution obtained is the best one only amongst the neighboring solution candidates, which may or may not be best in the entire solution space. An example is shown in Fig. 1.

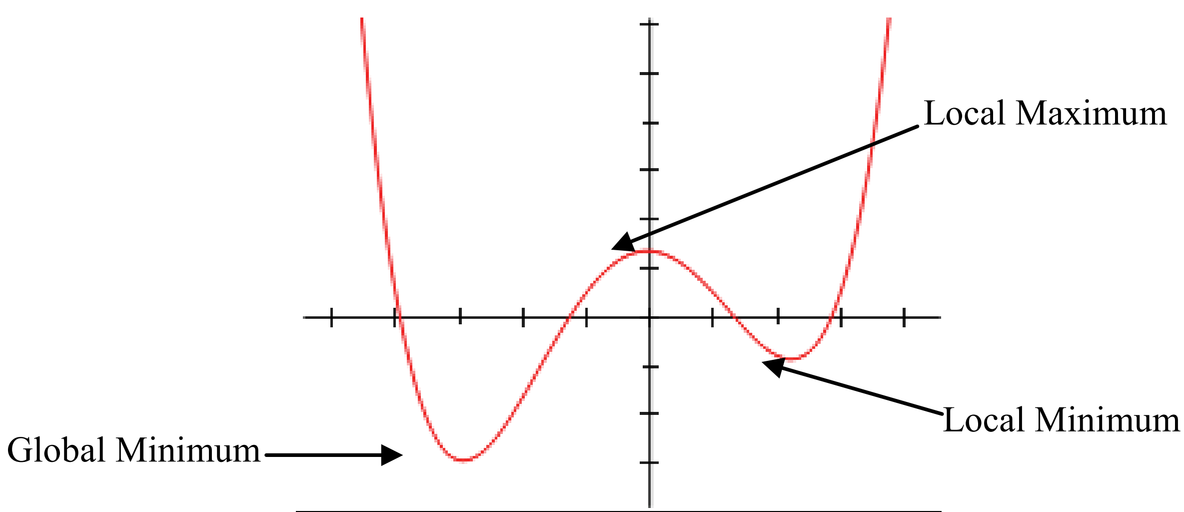


Fig. 1. Global and local optimum [15].

1.5 Metaheuristics

Heuristics are basically approximation algorithms which are employed in cases where exact solution is either not available or where classical methods take too long to reach solution. It has been applied to a wide range of disciplines [1, 7, 22, 24, 31]. It aims at finding solution of problem in hand in a reasonable timeframe that is good or good enough but may not be the best. Thus, guaranteed optimal solution is traded off with less computational effort and time to reach the solution. Heuristics have its limitation: a) Problem specific information is required for working of heuristic; b) Single run algorithms provide limited solutions and terminates search on reaching local optima. The problem has not been addressed despite incorporating the improvement that runs the heuristic various times [12]. In order to address limitations of heuristics, concept of metaheuristic has been introduced. A metaheuristic is a set of algorithmic concepts that can be used to define heuristic methods applicable to a wide set of different problems [12]. Thus, specific problem would require incorporation of only few modifications. Alternatively, it can be said that a metaheuristic is basically higher-level heuristic built upon lower level heuristic. Metaheuristics also tend to bypass the problem of entrapment in local optima. It provides good or good enough solution to optimization problems in a quick time frame with limited computation capacity. Thus, these can be referred to as soft computing techniques capable of solving hard problems. In metaheuristic, first a set of solutions is sampled. This set of solutions is searched to find feasible solution and ultimately the good solutions. In certain cases, few assumptions are made especially where information is not perfect or complete. This flexible property enhances utilization domain of metaheuristic for diverse categories of problems [9]. Available metaheuristics differ in two aspects, the way in which they avoid entrapment in local optima and the way in which solution space is searched [12].

1.6 Conceptual Framework

ACO employed in this study is metaheuristic-based and has ability to converge to global optima. ACO algorithm refers to ‘swarm intelligence’: a concept used in artificial intelligence inspired from natural systems. Swarm means collection of agents. These agents demonstrate alike behavioral traits, and mostly adhere to certain rules on collective level. Their interplay results in collective intelligence that would not be possible individually or at non-group level. Examples include honey bee, ant colonies, etc. [5, 13, 23].

The ACO heuristic has been developed based on behavioral pattern of ants foraging for food. The ants search for food more or less follows a certain pattern [3]. Initially ants roam randomly looking for food source. Once food source is found, ants return to their colony. On their way back, these ants lay a chemical called ‘pheromone’. Consequently a pheromone trail is formed, which can be sensed by other ants. Once other ants wandering for food sense a pheromone trail, they no longer move randomly rather follow the pheromone trail. While returning from the food source, these ants reinforce the pheromone trail. However, pheromone trail also evaporates with time resulting in decrease of attractive strength, and on a longer path pheromones strength decrease due to evaporation. If a shorter path is detected, then more ants follow this shorter path. This results in increase of pheromone level on this shorter path. This positive feedback eventually leads to following a single short path. Advantage of pheromone evaporation is that longer paths are avoided due to decreased pheromone level. In the absence of evaporation phenomenon, new shorter paths would not have been possible [3]. In ACO algorithm, the artificial ants replicate the real ants. “Simulated ants" imitate natural ants behavior in search of optimal solution. Its working is based on the principle of indirect communication between ants (stigmergy) through pheromone. The artificial ants deposit pheromone, which is taken as numerical information. Based on this information, probabilities are calculated which directs the artificial ants to shortest paths. The pheromone trails are reinforced or evaporated as per the experience of the artificial ants.

In ACO, metaheuristic delineates: a) Problem in general terms; b) Objective function and behavior of the ants in relation to the objective function and solution construction procedure i.e. rules for solution built up; c) Pheromone Update i.e. rules for pheromone updating; and d) Daemon actions i.e. actions prohibited for single ant. The simulated ants move asynchronously at same time and solution is built up incrementally. Better solutions are given higher pheromone level thus probability of artificial ants to converge to better solutions increase iteratively. This knowledge is updated for entire community and is used in solution building. Desired preferences are heuristically defined, which adds remaining components to the solution [23]. Thus, provision of correct heuristic information is pivotal in reaching solution of excellent quality. The algorithm follows three basic steps i.e. construct ant's solutions, update pheromones and daemon actions. Updating the pheromones describes how pheromone levels will be increased (reinforced) or decreased (evaporated). Daemon actions are related to collective wisdom of colony, which is not possible individually. The three procedures are managed through schedule activities construct; however scheduling is left on the discretion of designer to specify as per the requirements of problem [12].

2. METHODS

2.1 Sample and Data

The data included in this study presents a real workplace JSSPs that are encountered in Burq Air Services (Pseudo name for general aviation organization) that owns a large fleet of XT-10 Helicopters (Pseudo name), which has an elaborative maintenance setup is in place. Purposive sampling technique i.e. total population sampling was espoused, and entire fleet of 60 XT-10 helicopters was included in the study. These helicopters have to undergo season change inspection on three workstations. As a standard, each inspection takes 120 man-hours however it fluctuate towards upper side due various factors, major one being unscheduled tasks detected during inspection. The helicopters were divided into five batches of 12 helicopters, which is a multiple of number of workstations (3 workstations) as well as a factor of total population (60 helicopters). This way inspection of the complete population can be completed in 5 steps with the critical advantage that synchronous loading of the three workstations remains a candidate solution. An additional advantage that can be accrued from this number is that once inspection of a batch is completed, its data can be incorporated as past data for the next set of jobs. In order to have substantial amount of data for comparison and testing of algorithm, secondary data for period covering last six seasons was selected. Data set for each season was taken as separate problem set. Data was collected using job

Table 1. Data for problem set 1.

Workstation 1				Workstation 2				Workstation 3			
Helicopter	Airframe	Engine	Avionics	Helicopter	Airframe	Engine	Avionics	Helicopter	Airframe	Engine	Avionics
	Hours	Hours	Hours		Hours	Hours	Hours		Hours	Hours	Hours
1	103	22	4	1	100	23	5	1	99	22	4
2	100	23	4	2	100	22	4	2	110	20	5
3	101	25	4	3	105	25	4	3	106	21	4
4	103	21	5	4	115	22	6	4	96	22	6
5	96	25	4	5	96	28	4	5	107	20	4
6	100	22	4	6	115	21	5	6	99	20	4
7	104	28	5	7	100	22	4	7	109	21	5
8	99	21	4	8	103	20	6	8	96	22	4
9	96	22	7	9	110	20	4	9	101	22	4
10	103	24	4	10	96	21	5	10	103	24	6
11	96	23	5	11	96	22	4	11	120	20	6
12	110	20	4	12	96	21	4	12	96	27	4

Note: Hours means man-hours

cards and process sheets for analysis as well as for comparison and validation of results.

In Burq Air Services, a job card is a controlled document used for every work order. It records all the materials / spares used during the job, details of the technicians & inspectors who have worked on the job and the time taken besides other auxiliary details. Similarly, the process sheet is another controlled document, which includes all inspection steps, reference for the technical manual, names & signatures of technicians and inspectors, standard time allowed against each step, time actually taken at each step and various other auxiliary details. Data set for each season was considered a separate problem set thus presenting total of six problems. A sample data set is shown in Table 1.

2.2 Problem and Constraints

Various inspections mandated by the OEM (Original Equipment Manufacturer) of the helicopter are carried out to ensure safe operations. This study focused on the season change inspection of XT-10 helicopters. This inspection is performed at the onset of every season i.e. summer and winter. Thus all 60 helicopters become due for inspection at the same time. This context exactly matches the generic assumption of JSSP thus enabling true modeling of the actual scenario. To undertake these inspections three maintenance setups are available i.e. workstation 1, workstation 2 and workstation 3. Each inspection comprises three trade examinations called Airframe, Engine and Avionics, and every workstation is capable to undertake the complete inspection and all three workstations work in parallel.

All the scheduling solutions are based on certain assumptions/constraints:

- There are *n* jobs that are required to be assigned to workstations. Number of these jobs remains fixed and does not vary for the problem under consideration.
- At the start, say (time zero) all the *n* jobs are ready to be assigned to the workstations, and very individual job consists of one operation as a whole.
- There are *w* workstations and all are ready to take jobs. Number of these workstations remains fixed and does not vary for the problem being considered. Handling capacity of any of the workstation is limited to one job at one time. In other words one workstation cannot undertake two or more jobs simultaneously.
- Each workstation has known benchmark-processing times that are fixed. There are no breakdowns and setup time for any of the workstation and any job can be assigned to any of the workstation initially and subsequently.
- There is no transportation time between workstations.
- Enough workers are available at each of the workstation to undertake the job, and all workers are equally proficient to undertake the job.
- Splitting any job and retrieving any job incomplete and assigning to other workstation is prohibited.

2.3 Instruments

The complete scenario was modeled mathematically based on the generic assumptions of JSSP in the form of equations, expressions and inequalities that presented constraints for the model. Objective function was set as makespan minimization. It is important to highlight that all the generic assumptions of JSSP were consistent with the problem considered except: a) zero set up time; and b) fixed processing times. The former was addressed by adding set up times in the processing times while the latter inconsistency was addressed by deriving processing times from past data instead of minimum standard bench mark time of 120 hours. Processing time at each workstation was worked out according to PC Hu rule [18]:

$$t_k = X_k + \frac{Y_k}{Z_k} \qquad \text{(a)}$$

For makespan minimization, the ant (job) should be assigned to nearest food source (workstation) based upon the shortest path (processing time). The probability of an ant *j* to converge to a food source *k* was expressed as under:

$$P_{j,k} = \frac{[T_{k,j}]^{\alpha}[\eta_{k,j}]^{\beta}}{\Sigma[T_{k\prime,j}]^{\alpha}[\eta_{k\prime,j}]^{\beta}}$$ (Dorigo & Stutzle [12], pp. 171-172 (b)

Where:

$k'j$ represented job not assigned to workstation k

$T_{k,j}$ represented desirability of assigning job j to workstation k based on probability

Here it is the pheromone value or pheromone level. It is given by the expression: -

$T_{k,j} = T + \Delta T$

Whereas T= Present pheromone level and

$\Delta T =$ Evaporation or reinforcement

$\eta_{k,j}$ is the reciprocal of the heuristic function defined in equation (a). So it can be written as: -

$\eta_{k,j}$ = 1/ heuristic function =

$$\frac{1}{t_{kj}} = 1/(X_{kj} + \frac{Y_{kj}}{Z_{kj}}) \quad \text{(c)}$$

The α and β are sensitivity factors that represent the boundary level condition for this probability.

Initial triggering is independent of the evaporation or reinforcement so

$$[T_{k,j}]^{\alpha} = 1 \ \& \ [T_{k\prime,j}]^{\alpha} = 1$$

The jobs are assigned to workstations as per heuristic function of minimum processing times as per equation (b). Total flow time at workstation k (k = 1, 2 or 3) is the sum of processing times of all the assigned jobs to workstation k while makespan is the maximum of the total flow time at workstation k.

Based on above, algorithm was defined in Matlab seeking for output schedule of the jobs with minimum makespan. The problem data sets were analyzed using this ACO based algorithm, which provided output schedules of jobs with minimum makespan.

2.4 Construct Validity and Reliability

Aftab et al. [4] applied the similar technique to well-known set of 100 problems and validated it through comparison of results with exhaustive search and genetic algorithm. The standard for the comparison was average deviation from the best solution through exhaustive search and time taken to solve these problems [4]. Moreover, the method has also been tested and validated by prior researchers, and as compared with others this method produced better solutions [e.g., 1, 34]. Umer et al. [34] validated neural augmented ACO technique (NaACO) through application on set of 100 problems. The heuristic information in both the cases was same as in this study i.e. processing times from past data although the applications scenarios were different. Thus defined ACO metaheuristic and its Matlab code are validated. Likewise, both the documents (Job card and process sheet) used for data collection are controlled documents of Burq Air Services. These documents form the basis for all the planning and control mechanism of the organization. Job cards, in addition, have information on materials used and thus have

Table 2. Summary of results.

Problem #	Jobs Assigned to Workstations			Makespan (Hours)
	WS 1	WS 2	WS 3	
Problem 1	4 (3,5,9,11)	4 (2,7,10,12)	4 (1,4,6,8)	412.4500
Problem 2	4 (2,7,10,12)	4 (1,4,6,8)	4 (3,5,9,11)	411.4500
Problem 3	4 (1,4,9,11)	4 (3,5,7,10)	4 (2,6,8,12)	412.9167
Problem 4	4 (2,5,8,11)	5 (1,4,7,10,12)	3 (3,6,9)	512.5000
Problem 5	4 (1,6,8,12)	4 (3,5,7,10)	4 (2,4,9,11)	428.7429
Problem 6	4 (3,5,8,12)	4 (1,4,7,10)	4 (2,6,9,11)	419.0833

Note: Time taken to solve the problem = 0.017800seconds financial aspects, which add to the accuracy of the document. In other words, these two are the most authentic documents for man-hours data collection. Therefore, data collected for this study is considered reliable and accurate.

3. RESULTS

ACO metaheuristic has been applied for scheduling of jobs (helicopters) to workstations and calculations of makespan. Each data table concerns a separate problem, thus presenting a total of six separate problem sets requires scheduling. These have been labeled as problem 1 to 6. Summary of study results is presented in Table 2.

Table 2 elucidated that workstations have been loaded with equal number of jobs for all the problem sets except for problem set 4 where workstation 2 has been assigned 5 jobs, while workstation 3 has been assigned 3 jobs. Analysis of data for problem set 4 revealed that workstation 2 was taking times close to the minimum benchmark time while workstation 3 was taking times much more than the minimum benchmark time. Thus workstation 2 was able to complete the 4th job prior to the end of 3rd job by the workstation 3, and 4th job by workstation 1. Therefore workstation 2 was loaded with the 5th job as shown graphically in Fig. 2. Investigation revealed two major factors were attributable to lower efficiency of workstation 3 at that time. First, severe deficiency of manpower at workstation 3 at that point of time, and second was breakdown of specific equipment required for the inspection at that point of time. It can be inferred from the results that scheduling is not only confined to the efficient utilization of the resources, rather it also signals some underlying problems demanding investigations and subsequent corrective measures in such like scenarios. It is worth mentioning that all these loading are probabilistic in nature derived from the past data. However, in practice variations are expected due to factors attributable to jobs and/or workstations. Job related factors might include some unscheduled works required on specific jobs. Workstation factors may include non-availability of required equipment, workers availability, workers proficiency, etc. Minimum makespan for all the problem sets are shown in Table 2 indicating that jobs would be completed in this time if matching assignment of jobs to workstations were adopted.

Amongst all possible scheduling possibilities, the projected sequence presented minimum time to complete all the jobs. In other words this sequence enables minimum makespan. Assuming uniform loading of the workstations, i.e., 4 jobs each to all the 3 workstations and calculation of corresponding makespan would enable uniform comparison. Based on past data, the makespan has been worked out by calculating average time taken to complete 4 jobs at each workstation, and choosing the maximum one. This comparison is shown in Table 3. Results indicated improvements up to 25% through application of ACO technique.

4. DISCUSSION

The study presents 1,714,233,849 scheduling possibilities if simple exhaustive search is applied.

Table 3. Comparison of results with empirical data.

Problem#	Average Time to Complete 4 Jobs			Makespan*	Results Achieved through ACO	% Improvement
	WS1	WS2	WS3			
1	514	518	520	520	412	21
2	518	520	510	520	411	21
3	517	533	547	547	413	24
4	553	556	580	580	513	12
5	574	574	571	574	429	25
6	536	538	540	540	419	22

*Makespan for the empirical data is the maximum of the time to complete 4 jobs by WS1, WS2 and WS3.

However, the proposed model established that the number of jobs 'n' assigned to a workstation 'w' matters and not their sequence. It means that processing jobs assigned to workstations in any order would not result in change of flow times and makespan. For example, if 4 jobs (say job # 1, 5, 7 and 9) are assigned to workstation 1, the processing of these jobs in any sequence would result in same flow time and makespan. This would hold true even for the scenarios where all the jobs are not available at time zero provided they are made available before completion of the preceding job. It reduces the possibilities to 91 for each problem set, thus presenting a total of 546 possibilities for all the six problem sets under consideration. These possibilities are explored heuristically for optimal solutions to bring a single optimal schedule. A worth mentioning achievement of this work is that all the constraints and limitations posed by the problem have been well catered for while formulating the problem. No unnecessary and unreal simplifications have been made to the problem in hand. Thus, the developed model encompasses the problem in its real form. In line with prior researchers [e.g., 1, 10, 13, 20, 22, 23, 24, 27], this work has come up with a customized yet efficient way to optimize the scheduling problem in a real workplace setup.

4.1 Practical Application

Aviation maintenance is not only critical from safety point of view but also equally important from financial perspective due to intense competition. Careful application of the metaheuristic technique presented in this study can provide scheduling efficiencies, which has direct implications for financial gains, and indirect implications for safety enhancement. Being the first attempt, this study has addressed the scheduling problem that is encountered in an aviation maintenance setup to schedule upcoming inspections using artificial intelligence through ACO metaheuristic. It has provided a scheduling technique, which is not merely based upon theory or standard times, but provided its practical application based on the actual past data. The findings are important for systems where processing times are not fixed rather varies due to various factors like unscheduled works, different proficiencies of workers, breakdowns etc. These scenarios are normally encountered in aviation setups and pose serious challenges to optimization researchers. Though this work addressed scheduling problem in job-shop setting, with little modification it can be easily extended to various other variants like open shop scheduling problem (OSSP), group shop scheduling problem (GSSP) etc. This effort addresses scheduling problem of twelve jobs assignments to three workstations. With minor changes, it can be tailored to any number of workstations and jobs. An important aspect is that this work can be applied in service setup as well as workstations and jobs. This work can be applied in other allied areas of social and medical sciences for solving optimization problems with little or no modification. This study enabled futuristic loading of workstations based on past data as a forecasting tool. With these results in hand, it can be posited that the proposed ACO algorithm will help in better forecasting based on past data in variety of situations. The Burq Air Services and similar setups can accrue various advantages from this research for improved efficiency through better loading of workstations, better forecasting of future scenarios and indication of low efficiency workstations warranting investigations and subsequent corrective actions.

4.2 Limitations

This research attempted to cover all major facades of JSSP; however, there are certain limitations. First, the presented algorithm is not a generic tool; rather it is specific to the problem discussed in this study. It implies that the algorithm presented in this study requires customization to match other problems. Second, the developed tool is not simple, and its application is adjunct to method that alters work dynamics with technical arbitration [38]. It requires basic understanding of the working of algorithm and certain aspects of the problem in hand. Therefore, training is required prior to the application of the presented technique. Third, the technique does not incorporate pheromone up gradation because the problems set were small. The developed technique faces problem in assigning jobs to workstations when heuristic information (minimum processing time) is the same for two or more workstation. The simulated ants (jobs)

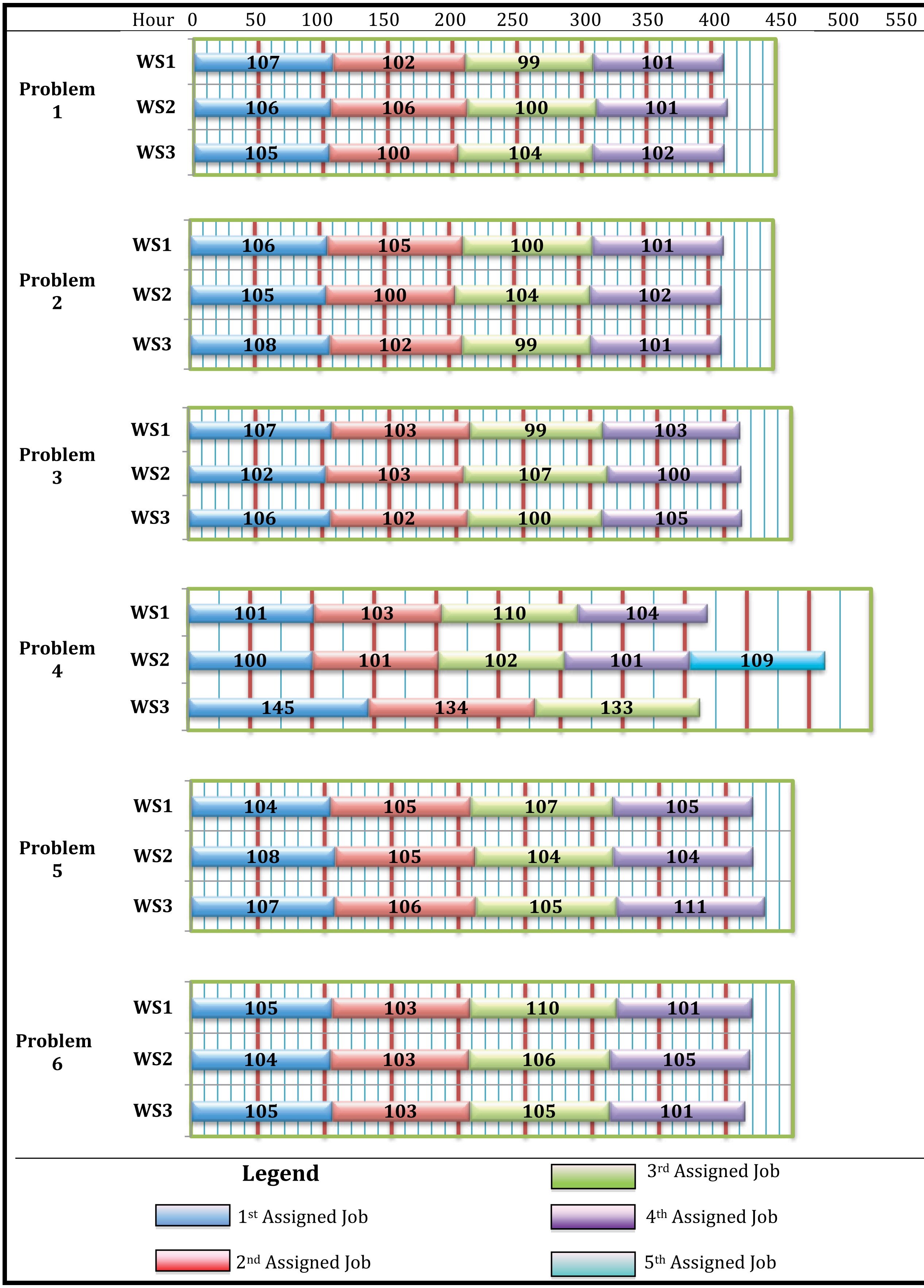


Fig. 2. Jobs assignment and Makespan of all six problems.

get confused in deciding which food source (workstation) to choose, as the probabilities are equal. However, it is a rare possibility as heuristic information is checked till the last decimal place. Fourth, the algorithm may present deviated results in cases where past data has huge dispersion with regards to the processing times. In such scenarios data should first be adjusted manually. Last, the work does not address worker's assignment of the problem, which can further reduce the makespan.

4.3 Future Prospects

Although this study extended the existing body of research on the topic, it has remained confined to the specific problem. Various avenues are opened for the future researchers in the form of extension and modification of the proposed technique. Firstly, formulation of generic tool capable of addressing wide variety of scheduling problems using ACO is an important area for future research. Secondly, improvement of technique through incorporation of pheromone up gradation functions is a promising area for future research. Thirdly, the study addressed problem sets involving only three machines or workstations and twelve jobs only. Problems with more workstations and jobs would pose interesting scenarios for future researches. Fourthly, improvement of the proposed technique to overcome equal probability problem is possible through EAS (Elitist Ant System). Besides, the improvement of the proposed technique various other problems can be modeled and solved through this variant of ACO. Thus, it is another favorable area for future researchers. Fifthly, the proposed algorithm does not address workers assignment problem. Incorporation of the same in the proposed technique using ACO is a potential area of research. Sixthly, present study focused only one particular type of inspection for XT-10 helicopter. Nonetheless, there are various other forms of inspections that are performed on XT-10 helicopters. Formulation of tool capable of addressing scheduling as well worker assignment problems for all sorts of inspections of XT-10 helicopters would be a challenging domain for the future researches. Similarly, domain of such optimization problems is strictly restricted to a certain setup without taking into account the upstream and downstream players i.e. suppliers and customers. Scenarios requiring optimization of complete supply chain would require intricate programming from optimization researchers. Lastly, application of this technique in service industry would be fascinating and challenging. Its application in hospital, food and transportation disciplines is strongly recommended for future studies in this field.

5. CONCLUSIONS

This study adds to the existing body of knowledge in the realm of optimization. It covered both the theoretical and practical application of optimization. All the constraints and limitations posed by the problems were catered for, and no unrealistic simplification was made. This work has come up with a customized yet efficient way of optimizing the scheduling problem in a real aviation maintenance job-shop setup to schedule the upcoming inspections using artificial intelligence through ACO metaheuristic which has not been attempted earlier. Application of presented ACO algorithm can help in better forecasting based on past data in a variety of situations. Aviation organizations can accrue various advantages from this research aimed at improving efficiency through better loading of workstations, forecasting the future scenarios, detecting the underperforming workstations, and taking the corrective actions. This work can also be taken as a reference source in future research to extend its application in other fields with requisite modifications.

6. REFERENCES

1. Abdullah, S., H. Turabieh, B. McCollum, & P. McMullan. A hybrid metaheuristic approach to the university course timetabling problem. *Journal of Heuristics* 18: 1-23 (2012).
2. Abel, R. *Man is the Measure*: *A Cordial Invitation to the Central Problems in Philosophy*. The Free Press, NY, USA (1976).
3. Adnan, M.A., M.A. Razzaque, I. Ahmed, & I.F. Isnin. Bio-Mimic Optimization Strategies in Wireless Sensor Networks: A Survey. *Sensors* 14: 299-345 (2013).
4. Aftab, M.T., M. Umer, & R. Ahmad. Jobs Scheduling and Worker Assignment Problem to Minimize Makespan using Ant Colony Optimization Metaheuristic. *World Academy of Sciences,*

Engineering and Technology 6: 12-29 (2012).

5. Alzaqebah, M., & S. Abdullah. An adaptive artificial bee colony and late-acceptance hill-climbing algorithm for examination timetabling. *Journal of Scheduling* 17(3): 249-262 (2014).
6. Barbulescu, L., J.P. Watson, L.D. Whitley, & A.E. Howe. Scheduling space–ground communications for the air force satellite control network. *Journal of Scheduling* 7(1): 7-34 (2004).
7. Bianco, L., P. Dell'Olmo, & S. Giordani. Scheduling models for air traffic control in terminal areas. *Journal of Scheduling* 9(3): 223-253 (2006).
8. Blazewich, J., W. Domschke, & E. Pesch. The jobshop scheduling problem: Conventional and new solution techniques. *European Journal of Operational* Research 93: 1-33 (1996).
9. Blum, C., & A. Roli. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys* 35 (3): 268-308 (2003).
10. Chelouah, R., & C. Baron. Ant colony algorithm hybridized with tabu and greedy searches as applied to multi-objective optimization in project management. *Journal of Heuristics* 13(6): 640-650 (2007).
11. Dorigo, M. *Optimization Learning and Natural Algorithms*. PhD Thesis,Politecnico di Milano, Italy (1992).
12. Dorigo, M., & T. Stutzle. *Ant Colony Optimization*. MIT Press: Cambridge, MA, USA (2004).
13. Ghiani, G., D. Laganà,G. Laporte, & F. Mari. Ant colony optimization for the arc routing problem with intermediate facilities under capacity and length restrictions. *Journal of heuristics* 16(2): 211-233 (2010).
14. Gloag, E.S., M.A. Javed, H. Wang, & M. L. Gee. Stigmergy: A key driver of self-organization in bacterial biofilms. *Communicative &Integrative Biology* 6 (6): 11541-11546 (2013).
15. Global Optimum. (n.d.). Wikipedia: http:// en.wikipedia.org/wiki/Global_optimum (Accessed on March 10, 2015).
16. Grandinetti, L., F. Guerriero ,L.D. Pugliese,& M. Sheikhalishahi. Heuristics for the local grid scheduling problem with processing time constraints. *Journal of Heuristics* 21(4): 523-547 (2015).
17. Groover, M.P. *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*. John Wiley & Sons (2007).
18. Hu, P.C. Minimizing total flow time for the worker assignment scheduling problem in the identical parallel-machine models. *The International Journal of Advanced Manufacturing Technology* 25 (9-10): 1046-1052 (2005).
19. Huang, R. H., C.L. Yang, & W.C. Cheng. Flexible job shop scheduling with due window- a two-pheromone ant colony approach. *International Journal of Production Economics* 141 (2): 685-697 (2013).
20. Jair, C.D. Paternina-Arboleda,V. Cantillo, & J.R. Montoya-Torres. A two-pheromone trail ant colony system—tabu search approach for the heterogeneous vehicle routing problem with time windows and multiple products. *Journal of Heuristics* 19(2): 233-252 (2013).
21. Jing, T., & M. Tomohiro. Multi-objective flexible job shop scheduling with uncertain processing time and machine available constraint based on hybrid optimization approach. In: *Automation and Logistics (ICAL), 2010 IEEE International Conference on Automation and Logistics*p. 581-586. IEEE (2010).
22. Kim, Y.D., B.J. Joo, & J.H. Shin. Heuristics for a two-stage hybrid flowshop scheduling problem with ready times and a product-mix ratio constraint. *Journal of Heuristics* 15(1): 19-42 (2009).
23. Liang, Y.C., Z.H. Lee, & Y. S. Chen. A novel ant colony optimization approach for on-line scheduling and due date determination. *Journal of Heuristics* 18(4): 571-591 (2012).
24. Lin, S.W., S.Y. Chou, & S. C. Chen. Meta-heuristic approaches for minimizing total earliness and tardiness penalties of single-machine scheduling with a common due date. *Journal of Heuristics* 13(2): 151-165 (2007).
25. Mauguière, P., J.C. Billaut, & J.L. Bouquard. New single machine and job-shop scheduling problems with availability constraints. *Journal of Scheduling* 8(3): 211-231 (2005).
26. Mehmood, N., M. Umar, & R. Ahmad. A survey of recent developments for JSSP and FJSSP using ACO. *Advanced Materials Research* 816: 1133-1139 (2013).
27. Monteiro, M.S., D.B. Fontes, & F.A. Fontes. Concave minimum cost network flow problems solved with a colony of ants. *Journal of Heuristics* 19(1): 1-33 (2013).
28. Ravizza, S., J.A. Atkin, & E.K. Burke. A more realistic approach for airport ground movement optimization with stand holding. *Journal of Scheduling* 17(5): 507-520 (2014).
29. Sada, A.N. & A. Maldonado. Research Methods In Education. *British Journal of Educational Studies* 55(4): 469-470 (2007).
30. Sumathi, S., & T. Hamsapriya. *Evolutionary Intelligence an Introduction to Theory and Applications with Matlab.* Springer, Berlin (2008).
31. Thevenin, S., N. Zufferey, & M. Widmer. Metaheuristics for a scheduling problem with rejection and tardiness penalties. *Journal of Scheduling* 18(1): 89-105 (2015).
32. Tseng, S.P., C.W. Tsai,J.L. Chen, M.C. Chiang, & C.S. Yang. Job shop scheduling based on ACO with a

hybrid solution construction strategy. *Fuzzy Systems (FUZZ), 2011 IEEE International Conference on Fuzzy Systems* p. 2922-2927.IEEE (2011).

33. Umer, M., R. Ahmad, & S.I. Butt. Intelligent pheromone up gradation mechanism through Neural augmented Ant Colony Optimization (NaACO) metaheuristic in machine scheduling. *Scientia Iranica* 21 (5): 1726-173 (2014).
34. Umer, M., R. Ahmad, & I. Chaudhry. Unsupervised Artificial Neural Networks (ANNs) For Intelligent Pheromone up Gradation. Further Evolution of Neural Augmented Ant Colony Optimization (NaACO). *Life Science Journal* 10(3): 318-327 (2013).
35. Vidal , T., T.G. Crainic, M. Gendreau,& C. Prins. Time-window relaxations in vehicle routing heuristics. *Journal of Heuristics* 21 (3): 329-358 (2015).

36 Zhang, X., & S. van de Velde. Two-machine interval shop scheduling with time lags. *Journal of Scheduling* 1-10 (2013).

37. Zhang, X., X. Chen, & Z. He. An ACO-based algorithm for parameter optimization of support vector machines. *Expert Systems with Applications* 37 (9): 6618-6628 (2010).
38. Alam, M.A. Techno-stress and productivity: Survey evidence from the aviation industry. *Journal of Air Transport Management* 50: 62-70 (2016).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 145–162 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Practices for Implementation of the Critical Success Factors in Software Outsourcing Partnership from Vendors' Perspective: A Literature Review

Sikandar Ali[1,2], and Siffat Ullah Khan[1,*]

[1]Department of Computer Science & IT, University of Malakand, Malakand, Pakistan
[2]Department of Computer Science & Software Technology, University of Swat, Mingora, Swat, Pakistan

Abstract: Software outsourcing partnership (SOP) is a trustful relationship between client and vendor organizations for shared goals. A SOP is different to ordinary software development outsourcing relationship. Usually a successful outsourcing relationship may lead to SOP. Software companies currently use a wide variety of mechanisms to outsource software development work. Beside all, SOP is an emerging strategy. The development of SOP depends on the proper implementation of various critical success factors (CSFs) like 'mutual trust', 'effective and timely communication', 'mutual interdependence and shared values', 'organisational proximity', 'quality production' and '3C (coordination, cooperation and collaboration)'. Moreover, the practices of SOP in the software industry are very little. The objective of this research is to identify practices for implementing CSFs in SOP. We have identified a list of 142 'practices' for various CSFs through SLR methods. Initially we have found 375 research articles through SLR process and finally 65 papers were selected. In total 142, practices identified through SLR for 14 CSFs.

Keywords: Systematic literature review, software outsourcing partnership, practices/solutions

1. INTRODUCTION

Software outsourcing partnership is a global software engineering (GSE) paradigm for developing high quality software at reduced cost. Software outsourcing partnership is different than software outsourcing. This is because software outsourcing is a contract-based relationship between client-vendor organisations whereas software outsourcing partnership is a collaborative relationship beyond organisational boundaries. Client-vendor relation in this fashion often crossing the traditional contractual limits. Here risks and benefits, investments and work load of joint labours are equally divided among the collaborative members. Companies achieve competitive advantages through inter and intra-organisational collaboration. Long term working relationships are developed based on bidirectional trust, mutual interdependence and win-win mind-set between partners. Companies usually develop collaboration to decrease the costs of obtaining appropriate information/understanding and capabilities or competencies needed for well-organized professional processes. Collaborative relationships usually are in the form of joint ventures, alliances, association or partnerships [1].

However, developing a fruitful long term cooperative relationship based on collaboration between two diverse businesses are more challenging and complex than commonly estimated. In view of Kelly et al. [2] disappointment proportion for collaborative relationships (like associations, alliances, joint ventures or partnerships) varied from 50% to 60%. Bamford et al. [1] reported in his research article that success rate was only 53%. When collaboration like partnership is in developing stage, the focus might very often be in financial and legal aspects. Beside all the complexities, collaboration still happens amongst

Received, November 2015; Accepted, May 2016
* Corresponding author: Siffat Ullah Khan; Email: siffatullah@uom.edu.pk

organisations. Focusing on social aspects beside legal and financial, might lead to improved and more lasting results when developing partnership [3].

Building a successful inter-organisational partnership is a multi-dimensional and iterative process in which legal, psychosocial, economical sub processes are concurrently taking place [4]. Shared goals and ownership, mutual interdependence, mutual trust, long term commitment, effective and timely communication, quality production and partner's proximity are constituent parts of a successful partnership [3].

Due to mega economic changes, internationalisation, competition from low wage emerging countries, and improvements in information and communication technology (ICT) from 1980 to 1990 many different kinds of organisations have been created including multiple vendor agreements, strategic networks, different types of conglomerates, alliances and joint ventures etc [5]. Different types of organisations having different kinds of needs, therefore different kinds of relationships are needed. Numerous of the aforementioned relationships, at least at the start, are just intended for cost savings. Initially the core activities were performed in-house and relationship is formed only to buy non-core activities from other organisations. Now relationship is usually formed to acquire the information/understanding of new technologies, knowledge and skills beyond the organisation's boundaries and competence. So interest in closer relationships (a partnering relationship) has grown, because partnership provides the opportunity to enter into the knowledge of new technologies. Other possible motives include access to skilled human resources, acquiring complementary skills, entrance to new markets, more focus on company core competence, focus on strategic issues, boosting innovation, to reduced time to-market, increasing product or service quality, transferring static costs to variable costs, and improving competitiveness [6, 7].

Software companies currently use a wide variety of mechanisms to source software development; they outsource development work, develop insource, expand insource capability through acquisitions, and build partnerships and joint ventures with counterpart organisations [8]. Four of the strategies are highlighted by Moe et al [8], including insource, outsource, separate profit centre and strategic partnership. Similarly Roy [9] present four internal outsourcing strategies, including internal governance, recuperation, outsourcing and partnership.

According to Kishore [10], outsourcing relationship can be categorised into four categories. These are support, alignment, reliance and alliance. Alliance is a relation with high trust and low contractual control. Outsourcing partnership is a types of an alliance relationship [11].

1.1 Outsourcing Partnership – What it is?

Outsourcing partnership is a widely used terminology in the literature but still no precise definition exists for it. It is a relationship composed of two words outsourcing and partnership and therefore thoughtful understanding of individual terms is desirable for its definition.

Outsourcing is the contracting of various system's sub-functions, programming, data entry, facilities management, maintenance operation, system integration, disaster recovery, data centre management, and telecommunication by client firm to external vendor.

According to Oxford English Dictionary [12] outsourcing is defined as "the obtaining of goods or service or components from an outside or foreign supplier, especially in place of an internal source" In view of Kinnula et al [6] it is "the process of transferring the responsibility for a specific business function from an employee group to a non-employee group".

The main reasons for outsourcing are cost savings, increased flexibility in bidirectional decision making, access to specialist expertise, improved quality of service, free management time when there is lack of resources, improved financial control [13]. According to Brinkerhoff and Jennifer [14] the reasons for outsourcing, includes marked pressure on organisation to reduce costs, increase core competencies, and to provide specialized expertise more effectively.

In the management literature the partnership type correlation between companies has been studied extensively [4]. For example, inter-firm cooperation has examined in the marketing discipline, partnering between manufacturers and distributors , manufacturers and sales agents , buyers and sellers as well as auditors and clients . While in computer literature empirical literature survey on the partnership relationship between outsourcer and outsourcee started to grow after 2000 in the Europe, US and Asia [4].

Lambert et al. [15] have the view as "a tailored business relationship based upon mutual trust, openness, shared risks, and shared rewards that yield a competitive advantage, resulting in business performance greater than that would be achieved by the firms individually". Lee et al. [16] define outsourcing partnership as "an inter-organisational relationship to achieve the participants shared goals".

In nut shell 'an outsourcing partnership' is a commonly used word with no clear-cut definition. It is used quite charitably by the academics, without proper definition. However, obliquely it is conceivable to develop a universally acceptable understanding of the use of the term: an outsourcing partnership is a partnering relationship resulting from the outsourcing process, unlike to the other types of relationships that can be engendered from the outsourcing process. It is a long term inter-organisational software development relationship between client and vendor organisations with mutual adjustment and renegotiations of tasks and commitment that exceeds mere contractual obligations stated in an initial phase of the collaboration.

In this research paper, we consider software outsourcing partnership as "a strategic partnering relationship resulting from a process of transferring the responsibility of developing software for a specific business function from an employee group to a non-employee group, including transfer of assets, such as personnel". SOP is a mutually beneficial, continuous and long term relationship, in which future plans, visions and confidential information is shared with partner organisations proactively and willingly, with the aim to help each other, in concentrating their skills and resources towards the right track.

1.2 Difference Between Ordinary Outsourcing and Partnership Outsourcing

Ordinary software development outsourcing (SDO) relationship is different than SDO partnership. This is because, in ordinary outsourcing relation a client contracts software development work to a an external vendor who provides development services for payment while outsourcing partnership is the superior form of ordinary outsourcing relationship [17-19]. SOP is a relation for long time based on the renegotiations of mutual adjusted task and commitment that supersede the initially agreed contractual terms and conditions that are mentioned as the start of the association [19]. It is flexible, long term and based on sharing of risks and benefits and future goals and visions. In practice only a fruitful outsourcing relationship is eligible to promote to outsourcing partnership [10]. It cannot be instantly developed, but rather, it shapes with the passage of time [6]. A key difference is in the level of depth; SOP is deeper relationship in which many traditional border line between companies are wrecked [6] . A relationship is said to be SOP, where the parties share confidential information about future plans, work together, combine resources, and share ownership, risks and benefits [5] and take joint decisions to undertake mutually beneficial business [20]. Outsourcing partnership is a good tool to overcome technological uncertainty, because outsourcing partnership is the unique type of outsourcing relationship where partners share information of unexpected events [21]. Here both the parties share tacit information, human resources, and work load, to achieve mutual goals [20]. The main difference between partnership and contractual relationship is that, in partnership relationship the stress is given on trust and achieving general business goals while in contractual relationship the stress is given on the obligation of formal contract and on achieving specific business goals. In summary Partnerships are about relationships, not contracts [6]. In order to understand SOP practices in SOP from vendor's perspective, we have formulated the following research question (RQ).

RQ1. What are the solutions/practices, as identified in the literature for proper implementation of success factors in Software Outsourcing Partnership from vendor's perspective?

Here we have tried to find out the possible solutions through SLR in the related research articles for the identified critical success factors of software outsourcing partnership. The findings of SLR have been validated in the software industry through questionnaire survey. The rest of the paper is structured as. Related work is presented in section 2, section 3 presents the methodology. Section 4 describes the study results. Section 5 summaries and discusses the paper. Section 6 deals with limitations while Section 7 is the conclusions and future work.

2. BACKGROUND AND ASSOCIATED WORK

A number of approaches exist for collaboration, such as sub-contracting, partnership, alliance, reliance and joint ventures, etc. Kinnula [6] has presented a summary of the research areas of partnership in the context of software development outsourcing. These include (1) motivation towards partnership (2) performance evaluation of partnership (3) Scope of partnership (4) success of partnership and (5) decision making frameworks for partnership.

Ellram and Edis [22] explain how traditional outsourcing relationship is moved towards partnering relationship. Previously alliance has been highlighted with opportunism, doubt and distrust, contracts for single projects, strictly watched over communication between client and vendor, limited objectivity, restraint access of organisational resources, retribution for slip-ups, blame and distance and connection for specific project only. This type of outdated mind-set is not fruitful to an outsourcing partnership relationship. To bring the relationship on right way, a key change in approach is required. In partnership type relations shared aims and objectives, mutual trust, openness and honesty in dealings, effective and in time communication, objective critique, long-term commitment, innovative and supportive work place, organisational access to new technology, complementary skills and market, knowledge and resources sharing, teamwork, complete company engrossment at every levels of contacts and organisational proximity provide foundation for the partnership relationship formation.

Bowersox et al. [20] state that, partnership is formed in order to achieve shared benefits greater than that firms would achieve individually. It is a long term process in which partners with mutual goals makes joint decisions, work closely together, share information, ownership, benefits, risks, resources and achieve mutual beneficial results.

Different studies have been conducted to identify the various factors related to software outsourcing. Khan et el. [23] have conducted a systematic literature review and have identified 22 success factors regarding selection of software outsourcing vendors. They have briefly analysed the identified success factors across the different continents, organisations size and in different decades. They have also conducted an empirical study to address the solutions for their identified success factors. The same kind of study has been conducted by Niazi et al. [24] find out the different critical success factors (CSFs) through empirical study regarding software process improvement.

A number of other researchers have tried to address some of the issues of SOP, e.g (Bowersox et al. [20], Sehic et al. [25], Millson et al. [26], Kinnula [6], Alexandrova [27], Dominguez [28], Ellram and Edis [22], Mohr et al. [29], Hossain et al. [30], Bruce et al. [31], Mishra [32]. Summary of some of these research works are presented below:

A research study was carried out in USA on factors affecting partnership formation [33], the main results of the investigation were mutual trust and cultural differences. A comparable study was carried out by Kinnula [6] to investigate the formation of outsourcing partnership and has proposed outsourcing partnership life cycle model. Sehic et al. [25] proposed a strategic partnership model (SPM) and have identified various external factors (such as social, political, competitive and technology) and internal factors (such as organisational perspective, cost, resource, history and competitiveness).

Dominguez [28] argues the partnership as a manifestation of trust. The need for partnering relationships arises in case where countless and faster co-operation is demanded. One of the constituent's elements of partner type relationship is the provision of trustful atmosphere between the outsourcing client and vendor. Open communication, information sharing and mutual goals are all tools for getting partner position.

Bruce et al. [31] in his paper 'complexities of collaborative product development' present the following success factors, trust and flexibility, communication, equivalence in power, benefits and contribution, commitment and strong personal relationships. Millson et al [26] identified success factors like mutual understanding of strengths and weaknesses, goals related to strategy, intellectual property rights, information sharing and exit strategies.

Mohr et al. [29] have identified various critical factors such as coordination, communication quality, commitment, trust, information sharing, active participation, honesty and openness, and joint problem solving in partnership formation [29]. Similarly other identified factors include bidirectional information sharing, shared goals, trust, early communication with client, distinct value addition by vendor, top management support, mutual commitment and mutual understanding [34].

Hossain et al. [30] conducted a study about SCRUM in Global Software Engineering (GSE) and found that agile practices shown to be enormously favourable in GSE projects. They found that some of the outcomes that the authors conceived could enhance GSE projects in general.

Smite et al. [35] conducted SLR which provides detail GSE practices and techniques. It provides the seven most important discussed practices in literature that is particularly significant for practitioners.

Although multiple different collaboration approaches exist, such as sub-contracting, partnering, joint ventures, etc. our study focus solely on outsourcing partnership as the empiric case used in this study. It is observed that no SLR has been conducted so far for finding the critical success factors (CSFs) in SOP. Critical factors are those impacting either positively in partnership formation. Usually these are listed under "success factors" or "factors leading to successful partnership" in the relevant literature. It is observed that no SLR has been conducted so far for finding the solutions for identified CSFs. We consider all those factors as critical, which have been cited by >=30% in research articles. This paper is one component of our proposed software outsourcing partnership model (SOPM) [36, 62]. This paper concentrates on the solution of the identified CSFs of SOPM. We have conducted SLR and have identified 26 success factors of SOP, out of which 14 are CSF [19]. Furthermore, we have also conducted a separate SLR for extraction of the solutions from the literature.

3. RESEARCH METHODOLOGY

For the identification of practices for proper implementation of critical success factors (CSFs) we had used systematic literature review (SLR) process. For this process we follow the guiding principle of Kitchenham and Charters [37-39]. A similar approach has also been used by other researchers [19, 40], we also studied these approaches and get guidance from it. We presented the research methodology of our paper in Fig.1. Khan and Niazi [41] has also used a similar approach. In first phase of SLR, research questions are defined. In the second phase general literature review is conducted. The selection of relevant literature is carried out on the bases of title and abstract in the third phase. In the fourth phase data

Fig. 1 Systematic literature review process.

are extracted from the relevant papers and also we synthesis these data into different categories. Finally, we classify these categories and identify practices for SOP.

By using systematic literature review our study will confine the missing practices in SOP. Our research in the area of SOP is expected to give the right and useful approach in order to views outsourcing practitioners about the awareness of SOP. The novelty of our research shows that nobody has conducted SLR in the domain of SOP to find out practices for proper implementation of CSFs in the context of SOP formation. This paper helps the SDO vendor organisations to use the practices in order to implement the various factors. This paper wills also improving the readiness of SDO vendors toward SOP strategy.

The ultimate goal of our project software outsourcing partnership model is under process. However, we have published methodology and design of our proposed model [17, 19, 36]. This paper contributes to the conduction and findings of the SLR. The major steps in our methodology are:

3.1 Search Strategy and Search

The search strategy and search is available in the protocol [17]. A manual search was conducted for the determination of resources to be searched. In this phase we initially develop a trial search string which was used in different digital libraries. The available different digital libraries are IEEE Xplore (http://ieeexplore.ieee.org/), ScienceDirect (http://www.sciencedirect.com), ACM (http://dl.acm.org/), CiteSeer (http://citeseerx.ist.psu.edu) and Springer Link (http://link.springer.com/).

3.2 Publication Selection

3.2.1 Inclusion Criteria and Exclusion Criteria

The inclusion criterion is available in the protocol [17, 18].

3.3 Selecting Primary Sources

The selection process had two parts: a primary selection from the search results of papers that could plausibly satisfy the selection criteria, based on a reading of the title and abstract of the papers; followed by a final selection from the primarily selected list of papers that assure the selection criteria, based on a reading of the whole papers. The inter-rater reliability test was performed to reduce the researcher's biasness. However, no variances were found. Only 74 papers out of 2550 qualify the inclusion/ exclusion criteria. Finally the duplication was removed by excluded 09 papers from the final list of papers which repeated across different digital library, and we get a final total of 65 papers.

3.4 Publication Quality Assessment

Details are available in the protocol [17, 18]. The results of the study quality assessment were used in the selection of publications. After applying the quality assessment criterion, 65 papers remained included in the final list.

3.5 Data Extraction

The review was undertaken in a team by the researchers (authors), who were responsible for the data extraction. The inter-rater reliability test was performed after the data extraction process and no disagreements were found. The following data was extracted from each publication: date of review, title, authors, reference, database, critical factors, methodology (interview, case study, report, survey etc), target population, sample population, publication quality description, organisations type (software house, university, research institute etc), company size (small, medium, large), spi certification, country/location, year and practices for proper implementation of CSFs.

3.6 Classification of Practices

After identifying the practices for proper implementation of SFs in SOP SLR, we classified few practices as critical practices. The classification of critical success practices was based upon criteria, such as: those practices will be considered as critical whose frequency was >= 30.

4. RESULTS

This section demonstrates the outcomes of the SLR i.e the practices/solutions for implementing CSF for SOP. In the following tables, we have used the term 'CSF' for short to represent 'critical Success factors'. The subsequent sections represent 14 CSFs and their respective identified practices. We have identified 142 practices in total.

4.1 Mutual Interdependence and Shared Values

'Mutual interdependence and shared values' is the most important factor identified in our research study. By mutual interdependence and shared values, we mean common aims and objectives, sharing risk, benefits and shared ownership. According to Alexandrova [27] it is considered as important factor of the outsourcing partnership as it presumes "goal symmetry" between the outsourcer and vendor organizations. Table 1 presents our identified list of 7 practices to implement CSF 'mutual interdependence and shared values'.

4.2 Mutual Trust

'Mutual trust' is the second most cited success factor in our findings. Mishra [32] defined mutual trust as "…one party's willingness to be vulnerable to another party based on the belief that the latter party is 1) competent, 2) open, 3) concerned, and 4) reliable". Mutual trust and transparency leads to the establishment of long-term relationship between client and vendor organisations [42]. The degree of trust between the partners compensates any potential drawbacks of the formal contracting and the lack of strong defenses clauses in the outsourcing agreement [27]. Table 2 presents our identified list of 11 practices to implement CSF 'mutual trust'.

4.3 Effective and Timely Communication

By 'effective and timely communication' we mean exchanging status of the efficiency and effectiveness between partners. According to Webb [43] 'effective and efficient communication' between client and vendor organisations gives them an opportunity for the development of mutual understanding, respect and qualities, which can significantly increase the permanence of an outsourcing relationship. In view of Berger and Lewis [44] effective communication between outsourcing partners is assumed to be of crucial importance for the successful relationship. Table 3 presents our identified list of 11 practices to implement CSF 'effective and timely communication'.

4.4 Quality Production

'Quality production' by vendors can lead towards partnerships with their clients. By quality production, we mean delivery of high quality products, by using up to the mark capability and expertise, up-to-date technology and core competencies of vendor's in providing the required service quality. Due to the outstanding evolution in free marketplaces under the conditions of globalization and improvements in ICT, organisations have to consider outsourcing strategies, not for utilisation of the cost advantages but also to take benefits from the enhanced quality that counterpart vendors offer [14]. Table 4 presents our identified list of 13 practices to implement CSF 'Quality production'. Indian software companies have been reported to provide high quality software at low cost [45]. This is the reason that in the software export market, India is a dominant software outsourcing provider [46]. These trends show that 'quality production' is used as one of the criteria in the promotion/conversion of software development outsourcing vendors.

4.5 Organisational Proximity

By organisational proximity we mean strategic compatibility, business and technology understanding and language symmetry (refer to situations where both partners speak the same language). Organisational proximity is defined as "belonging to the same space of references" and manifested by shared representations, norms, standards and work practices [47]. Table 4 presents our identified list of 9 practices to implement CSF 'organisational proximity'.

4.6 Coordination, Cooperation and Collaboration

Coordination is the harmonization and combination of responsibilities, activities and command and control structures to efficiently use resources of an organization, in order to achieve predefined objectives. Cooperation is the actions of common efforts or association for mutual benefits. Collaboration is a cooperative working arrangement

Table 1. Practices for implementing mutual interdependence and shared values.

S. No.	Practices/Solutions for implementing 'mutual interdependence and shared values', identified through SLR	% of Practices via SLR (N=65)
1	Collaborate with client in decision making process and engage client in the development phases updates	22%
2	Respect for mutual obligations and recognition of dependence	8%
3	Establish collaboration in the form of Sharing risks, benefits and burden with the client	18%
4	Set up common goals, vision, expectation and ownership	25%
5	Establish frequent communication in different modes (site visits, synchronous and asynchronous communication using online tools)	5%
6	Develop complimentary assets and skills.	6%
7	Provide competitive quality of service, skills and resources	5%

Table 2. Practices for implementing mutual trust.

S. No.	Practices for implementing 'mutual Trust', identified through SLR	% of Practices via SLR (N=65)
1	Properly define role and responsibility	6%
2	Maintain reputation and good track records of the previous project	6%
3	Provide long term cooperation	9%
4	Openly share knowledge among the team members	14%
5	Follow the agreed time schedule strictly	11%
6	Deliver high quality products as per requirements	11%
7	Fulfill the client expectation as stated in the service level agreement	3%
8	Collaborate with the client in the form of joint investments, joint execution and jointly management	14%
9	Follow professional ethics in dealings	6%
10	Arrange site visits	5%
11	Maintain confidentiality and security of client's information including intellectual property rights	9%

Table 3. Practices for implementing effective and timely communication.

S. No.	Practices for implementing 'Effective and timely communication, identified through SLR	% of Practices via SLR (N=65)
1	Encourage both asynchronous and synchronous communication	22%
2	Establish communication guidelines and ICT infrastructure	8%
3	Create and offer shared cyber space	5%
4	Encourage frequent communication through latest technologies	8%
5	Communicate project status on daily basis	9%
6	Arrange ICT training sessions for the team members	5%
7	Establish open communication between stakeholders through face to face meetings and onsite visits	11%
8	Adjust communication barrier through the use of middleman with efficient communication skills and domain knowledge	9%
9	Design special interfaces between client and offshore vendor employees such as EDI(Electronic Data Interchange) link	6%
10	Encourage frequent formal and informal communication among team members	6%
11	Communicate well and according to a plan by using a team calendar for who needs to know what and when	9%

in which two or more parties (which may or may not have any previous relationship) work jointly towards a common goal. Literature reveals that the current inter-organisational trend is changing from competition to coordination, cooperation and collaboration [48]. Table 6 presents our identified list of 7 practices to implement CSF 'coordination, cooperation and collaboration'.

4.7 Flexible Service Level Agreements (SLA)

SLA is an officially written agreement jointly developed between client and the vendor organisation. It specifies a service provision or product development at such level to achieved business aims and objectives. It is an outsourcing agreement in the form of written contract between outsourcee (service provider) and outsourcer (a service receiver). SLA covers all of the terms and conditions of the corporate association, including services to be provided and service provider fees, time and schedule etc. In practice SLA is a constituent part to the outsourcing agreement. The SLA governs the quality and availability of the service, covering areas listed in the response to the prior question [49]. Table 7 presents our identified list of 4 practices to implement CSF 'flexible service level agreements (SLA)'.

4.8 Bidirectional Transfer of Knowledge (BTK)

BTK emerges when optimal (in terms of quantity and quality) information necessary for the realization of the service is provided through the channels of effective communication between the partners. The knowledge could have two forms: implicit i.e. informal, tacit, and explicit i.e. formal [50]. Special attention should be put on the way in which organizations "learn" from their partners, as this appears to be one of the means for the development of key competences [27]. Table 8 presents our identified list of 6 practices to implement CSF 'BTK'.

4.9 Long-Term Commitments

Commitment is the willingness of the parties to devote resources and exert effort in order to sustain an ongoing relationship [51]. It has a future orientation [29] with long-term perspective [52]. It reflects the partner vision that the relationship will be sustained over time [53], and has been characterized as "an enduring desire to remain in a valued relationship" [54]. Table 9 presents our identified list of 4 practices to implement CSF 'Long-term commitments'.

4.10 Joint Management Infrastructure

Joint management infrastructure is an inclusive term used to describe all of the structures, and substructures, used for the joint management. Once a trust over vendor is developed as a partner, then the relationship with the client typically moves toward joint management infrastructure, such as joint investment and jointly managing assets used in the relationship (e.g., vehicles, human resources and machinery, etc.). Table 10 presents our identified list of 5 practices to implement CSF 'joint management infrastructure'.

4.11 Cross Cultural Understanding and Sensitivity

Many cross cultural software development relationship failures have been endorsed to a cultural differences and lack of capability to boost 'cross cultural understanding and sensitivity' [55]. It is degree of understanding of behaviour patterns, values and norms between partners. Increased globalization in the political, economic, and social fields has developed greater interpersonal cross-cultural contact. Cross-cultural training has been suggested by many scholars as a means of facilitating more effective interaction, because much of cross-cultural contact has not been successful [56]. Table 11 presents our identified list of 6 practices to implement CSF 'cross cultural understanding and sensitivity'. Sixty nine percent of all the outsourcing projects failed either completely or partial due to intercultural incompatibilities among the client and vendor organisations and poor relationship management [57].

4.12 Success Stories of Previous Projects

Another key factor in the model is the degree of achievement of results set as contract goals in the previous outsourcing agreement. This degree should reflect the divergence between the actual benefits and the relative costs that the client organisation would have to spend without the realization of the particular partnership [58]. Every organisation has a working story/history. These success stories pro-

Table 4. Practices for implementing Quality production.

S. No.	Practices for implementing 'Quality production', identified through SLR	% of Practices via SLR (N=65)
1	Improve capability of vendors by implementing SPI certification such as CMM, CMMI, COPC-2000, ISO and LEAN etc.	14%
2	Improve quality of product through proper monitoring	6%
3	Improve competitiveness in service provision	12%
4	Provide ways for proper interaction between team members and sharing of tacit knowledge	18%
5	Acquire employees with good job-based knowledge skills (qualification , project management and IT skills)	14%
6	Provide good service design and execution	5%
7	Strictly follow development time schedule	12%
8	Establish mutual trust	11%
9	Offer trainings related to quality management	5%
10	Conduct requirements engineering phase thoroughly by using standard RE models	8%
11	Improve client-vendor coordination	5%
12	Ensure to high extent the project met the following client requirements.1) Response time. 2) Flexibility. 3) Usability. 4) Reliability.	9%
13	Ensure availability of global innovative talent, and reliance on world class delivery models	6%

Table 5. Practices for implementing Organisational proximity.

S. No.	Practices for implementing 'Organisational proximity' , identified through SLR	% of Practices via SLR (N=65)
1	Achieve mutual understanding by active cross-cultural communication through short visits of employees and face-to-face meetings when and where possible	15%
2	Offer different skills trainings such as formal communication language, client-specific and domain-specific, analytical, logical reasoning	22%
3	Understand and respect the differences in norms and values	8%
4	Temporarily relocate selected members to client's site	6%
5	Adjust communication barrier through the use of middlemen.	5%
6	Thoroughly understand provider's business (e.g., core competencies, values, philosophy, work culture)	15%
7	Use time management (schedule) to mitigate time zone difference	5%
8	Use a common set of development techniques/policies and tools between the organisations to minimize differences in understanding.	11%
9	Establish a formal communication protocol so that the communication lapses and barriers are avoided	6%

Table 6. Practices for implementing coordination, cooperation and collaboration.

S. No.	Practices for implementing 'Coordination, Cooperation and Collaboration' , identified through SLR	% of Practices via SLR (N=65)
1	Utilize the time zone differences by managing the working hours between the two sites in such a way that can lead towards 24 hours development	9%
2	Promote visits and exchanges among sites	9%
3	Provide active cooperation at all stages and respect for mutual obligations	12%
4	Escalate formal and informal cooperation and coordination through formal and informal meetings	12%
5	Regularly share data about each phase of development	9%
6	Use special collaborative technology like: Software configuration management applications are used to manage different versions of the components of a software system.	14%
7	Increase dependence between the partners	5%

vide in depth information about organization. Table 12 presents our identified list of 5 practices to implement CSF 'success stories of previous projects'.

4.13 Access to New Markets, Technologies and Complementary Skills

With help of partnership both parties gets access to new markets, technologies and complementary skills, for undertaking complimentary activities to achieve mutual benefits [59]. Table 13 presents our identified list of 3 practices to implement CSF 'access to new markets, technologies and complementary skills'.

4.14 Governance and Control

Governance in outsourcing partnership arena generally refers to the processes, mechanisms and relations by which partners are directed and controlled [60]. It includes the sharing of ownership and rights, benefits and responsibilities among different parties involved in the partnership. It also includes processes through which outsourcing partnership objectives are set and pursued, in the context of the social, regulatory and market environment [61]. Governance in outsourcing partnership is concerned with the resolution of shared action and problems among distributed allies and the settlement of conflicts of concern between numerous outsourcing partners [62]. Table 14 presents our identified list of 4 practices to implement CSF 'governance and control'.

5. SUMMARIES AND DISCUSSION

We have identified 142 practices in total, through SLR, for implementing CSFs which can lead vendor towards partnership with client organisation. The SDO vendor organisations can also get help from these practices in order to know that how can they solve the problems of their clients. Afetr applying the criterion that a practice should be considered only if it having %age >=5 in the SLR. We have found 7 practices for implementing CSF 'mutual interdependence and shared values'. We have identified through SLR, those most suitable practices/solutions for implementing mutual interdependence and shared values are the following two practices (% ≥ 22) as shown in Table 1:

i. Set up common goals, vision, expectation and ownership. Mutual interdependence increases as both sides would understand each other's needs and expectation, goals and vision, ownership and responsibility especially for long term partnership.

ii. Collaborate with client in decision making process and engage client in the development phases updates.

For implementing the 'mutual trust' CSF our SLR study finds 11 practices. Table 2 noted that mutual trust can be best established by follow the following two practices (% ≥ 14):

i. Openly share knowledge among the team members.

ii. Collaborate with the client in the form of joint investments, joint execution and jointly management.

For implementing the CSF 'effective and timely communication' our SLR study also finds out 11 practices. Table 2 illustrious that lack of effective communication can be achieved by following practice (% ≥ 22):

i. Encourage both asynchronous and synchronous communication.

Adopt both synchronous (voice) and asynchronous (text) tools like: instant messaging, telephone, wiki, fax, e-mail, voicemail, internet, mailing lists, shared databases, IRC, Skype, messenger, net meeting, chat, phone, virtual white boards, change management system, team intranet websites, photo gallery, group calendars, power point presentations, nor-real-time database, blog, camel, team space and next move

We have found 13 practices for achieving 'quality production' through SLR. From the Table 4 we have noted that most suitable practices/solutions (% ≥ 14) for 'quality production' are:

i. Provide ways for proper interaction between team members and sharing of tacit knowledge.

ii. Acquire employees with good job-based knowledge skills (qualification , project management and IT skills)

For implementing CSF 'organisational proximity'

Table 7. Practices for implementing flexible service level agreements (SLA).

S. No.	Practices for implementing 'flexible service level agreements (SLA)' , identified through SLR	% of Practices via SLR (N=65)
1	Effectively manage changes to ensure SLA flexibility	8%
2	Appoint Service level manager who is responsible that the content of the SLA is continuously aligned with the business requirements.	5%
3	Establish mechanism for proper negotiations and mutual consensus on SLA specifications (e.g scope, price, schedule, resource requirements, security provisions, intellectual property rights, and penalties and escalation processes)	28%
4	Establish informal contracting mechanisms, such as trust and relationship specificity, as it can serve as safeguards.	5%

Table 8. Practices for implementing bidirectional transfer of knowledge (BTK).

S. No.	Practices for implementing 'bidirectional transfer of knowledge (BTK)'	% of Practices via SLR (N=65)
1	Explicitly share of information, expectations, and work related concerns during site-visits, conference calls, e-mail exchanges or teleconferencing by using specifications, blueprints and prototypes.	8%
2	Use latest technology and processes for knowledge sharing and management i.e electronic scheduling, groupware, and shared knowledge databases and repositories, intranets, collaborative technologies and social media, Shared cyber space and TMS.	29%
3	Use variance analysis to plan which knowledge is required and where it is required	6%
4	Establish mechanism for knowledge creation and dissemination. e.g how new knowledge can be created and how knowledge of experienced staff can be utilized for new team members	5%
5	Convert tacit knowledge to explicit knowledge by documentation and process description.	5%
6	Conduct domain specific and technical trainings and update skill database and monitor the skill profile	22%

Table 9. Practices for implementing Long-term commitments.

S. No.	Practices for implementing 'Long-term commitments' , identified through SLR	% of Practices via SLR (N=65)
1	Focus on developing trustful relationship with client	9%
2	Offer additional services that will contribute to the development of a mutually beneficial partnership	5%
3	Exert effort and devote resources in order to sustain an on-going relationship	5%
4	Always discuss and share long range plan with client and create "future orientation"	5%

Table 10. Practices for implementing Joint management infrastructure.

S. No.	Practices for implementing 'joint management infrastructure' , identified through SLR	% of Practices via SLR (N=65)
1	Take joint mutually beneficial decisions with client in most problematic circumstances	8%
2	Use of inter-organizational systems such as EDI	5%
3	Make joint investments, such as jointly managing assets used in the relationship (e.g., machinery, vehicles, etc.).	22%
4	Establish a joint configuration management infrastructure.	5%
5	Update the existing steering boards and add some members from the client's company	6%

Table 11. Practices for implementing cross cultural understanding and sensitivity.

S. No.	Practices for implementing 'cross cultural understanding and sensitivity' , identified through SLR	% of Practices via SLR (N=65)
1	Face-to-face meetings are recommended when and where possible, ideally at the start of the project and/or when a new member joins	6%
2	Offer language skills training	12%
3	Understand differences in norms and values	22%
4	Temporarily relocate selected members to client's site	8%
5	Reduce inter-organizational differences	5%
6	Build mixed teams with memberships from different cultural backgrounds by integrating your team with client team	8%

Table 12. Practices for implementing success stories of previous projects.

S. No.	Practices for implementing 'success stories of previous projects' , identified through SLR	% of Practices via SLR (N=65)
1	Maintain reputation and good track record of the previous projects	9%
2	Hire experienced staffs with relevant skills	5%
3	Learn from your past experiences	5%
4	Undertake a pilot project	5%
5	Acquire required licenses and certifications	5%

Table 13. Practices for implementing access to new markets, technologies and complementary skills.

S. No.	Practices for implementing 'access to new markets, technologies and complementary skills' , identified through SLR	`% of Practices via SLR (N=65)
1	Use state-of-the art IT infrastructure (servers, broadband, routers, modems, voice and data circuits etc.)	3%
2	Acquire CMMI certification that helps you to compete better	5%
3	Develop complementary resources and capabilities	17%

Table 14 : Practices for implementing governance and control.

S. No.	Practices for implementing 'Governance and control' , identified through SLR	% of Practices via SLR (N=65)
1	Clearly define roles and responsibilities of the stakeholder as per their competencies	5%
2	Use both formal and informal governance mechanisms	5%
3	Adopt proper mechanism for performance monitoring and incentive, corrective action, and penalty rewards systems	8%
4	Collect performance data from multiple sources and stakeholders and measure employee performance holistically through data, reports, graphs and charts	6%

our SLR study finds out 9 practices. Table 5 memorable that 'organisational proximity' can be best achieved by ensuring the following three practices (% ≥ 15):

i. Offer different skills trainings such as formal communication language, domain-specific and client-specific, logical and analytical reasoning.

ii. Achieve mutual understanding by active cross-cultural communication through short visits of employees and face-to-face meetings when and where possible.

iii. Thoroughly understand provider's business (e.g. work culture, values, core competencies and philosophy).

Table 6 represents 7 practices for CSF 'coordination, cooperation and collaboration'. From the Table 6 it is clear that most suitable practice/solution for addressing 'coordination, cooperation and collaboration' is the following practice (% ≥ 14):

i. Use special collaborative technology like: software configuration management (SCM) applications are used to manage different versions of the components of a software system.

We have found 4 practices for implementing 'flexible service level agreements (SLA)' through SLR. From the Table 7 we have noted that most suitable practice/solution (% ≥ 28) for implementing flexible SLA is:

i. Establish mechanism for proper negotiations and mutual consensus on SLA specifications (e.g scope, price, schedule, resource requirements, security provisions, intellectual property rights, and penalties and escalation processes).

For implementing the 'bidirectional transfer of knowledge (BTK)'CSF our SLR study finds out 6 practices. Table 8 eminent that efficient and effective bidirectional transfer of knowledge can be best achieved by follow the following two practices (% ≥ 22):

i. Use latest technology and processes for knowledge sharing and management i.e electronic scheduling, groupware, and shared knowledge databases and repositories, intranets, collaborative technologies and social media, shared cyber space and TMS.

ii. Conduct domain specific and technical trainings and update skill database and monitor the skill profile. We have found this practice in 14 papers symbolize 22 %.

Table 9 represents 4 practices for CSF 'long-term commitments'. From the Table 9 we noted that most suitable practices/solutions for creating 'Long-term commitments' is the practice (% ≥ 09):

i. Focus on developing trustful relationship with client

We have found 5 practices for CSF 'joint management infrastructure' through SLR study. From the Table 10 we have noted that most suitable practice/solution (% ≥ 22) for implementing joint management infrastructure is:

i. Make joint investments, such as jointly managing assets used in the relationship (e.g., machinery, vehicles, etc.).

For CSF 'cross cultural understanding and sensitivity' our SLR study finds out 6 practices. Table 11 shows that 'cross cultural understanding and sensitivity' can be best developed by keep an eye on the following practice (% ≥ 22):

i. Understand differences in norms and values.

Table 12 represents 5 practices for 'success stories of previous projects' as CSF. From the Table 12 it is clear that through SLR the most suitable practices/ solutions found for 'success stories of previous projects' are the following practice (% ≥ 9):

i. Maintain reputation and good track record of the previous projects

Glancing on Table 13 it can be easily find out that out of 3 practices most suitable practice/ solution for achieving 'Access to new technologies, markets and complementary skills' is :

ii. Develop complementary resources and capabilities. Having reporting in SLR study (% ≥ 17).

Table 14 is the last table presenting 4 practices, it illustrates that for achieving good 'governance and control' we have to:

i. Adopt proper mechanism for performance monitoring and incentive, corrective action, and penalty rewards systems. The above practice having (% ≥ 8) in SLR.

6. STUDY LIMITATIONS

In this section, the threats of validity concerning the SLR study have been discussed. By using our systematic literature review, we extracted data about the practices/solutions for implementing CSF in SOP, but how valid are our findings? One possible threat to internal validity is that for any practice/solution in an article, the author may not in fact have described the underlying cause of practices for CSFs implementation in SOP and also for any specific response; the respondent does not provide the reasons to report practice/solution. We are not able to independently control this treat. With the increasing number of papers in SOP, our SLR process may have missed out some relevant papers. However, like other researchers of SLR this is not a systematic omission [30]. In the present paper we have identified the practices through SLR, which are validated through empirical investigation in outsourcing industry [63].

7. CONCLUSION AND FUTURE WORK

Initially we have identified 142 practices, in total, through SLR for implementing CSF by vendors in SOP relationships. After applying the criterion it was reduced to 88. Our results reveal that focusing on these practices/solutions can help vendor organisations in order to promote their ordinary contractual relationship to outsourcing partnership.

Beside all stated limitations, we are confident in that our study will contribute in academia and industrial domains. This study will:

- Provide SDO vendors, a guiding knowledge that can assist them to implement and design successful outsourcing partnership initiatives. Our results recommend that SDO vendors should adopt all of the reported practices for CSFs especially those reported with greater percentage, in order to gain partner position.
- Increase partnership cohesiveness, as it will guide both sides to understand each other's requirements and goals, in order to sustain long term commitment.
- Provide guidance to SDO client, in making factual choices in terms of continuing, renewing, or terminating their agreements with their current vendor.
- Provide assistance in well understanding of CSFs practices for SOP, to ensure successful partnership.

We have noted the following points, as a future plan, from the findings of this study:

- Will validate the practices identified through SLR by conducting empirical investigation in the outsourcing industry.
- The practices/solutions in SOP relationships from client's perspectives will be identified and analysed.
- To analyses the critical risk in the conversion to, or formation process of SOP from vendor perspective.
- To determine the underlying reasons of why some factors are not important for specific group of SDO organisations.

Our future work will focus on the development of a Software Outsourcing Partnership Model (SOPM). This paper provides input for the development of the second phase of the SOPM, such as the identification of various practices CSFs through SLR. The SOPM will assist SDO vendors in promoting their existing contractual SDO relationship into SOP with client organization. The SOPM will provide guidance and boost the work that has been undertaken on frameworks and models development for outsourcing partnership.

8. ACKNOWLEDGMENTS

We are thankful to software engineering research group (SERG_UOM) at University of Malakand for providing assistance. We are also grateful to anonymous reviewers of the MDRC-2015 conference, for the review and their valuable comments.

9. REFERENCES

1. Bamford, J., D. Ernst, & D.G. Fubini. Launching a world-class joint venture. *Harvard Business Review* 82(2): 90-100(2004).
2. Kelly, M.J.,J.L. Schaan & H. Joncas. Managing alliance relationships: Key challenges in the early

stages of collaboration. *R&D management* 32: 11-22 (2002).
3. Ylitalo, J., E. Mäki, & K. Ziegler. *Building Mutuality and Trust in Strategic Partnership. Meaning of Early Stages in Relationship Formation: A Case Study*. Helsinki University of Technology, BIT Research Centre, Espoo, Finland (2004).
4. Lane, M. S. & W.H. Lum. Examining client perceptions of partnership quality and the relationships between its dimensions in an IT outsourcing relationship. *Australasian Journal of Information Systems* 17: 5-34 (2011).
5. Kinnula, M. & V. SeppAnen. *Information Technology Requirements in an Outsourcing Partnership. Frontiers of e-Business Research Conference*, Tampere, Finland, p. 493-502 (2003).
6. Kinnula, M, V. Seppanen, J. Warsta & S. Vilminko. *The Formation and Management of a Software Outsourcing Partnership Process*. In: *40th Hawaii International Conference on System Sciences.* Washington, USA, p. 1530-1605 (2007).
7. Tuten, T.L. & D.J. Urban, An Expanded Model of Business-to-'Business Partnership Formation and Success. *Industrial Marketing Management* 30(2):149-164(2001).
8. Moe, N. B, Mite, S. Darja, & G. K, Hanssen. From offshore outsourcing to insourcing and partnerships: four failed outsourcing attempts. *Empirical Software Engineering* 19: 1225-1258 (2014).
9. Roy,V & B. A. Aubert. A resource-based analysis of IT sourcing. *ACM Special Interest Group on Management Information Systems* 33(2): 29-40 (2002).
10. Kishore, R., H.R. Rao., K. Nam., S. Rajagopalan & A. Chaudhury. A relationship perspective on IT outsourcing. *Communications of the ACM - Mobile Computing Opportunities and Challenges* 46(12):86-92 (2003).
11. Srinivasan, R., T.H. Brush. Supplier Performance in Vertical Alliances: The Effects of Self-Enforcing Agreements and Enforceable Contracts. *Organization Science* 17(4): 436-452 (2006).
12. Dictionary, O. E. *Oxford English dictionary online*: Oxford University Press, Oxford, UK http://www.oed.com/ (2008).
13. Bocij, P., D. Chaffey., A. Greasley & S. Hickie. *Business information systems: technology, development and management for the E-Business* 4th ed. FT Press (2009).
14. Brinkerhoff, J.M. Government non-profit partnership: A defining framework. *Public Administration and Development* 22:19-30 (2002).
15. Lambert, D.M., M.A. Emmelhainz & J.T. Gardner. Building Successful Logistics Partnerships. *Journal of Business Logistics* 20:165-181(1999).
16. Lee, J.N. & Y.G. Kim. Understanding outsourcing partnership: a comparison of three theoretical perspectives. *IEEE Transactions on Engineering Management* 52(1): 43-58 (2005).
17. Ali, S. & S.U. Khan. Systematic Literature Review Protocol for Software Outsourcing Partnership (SOP). *IOSR Journal of Computer Engineering* 2(1): 08-18 (2012).
18. Ali, S. & S.U. Khan. Software outsourcing partnership (SOP): A systematic literature review protocol with preliminary results. *International Journal of Hybrid Information Technology* 7: 377-392 (2014).
19. Ali, S. & S.U. Khan. Critical Success Factors for Software Outsourcing Partnership (SOP): A systematic literature review, In: *9th International Conference on Global Software Engineering,* ICGSE-14, Shanghai China, p. 153-162 (2014).
20. Bowersox, D, D. Closs & T. Stank. How to master cross-enterprise collaboration. *Supply Chain Management Review* 7: 18–27 (2003).
21. Verwaal, E. and M. Hesselmans, Drivers of supply network governance: An explorative study of the Dutch chemical industry. *European Management Journal* 22(4):442-451(2004).
22. Ellram, L.M. & O.R. Edis, A case study of successful partnering implementation. *Journal of Supply Chain Management* 32(4): 20-28 (1996).
23. Khan, S. U, M. Niazi & R. Ahmad. Empirical investigation of success factors for offshore software development outsourcing vendors. *Institution of Engineering and Technology Software* 6(1): 1-15 (2010).
24. Niazi, M., D. Wilson, & D. Zowghi. Critical success factors and critical barriers for software process improvement: An analysis of literature, In: *Proceedings of Australasian Conference on Information Systems* (ACIS03), Perth, Australia (2003).
25. Sehic, I. & H. Gilani. *IT Outsourcing: A Strategic Partnership between Buyer and Seller Organization*. PhD dissertation, International Business School, Jönköping University, Jönköping, Sweden. 105 pp. (2010).
26. Millson, M.R., S. Raj, & D. Wilemon, Strategic partnering for developing new products. *The Journal of Product Innovation Management* 14(1): 59-60 (1997).
27. Alexandrova, M. IT outsourcing partnerships: Empirical research on key success factors in Bulgarian organizations. *Journal of Contemporary Management Issues* 17: 31-50 (2012).
28. Dominguez, L.R. *The Manager's Step-by-step Guide to Outsourcing*. McGraw Hill Professional. 226 (2006).
29. Mohr, J. & R. Spekman. Characteristics of partnership success: partnership attributes,

communication behavior, and conflict resolution techniques. *Strategic Management Journal* 15: 135-152 (1994).

30. Hossain, E., M. Ali-Babar, & P. Hye-young. Using scrum in global software development: A systematic literature review. In: *IEEE International Conference on Global Software Engineering, ICGSE09.* Lero, Limerick, Ireland (2009).
31. Bruce, M., F. Leverick, & D. Littler, Complexities of collaborative product development. *Technovation* 15(9): 535-552 (1995).
32. Mishra, A.K., Organizational responses to crisis: The centrality of trust. In: *Trust in Organizations*. Newbury Park, CA, p. 261-287 (1996).
33. Oza, N.V. *An Empirical Evaluation of Client - vendor Relationships in Indian Software Outsourcing Companies. PhD* dissertation, School of Computer Science: University of Hertfordshire, UK, 231 pp. (2006).
34. Virolainen, V.M. *Motives, Circumstances and Success Factors in Partnership Sourcing:* PhD dissertation, Lappeenranta University of Technology, Lappeenranta, Finland, 232 (1998).
35. Smite, D., C. Wohlin., T. Gorschek. & R. Feldt. Empirical evidence in global software engineering: A systematic review. *Empirical Software Engineering* 15(1): 91–118 (2010).
36. Ali, S. & S.U. Khan. Software Outsourcing partnership model. *Science International, Lahore* 26(4): 1437-144 (2014).
37. Kitchenham, B. & P. Brereton. A systematic review process research in software engineering. *Information and Software Technology*, 55(12): 2049-2075 (2013).
38. Steinmacher, L., M. A. G. Silva., M .A. Gerosa., & D.F. Redmiles. A systematic literature review on the barriers faced by newcomers to open source software projects. *Information and Software Technology* 59: 67-85 (2010).
39. Leitner, M. & S. Rinderle-Ma. A systematic review on security in process-aware information system and constitution, challenges and future directions, *Information and Software Technology* 56(3): 273-293 (2014).
40. Bozzelli, P., Q. Gu, & P. Lago. A systematic literature review on green software metrics. *International Journal, University of Tampere*, Amsterdam 1-25 (2013).
41. Khan, S.U., M. Niazi, & A. Rashid. Factors influencing clients in the selection of offshore software outsourcing vendors: an exploratory study using a systematic literature review. *Journal of Systems and Software* 84(4): 686-699 (2011).
42. Niazi, M, N. Ikram., M. Bano., S. Imtiaz. & S.U. Khan. Establishing trust in offshore software outsourcing relationships: an exploratory study using a systematic literature review. *IET Software* 7(5): 283-293 (2013).
43. Webb, L. & J. Laborde, Crafting a successful outsourcing vendor/client relationship. *Business Process Management Journal* 11(5): 437-443 (2005).
44. Berger, H. & C. Lewis. Stakeholder analysis is key to client–supplier relationships of global outsourcing project success. *International Journal of Information Management* 31: 480-485 (2011).
45. Bhatnagar, S.C. & S. Madon, The Indian software industry: moving towards maturity. *Journal of Information Technology* 12(4): 277-288 (1997).
46. Niazi, M. Do systematic literature reviews outperform informal literature reviews in the software engineering domain? An initial case study. *Arabian Journal for Science and Engineering* 40(3): 845-855 (2015).
47. Torre, A. & A. Rallet. Proximity and Localization. *Regional Studies* 39(1): 47-59 (2005).
48. Kumar, K., & H. G. V. Dissel Sustainable collaboration: Managing conflict and cooperation in inter-organizational systems. *MIS Quarterly* 20(3): 279-300 (1996).
49. Mingay, S. & M. Govekar. *ITIL's service-level management strength is in integration*. Research Note, Gartner Group, Tactical Guidelines, TG-15-349 (2002).
50. Nonaka, I. & H. Takeuchi, *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation*: Oxford University Press, UK (1995).
51. Fontenot, R.J. & E.J. Wilson, Relational exchange: a review of selected models for a prediction matrix of relationship activities. *Journal of Business Research* 39(1): 5-12 (1997).
52. Anderson, J.C. & J.A. Narus. A model of distributor firm and manufacturer firm working partnerships. *Journal of Marketing* 54(1): 42-58 (1990).
53. Henderson, J.C. Plugging into strategic partnerships: the critical IS connection. *Sloan Management Review* 31(3): 7-18 (1990).
54. Simpson, J.T. & D.T. Mayo. Relationship management: a call for fewer influence attempts? *Journal of Business Research* 39(3): 209-218 (1997).
55. Johnson, J.P., T. Lenartowicz, & S. Apud, Cross-cultural competence in international business: Toward a definition and a model. *Journal of International Business Studies* 37(4): 525-543 (2006).
56. Black, J.S. & M. Mendenhall, Cross-cultural training effectiveness: A review and a theoretical framework for future research. *Academy of Management Review* 15(1): 113-136 (1990).
57. Shah, H., N. J. Nersessian., M. J. Harrold., & W.

Newstetter. Studying the influence of culture in global software engineering: thinking in terms of cultural models. In: *Proceedings of the 4th International Conference on Intercultural Collaboration*. Bengaluru, India, p. 77-86 (2012).
58. Anderson, J.C., & J. A. Narus. A model of distributor firm and manufacturer firm working partnerships. *Journal of Marketing* 54(1): 42-58 (1990).
59. Lee, S. & G.G. Lim, The impact of partnership attributes on EDI implementation success. *Information & Management* 42(4): 503-516 (2005).
60. Kang,H., M. Cheng, & S.J. Gray, Corporate governance and board composition: Diversity and independence of Australian boards. *Corporate Governance*: *An International Review* 15(2): 194-207 (2007).
61. Jesover, F. & G. Kirkpatrick, The revised OECD principles of corporate governance and their relevance to non-OECD countries. *Corporate Governance: An International Review* 13(2): 127-136 (2005).
62. Becht, M., P. Bolton, & A. Rajell, Corporate governance and control, In: *Handbook of the Economics of Finance*, p. 1-109 (2003).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 163–176 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Design, Modeling and Simulation of Six Degree of Freedom Machining Bed

Muhammad F. Shah*, Zareena Kausar, and Faizan K. Durrani

Department of Mechatronics Engineering, Air University,
Islamabad, Pakistan

Abstract: Machining is a process in which a raw material is cut into desired shape and size after various material removal processes. As application becomes more complex and size shrinks, machines need to be designed to meet the desired precision and accuracy. To date researchers and practitioners generate 6 degree of freedom for the tool which affects the accuracy and precision of machining. In this paper it is proposed to produce the desired degrees of freedom from the bed replacing a machine bed with a parallel manipulator having six legs of variable lengths. This will give work piece a 6 DOF motion. The paper presents design for the proposed machine bed, kinematic model for the machine bed and dynamics model of the linear actuator. The linear actuator is used to change the lengths of legs. Each leg will consist of a linear actuator. The physical model of the proposed bed is generated in Solid Works, software used to model the engineering systems. Six degree of freedom motion for the proposed machining bed is simulated in the Solid Works. Results which verify that the proposed design is capable to produce motion in six directions in space are presented. Results in tabular form are presented which verify kinematic modelling of the platform while the dynamics model of linear actuator is verified through MATLAB® Simulations. All simulation results support the hypothesis that 6 DOF may be produced by a machining bed.

Keywords: Machining, Stewart platform, DOF, linear actuator; machine bed

1. INTRODUCTION

Machining is a term used to describe the removal of material from a work piece to get the desired shape. Machining process is usually categorised as drilling, milling and turning [1]. Machining is necessary where constricted tolerances on dimensions and finishing are essential. Present-day spindle technologies used for material removal give very high rotational speed and power ratings so this enables the material to remove very quickly. Unfortunately, the performance envelope for conventional machine structures limits the application of higher performance spindles. The high speed tools are also vulnerable to external and internal vibrations. In the presence of vibrations the cuts made and the parts created will be dimensionally inaccurate and imprecise. A solution of these problems is to restrict the tool motion in a single dimension and instead give 6-DOF to the machine bed. For this purpose a parallel robotic mechanism is proposed [2] to be used as a bed for micro machining. This research also proposes for machining to replace the existing beds with a 6-DOF manipulator.

There are different types of manipulators serial as well as parallel. For tackling this problem we proposed a parallel manipulator as it offers high stiffness, high dynamics and less position accuracy errors [3]. Due to their high accuracy, high control and high load capacity parallel manipulators have received a lot of attention from industries and robotics community as well [4]. However, parallel manipulators have some drawbacks of difficult forward kinematics and relatively small workspace. Generally, forward kinematics of a parallel manipulator is difficult to solve. Liu et al. [5] proposed a numerical algorithm

Received, December 2015; Accepted, May 2016
*Corresponding Author: Muhammad Faizan Shah; Email: shahfaizan250@gmail.com

which provided the kinematic solutions based on a set of three nonlinear simultaneous equations. On the other hand, dynamics of a parallel manipulator, which is very important to develop a controller, tends to be complicated. Most studies on dynamics have been performed based on the Lagrangian formulation [6], the Newton–Euler formulation [7] and the principle of virtual work [8]. These dynamic equations were accurate by including actuating legs having closed-loop connections to each other in its formulation, but computation of the dynamics was very time-consuming. Also in real-time calculation was not easy to achieve in the control system. Fichter [9] reported simulation-based control analysis but the ways of dealing with computational burden in real-time implementation were not mentioned. In case where the platform was used as a tool holder, an accurate modelling of dynamics, including legs, is required for control performance [10].

Proposed parallel manipulator is one of the types of Stewart platform in which the tool is fixed and all the motions are generated by bed [11]. With the flexibility of a Parallel Mechanism, tool may not need to be set again and again, whereas, the conventional robots need adjustments to these stations often require shutting down the entire line to reset tool for a different part to be done. Many manufacturers have developed commercial machining centres based by using Stewart Platform architecture but the literature [12, 13] lacks research in this area. The proposed 6 DOF include three linear movements (lateral, longitudinal and vertical), and the three rotations pitch, roll, and yaw.

The mechanism of the bed of the platform is presented in section 2. Solid works model to explain the motion of the bed is presented and its Kinematics is explained in the form of inverse kinematics in the following section while Dynamics model of linear actuator is presented in section 4. The simulation results which verify the proposed design and model are presented in section 5.

2. MECHANISM DESIGN AND ANALYSIS

The mechanism to produce six degrees of freedom for the work piece is proposed to be from the family of Stewart platforms which give inventive solutions to complex motion applications requiring high load capacity and accuracy in up to six independent axes [14].There are several configurations of the 6 DOF parallel manipulators depending upon the distribution of six points on the base and top plates of the platform. The configuration proposed for the machining bed is 3-6 configuration. In this configuration the base plate is divided into three sections. At each section two legs are joined to the plate at same or closest points while the top plate has six sections and each section has joint of one leg.

A schematic of the proposed mechanism for machining bed is shown in Fig. 1. This consists of a movable top plate which is connected with a fixed base plate with six independent linear actuators [15]. These linear actuators act as a kinematic chain between upper rigid plate and lower fixed plate of the platform. Linear actuators at the base plate are attached through universal joints each giving motion in 2 DOF whereas linear actuators are connected with the help of ball and socket joint with the top moving plate. Each ball and socket joint give motion in 3 axes. Anything placed on the top plate will be able to move in six DOF with the help of these linear actuators as the leg lengths of these actuators is variable.

The reachable workspace is the collection of all points $\{x\ y\ z\}^T$ that can be reached by the manipulator in at least one orientation. A subset of reachable workspace is dexterous workspace that is the volume of space that can be reached by the parallel manipulator in all orientations consequently. Dexterous workspace is null for the proposed parallel manipulator machining bed, as this cannot reach all orientations at any position in the reachable workspace. The workspace of the machining bed is defined based on assumptions that there is no actuator constraint, leg interference and singularities.

3. KINEMATICMODELINGOF THE MACHINING BED

Kinematic analysis of the machining bed is considered without regard of the forces. In kinematic analysis velocity and position are analysed for the manipulator [16]. There are two categories of kinematic analysis: Forward and inverse. In this research we propose inverse kinematics for the manipulator as the final position of the end effector is known to us and we owe to find the orientation of the legs to reach the end point. The end point is assumed to be the final position of the centre of the top moving plate of the bed. The solution obtained is unique, and can be simply determined. Symbols used in modelling of the system are described in Table 1. Forward Kinematics is not recommended for this study as forward kinematics gives us the final position of the end effector, whereas, we need to find out the displacement of actuators.

Table 1. Description of symbols / nomenclature used.

Symbol	Description
V_b	Back Electromotive force (emf)
$\omega_m(t)$	Angular velocity of motor
k_b	Back emf constant
$i_a(t)$	Armature current
$e_a(t)$	Applied armature voltage
$T_m(s)$	Torque developed by the motor
K_t	Motor Torque Constant
D_m	Equivalent viscous damping of motor
J_m	Equivalent inertia of motor
l	Lead of screw of linear actuator
$\|I_i\|$	Length of each linear actuator
T	Translation Vector
P_I	Vector defining the coordinates of the anchor point with respect to the Platform
B_i	Vector defining the coordinates of the lower anchor point of the base of the platform

The machining bed's two plates are connected by six variable length legs. The reference framework is considered to be at the base plate [17]. In Fig. 2 schematic view of *i*th leg of parallel manipulator machining bed is shown, where T is the translation vector and Pi is the vector which defines upper anchor point with respect to the platform frame work. Translation vector T and P_i are added with each other by head to tail rule where vector G_i is the resultant vector. Vector l_i represents length of a linear actuator and B_i is the vector that defines the coordinates of the lower anchor point. These two vectors are also added by head to tail rule and the resultant vector is G_i.

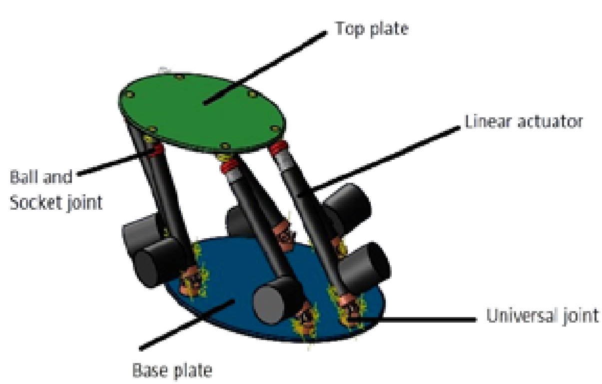


Fig. 1. Schematic diagram of the parallel manipulator bed.

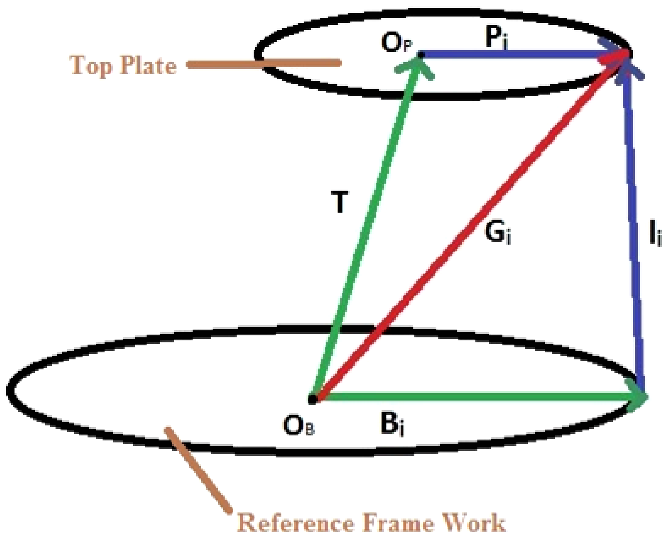


Fig. 2. Schematic view of the $\boldsymbol{i_{th}}$ leg of the parallel manipulator machining bed.

From Fig. 2, using the geometric relations and vector addition the equations (1-5) are established.

$$\vec{G_i} = \vec{T} + \vec{P_i} \tag{1}$$

$$\vec{G_i} = \vec{T} + {}^{P}_{B}R\vec{P_i} \tag{2}$$

$$\vec{G_i} = \vec{I_i} + \vec{B_i} \tag{3}$$

$$\vec{I_i} + \vec{B_i} = \vec{T} + {}^{P}_{B}R\vec{P_i} \tag{4}$$

$$\vec{I_i} = \vec{T} + {}^{P}_{B}R\vec{P_i} - \overrightarrow{B_i} \tag{5}$$

where,

$$Length\ of\ actuator = |I_i| \tag{6}$$

And i vary from 1 to 6.

Moreover, T is the translation vector, giving the positional linear displacement of the origin of the moving plate frame with respect to the base reference framework, and $\mathbf{P}_i$ is the vector defining the coordinates of the anchor point with respect to the moving plate framework and $\mathbf{B_i}$ is the vector defining the coordinates of the lower anchor point. These 6 equations give the lengths of the six legs to achieve the desired position of the platform.

In Fig. 3 and 4, distribution of the joint positions is given. The top moving plate is divided into six equally distant positions. Therefore, angle between each joint position is 60 degrees. On the bottom stationary plate the circular plate is divided into three parts: 120 degrees apart from each other. In each parttwo joints (say B_1 and B_2) are placed which are 35 degrees away from each other.

From Fig. 3, the equations (7-12) are established for positions of joints at top moving plate:

$$P_1 = [r \quad 0 \quad 0] \tag{7}$$

$$P_2 = [rcos60 \quad rsin60 \quad 0] \tag{8}$$

$$P_3 = [-rcos60 \quad rsin60 \quad 0] \tag{9}$$

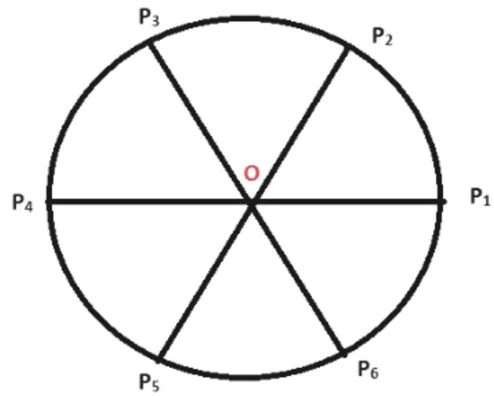


Fig. 3. Angle distribution for positioning of joints on top moving plate.

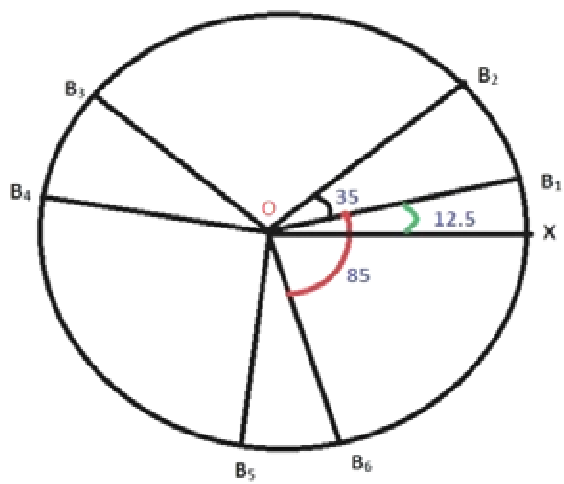


Fig. 4. Angle distribution for positioning of joints on bottom stationary plate.

$$P_4 = [-r \quad 0 \quad 0] \tag{10}$$

$$P_5 = [-rcos60 \quad -rsin60 \quad 0] \tag{11}$$

$$P_6 = [rcos60 \quad -rsin60 \quad 0] \tag{12}$$

Where *r* is the radius of the top moving plate.

Similarly from Fig. 4, the equations (13-18) are established for positions on stationary base plate:

$$B_1 = [Rcos12.5 \quad Rsin12.5 \quad 0] \tag{13}$$

$$B_2 = [Rcos47.5 \quad Rsin47.5 \quad 0] \tag{14}$$

$$B_3 = [-Rcos47.5 \quad Rsin47.5 \quad 0] \tag{15}$$

$$B_4 = [-Rcos12.5 \quad Rsin12.5 \quad 0] \tag{16}$$

$$B_5 = [-Rcos17.5 \quad -Rsin17.5 \quad 0] \tag{17}$$

$$B_6 = [Rcos17.5 \quad -Rsin17.5 \quad 0] \tag{18}$$

Where, R is the radius of the bottom stationary plate.

The rotation matrix for rotation from O_B to O_P is calculated as given in (19).

$$
\begin{aligned}
{}^P_BR &= R_Z(\psi) + R_Y(\theta) + R_X(\varphi) \\
&= \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & sin(\varphi) & \cos(\varphi) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\psi)\cos(\theta) & -\sin(\psi) & \cos(\psi)\sin(\theta) \\ \sin(\psi)\cos(\theta) & \cos(\psi) & \sin(\psi)\sin(\theta) \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & sin(\varphi) & \cos(\varphi) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\psi)\cos(\theta) & -\sin(\psi)\cos(\varphi) + \cos(\psi)\sin(\theta)\sin(\varphi) & \sin(\psi)\sin(\varphi) + \cos(\psi)\sin(\theta)\cos(\varphi) \\ \sin(\psi)\cos(\theta) & \cos(\psi)\cos(\varphi) + \sin(\psi)\sin(\varphi)\sin(\theta) & -\cos(\psi)\sin(\varphi) + \sin(\psi)\sin(\theta)\cos(\varphi) \\ -\sin(\theta) & \cos(\theta)\sin(\varphi) & \cos(\theta)\cos(\varphi) \end{pmatrix}
\end{aligned}
\tag{19}
$$

Where, ψ, θ and Φ are angles of rotation about X, Y and Z axis respectively.

In conclusion, the leg length of each leg of the bed is calculated using (5) knowing the orientation of center of the top plate of the bed (where work piece needs to be moved). From known orientation T is calculated and value of rotation matrix from (19), value of Bi from (13-18) and Pi from (7-12) are calculated and substituted in (5).

4. DYNAMICS MODELLING OF LINEAR ACTUATOR

A linear actuator is an actuator that creates motion in a straight line, in contrast to the circular motion of a conventional electric motor. From Fig. 5 it can be seen that current carrying armature is rotating in a magnetic field, its voltage is proportional to speed [18]. Thus

$$v_b(t) = k_b \frac{d\theta_m(t)}{dt} \tag{20}$$

Where,

$v_b(t)$is back electromotive force ,k_b is constant of proportionality (back emf constant) and $\frac{d\theta_m(t)}{dt} = \omega_m(t)$ is the angular velocity of the motor.

The relationship between the armature current $i_a(t)$, the applied armature voltage $e_a(t)$ and the back emf $v_b(t)$ is given by writing equation (21) for a loop transformed armature circuit in Laplace form

$$R_aI_a(s) + L_a(s)I_a(s) + V_b(s) = E_a(s) \tag{21}$$

Torque developed by the motor is proportional to the armature current

$$T_m(s) = K_tI_a(s) \tag{22}$$

Where

$T_m(s)$ is Torque developed by the motor and $K_t = \text{Motor torque constant}$ which depends on the motor and magnetic field constants. Fig.6 shows a typical equivalent mechanical loading on a motor. J_m is the equivalent inertia. D_m is the equivalent viscous damping at the armature and includes both the armature viscous damping . From Fig.6:

$$T_m(s) = (J_ms^2 + D_ms)\theta_m(s) \tag{23}$$

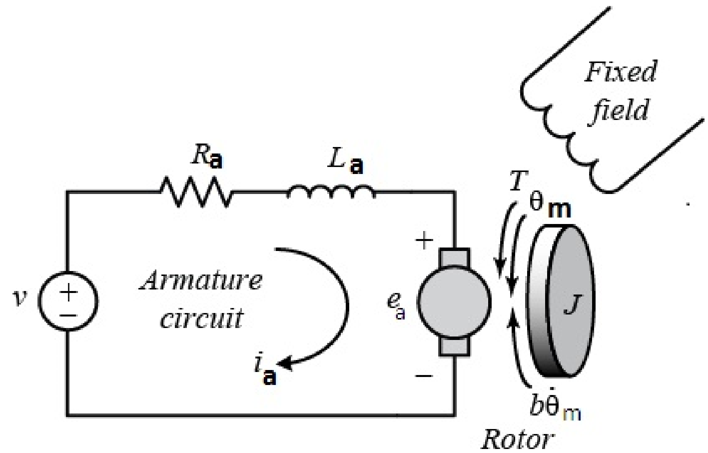


Fig. 5. Schematic diagram of the DC motor.

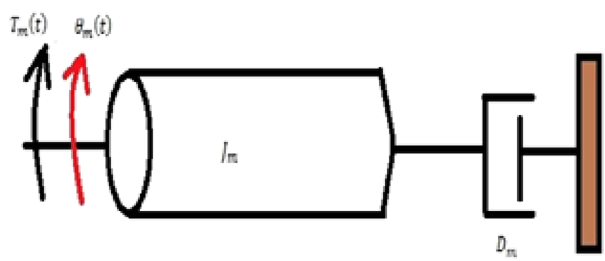


Fig. 6. Representative equivalent mechanical load in a motor.

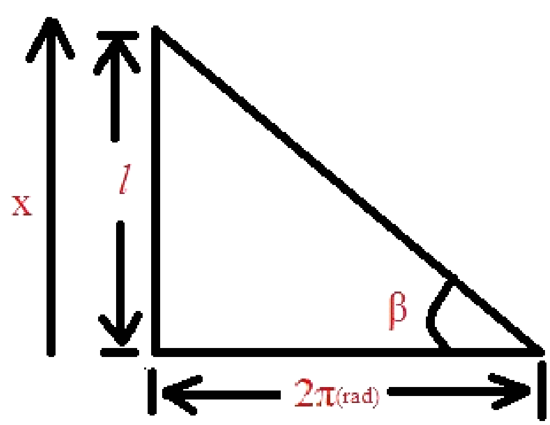


Fig. 7. Relation between angular and linear movement.

For six linear actuators above equation will be:

$$\sum_{i=1}^{6}\big(T_m(s)\big)_i = \sum_{i=1}^{6}\big((J_m s^2 + D_m s)\theta_m(s)\big)_i \tag{24}$$

We will solve here for one linear actuator.

Solving equations:

$$\frac{(R_a+L_a(s))(J_m s^2+D_m s)\theta_m(s)}{K_t} + K_b s\theta_m(s) = E_a(s) \tag{25}$$

Assuming that L_a is small as compared to R_a which is usual for a dc motor, so above equation becomes

$$\left[\frac{R_a}{K_t}(J_m s + D_m) + K_b\right] s\theta_m(s) = E_a(s) \tag{26}$$

Simplifying above equation yields,

$$\frac{\theta_m(s)}{E_a(s)} = \frac{K_t}{R_a J_m s^2+(R_a D_m+K_t K_b)} \tag{27}$$

Fig. 7 shows the relationship between the angular advance caused by the motor and the linear advance after the ball screw [19], in this case β represents the angle of the ball-screw lead; *l* represents the step of the lead and x(t) the linear advance [20].

$$x(t) = \frac{l}{2\pi}\theta_m(t) \tag{28}$$

Where$\theta_m(t)$ is given in radians and represents the angular advance generated by the motor.

So we have

$$\theta_m(t) = \frac{2\pi\, x(t)}{l} \tag{29}$$

And,

$$\frac{X(s)}{E_a(s)} = \frac{lK_t}{2\pi[R_a J_m s^2+(R_a D_m+K_t K_b)s]} \tag{30}$$

A free body diagram of the linear electric actuator is shown in Fig. 8 for a clarity.

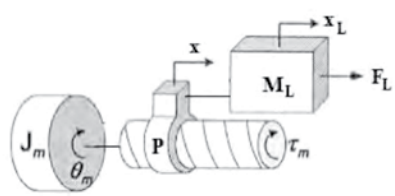


Fig. 8. Free body diagram of linear electrical actuator [18].

```
a = input('angle w.r.t x axis')
b = input('angle w.r.t y axis')
c = input('angle w.r.t z axis')
T= input('enter the coordinates')
r=4;%%upper platform radius
R=5.5;%%lower base radius
```

Fig. 9. Initial part of Matlab code to enter user defined data.

5. SIMULATION RESULTS

In this section simulation results are presented. These results include kinematic analysis, actuator's dynamics analysis and motion analysis of the bed.

5.1 Kinematic Analysis of the Bed

From the kinematics study done in section 3,a code was developed in MATLAB® for calculating lengths of actuators from known coordinates of the work piece. The initial data is entered by the user and code for it is as shown in Fig. 9. Using the initial data entered by the user and the equations (7) to (19) rotation matrix, B and P matrices are calculated. The code used for this purpose is given in Fig. 10. The final lengths of the legs are calculated using (5). The Matlab code of this is presented in Fig. 11.

In order to verify the inverse kinematic model of the machining bed, simulations were carried out on a work piece shown in Fig. 12. Four different positions of the work piece were selected as marked in Fig.12. The Matlab code calculated the desired lengths of all six legs for respective positions on the work piece. Results obtained from MATLAB® simulations are given in Table 2. In Table 2, **ϕ, Θ, Φ** are rotations about X, Y and Z axis respectively whereas L1, L2, L3, L4, L5, L6 are lengths of linear actuators. All the lengths and coordinates are in mm.

```
PRB=[cos(b)*cos(a) -sin(b)*cos(c)+cos(b)*sin(a)*sin(c)sin(b)*sin(c)+cos(b)*sin(a)*cos(c);
   sin(b)*cos(a) cos(b)*cos(c)+sin(b)*sin(a)*sin(c) -cos(b)*sin(c)+sin(b)*sin(a)*cos(c);
   -sin(a) cos(a)*sin(c) cos(a)*cos(c)]

 b1=[R*cos(12.5);R*sin(12.5);0];
 b2=[R*cos(47.5);R*sin(47.5);0];
 b3=[R*cos(47.5);R*sin(47.5);0];
 b4=[-R*cos(12.5);R*sin(12.5);0];
 b5=[-R*cos(17.5);-R*sin(17.5);0];
 b6=[R*cos(17.5);-R*sin(17.5);0];

p1=[r;0;0];
p2=[-r*cos(60);r*sin(60);0];
p3=[-r*cos(60);r*sin(60);0];
p4=[-r;0;0];
p5=[-r*cos(60);-r*sin(60);0];
p6=[r*cos(60);-r*sin(60);0];
```

Fig. 10. Middlepart of Matlab code to calculate matrices from user defined data.

```
L1 = T+PRB*p1-b1;
length1 = norm(C1);
L2 = T+PRB*p2-b2;
length2 = norm(C2);
L3 = T+PRB*p3-b3;
length3 = norm(C3);
L4 = T+PRB*p4-b4;
length4 = norm(C4);
L5 = T+PRB*p5-b5;
length5 = norm(C5);
L6 = T+PRB*p6-b6;
length6 = norm(C6);
```

Fig. 11. Part of Matlab code calculating lengths of six legs required to move to the desired positions.

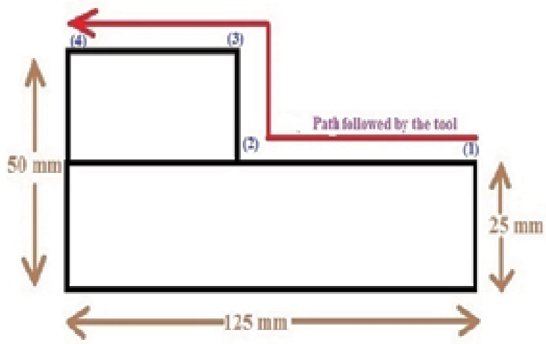


Fig. 12. Work piece used for machining.

Table 2. Length of linear actuators calculated for different coordinates.

Φ (deg)	Θ (deg)	Φ (deg)	Coordinates (mm)	L_1 (mm)	L_2 (mm)	L_3 (mm)	L_4 (mm)	L_5 (mm)	L_6 (mm)
0	0	0	0,0,430	53	74	74	53	60	60
0	0	0	0,75,430	54	77	77	54	60	60
0	0	0	0,75,450	74	96	96	74	79	79
0	0	0	0,125,450	81	105	105	81	85	85

5.2 Dynamics Analysis of Linear Actuator

In order to verify the dynamics of the linear actuators an open loop model was created in Simulink, a tool box of Matlab, where the transfer function of the motor was implemented. Substituting values of constants in (30) following transfer function is obtained and this is embedded in the Simulink model:

```
Transfer function:
               0.0015
----------------------------------
0.000338 s^3 + 42.48 s^2 + 0.045 s
```

To verify the dynamics of the motor, a step response of the open loop system was observed as shown in Fig. 13-a. On the graph time in seconds is on horizontal axis and Amplitude of the displacement is on vertical axis. From open loop response it was expected that the length of linear actuator increases with increase in applied voltage and this behaviour can be seen in Fig. 13-a.The length of linear actuator is increasing linearly but approaching to infinity as time approaches to infinity but the system requirement is to achieve a desired length.

In order to achieve the desired length the system needs to be closed loop. A proportional controller is designed, therefore, and tuned. As a result of auto tuning a proportional gain of 43 is chosen which gives rise time 0.395 seconds and settling time 0.676 seconds with no overshoot. A response of the closed loop system is shown in Fig. 13-b. The closed loop system is now stable and each motor generates the desired input to the lead screws of the linear actuators to give desired lengths of legs within half a second without overshoot of the top plate of the machining bed. More benefits of closed loop response are that precise results are obtained, error is minimized and information about the position of end effect or may be obtained.

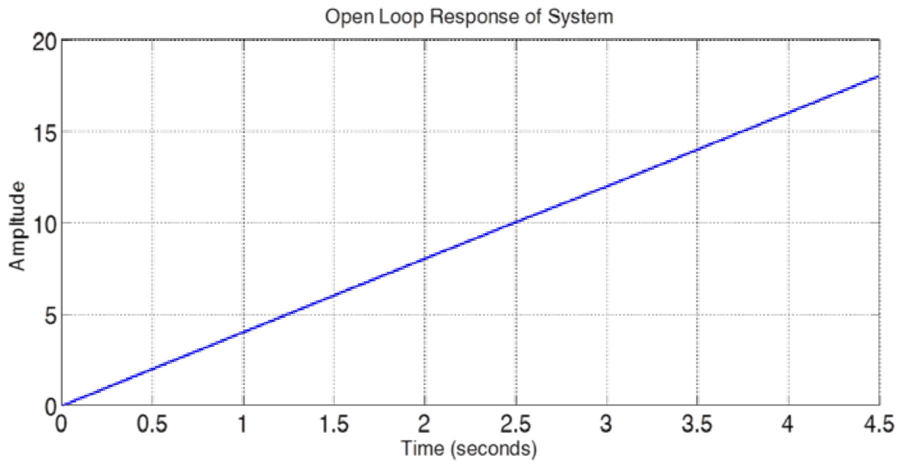


Fig. 13-a. Open loop response of the actuator of machining bed system.

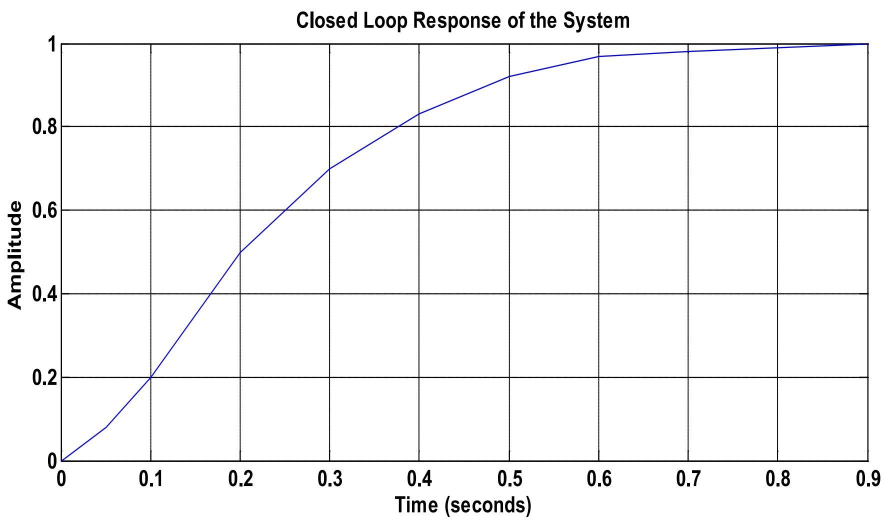


Fig. 13-b. Closed loop response of the actuator of machining bed system.

5.3 Motion Analysis of Bed

For simulations of motion of the proposed six degree of freedom machining bed software called Solid work® was used. The leg lengths were assumed to be same (380 mm: distance between bottom and top plates when the linear actuators are closed) initially. Simulation was performed in Solid Works and movement of bed was tested for 6 degree of freedom by giving different lengths to linear actuators. Fig. 14-a and14-b show shots of the movement of bed with different actuator leg length. The motion presented in Fig. 14-a was generated in x-axis only while Fig. 14-b is generated in three dimensions. The movements were observed as per expectations which verify that the design is acceptable for a parallel manipulator to be used as a 6 DOF machining bed. The leg lengths calculated in section 5.1 for the work piece shown in Fig. 12 were used to simulate motion of the machining bed in Solid Works. The resulting values of position of the centre of top plate and the position values (reference values) for which leg lengths were calculated are plotted for two different cases of motion as shown in the Fig. 15a and Fig. 15b. The difference in plots is error in position. This error occurred due to losses in lead screw or friction in the joints and may be removed in future using position feedback and a controller to make the overall system as closed loop system.

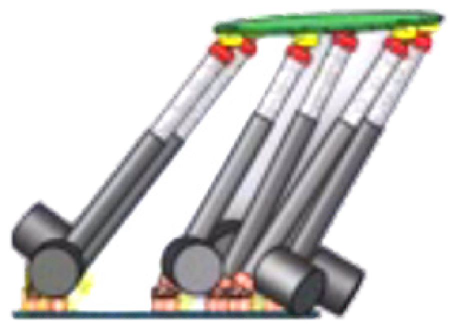


Fig. 14-a. Motion of top plate of machining bed in x-axis.

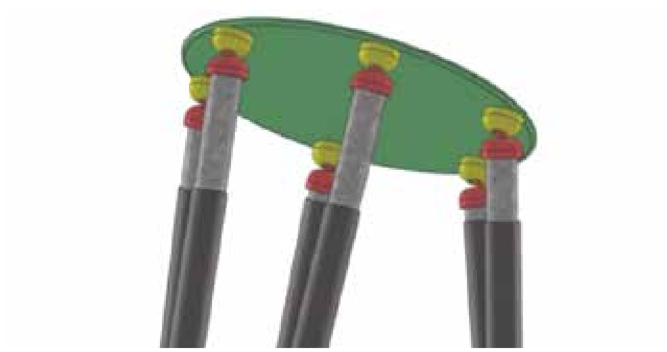


Fig. 14-b. Motion of top plate of machining bed in three axes.

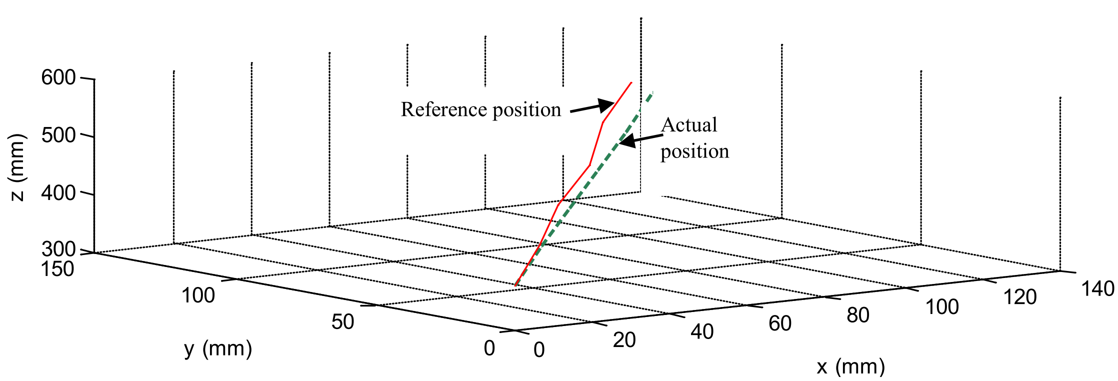


Fig. 15-a. Reference and actual positions of centre of top plate of the machining bed when variation is in x, y & z axis.

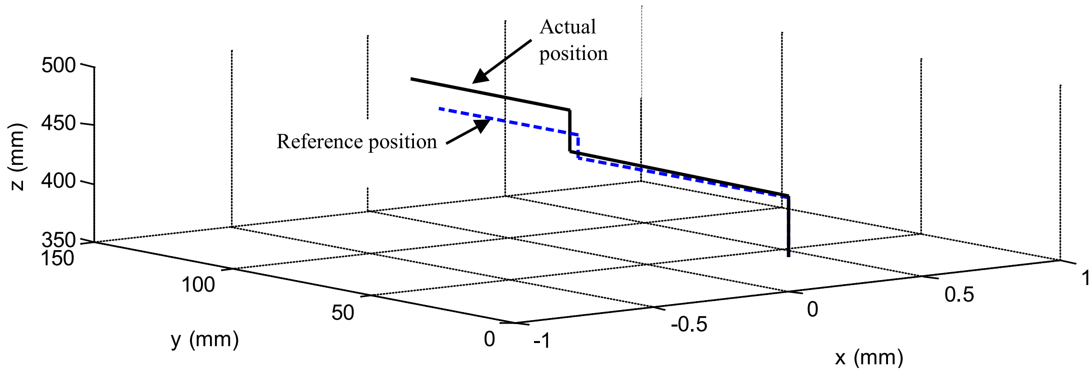


Fig. 16-b. Reference and actual positions of centre of top plate of the machining bed when variation is in y and z axis only.

6 CONCLUSIONS

We have presented a design for a parallel robotic mechanism for a machining bed, the inverse kinematics of the parallel robot and dynamics model of linear actuator. We used Solid Works to simulate the behaviour of the design of bed, dynamics model for linear actuator was verified using MATLAB® and kinematic study for the platform is used to find the lengths of linear actuator for different coordinates in MATLAB®. It is concluded that the use of a 6 DOF parallel mechanism as a machining bed provides the desired results and the verified model of the proposed mechanism may be used for the bed control design and the development of the bed for machining. A control design and implementation will produce precision and accuracy.

7. REFRENCES

1. Sutherland, J.W. *Lecture Notes on Manufacturing*. *Department of Mechanical Engineering, Michigan Technological University,* Houghton, MI, USA. www.mfg.mtu.edu/cyberman/machining.html (Accessed on March 03, 2015).
2. Kausar, Z. M., A. Irshad & S. Shahid. A parallel robotic mechanism replacing a machine bed for micro-machining. In: *Scientific Cooperations International Workshops on Electrical and Computer EngineeringSubfields,* 22-23 August 2014, Istanbul, Turkey, p. 228-233 (2014).
3. Stewart, D. A platform with six degrees of freedom. *Proceedings of Institution of Mechanical Engineers* 180(1): 371–386 (1965).
4. Lee, S. H., J.B. Song, W.C. Choi & D. Hong. Position control of a Stewart platform using inverse dynamics control with approximate dynamics. *Mechatronics* 13(6): 605–619(2003).
5. Liu, K., J. Fitzgerald, & F.L. Lewis. Kinematic analysis of a Stewart platform manipulator. *IEEE Transactions of Industrial Electronics* 40(2):282–93 (1993).
6. Lebret, G., K. Liu, & F.L. Lewis. Dynamic analysis and control of a Stewart platform manipulator. *Journal of Robotic Systems* 10(5):629–655 (1993).
7. Dasgupta, B, & T.S. Mruthyunjaya. Closed-form dynamic equations of the general Stewart platform through the Newton–Euler approach. *Mechanism and Machine Theory* 33(7):993–1012 (1998).
8. Zhang, C., & S. Song. An efficient method for inverse dynamics of manipulators based on the virtual work principle. *Journal of Robotic Systems* 10(5):605–627 (1993).
9. Fichter, E.F.A Stewart platform-based manipulator: general theory and practical construction. *International Journal of Robotics Research* 5(2):157–182 (1986).
10. Zhiming, J. Study of the effect of leg inertia in Stewart platforms. In: *International Conference on Robotics and Automation*, p. 121–126 (1993).
11. Wang, Y. *Symbolic Kinematics and Dynamics Analysis and Control of a General Stewart Parallel Manipulator*. MS thesis, Department of Mechanical and Aerospace Engineering, State University of New York, Buffalo, New York, USA (2008).
12. Valenti, M. Machine tools get smarter. *ASME Mechanical Engineering* 117(11): 70–75 (1995).

13. Kim, J.W., et al. Performance analysis of parallel manipulator architectures for CNC machining. In: *Proc. ASME International Mechanical Engineering Congress and Exposition Symposium on Machine Tools*, Dallas, TX, USA (1997).
14. Dwarakanath, T.A., B. Dasgupta & T.S. Mruthyunjaya. Design and Development of a Stewart platform based Force-Torque Sensor. *Mechatronics* 11(7): 793-809 (2001).
15. Zaiter, A.A., E.L. Ng, S. Kazi & M.S.M. Ali**.** Development of Miniature Stewart Platform Using TiNiCu Shape-Memory-Alloy Actuators. *Advances in Materials Science and Engineering*, Vol. 2015, 928139, (2015).
16. Wulfsberg, J.P, T. Redlich, & P. Kohrs. Square foot manufacturing: a new production concept for micro manufacturing. *Production Engineering* 4(1): 75–83 (2010).
17. Bingul, Z., & O. Karahan. Dynamic Modeling and Simulation of Stewart Platform. In: *Serial and Parallel Robot Manipulators-Kinematic, Dynamics, Control and Optimization.* INTECH, Croatia, p. 19-44 (2012).
18. Norman, S.N. *Control Systems Engineering*. John Wiley & Sons, p. 79-85 (1995).
19. Ruiz-Rojas, E.D., J.L. Vazquez-Gonzalez, et.al. Mathematical Model of a Linear Electric Actuator with Prosthesis Applications. In: IEEE *International Conference on Electronics, Communications and Computers* 3-5 March, Puebla, p.182 - 186(2008).
20. Kim, M.S., & S.C. Chung. A systematic approach to design high-performance feed drive system. *International Journal of Machine Tools & Manufacture* 45(12-13): 1421-1435 (2005).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 177–185 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Application of Computational Flow Dynamics Analysis for Surge Inception and Propagation for Low Head Hydropower Projects

M. Mohsin Munir[1 *], Taimoor Ahmed[2], Javed Munir[3], and Usman Rasheed[4]

[1,2,3]Water Resources Division, National Engineering Services, Pakistan, Pvt, Ltd, Lahore, Pakistan
[4]Civil Engineering Department, University of Management and Technology, Lahore, Pakistan

Abstract: Determination of maximum elevation of a flowing fluid due to sudden rejection of load in a hydropower facility is of great interest to hydraulic engineers to ensure safety of the hydraulic structures. Several mathematical models exist that employ one-dimensional modeling for the determination of surge but none of these perfectly simulate real-time circumstances. The paper envisages investigation, inception and propagation of surge for a Low-Head Hydropower project using Computational Fluid Dynamics (CFD) analysis in FLOW-3D software. The fluid dynamic model utilizes Reynolds' Averaged Navier-Stokes Equations (RANSE) for surge analysis. The CFD model is designed for a case study at Taunsa Hydropower Project in Pakistan which has been run for various scenarios keeping in view the upstream boundary conditions. The prototype results were compared with the results of physical model and proved quite accurate and coherent. It is concluded that CFD Model gives an insight of the phenomenon which are not apparent in physical model and shall be adopted in future for the similar low head projects. Its application will be helpful in limiting delays and cost incurred in the physical model testing.

Keywords: Surge, FLOW-3D, numerical model, Taunsa, RANSE

1. INTRODUCTION

Maximum elevation that the flow will achieve due to a sudden rejection of load in a hydropower facility is an important parameter to hydraulic engineers. The information is required in setting the maximum height of side walls to prevent overflow in the headrace channel as well as to understand the surge propagation upstream in headrace channel to schedule the opening of gates of the main barrage and balance the discharge through the barrage and power channel.

In Pakistan, physical model studies are the only practical medium available to understand and analyze the three dimensional and time-dependent complexities of the fluid flow phenomenon. Physical models can only be setup at the final design stage and their execution is expensive. However, the avant-garde computational flow dynamics has emerged not only as an alternative analysis and design tool but also an approach to analyze phenomenon that are not possible to evaluate with physical testing.

2. LITERATURE REVIEW

In a hydropower project, a monoclinal wave exhibiting a rapidly varying flow may generate due to sudden closure or opening of the powerhouse control structure such as the sluice gates or wicket gates. This hydraulic transient analysis is of immense importance in hydropower projects.

Audrius et al. [2] concluded in his study that recent advances in Computational Fluid Dynamics (CFD) has made it possible to simulate highly turbulent multiphase flows in reasonable time with good accuracy. The study explored recent development in hydraulic design of Pelton and Turgo Impulse turbines and highlights the opportunities for future expansion. Shojaeefard

Received, January 2016; Accepted, May 2016
*Corresponding Author: M. Mohsin Munir:Email: mohsin_uett@yahoo.com

et al. [3] investigated 3D behavior of axial-flow type microhydro turbines by utilizing an open source computational fluid dynamic (CFD) code, to investigate the rotor-stator interaction and losses occurring in the turbine.

Several mathematical methods exist for one-dimensional analysis, e.g., Saint Venant Equations and the Johnson method. The first method is simplification of the original problem and is based on several assumptions. These methods do not truly represent the original problem and these conditions are rarely met in practice. Similarly, the Johnson method can be arduous as the computation proceeds because it involves the production and propagation of numerous surges. Hence it becomes inaccurate and difficult to assess the surge using these methods.

Other time-dependent analysis methods increase the complexities of the problem by introducing additional independent variable of time since the resulting equations become partial differential equations instead of ordinary differential equations. Method of Characteristics and Implicit and Explicit Finite-Difference Methods have better accuracy and rigor but are time-consuming and cannot be applied appropriately to original conditions. Implicit Methods are preferred in special cases, such as analysis of hydropower projects comprising of open headrace channel with tailrace tunnel (Closed conduit). Such complexities can only be modeled using a contemporary software like FLOW-3D, based on numerical solution schemes that can accurately predict fluid flow using the concept of fluid volume tracking.

FLOW-3D was developed at the Los Alamos National Laboratory in the 1960s and 1970s as a general purpose computational fluid dynamics simulation package. FLOW-3D uses an Eulerian framework in which volume tracking technique models the free surface. It is based on volume of fluid method in which fraction analysis of fluid in each cell is carried out. FLOW-3D is a robust software that can handle the breakup and coalescence of fluid masses.

FLOW-3D uses several models to numerically simulate turbulent flows. For this case study, the Renormalization group (RNG) k-epsilon turbulence model was used with a no slip or partial slip wall shear boundary condition. The RNG turbulence model uses statistical models to solve the turbulent kinetic energy (k) and the turbulent kinetic energy dissipation rate (Є), renormalizing the Navier Stokes Equations to cater for the effects caused by smaller scale motion.

3. CASE STUDY OF 135MW TAUNSA HYDROPOWER PROJECT

The Taunsa Hydropower Project has been sited along right bank of Taunsa Barrage across Indus River in Pakistan. It is a run of the river, low head hydropower scheme envisaged to provide power to the national grid. Table 1 summarizes the salient

Table 1. Salient features of 135MW Taunsa HPP.

Type of Turbine	Horizontal Bulb
Mode of Operation	Run of the River
Installed Capacity	135 MW
Gross Head	6.0 m
Rated Head	5.8 m
Design Discharge	3,155.5 m^3/sec
Turbine Units	9 Units
Headrace Width	203 m
Headrace Length	1100 m
Vertical Gate Height	16 m

features of the project.

FLOW-3D analysis is often limited by computation power available. For numerical simplicity in this case, the three-dimensional model was setup based on only one hydropower unit and not the complete power-station. FLOW-3D can effectively use the symmetrical approach conditions of the project layout without affecting the accuracy of the surge analysis. A 3D model was used since flow has three-dimensional characteristics.

3.1 Model Preparation

The model for Taunsa hydropower project was prepared in AutoCAD and imported in stereo-lithographic format which is used for rapid prototyping, 3D printing and computer-aided manufacturing. Three dimensional models were prepared for the headrace, transitions, powerhouse and tailrace. Fig. 1 illustrates the 9-unit powerhouse with gates open. Fig. 2 shows the detail of transitions of a single unit with bulb turbine housed inside.

3.2 Meshing

Meshing is a consequential part of the analysis process which not only determines the numerical accuracy of the model but also the memory and time required for the simulation. Meshing was done and refined in FLOW-3D model setup options. FLOW-3D has an advantage of FAVOR; the Fractional Area/Volume Obstacle Representation Method, which allows modeling of complex geometries based on equations formulated as functions of the area and volume porosity.

As specified before, for simplicity as well as increased accuracy, meshing was done for one unit. This is a common practice since for large models no of total active cells are limited by computer memory. For Taunsa Hydropower Project, the

Table 2. Summary of meshing in FLOW-3D.

Total active cells	2320000		
Dimension	**Extent (m)**	**No of cells**	**Cell size (m)**
X	400	400	1.00
Y	18.5	50	0.37
Z	29	116	0.25

Table 3. Flow conditions for Taunsa HPP.

Flow Condition	U/S water level (m)	D/S water level (m)
Normal Flow	135.94	130.25
Flood Flow	136.25	133.26

Table 4. Summary for normal conditions.

Location	Max Surge	Surge Elevation	Time (sec)
Unit Entrance	2.25 m	138.19 m	7
85m U/S	2.21 m	138.15 m	14
Headrace Entrance	0.53 m	136.47 m	260

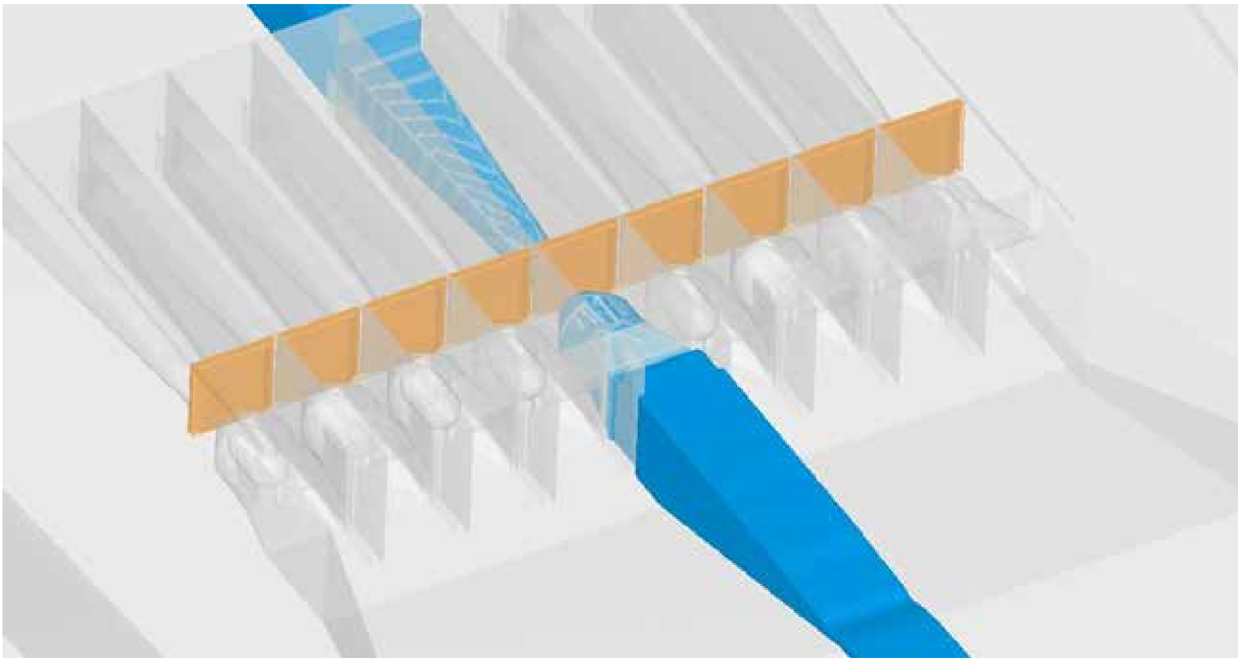


Fig.1. Isometric view of Taunsa HPP model.

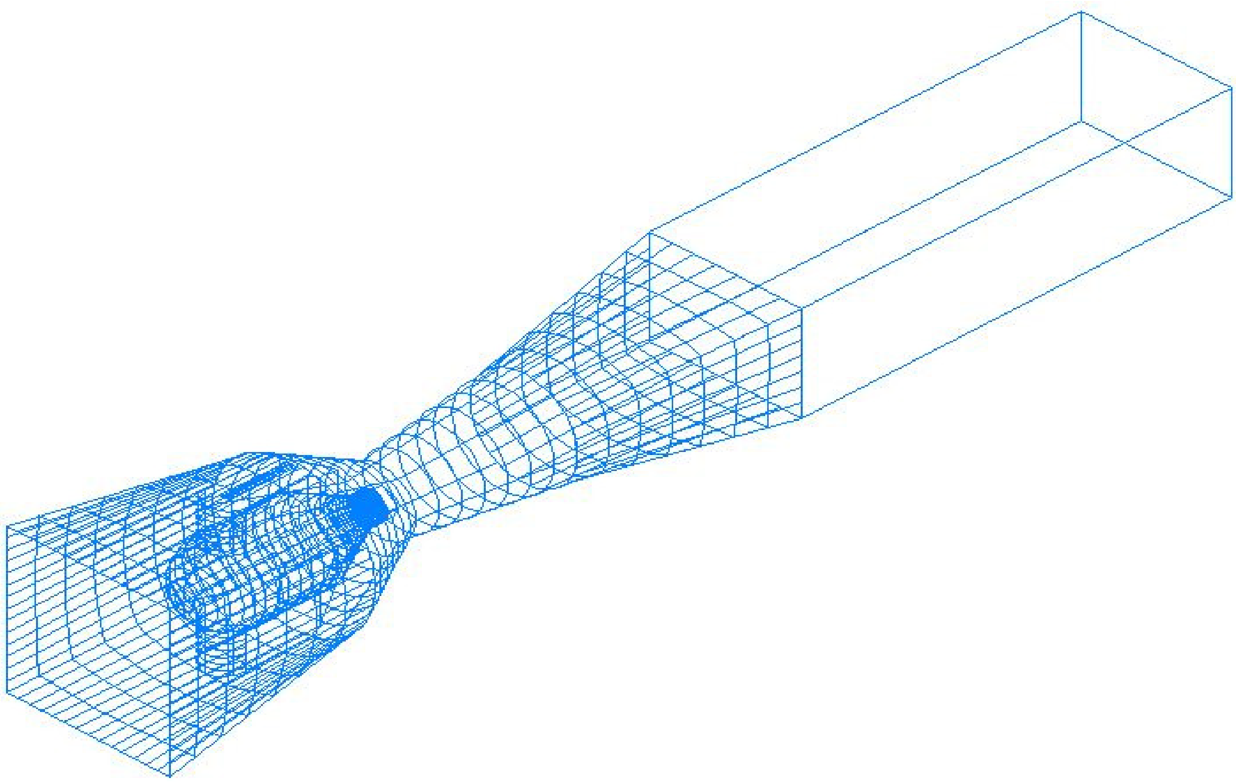


Fig.2. Wireframe of a single in-take and draft tube with bulb turbine.

extents, no of cells and cell sizes are tabulated in Table 2.

Cell sizes for the mesh were selected based on two factors. First consideration in this regard is the model accuracy. In case of surge analysis, height of surge (z-direction) is the most important. Hence, smallest cell size has been selected for z axis. Second consideration is CPU memory and computation time. Increasing the number of cells requires greater computation power and longer simulation time. Average simulation time for Taunsa HPP was 15 hours. Keeping in view the aforementioned two factors, minimum cell size was selected so that full computation power can be employed.

3.3 Boundary Conditions

Boundary conditions have a huge impact on the final results of the simulation. It is necessary to assess the boundary conditions that best replicate the real-time conditions and actual simulation. Boundary conditions applied for the problem are specified below:

X minimum	Specified Pressure Boundary
X maximum	Outflow Boundary
Y minimum	Symmetry Boundary
Y maximum	Symmetry Boundary
Z minimum	Wall Boundary
Z maximum	Symmetry Boundary

The upstream boundary condition consisted of a specified pressure to maintain a prescribed reservoir elevation. The downstream boundary utilized the FLOW-3D outflow boundary condition. A symmetry boundary condition was applied along right and left side of the mesh section to take advantage of the inherent symmetry in the problem and thereby decrease computational time while maximizing spatial resolution. Boundary conditions used have been explained below:

3.3.1. Symmetry Boundary

There is no mass flux (flow) through a symmetry boundary. There is also no shear stress or heat transfer applied at this boundary type. It is useful for reducing the size of a simulation when symmetry exists by cutting the simulation at the symmetry plane.

3.3.2. Wall Boundary

The wall boundary is similar to the symmetry boundary in which mass flux across the boundary is not allowed. However, with a wall boundary, heating and viscous stresses can be applied.

3.3.3. Specified Pressure Boundary

This boundary type sets a pressure condition. The pressure can be constant by setting a value in the dialog box, or time dependent by selecting the pressure button.

3.3.4. Outflow Boundary

This boundary type is useful for surface waves because they are able to leave the flow region without reflecting back into the domain. This boundary type looks at flow conditions just inside the mesh and matches them to allow fluid to freely dissipate through the mesh extent.

3.4 Case Study Scenarios

Two flow scenarios have been analyzed for the case study as tabulated in Table 3. During normal flow operation while all 9units are in operation, initial water level elevation of 135.94 m will be maintained in the headrace channel and 130.25m in the tailrace channel. The model is allowed to operate freely for 30 seconds after which the gates are shut at 1.4 m/s to close the orifice in 5 seconds in order to replicate sudden closure conditions. Type B surge or the rejection surge occurs as a result of sudden decrease in power output. Similarly during flood flow operation, initial water level elevation of 136.25 m will be maintained in the headrace channel and 133.26 m in the tailrace channel. Other conditions are kept same as in normal flow conditions.

4. RESULTS

The rejection surge occurs as a result of sudden decrease in power output. Fig. 3, 4 and 5 illustrate the surge inception and propagation at different times after load rejection in the extent just upstream of the gates. The points of flow direction reversal

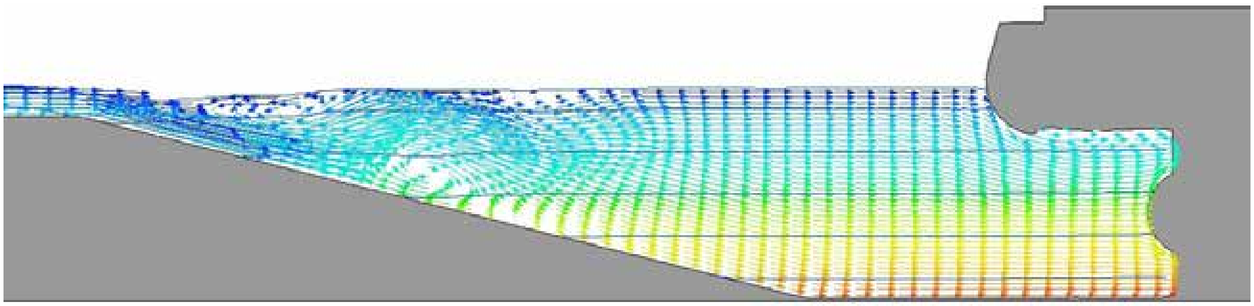


Fig. 3. Flow through the power unit at full operation.

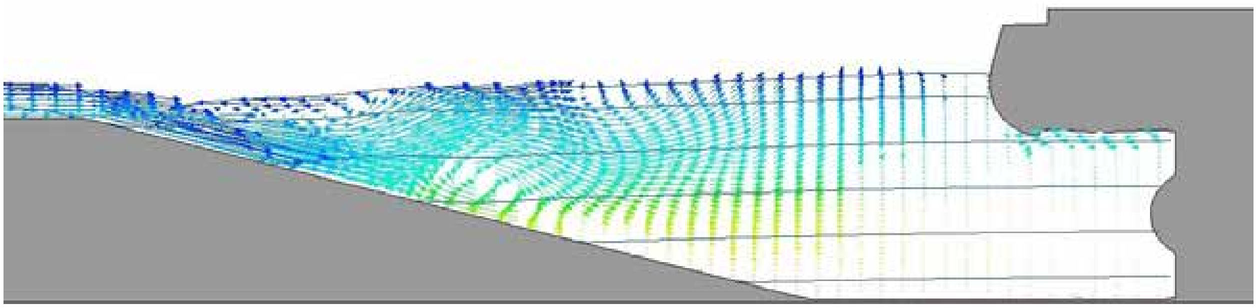


Fig. 4. Surge inception at time of load rejection.

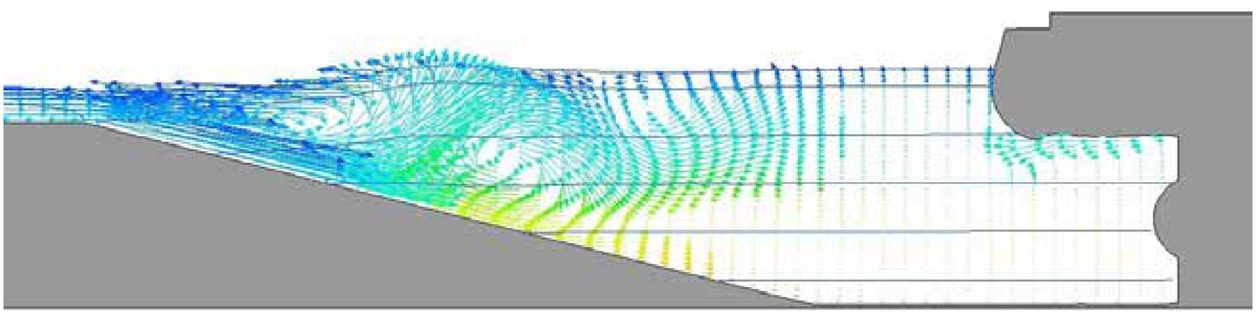


Fig. 5. Flow reversal after 10 seconds of load rejection.

can easily be tracked for the upstream advancing surge. This is a feature that is only possible with computational flow dynamics analysis since it is not possible to observe these phenomena in physical model testing.

In Fig. 3, flow vectors at full flow conditions are illustrated by arrows. This shows an uninterrupted movement of fluid. After load rejection, a *region of fluid immobility* starts to develop as observed from Fig. 4. This region acts a cushion against incoming discharge and reverses its direction which results in surge inception. Fig. 5 shows the region of fluid immobility develops upstream with time as surge wave begins to achieve greater elevation. The phenomenon holds true for both normal flow conditions and flood flow conditions.

4.1 Normal Flow Conditions

As the surge moves upstream, change in flow depth will ensue. FLOW-3D can track the change in elevation of free surface efficiently for any spatial location of the model. Changes in surface elevation with time were observed for three locations, i.e., at turbine unit entrance, 85 m upstream of the entrance and at headrace entrance.

The temporal and spatial changes in free surface elevation at 3 locations have been shown in Fig. 6. Table 4 shows the summary of results which depict a maximum surge of 2.25 m at unit entrance. Time of arrival of first wave has also been tabulated. The surge is expected to reach the headrace entrance in 260 seconds after gate closure with an average velocity of 4.84 m/s. As the surge wave moves upstream, the height of wave dampens.

4.2 Flood Flow Conditions

In case of flood flows in Taunsa hydropower, an upstream water level of 136.25 m is maintained in the headrace with a corresponding tail-water level of 133.26 m in tail-race. Other conditions are kept same as in normal flow conditions. The surge is expected to reach the headrace entrance in 206 seconds with an average surge velocity of 6.7 m/s. Maximum surge elevation of 138.4 m (2.21 m) is observed at the power unit entrance. The results have been tabulated in Table 5. Fig. 7 shows the plot between the change in free surface elevation at the three selected locations and time after load

Table 5. Summary for flood conditions.

Location	Max Surge	Surge Elevation	Time (sec)
Unit Entrance	2.21 m	138.39 m	6
85 m U/S	2.00 m	138.20 m	12
Headrace Entrance	0.32 m	136.52 m	206

Table 6. Salient features of physical model.

Features	Symbol	Conversion Factors	Scale Ratio
Length	L_r	L_r	1/36
Depth	Y_r	Y_r	1/36
Area	A_r	$L_r Y_r$	1/1296
Time	T_r	$L_r / Y_r^{1/2}$	1/6
Velocity	V_r	$Y_r^{½}$	1/6
Discharge	Q_r	$L_r Y_r^{3/2}$	1/7776
Roughness Co-efficient	n_r	$Y_r^{2/3} / L_r^{1/2}$	1/1.8193

Table 7. Comparison of parameters between numerical and physical model.

Parameters	Max Surge	Surge Elevation	Time (sec)
Numerical Model	2.25 m	138.19 m	260
Physical Model	2.05 m	138.00 m	207

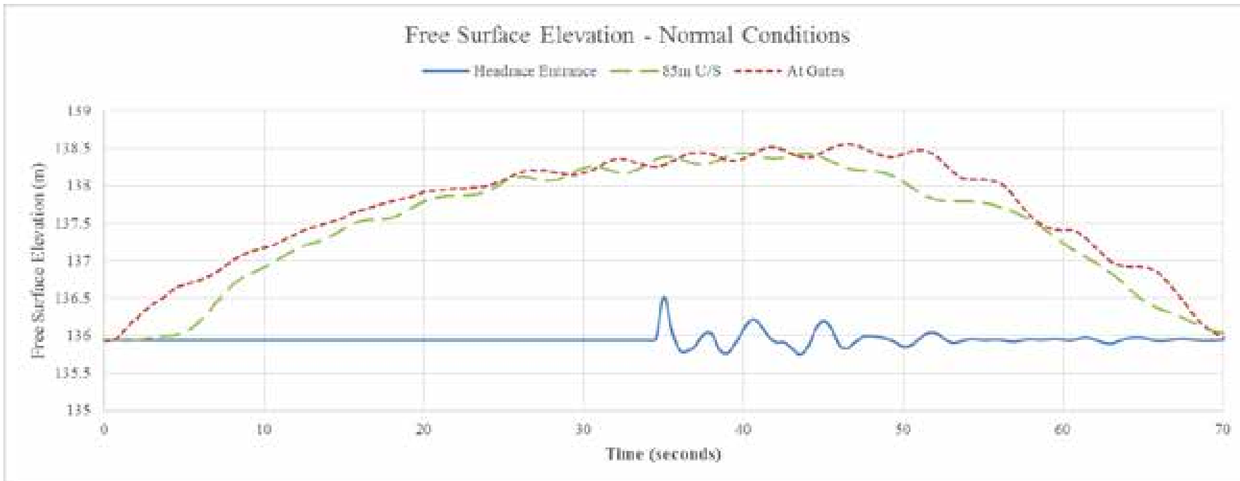


Fig. 6. Change in water elevation vs time after load rejection for normal flow conditions Taunsa headrace.

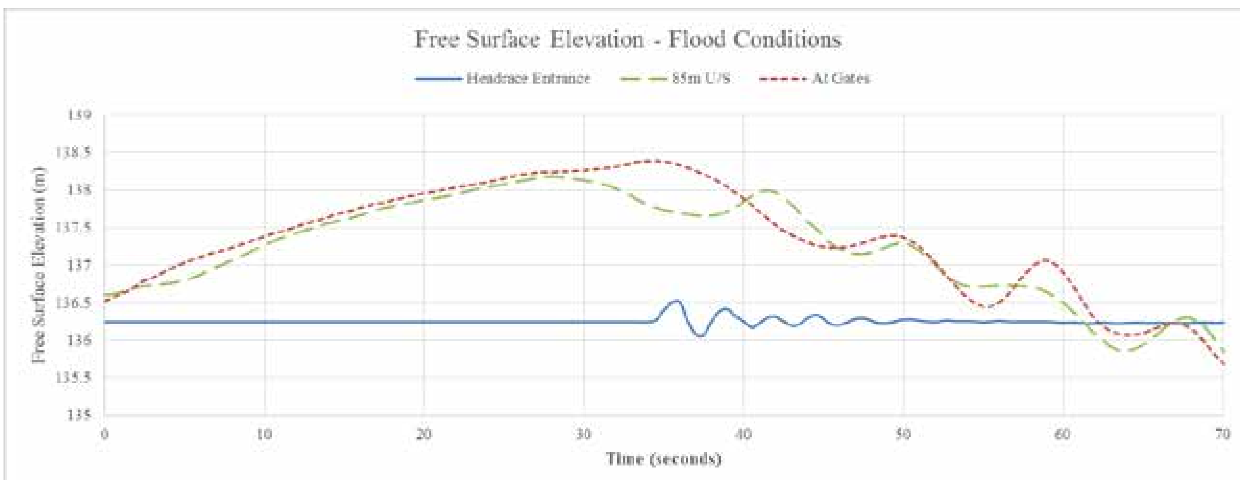


Fig. 7. Change inwater elevation vs time after load rejection for flood flow conditions Taunsa headrace.

rejection from the power unit.

It is observed that the surge travels faster and achieves a lower maximum free surface elevation for the flood flow conditions compared to the Normal Flow Conditions.

4.3 Physical Model Testing

Geometrical model of the proposed power channel along with appurtenant structures were constructed at a scale of 1:36 to compute maximum surge height and travel time to the start of headrace channel. The salient features of the model tray are presented in Table 6. The simulation results were compared with the mathematical formulae mentioned above as well as the physical testing model carried out at Irrigation Research Institute, Nandipur, Punjab. The results of the physical model testing were within ±10% of the results achieved with simulations in FLOW-3D and presented in Table 7.

5. CONCLUSIONS

Observations of the surge inception and propagation are in line with the theory and assumptions specified by VenTe Chow and other authors. For upstream advancing surge, when the surge wave reaches any point in the headrace, the water elevation behind the wave approximately equals the maximum elevation of the surge wave.

The simulation results were compared with the mathematical formulae mentioned above as well as the physical testing model carried out at Irrigation Research Institute, Nandipur, Punjab. The surge height computed at physical model is 2.05 m with maximum surge elevation of 138.00m. The time at which surge wave enters the pond is computed as 207 sec. The results obtained from physical model are found within ±10% of the results achieved with simulations in FLOW-3D. This has built greater confidence in modeling using computational flow dynamics especially for modeling in field of power generation.

Applications of computational flow dynamics have been increasing in engineering applications over recent times. The computer numerical models are a cost-effective alternative to physical modeling techniques offering more flexibility during design and analysis. However, CFD analyses limited by computational power. The present 3D detailed analysis of a single unit required more than 24 hours of computation per simulation to examine a 70-second time-history. A longer time-history is often more desirable.

The analysis of Taunsa hydropower headrace channel using computational flow dynamics is a step-forward in supplementing the results from mathematical modeling and conventional physical model testing in Pakistan. It is not only a robust analysis software solution but can be employed as efficacious tool in the design process of hydraulic structures where it can offer extensive flexibility to assess and compare different proposed designs and their efficiencies. This design support can save a lot of time and money by the optimization process before final designs are physically tested.

6. REFERENCES

1. Chow, VenTe. *Open Channel Hydraulics*. McGraw-Hill Book Company, USA, p. 561-565 (1959).
2. Zidonis, A., D.S. Benzon, & G.A. Aggidis. Development of hydro impulse turbines and new opportunities. *Renewable and Sustainable Energy Reviews* 51: 1624-1635 (2015).
3. Shojaeefard, M.H., A. Mirzaei, M.S. Abedinejad, & Y. Yassi. Fluid flow modeling through an axial-flow microhydro turbine, Tehran, Iran. *Technical Gazette 22*(6): 1517-1526 (2015).
4. Rich, G.R. *Hydraulic Transients*, 2nd ed. Dover Publications, New York, NY, USA (1951).
5. Chaudhry, M.H. *Applied Hydraulic Transients*, 3rd ed. Springer, London: p. 450-474 (2014).
6. Yakhot, V., S,A. Orszag, S. Thangam, T.B. Gatski, & C.G. Speziale. Development of turbulence models for shear flows by a double expansion technique. *Physics of Fluids A*. 4 (7): 1510-1520 (1992).
7. *FLOW-3D®SoftwareUser's Manual Version 9.3.* Flow Science Inc. Los Alamos, New Mexico, USA (1997).
8. Hirt, C.W. & B.D. Nichols. Volume of fluid (VOF) method for the dynamics of free boundaries. *Journal of Computational Physics* 39(1): 201-225. (1981).
9. Teklemariam, E., B.W. Korbaylo, J.L. Groeneveld, & D,M. Fuchs. Computational fluid dynamics: Diverse applications in hydropower project's design and analysis. In: *Proceedings of the CWRA 55th Annual Conference, Winnipeg, Man,* p. 11-14 (2002).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 187–201 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Some Local-Value Relationships for the Recurrence Relation Related to the Tower of Hanoi Problem

Abdullah-Al-Kafi MAJUMDAR*

Ritsumeikan Asia-Pacific University, 1–1 Jumonjibaru, Beppu 874-8577, Japan

Abstract: Motivated by the recurrence relation satisfied by the 4-peg tower of the Hanoi Problem, Matsuura [1] has considered the generalized recurrence relation of the form

$$T(n,\alpha,\beta) = \min_{1 \le t \le n} \left\{ \alpha T(n-t,\alpha,\beta) + \beta S(t,3) \right\},$$

where α and β are natural numbers, and $S(t,3) = 2^t - 1$ is the solution of the 3-peg Tower of Hanoi problem with t discs. This paper studies more closely the above recurrence relation and gives some new relationships, including some local-value relationships. The Reve's puzzle is a particular case of the above recurrence relation with $\alpha = 2$.

Keywords : Tower of Hanoi, recurrence relation, local-value relationships

1. INTRODUCTION

The generalized 4-peg Tower of Hanoi problem (commonly known as the Reve`s puzzle), posed by Dudeney [2] is as follows : There are 4 pegs, designated as S, P_1, P_2, and D, and n discs D_1, D_2, …, D_n of different sizes, where D_1 is the smallest disc, D_2 is the second smallest, and so on with D_n being the largest. Initially, the discs rest on the source peg, S, in a tower in increasing order, with the largest disc D_n at the bottom, the second largest disc, D_{n-1}, above it, and so on, with the smallest disc D_1 at the top. The objective is to move the tower from the source peg S to the destination peg, D, in minimum number of moves, under the conditions that a move can transfer only the topmost disc from one peg to another such that no disc is ever placed on top of a smaller one.

Denoting by $M(n,4)$ the minimum number of moves required to solve the 4-peg Tower of Hanoi problem, the scheme followed is :

Step 1 : Move (optimally) the topmost n – k (smallest) discs from the source peg, S, to one of the auxiliary peg, say, P_1, using the 4 pegs, in $M(n-k, 4)$ moves,

Step 2 : Move (optimally) the tower of the largest k discs from the peg S to the destination peg, D, using the 3 pegs available, in M(k, 3) moves,

Step 3 : Move (optimally) the $n-k$ discs from the peg P_1 to the peg D, in $M(n-k, 4)$ moves,

where k is to be determined such that the total number of moves is minimum.

The above scheme leads to the following dynamic programming equation :

$$M(n,4) = \min_{0 \le k \le n-1} \left\{ 2M(k,4) + M(n-k,3) \right\}, \tag{1.1}$$

$$M(0,4) = 0;\ M(n,3) = 2^n - 1,\ n \ge 0. \tag{1.2}$$

Motivated by the recurrence relation (1.1), Matsuura [1] considers the following recurrence relation

Received, July 2015; Accepted, March 2016
*Email: aakmajumdar.gmail.com

$$T(n,\alpha,\beta) = \min_{1 \le t \le n} \{ \alpha T(n-t,\alpha,\beta) + \beta S(t,3) \}, \tag{1.3}$$

$$T(0,\alpha,\beta) = 0;\ S(t,3) = 2^t - 1,\ t \ge 0, \tag{1.4}$$

where α and β are natural numbers. It may be mentioned here that, the particular case $T(n,2,1) = M(n,4)$.

This paper gives some local-value relationships involving $T(0,\alpha,\beta)$, in connection with the recurrence relation (1.3). They are given in Section 3. In Section 2, we give some preliminary results, that would be required to prove the results in Section 3. We conclude the paper with some remarks in Section 4.

2. SOME PRELIMINARY RESULTS

The result below has been established by Matsuura [1].

Lemma 2.1 : For any natural numbers α and β,

$$T(n,\alpha,\beta) = \beta T(n,\alpha,1),\ n \ge 1.$$

In view of the Lemma 2.1, it is sufficient to consider $T(n,\alpha,1)=T(n,\alpha)$, say, where $T(n,\alpha)$ satisfies the following recurrence relation :

$$T(n,\alpha) = \min_{1 \le t \le n} \{ \alpha T(n-t,\alpha) + S(t,3) \}, \tag{2.1}$$

$$T(0,\alpha) = 0. \tag{2.2}$$

An equivalent form of (2.1) is the following :

$$T(n,\alpha) = \min_{0 \le k \le n-1} \{ \alpha T(k,\alpha) + S(n-k,3) \}, \tag{2.3}$$

The following result deals with the special case $T(n,1)$.

Proposition 2.1 : For any $n \ge 1$, $T(n,1) = n$.

Proof : The proof is by induction on n. By (2.1), $T(1,1) = 1$, so that the result is true for n=1. So, we assume that the result is true for some n (so that $T(i,1)=iS(1,3)$ for all $1 \le i \le n$). Now, using the induction hypothesis, we get

$$T(n+1,1) = \min_{1 \le t \le n+1} \{ T(n+1-t,1) + S(t,3) \},$$

$$= \min_{1 \le t \le n+1} \{ (n+1-t)S(1,3) + S(t,3) \}.$$

Now, the sequence of numbers $\{(n+1-t) + S(t,3)\}_{t=1}^{n}$ is strictly increasing in t, since

$$(n+1-t-1) + S(t+1,3) > (n+1-t) + S(t,3)$$

if and only if

$$S(t+1,3) - S(t,3) > 1,$$

which is true for all $t \ge 1$. Thus, $T(n+1,1)$ is attained at $t=1$, so that

$$T(n+1,1) = n+1,$$

completing induction.

From Proposition 2.1, we see that, it is sufficient to consider the case when $\alpha \ge 2$. Since the case x =2 has been treated in more detail in Majumdar [3], it is, in fact, sufficient to concentrate our attention to the case when $\alpha \ge 3$.

For small values of n, the explicit forms of $T(n,\alpha)$ can be derived. This is done in the following Lemmas 2.2 – 2.7.

Lemma 2.2 : $T(2,\alpha)=3$ for all $\alpha \geq 2$.

Proof : By definition,

$$T(2,\alpha) = \min_{1\leq t\leq 2} \{\alpha T(2-t,\alpha)+S(t,3)\}$$

$$= \min\{\alpha T(1,\alpha)+S(1,3),\ S(2,3)\}$$

$$= \min\{\alpha+1,\ 3\},$$

from which the result follows immediately.

Lemma 2.3 : $T(3,\alpha)=\begin{cases}\alpha+3, & \text{if } 2\leq\alpha\leq 4\\ 7, & \text{if } \alpha\geq 4\end{cases}$

Proof : Using the definition, together with Lemma 2.2, we get,

$$T(3,\alpha) = \min_{1\leq t\leq 3} \{\alpha T(3-t,\ \alpha)+S(t,\ 3)\}$$

$$= \min\{\alpha T(2,\alpha)+S(1,3),\ \alpha T(1,\alpha)+S(2,3),\ S(3,3)\}$$

$$= \min\{3\alpha+1,\ \alpha+3,\ 7\}\}.$$

Now, for any $\alpha\geq 2$, $3\alpha+1 > \alpha+3$; moreover,

$\alpha+3 > 7$ if and only if $\alpha\geq 4$.

All these establish the lemma.

Corollary 2.1 : $T(3,\alpha)-T(2,\alpha)=\begin{cases}\alpha, & \text{if } 2\leq\alpha\leq 4\\ 4, & \text{if } \alpha\geq 4\end{cases}$

Lemma 2.4 : $T(4,\alpha)=\begin{cases}\alpha+7, & \text{if } 2\leq\alpha\leq 8\\ 15, & \text{if } \alpha\geq 8\end{cases}$

Proof : By definition, together with Lemma 2.2 and Lemma 2.3, we have

$$T(4,\alpha) = \min_{1\leq t\leq 4} \{\alpha T(4-t,\alpha)+S(t,3)\}$$

$$= \min\{\alpha T(3,\alpha)+S(1,3),\ \alpha T(2,\alpha)+S(2,3),\ \alpha T(1,\alpha)+S(3,3),\ S(4,3)\}$$

$$= \min\{\alpha T(3,\alpha)+1,\ 3(\alpha+1),\ \alpha+7,\ 15\}\}.$$

Now, for any $\alpha\geq 2$, $3(\alpha+1) > \alpha+7$; and $\alpha+7 > 15$ if and only if $\alpha\geq 8$.

Also, for any $\alpha\geq 2$,

$\alpha(\alpha+3)+1 > \alpha+7$, $7\alpha+1 > \alpha+7$.

Thus, we get the desired expression for $T(4,\alpha)$.

Corollary 2.2 : $T(4,\alpha)-T(3,\alpha)=\begin{cases}4, & \text{if } 2\leq\alpha\leq 4\\ \alpha, & \text{if } 4\leq\alpha\leq 16\\ 8, & \text{if } \alpha\geq 16\end{cases}$

Lemma 2.5 : $T(5,\alpha)=\begin{cases}3\alpha+7, & \text{if } 2\leq\alpha\leq 4\\ \alpha+15, & \text{if } 4\leq\alpha\leq 16\\ 31, & \text{if } \alpha\geq 16\end{cases}$

Proof : By definition, together with Lemma 2.2 – Lemma 2.4, we have

$$T(5,\alpha) = \min_{1 \le t \le 5} \{\alpha T(5-t,\alpha) + S(t,3)\}$$

$$= \min\{\alpha T(4,\alpha) + S(1,3), \alpha T(3,\alpha) + S(2,3), \alpha T(2,\alpha) + S(3,3), \alpha T(1,\alpha) + S(4,3), S(5,3)\}$$

$$= \min\{\alpha T(4,\alpha) + 1, \alpha T(3,\alpha) + 3, 3\alpha + 7, \alpha + 15, 31\}$$

$$= \min\{3\alpha + 7, \alpha + 15, 31\},$$

where the last equality follows by virtue of Corollary 2.1 and Corollary 2.2.
Now,

$$3\alpha + 7 \ge \alpha + 15 \text{ if and only if } \alpha \ge 4,$$

and

$$S(5,3) = 31 \le \alpha + 15 \text{ if and only if } \alpha \ge 16.$$

All these complete the proof of the lemma.

Corollary 2.3 : $T(5,\alpha) - T(4,\alpha) = \begin{cases} 2\alpha, & \text{if } 2 \le \alpha \le 4 \\ 8, & \text{if } 4 \le \alpha \le 8 \\ \alpha, & \text{if } 8 \le \alpha \le 16 \\ 16, & \text{if } \alpha \ge 16 \end{cases}$

Lemma 2.6 : $T(6,\alpha) = \begin{cases} 3\alpha + 15, & \text{if } 3 \le \alpha \le 8 \\ \alpha + 31, & \text{if } 8 \le \alpha \le 32 \\ 31, & \text{if } \alpha \ge 32 \end{cases}$

Proof : Using Lemma 2.2 – Lemma 2.4, as well as Corollaries 2.1 – 2.3, we have

$$T(6,\alpha) = \min_{1 \le t \le 6} \{\alpha T(6-t,\alpha) + S(t,3)\}$$

$$= \min\{\alpha T(5,\alpha) + S(1,3), \alpha T(4,\alpha) + S(2,3), \alpha T(3,\alpha) + S(3,3), \alpha T(2,\alpha) + S(4,3), \alpha T(1,\alpha) + S(5,3), S(6,3)\}$$

$$= \min\{\alpha T(5,\alpha) + 1, \alpha T(4,\alpha) + 3, \alpha T(3,\alpha) + 7, 3\alpha + 15, \alpha + 31, 63\}$$

$$= \min\{3\alpha + 15, \alpha + 31, 63\}.$$

Now,

$$3\alpha + 15 \ge \alpha + 31 \text{ if and only if } \alpha \ge 8,$$

and

$$\alpha + 31 \ge 63 \text{ if and only if } \alpha \ge 32.$$

Hence the lemma.

Corollary 2.4 : $T(6,\alpha) - T(5,\alpha) = \begin{cases} 8, & \text{if } 3 \le \alpha \le 4 \\ 2\alpha, & \text{if } 4 \le \alpha \le 8 \\ 16, & \text{if } 8 \le \alpha \le 16 \\ \alpha, & \text{if } 16 \le \alpha \le 32 \\ 32, & \text{if } \alpha \ge 32 \end{cases}$

Lemma 2.7 : $T(7,\alpha)=\begin{cases}\alpha(\alpha+3)+15, & \text{if } 2\le\alpha\le 4\\ 3\alpha+31, & \text{if } 4\le\alpha\le 16\\ \alpha+63, & \text{if } 16\le\alpha\le 64\\ 127, & \text{if } \alpha\ge 64\end{cases}$

Proof : Using Lemma 2.2 – Lemma 2.4 and Corollaries 2.1 – 2.3, we have

$$T(7,\alpha)=\min_{1\le t\le 7}\{\alpha T(7-t,\alpha)+S(t,3)\}$$

$$=\min\{\alpha T(6,\alpha)+S(1,3),\ \alpha T(5,\alpha)+S(2,3),\ \alpha T(4,\alpha)+S(3,3),$$
$$\alpha T(3,\alpha)+S(4,3),\ \alpha T(2,\alpha)+S(5,3),\ \alpha T(1,\alpha)+S(6,3),\ S(7,3)\}$$

$$=\min\{\alpha T(6,\alpha)+1,\ \alpha T(5,\alpha)+3,\ \alpha T(4,\alpha)+7,\ \alpha T(3,\alpha)+15,\ 3\alpha+31,\ \alpha+63,\ 127\}$$

$$=\min\{\alpha T(3,\alpha)+15,\ 3\alpha+31,\ \alpha+63,\ 127\}.$$

Now, if $2\le\alpha\le 4$,

$$\alpha(\alpha+3)+15\le 3\alpha+31,$$

and if $\alpha\ge 4$,

$$7\alpha+15\ge 3\alpha+31.$$

Also,

$$3\alpha+31\ge\alpha+63 \text{ if and only if } \alpha\ge 16,$$

and

$$\alpha+63\ge 127 \text{ if and only if } \alpha\ge 64.$$

Thus, we get the desired expression for $T(6,\alpha)$.

Corollary 2.5 : $T(7,\alpha)-T(6,\alpha)=\begin{cases}\alpha^2, & \text{if } 3\le\alpha\le 4\\ 16, & \text{if } 4\le\alpha\le 8\\ 2\alpha, & \text{if } 8\le\alpha\le 16\\ 32, & \text{if } 16\le\alpha\le 32\\ \alpha, & \text{if } 32\le\alpha\le 64\\ 64, & \text{if } \alpha\ge 64\end{cases}$

In Corollaries 2.1 – 2.5, we give the expressions for $T(i+1,\alpha)-T(i,\alpha)$ for $1\le i\le 6$. In determining the values of $T(n+1,\alpha)$, these differences are vital, since

$$T(n+1,\alpha)=\sum_{i=1}^{n+1}[T(i,\alpha)-T(i-1,\alpha)]$$

In the next Section 3, we derive some local-value relationships involving the functions $T(n,\alpha)$, $T(n+1,\alpha)$ and $T(n+2,\alpha)$.

3. SOME LOCAL-VALUE RELATIONSHIPS

We start with the following lemma.

Lemma 3.1 : For any $\alpha\ge 2$ fixed, $T(n,\alpha)$ is strictly increasing in n, that is,

$$T(n+1,\alpha)-T(n,\alpha)>0 \text{ for all } n\ (\ge 1).$$

Proof : Let $T(n+1,\alpha)$ be attained at $t=t_1$, that is, let

$$T(n+1,\alpha) = \alpha T(n+1-t_1, \alpha) + S(t_1, 3).$$

We now consider the following two cases that may result :
Case (1) : When $t_1= n+1$.
Here,

$$T(n+1,\alpha) = S(n+1,3) < \alpha T(1,\alpha) + S(n,3) = \alpha + S(n,3)$$

$$\Rightarrow \alpha > S(n+1,3) - S(n,3) = 2^n. \tag{1}$$

We now want to show that $T(n, \alpha)$ is attained at $t = n$, that is, $T(n,\alpha) = S(n,3)$. The proof is by contradiction. So, let $T(n, \alpha)$ be attained at some $t = t_2$ with $1 \le t_2 < n$. Then,

$$T(n,\alpha) = \alpha T(n-t_2, \alpha) + S(t_2, 3) < S(n,3)$$

$$\Rightarrow \alpha\, T(n-t_2, \alpha) < S(n,3) - S(t_2,3) < 2^n,$$

and we reach to a contradiction by virtue of (1). Therefore, $T(n, \alpha) = S(n,3)$, so that

$$T(n+1,\alpha) - T(n,\alpha) = S(n+1,3) - S(n,3) = 2^n > 0.$$

Case (2) : When $1 \le t_1 < n+1$.

In this case, the proof is by induction on n. Since

$$T(2,\alpha) = S(2,3) > T(1,\alpha) = S(1,3) \text{ for any } \alpha \ge 2,$$

we see that the result is true for n=1. So, we assume that the result is true for some n. Then, we need only prove that the result is true for n+1 as well. Now, since

$$T(n,\alpha) \le \alpha T(n-t_1,\alpha) + S(t_1,3),$$

it follows that

$$T(n+1,\alpha) - T(n,\alpha) \ge \alpha[T(n+1-t_1,\alpha) - T(n-t_1,\alpha)] > 0,$$

where the last inequality follows by virtue of the induction hypothesis.
All these complete the proof of the lemma.

The corollary below follows immediately from Lemma 3.1, noting that

$$T(2,\alpha) - T(1,\alpha) = 2.$$

Corollary 3.1 : For any $\alpha \ge 2$ fixed, $T(n+1,\alpha) - T(n,\alpha) \ge 2$ for all n (≥ 1).

In course of proving Lemma 3.1, we also proved the following results.

Corollary 3.2 : For any $\alpha \ge 2$ fixed, if $T(n+1, \alpha)$ is attained at $t = n+1$, then $T(n, \alpha)$ is attained at $t = n$.

Corollary 3.3 : For any $\alpha \ge 2$ fixed, if $T(n, \alpha)$ is attained at $t = n$, then $\alpha > 2^{n-1}$ ($n \ge 2$).
Lemma 3.2 : For $n \ge 2$ and $\alpha \ge 3$, $T(n, \alpha)$ is not attained at $t = 1$.

Proof : We note that

$$\alpha T(n-1, \alpha) + S(1,3) > \alpha T(n-2, \alpha) + S(2,3)$$

if and only if

$$\alpha[T(n-1, \alpha) - T(n-2, \alpha)] > S(2,3) - S(1,3) = 2,$$

which is true (by Lemma 3.1) for any $n \ge 2$ and any $\alpha \ge 3$.

Lemma 3.3 : For any $\alpha \geq 3$ fixed, let $T(n, \alpha)$ be attained at $k = k_1$ and $T(n+1, \alpha)$ be attained at $k = k_2$. Then, $k_2 \geq k_1$.

Proof : Since (by (1.3)),

$$T(n, \alpha) = \alpha T(k_1, \alpha) + S(n - k_1, 3)$$
$$\leq \alpha T(k_2, \alpha) + S(n - k_2, 3),$$
$$T(n+1, \alpha) = \alpha T(k_2, \alpha) + S(n+1 - k_2, 3)$$
$$\leq \alpha T(k_1, \alpha) + S(n+1 - k_1, 3),$$

we get the following chain of inequalities :

$$2^{n-k_2} = S(n+1 - k_2, 3) - S(n - k_2, 3)$$
$$\leq T(n+1, \alpha) - T(n, \alpha)$$
$$\leq S(n+1 - k_1, 3) - S(n - k_1, 3) = 2^{n-k_1}. \tag{2}$$

Then, we must have $n - k_2 \leq n - k_1$, giving the result desired.

Corollary 3.4 : For any $\alpha \geq 3$ fixed, let $T(n, \alpha)$ be attained at $t = t_1$ and $T(n+1, \alpha)$ be attained at $t = t_2$. Then, $t_2 \leq t_1 + 1$.

Proof : follows immediately from Lemma 3.3, since $t_1 = n - k_1$, $t_2 = n + 1 - k_2$.

Lemma 3.4 : For any $\alpha \geq 3$ fixed,
(a) let $T(n, \alpha)$ be attained at $t = t_1$ and $T(n+1, \alpha)$ be attained at $t = t_2$; then, $t_2 \geq t_1$,
(b) for all $n \geq 1$,

$$T(n+1, \alpha) - T(n, \alpha) \leq T(n+2, \alpha) - T(n+1, \alpha)$$
$$\leq 2[T(n+1, \alpha) - T(n, \alpha)].$$

Proof : To prove part (b), we consider all the three possible cases that may arise.

Case (1) : When $T(n+2, \alpha) = S(n+2, 3)$.

In this case, by Corollary 3.2,

$$T(n+1, \alpha) = S(n+1, 3),\ T(n, \alpha) = S(n, 3),$$

so that

$$T(n+2, \alpha) - T(n+1, \alpha) = 2^{n+1} = 2[T(n+1, \alpha) - T(n, \alpha)].$$

Case (2) : When $T(n+1, \alpha) = S(n+1, 3)$.

Here, by Corollary 3.2, $T(n, \alpha) = S(n, 3)$, so that

$$T(n+1, \alpha) - T(n, \alpha) = 2^n.$$

Now, let $T(n+2, \alpha)$ be attained at $t = t_1$ for some $1 \leq t_1 \leq n+1$, that is, let

$$T(n+2, \alpha) = \alpha T(n+2 - t_1, \alpha) + S(t_1, 3) \leq S(n+2, 3).$$

Then, since

$$T(n+1, \alpha) \leq \alpha T(n+1 - t_1, \alpha) + S(t_1, 3),$$

we get the following chain of inequalities :

$$T(n+2, \alpha) - T(n+1, \alpha) \geq \alpha[T(n+2 - t_1, \alpha) - T(n+1 - t_1, \alpha)]$$
$$> \alpha > 2^n = T(n+1, \alpha) - T(n, \alpha)$$

where the last two inequalities follow by virtue of Corollary 3.1 and Corollary 3.3.
Also,

$$T(n+2,\alpha) - T(n+1,\alpha) \le 2^{n+1} = 2[T(n+1,\alpha) - T(n,\alpha)].$$

Case (3) : When $T(n+2,\alpha) \neq S(n+2,3)$ and $T(n+1,\alpha) \neq S(n+1,3)$.

This case and part (a) of the lemma is proved by induction on n. By Corollary 2.1,

$$T(3,\alpha) - T(2,\alpha) \ge 2 = T(2,\alpha) - T(1,\alpha) \text{ for any } \alpha \ge 3.$$

Thus, part (b) of the lemma holds true for $n = 1$. So, we assume the validity of the result for some n.

Now, let $T(n,\alpha)$ be attained at $t = t_1$ and $T(n+1,\alpha)$ be attained at $t = t_2$ with $t_1 > t_2$. Then,

$$T(n,\alpha) = \alpha T(n - t_1,\alpha) + S(t_1, 3) < \alpha T(n - t_2,\alpha) + S(t_2, 3),$$

and we get the following chain of inequalities :

$$\alpha[T(n+1 - t_2,\alpha) - T(n - t_2, \alpha)]$$
$$< T(n+1,\alpha) - T(n,\alpha) \le \alpha[T(n+1 - t_1,\alpha) - T(n - t_1, \alpha)],$$

which contradicts the induction hypothesis, since $n+1 - t_2 > n+1 - t_1$. Thus, $t_2 \ge t_1$, which we wanted to prove.

To complete the proof of part (b), let $T(n,\alpha)$ be attained at $k = k_1$, $T(n+1,\alpha)$ be attained at $k = k_2$ and $T(n+2,\alpha)$ be attained at $k = k_3$. By part (a) of the lemma and Lemma 3.3, we need to consider the following four cases :

Case (A) : When $k_1 = k_2 = k_3 = K$, say.

In this case,

$$T(n+2, \alpha) - T(n+1, \alpha) = 2^{n-K+1} = 2[T(n+1, \alpha) - T(n, \alpha)].$$

Case (B) : When $k_1 = k_2 = K, k_3 = K + 1$.
Here,

$$T(n+2, \alpha) - T(n+1, \alpha) > 2^{n-K} = T(n+1, \alpha) - T(n, \alpha).$$

Also,

$$T(n+2, \alpha) - T(n+1, \alpha) < 2^{n+1-K} = 2[T(n+1, \alpha) - T(n, \alpha)].$$

Case (C) : When $k_1 = K, k_2 = k_3 = K + 1$.

In this case,

$$T(n+2, \alpha) - T(n+1, \alpha) = 2^{n-K} \ge T(n+1, \alpha) - T(n, \alpha).$$

Again, since

$$T(n+1, \alpha) - T(n, \alpha) \ge 2^{n-K-1},$$

we see that

$$T(n+2, \alpha) - T(n+1, \alpha) \le 2[T(n+1, \alpha) - T(n, \alpha)].$$

Case (D) : When $k_1 = K, k_2 = K + 1, k_3 = K + 2$.

Here,

$$T(n+2, \alpha) - T(n+1, \alpha) = \alpha[T(K+2, \alpha) - T(K+1, \alpha)],$$
$$T(n+1, \alpha) - T(n, \alpha) = \alpha[T(K+1, \alpha) - T(K, \alpha)],$$

so that the result follows by virtue of the induction hypothesis.
All these complete the proof of the lemma.

From part (a) of Lemma 3.4 together with Corollary 3.1, we see that, if $T(n, \alpha)$ isattained at $t = t_1$, and $T(n+1, \alpha)$ is attained at $t = t_2$, then $t_1 \le t_2 \le t_1 + 1$. From the computational point of view, this allows to calculate recursively the value(s) of t where $T(n+1, \alpha)$ is attained, starting with the value(s) of t at which $T(n, \alpha)$ is attained. Part (b) of Lemma 3.4 shows that, for any $\alpha \ge 3$ fixed, $T(n, \alpha)$ is convex in n in the sense of the inequality. It also shows that for any $\alpha \ge 3$ fixed, $T(n+1, \alpha) - T(n, \alpha)$ is increasing (non-decreasing) in n.

Lemma 3.5 : For some $\alpha \ge 3$ and $\alpha \ge 3$, let $T(n, \alpha)$ be attained at the values $k = k_1$ and $k = k_2$. Then, $T(N, \alpha)$ is attained at all $k_1 \le k \le k_2$.

Proof : Let $T(n, \alpha)$ be attained at the values $k = k_1$ and $k = k_2$, so that

$$\alpha[T(k_2, \alpha) - T(k_1, \alpha)] = 2^{n-k_1} - 2^{n-k_2}.$$

There is nothing to prove if $k_2 = k_1 + 1$. So, let $k_2 \ge k_1 + 2$.
The proof is by contradiction. So, we assume that $T(n, \alpha)$ is not attained at $k = k_1 + 1$, so that

$$\alpha[T(k_1+1, \alpha) - T(k_1, \alpha)] > 2^{n-k_1-1}.$$

Let $k_2 = k_1 + m$ for some integer $m \ge 2$. Then, by part (b) of Lemma 3.4, together with the above inequality, we get the following chain of inequalities :

$$\begin{aligned} 2^{n-k_1} - 2^{n-k_2} &= \alpha[T(k_2, \alpha) - T(k_1, \alpha)] \\ &= \alpha[T(k_1+m, \alpha) - T(k_1, \alpha)] \\ &= \alpha \sum_{i=1}^{m} [T(k_1+i, \alpha) - T(k_1+i-1, \alpha)] \\ &\ge m\alpha[T(k_1+1, \alpha) - T(k_1, \alpha)] > m\, 2^{n-k_1-1}, \end{aligned}$$

which leads to a contradiction for $m \ge 2$. Thus, $T(n, \alpha)$ is attained at $k = k_1 + 1$.
Continuing the argument, we get the desired result.

Lemma 3.6 : For any $\alpha \ge 3$ fixed, $T(n, \alpha)$ is not attained at three (consecutive) values.

Proof : If possible, let $T(n, \alpha)$ be attained at the three values $k = K-1, K, K+1$. Then, using part (b) of Lemma 3.4, we get

$$\begin{aligned} 2^{n-K-1} &= \alpha[T(K+1, \alpha) - T(K, \alpha)] \\ &\ge \alpha[T(K, \alpha) - T(K-1, \alpha)] = 2^{n-K}, \end{aligned}$$

which is absurd.

From Lemma 3.5 and Lemma 3.6, we see that, for any $\alpha \ge 3$ fixed, $T(n, \alpha)$ is attained either at a unique k, or else at two (consecutive) values.

Lemma 3.7 : For some $\alpha \ge 3$, let $T(n, \alpha)$ be attained at $k = K, K+1$. Then,

$$\alpha[T(K+1, \alpha) - T(K, \alpha)] = 2^{n-K-1}.$$

Proof : The proof follows immediately from the fact that

$$T(n, \alpha) = \alpha T(K, \alpha) + S(n-K, 3) = \alpha T(K+1, \alpha) + S(n-K-1, 3).$$

Corollary 3.5 : If α is not of the form 2^i, then $T(n, \alpha)$ is attained at a unique k.

Proof : If $T(n, \alpha)$ is attained at two values $k = K, K+1$, then the result in Lemma 3.7 is violated if α is not of the form 2^i.

Lemma 3.8 : For any $\alpha \geq 3$ fixed (with $\alpha \leq n$), let $T(n,\alpha)$ and $T(n+1,\alpha)$ both be attained at $k = K$. Then,

$$T(n+1, \alpha) - T(n, \alpha) = 2^{n-K}. \tag{3}$$

Moreover, in such a case, if α is not of the form 2^i then

- **(a)** $T(n-1, \alpha)$ is not attained at $k = K$.
- **(b)** $T(n+2, \alpha)$ is not attained at $k = K$.

Proof : (3) follows from (2) (with $k_1 = k_2 = K$).

We prove parts (a) and (b) assuming that α is not of the form 2^i.

(a) Let $T(n-1, \alpha)$ be attained at $k = K$, so that

$$T(n-1, \alpha) = \alpha T(K, \alpha) + S(n-K-1, 3)$$
$$\leq \alpha T(K-1, \alpha) + S(n-K, 3).$$

Now, since

$$T(n+1, \alpha) = \alpha T(K, \alpha) + S(n+1-K, 3)$$
$$< \alpha T(K+1, \alpha) + S(n-K, 3),$$

we have the following chain of inequalities :

$$\alpha[T(K+1,\alpha) - T(K,\alpha)] > 2^{n-K}$$
$$\geq 2\alpha[T(K, \alpha) - T(K-1, \alpha)],$$

which contradicts part (b) of Lemma 3.4.

In this case, $T(n-1, \alpha)$ is attained at the (unique) point $k = K-1$, with

$$2^{n-K-1} < T(n,\alpha) - T(n-1,\alpha)$$
$$= \alpha[T(K,\alpha) - T(K-1,\alpha)] < 2^{n-K}.$$

(b) Let $T(n+2,\alpha)$ be attained at $k = K$. Then,

$$T(n+2,\alpha) = \alpha T(K, \alpha) + S(n+2-K, 3)$$
$$\leq \alpha T(K+1, \alpha) + S(n+1-K, 3).$$

Now, since

$$T(n,\alpha) = \alpha T(K, \alpha) + S(n-K, 3)$$
$$< \alpha T(K-1, \alpha) + S(n+1-K, 3),$$

we get

$$\alpha[T(K+1,\alpha) - T(K,\alpha)] \geq 2^{n+1-K}$$
$$> 2\alpha[T(K, \alpha) - T(K-1, \alpha)],$$

contradicting part (b) of Lemma 3.4.

Thus, $T(n+2, \alpha)$ is attained at the (unique) point $k = K+1$, satisfying the following chain of relations.

$$2^{n-K} < T(n+2,\alpha) - T(n+1,\alpha)$$
$$= \alpha[T(K+1,\alpha) - T(K,\alpha)] < 2^{n+1-K}.$$

All these complete the proof of the lemma.

Lemma 3.9 : Let, for some $\alpha \geq 3$, $T(n, \alpha)$ be attained at $k = K, K+1$. Then,

(a) $T(n-1, \alpha)$ is attained at $k = K$,

(b) $T(n+1, \alpha)$ is attained at $k = K+1$,

(c) $T(n, \alpha) - T(n-1, \alpha) = 2^{n-K-1} = T(n+1, \alpha) - T(n, \alpha)$.

Proof : Let $T(n, \alpha)$ be attained at $k = K, K+1$. Then,

$$T(n,\alpha) = \alpha T(K, \alpha) + S(n-K, 3)$$
$$= \alpha T(K+1, \alpha) + S(n-K-1, 3)$$
$$< \alpha T(K-1, \alpha) + S(n-K+1, 3).$$

(a) Let $T(n-1, \alpha)$ be attained not at $k = K$. Then, it must be attained at $k = K-1$. Thus,

$$T(n-1, \alpha) = \alpha T(K-1, \alpha) + S(n-K, 3)$$
$$< \alpha T(K, \alpha) + S(n-K-1, 3).$$

But, then

$$\alpha[T(K+1, \alpha) - T(K, \alpha)] = 2^{n-K-1}$$
$$< \alpha[T(K, \alpha) - T(K-1, \alpha)],$$

which contradicts part (b) of Lemma 3.4.

Thus, $T(n-1, \alpha)$ is attained at the (unique) point $k = K$.

(b) If, on the contrary, $T(n+1,\alpha)$ is attained at $k = K+2$, then

$$T(n+1, \alpha) = \alpha T(K+2, \alpha) + S(n-K-1, 3)$$
$$< \alpha T(K+1, \alpha) + S(n-K, 3).$$

Therefore,

$$\alpha[T(K+2, \alpha) - T(K+1,\alpha)] < 2^{n-K-1}$$
$$= \alpha[T(K+1, \alpha) - T(K, \alpha)],$$

and we are led to a contradiction to part (b) of Lemma 3.4.

Hence, $T(n+1, \alpha)$ is attained at the (unique) point $k = K+1$.

(c) follows from the proofs of parts (a) and (b).

For any $\alpha \geq 3$ and $n \geq 1$ fixed, let

$$T_\alpha(n, k) = \alpha T(k, \alpha) + S(n-k, 3) \text{ for } 0 \leq k \leq n-1.$$

Then,

$$T(n, \alpha) = \min_{0 \leq k \leq n-1} \{ T_\alpha(n, k) \}.$$

Lemma 3.10 : $T_\alpha(k)$ is convex in k in the sense that

$$T_\alpha(n, k+2) - T_\alpha(n, k+1) \geq T_\alpha(n, k+1) - T_\alpha(n, k) \text{ for all } 0 \leq k \leq n-2.$$

Proof : Since

$$T_\alpha(n, k+2) - T_\alpha(n, k+1)$$
$$= \alpha[T(k+2, \alpha) - T(k+1, \alpha)] - S(n-k-2, 3),$$

$$T_\alpha(n, k+1) - T_\alpha(n, k)$$
$$= \alpha[T(k+1, \alpha) - T(k, \alpha)] - S(n-k-1, 3),$$

we get

$$[T_\alpha(n, k+2) - T_\alpha(n, k+1)] - [T_\alpha(n, k+1) - T_\alpha(n, k)]$$
$$= \alpha[\{T_\alpha(n, k+2) - T_\alpha(n, k+1)\} - \{T_\alpha(n, k+1) - T_\alpha(n, k)\}]$$
$$+2^{n-k-2}.$$

The result now follows by virtue of part (b) Lemma 3.4.

4. DISCUSSION

When $\alpha = 2$, it can be proved that (see, for example, Majumdar [3])

$$T(n+1, 2) - T(n, 2) = 2^s \text{ for some integer } s \geq 1.$$

However, for $\alpha \geq 3$, such a relationship need not hold, as can be verified from the entries of Table 1, which gives the values of $T(n, \alpha)$ for $1 \leq n \leq 9$ and $3 \leq \alpha \leq 9$.

Table 1. Values of $T(n, \alpha)$ for $1 \leq n \leq 9$ and $3 \leq \alpha \leq 9$. In each cell, the number in parathesis gives the value(s) of k at which $T(n, \alpha)$ (in the formulation of (2.3)) is attained.

α \ n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
3	1 (0)	3 (0)	6 (1)	10 (1)	16 (2)	24 (2)	33 (3)	45 (4)	61 (4)
4	1 (0)	3 (0)	**7** (0, 1)	11 (1)	**19** (1, 2)	27 (2)	**43** (2, 3)	**59** (3, 4)	75 (4)
5	1 (0)	3 (0)	7 (0)	12 (1)	20 (1)	30 (2)	46 (2)	66 (3)	91 (4)
6	1 (0)	3 (0)	7 (0)	13 (1)	21 (1)	33 (2)	49 (2)	73 (3)	105 (3)
7	1 (0)	3 (0)	7 (0)	14 (1)	22 (1)	36 (2)	52 (2)	80 (3)	112 (3)
8	1 (0)	3 (0)	7 (0)	**15** (0, 1)	23 (1)	**39** (1, 2)	55 (2)	**87** (2, 3)	119 (3)
9	1 (0)	3 (0)	7 (0)	15 (0)	24 (1)	40 (1)	58 (2)	90 (2)	126 (3)

For $\alpha \geq 3$, it is an interesting problem to find the value(s) of n such that

$$T(n+1, \alpha) - T(n, \alpha) = 2^s \text{ for some integer } s \geq 1.$$

In this connection, we have the following result :

Lemma 4.1 : Let α be of the form 2^i (for some integer $i \geq 1$). Then, for all $n \geq 1$, the difference $T(n+1,\alpha) - T(n,\alpha)$ is of the form 2^s.

Proof : The proof is by induction on n. Corollary 2.1 shows that the result is true for $n = 2$. So, we assume the validity of the result for some n (so that the result is true for all i with $2 \leq i \leq n$). We have to show the validity of the result for $n + 1$.

To do so, let $T(n, \alpha)$ be attained at $k = K$. Then, one of the following two cases arises :

Case 1 : When $T(n + 1, \alpha)$ is attained at $k = K$.

In this case,

$$T(n+1,\alpha) - T(n, \alpha) = 2^{n-K}.$$

Case 2 : When $T(n + 1, \alpha)$ is attained at $k = K + 1$.

In this case,

$$T(n+1,\alpha) - T(n, \alpha) = \alpha[T(K + 1, \alpha) - T(K, \alpha)].$$

Then, by virtue of the induction hypothesis, $T(n+1, \alpha) - T(n, \alpha)$ is of the form 2^s for some integer $s \geq 1$.

If α be of the form 2^i, $T(n+1,\alpha) - T(n,\alpha)$ is of the form 2^s in the trivial case when $T(n + 1, \alpha) = S(n + 1, \alpha)$ (see Corollary 3.2). Thus,

$$T(2,\alpha) - T(1, \alpha) = 2 \text{ for all } \alpha \geq 3,$$

$$T(3,\alpha) - T(2, \alpha) = 2^2 \text{ for all } \alpha \geq 4.$$

Lemma 4.2 : Let α (≥ 3) be an integer, not of the form 2^i. Then $T(n+1,\alpha) - T(n,\alpha)$ is of the form 2^s (for some integer $s \geq 1$) if and only if $T(n+1, \alpha)$ and $T(n, \alpha)$ both are attained at the same $k = K$.

Proof : The "if" part of the lemma follows from Lemma 3.8. To prove the "only if" part, let $T(n, \alpha)$ be attained at $k = K$ and $T(n + 1, \alpha)$ be attained at $k = K+1$. Then,

$$T(n,\alpha) = \alpha T(K, \alpha) + S(n - K, 3)$$
$$< \alpha T(K + 1, \alpha) + S(n - K - 1, 3),$$
$$T(n+1,\alpha) = \alpha T(K + 1, \alpha) + S(n - K, 3)$$
$$< \alpha T(K, \alpha) + S(n + 1 - K, 3).$$

Then, we get the following chain of inequalities :

$$2^{n-K-1} < T(n+1, \alpha) - T(n, \alpha) < 2^{n-K}.$$

The above inequality shows that $T(n+1, \alpha) - T(n, \alpha)$ can not be of the form 2^s (for some integer $s \geq 1$).
Let α (≥ 3) be not of the form 2^i. Let for any α fixed, $T(n,\alpha)$ and $T(n + 1, \alpha)$ be such that $T(n+1,\alpha) - T(n, \alpha)$ is of the form 2^s (for some integer $s \geq 1$). From Lemma 3.8, coupled with Lemma 4.2, we see that, $T(n+2,\alpha) - T(n+1,\alpha)$ is not of the form 2^s. This raises the following question : Is there any integer N with $N > n$, such that $T(N+1,\alpha) - T(N,\alpha)$ is of the form 2^t (for some integer $t > s$)? The following lemma answers the question in the affirmative.

Lemma 4.3 : Let $\alpha = 2^i$ for some integer i (≥ 1). Let $T(n, \alpha)$ be attained at the two values $k = N, N + 1$. Let m be such that

$$m - n = n - N + i - 1. \quad (4)$$

Then, $T(m,\alpha)$ is attained at the two values $k = n-1, n$.

Proof : Let $T(n,\alpha)$ be attained at the two values $k = N, N+1$, so that by Lemma 3.9,

$$T(n,\alpha) - T(n-1,\alpha) = 2^{n-N-1} = T(n+1,\alpha) - T(n,\alpha).$$

Then,

$$\alpha[T(n,\alpha) - T(n-1,\alpha)] = 2^{m-n},$$

so that

$$\alpha T(n-1,\alpha) + S(m-n+1,3) = \alpha T(n,\alpha) + S(m-n,3).$$

Also, since

$$\alpha[T(n+1,\alpha) - T(n,\alpha)] = 2^{m-n} > 2^{m-n-1},$$

it follows that

$$\alpha T(n+1,\alpha) + S(m-n-1,3) > \alpha T(n,\alpha) + S(m-n,3).$$

Hence,

$$\begin{aligned} T(m,\alpha) &= \alpha T(n-1,\alpha) + S(m-n+1,3) \\ &= \alpha T(n,\alpha) + S(m-n,3) \\ &< \alpha T(n+1,\alpha) + S(m-n-1,3), \end{aligned}$$

which shows that $T(m,\alpha)$ is attained at the two values $k = n-1, n$.

Lemma 4.3 shows that the sequence $\{T(n,\alpha)\}_{n=1}^{\infty}$ contains an infinite number of functions of the form $T(m,\alpha)$, each of which is attained at exactly two values of k.

When $\alpha = 2$, it can be shown (as in Majumdar [3]) that exactly one of the following two alternatives holds true :

Case (1) $T(n+2,\alpha) - T(n+1,\alpha) = T(n+1,\alpha) - T(n,\alpha)$, (5)

Case (2) $T(n+2,\alpha) - T(n+1,\alpha) = 2[T(n+1,\alpha) - T(n,\alpha)]$. (6)

When $\alpha (\geq 3)$ is of the form 2^i $(i \geq 2)$ $T(n+2,\alpha)$, $T(n+1,\alpha)$ and $T(n,\alpha)$ still satisfy one of the above two relationships. We now look at this problem more closely. Let $T(n,\alpha)$ be attained at two values $k = N$, $N+1$. By Corollary 3.5, α must be of the form 2^s (for some integer $s \geq 1$). From the proof of Lemma 3.9, we see that the necessary conditions (that $T(n,\alpha)$ is attained at $k = N$, $N+1$) are

1. $T(n,\alpha) - T(n-1,\alpha) = 2^{n-N-1}$,
2. $T(n+1,\alpha) - T(n,\alpha) = 2^{n-N-1}$,
3. $\alpha[T(N+1,\alpha) - T(N,\alpha) = 2^{n-N-1}$.

Now, we consider the function $T(n+2,\alpha)$. One of the following two cases arises :

Case 1 : When $T(n+2,\alpha)$ is attained at $k = N+1$.
In this case,

$$T(n+2,\alpha) - T(n+1,\alpha) = 2^{n-N},$$

so that

$$T(n+2,\alpha) - T(n+1,\alpha) = 2[T(n+1,\alpha) - T(n,\alpha)].$$

Case 2 : When $T(n+2,\alpha)$ is attained at $k = N+2$.
Here,

$$\begin{aligned} T(n+2,\alpha) &= \alpha T(N+2,\alpha) + S(n-N,3) \\ &< \alpha T(N+1,\alpha) + S(n-N+1,3), \end{aligned}$$

and hence,

$$\alpha[T(N+2,\alpha) - T(N+1,\alpha)] > 2^{n-N}. \tag{7}$$

Then, $T(n+1,\alpha)$ is attained at $k = N+2$; otherwise,

$$T(n+1,\alpha) = \alpha T(N+1,\alpha) + S(n-N,3)$$
$$< \alpha T(N+2,\alpha) + S(n-N-1,3),$$

giving

$$\alpha[T(N+2,\alpha) - T(N+1,\alpha)] > 2^{n-N-1},$$

which, together with Lemma 4.1, contradicts (7).

Hence, $T(n+1,\alpha)$ is attained at $k = N+2$, and consequently,

$$T(n+2,\alpha) - T(n+1,\alpha) = 2^{n-N-1}.$$

When $\alpha(\geq 3)$ is not of the form 2^i $(i \geq 2)$, then $T(n+2,\alpha)$, $T(n+1,\alpha)$ and $T(n,\alpha)$ need not satisfy either of the two relationships (5) and (6). For example,

$$4 = T(4,3) - T(3,3) < 6 = T(5,3) - T(4,3) < 8 = 2[T(4,3) - T(3,3)].$$

We conjecture that, in such a case, neither of the relationships (4) and (5) holds true.

REFERENCES

1. Matsuura, A. Exact analysis of the recurrence relations generalized from the Towers of Hanoi. *SIAM Proceedings in Applied Mathematics* 129, 228 – 233 (2008).
2. Dudeney, H.E. *The Canterbury Puzzles.* Thomas Nelson and Son, London (1907) (4th ed. of 1919 published by Dover in 1958).
3. Majumdar, A.A.K. *The Classical Tower of Hanoi Problem and Its Generalizations,Vol. 1: Multi-Peg Generalization*. Lambert Academic Publishing, USA (2012).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 203–211 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

The Study of Accuracy and Efficiency of ODE Solvers While Performing Numerical Simulations of Terrestrial Planets

A. Amin, S. Rehman, and A. Pervaiz*

Department of Mathematics, University of Engineering and Technology,
Lahore, Pakistan

Abstract: The N-body problem is one of the well-known and most central computational problem. The N-body problem of the Solar System is not only a rich source of initial value problems (IVPs) for ordinary differential equations (ODEs), but is also very convenient to understand the orbital evolution of the Solar System; see, for example,[1, 2]. Wide range of numerical integrators have been developed and implemented for performing such N-body simulations. The main objective of this research paper is to analyze and compare the accuracy and efficiency of different ordinary differential equation (ODE) solvers applied to the Kepler's two-body problem for Terrestrial planets. Throughout this paper, the error growth is investigated in terms of global error in position and velocity, and the relative error in terms of angular momentum and total energy of the system. To quantify the quality of different ODE solvers, we performed numerical tests applied to the Kepler's two body problem for Terrestrial planets with local error tolerances ranging from 10^{-12} to 10^{-4}.

Keywords: Kepler's two body problem, N-body simulations, Terrestrial planets, ODE solvers

1. INTRODUCTION

In the dynamical astronomy and celestial mechanics, the numerical integration plays a vital role to investigate complex problems, for example, N-body problems. Computational astronomers make extensive use of accurate N-body simulations to investigate the orbital evolution of the planets, comets, asteroids, and other small celestial bodies in the Solar System. The numerical simulations of N-body problems are performed by first obtaining a set of second order ODEs for the acceleration of the N-bodies, and describing the positions and velocities of the N-bodies at the initial time $= t_0$. The initial value problem (IVP) we are considering is represented by:

$$y''(t) = f\big(t, y(t)\big),\ y(t_0) = y_0,\ y'(t_0) = y_0', \tag{1}$$

where, $y_0, y_0' \in R^k$ represent the initial positions and velocities, operator $'$ represents differentiation with respect to time, $f : R \times R^k \rightarrow R^k$ a smooth function, and k is the dimension of the above IVP. In the N-body problem, when N = 2 then the problem is in its simplest form, i.e., two-body problem. Large number of numerical integrators are developed and used to find numerical approximations of these types of problems; see, for example, Runge-Kutta [3, 4], ODE solvers for non-stiff problems [5, 6, 7].

2. MATERIALS AND METHODS

In celestial mechanics, the Kepler's two-body problem [10, 11] involves the motion of one body about another under the influence of their mutual gravitational attraction. The two-body problem is considered as the simplest problem in the dynamical astronomy, because the exact solution of the two-body problem

Received, April 2016; Accepted, May 2016
*Corresponding author: A. Pervaiz; Email: anjumpervaiz1@hotmail.com

exists. The best direct orbit calculations arise when the central body is largely heavier than the orbiting body, for example, the man-made satellite around the Earth and planetary orbits around the Sun. In this research work, we used Kepler's two-body problem for Terrestrial planets as the test problem. Terrestrial planets (inner planets), i.e., Mercury, Venus, Earth, and Mars are closest to the Sun [12, 13, 14]. The equations of motion of Kepler's two body problem are:

$$y_1'' = -\frac{y_1}{r^3}, \tag{2}$$

$$y_2'' = -\frac{y_2}{r^3}, \tag{3}$$

where, y_1,y_2 represent the $x-$ and $y-$components of one body corresponding to the other body, and $r = \sqrt{y_1^2 + y_2^2}$. The initial conditions are

$$y_1(0) = 1 - e, y_2(0) = 0, \quad y_1'(0) = 0, y_2'(0) = \sqrt{\frac{1+e}{1-e}}.$$

The parameter e is the orbital eccentricity and $0 \leq e < 1$. Since, the true solution of Kepler problem is available. Therefore, the Kepler's two body problem is quite useful for observing the accuracy of different ODE solvers over a short duration of time. The analytical solution of the Kepler's two body problem is given by

$$y_1 = \cos(\eta) - e, \qquad y_2 = \sqrt{1-e^2}\sin(\eta),$$

$$y_1' = -\sin(\eta)\left(1 - ecos(\eta)\right)^{-1}, \qquad y_2' = \sqrt{(1-e^2)}\cos(\eta)(1 - ecos(\eta))^{-1},$$

where, the eccentric anomaly η satisfies Kepler's equation $x = \eta - \text{e}\sin(\eta)$.

Throughout this paper, we have discussed different types of errors. To quantify the quality of numerical approximations obtained by different ODE solvers, the global error is of main concern. The main source of error, for the total error in the system, is the integration error. The integration error consists of two types of errors, namely, truncation and round-off error. Since, computer stores number to a certain arithematic precision. Therefore, for accurate N-body simulations, round-off error can contribute significantly to the global error.

Suppose, $y_{num}(t)$ and $y_{true}(t)$ are position vectors of the solutions obtained numerically and analytically, respectively, and $y_{num}'(t)$ and $y_{true}'(t)$ are velocity vectors of numerical and analytical solutions, respectively. The L_2 -norm of global errors in the position and velocity are defined as

$E_r(t) = \|y_{num}(t) - y_{true}(t)\|_2$, $E_v(t) = \|y_{num}'(t) - y_{true}'(t)\|_2$, where, $\|.\|_2$ is the L_2 -norm.

Physical systems usually have conserved quantities, for example, angular momentum $L(t)$ and total energy $H(t)$. Generally, $L(t)$ and $H(t)$ are not conserved accurately by numerical approximations. However, this digression provides evaluation to quantify the quality of numerical approximations. The total energy $H(t)$ of the two-body system is defined as

$$H(t) = \frac{y_1'^2 + y_2'^2}{2} - \frac{1}{\sqrt{y_1^2 + y_2^2}}.$$

The relative error in energy is defined as

$$H_{rel}(t) = \left|\frac{H_0 - H(t)}{H_0}\right|,$$

where, H_0 is the total energy at $t = t_0$.

The total angular momentum $L(t)$ is defined as

$$L(t) = y_1 y_2' - y_2 y_1'.$$

The relative error in angular momentum is defined as

$$L_{rel}(t) = \frac{\|L_0 - L(t)\|_2}{\|L_0\|_2},$$

where, L_0 is the angular momentum at $t = t_0$. Notice that, unlike the global error in position and velocity, the exact solution is not required to calculate $H_{rel}(t)$ and $L_{rel}(t)$. Hence, fewer computing resources are required to observe the quality of different ODE solvers. Since, $H_{rel}$and L_{rel} are scalar quantities. So, in order to obtain small error, $H_{rel}$and L_{rel} require only one constraint. Whereas, for E_r and E_v, being vector quantities, each coordinate of E_r and E_v has to be small.

3. ODE SOLVERS

A wide range of numerical integration techniques have been developed for the numerical approximations of ODEs that correspond to the continuous state of dynamic systems. MATLAB is a software that is used to solve an extensive variety of such problems. The MATLAB ODE suit is a set of codes for solving first order systems of ODEs for IVPs and plotting mathematical results of these problems [5]. The ODE solvers control the estimated local error for initial value problems. A local error tolerance is specified and, if the estimated local error is too large comparative to this specific tolerance. Then the time-step is rejected and a new attempt is made with smaller time-step. All ODE solvers in MATLAB use the same function interface, so it is very easy to try several methods on the same problem and observe their behavior [8]. The most important non-stiff solvers are ODE23, ODE45, and ODE113 [5].

3.1 ODE45 Solver

The ODE45 solver is a popular (4, 5) embedded pair of Dormand and Prince [7]. The ODE45 solver consists of six-stage embedded pair of Runge-Kutta methods of order 4 and 5. The ODE45 solver is a very attractive one step solver for the numerical approximations of non-stiff problems. The ODE45 advances the solution with 5th order method and the local error is controlled by taking the difference between the numerical approximations obtained by 5th order and 4th order methods. In order to compute $y(t_n)$, ODE45 solver requires only the solution $y(t_{n-1})$ at the immediately preceding time-step. Generally, the ODE45 solver is the best solver to implement as a "first choice" for most of the problems.

3.2 ODE23 Solver

The ODE23 solver is based upon explicit Runge-Kutta (2, 3) embedded pair of Bogacki and Shampine [6]. In the presence of mild stiffness and at crude tolerances, the ODE23 solver may be more efficient than ODE45 solver. The ODE23 solver is a one step solver which is frequently used for non-stiff problems. The ODE23 solver consists of four-stage embedded pair of explicit Runge-Kutta methods of order 2 and 3. The ODE23 advances the solution with 3rd order method and the local error is controlled by taking the difference between the numerical approximations obtained by 3rd order and 2nd order methods.

3.3 ODE113 Solver

The ODE113 solver is a variable order and variable time-step solver. The ODE113 solver uses Adams-Bashforth-Moulton predictor-corrector methods of order 1 to 13 [9]. When the ODE function is very expensive to evaluate then at stringent tolerances, the ODE113 solver may be more efficient than ODE45 solver. Normally, the ODE113 solver requires solution at several preceding time-steps to obtain the current solution values [5].

4. RESULTS AND DISCUSSION

In this section, we investigate the global error in position and velocity, and the relative error in angular momentum and energy for Terrestrial planets. We perform numerical experiments for different numerical ODE solvers applied to the Kepler'stwo body problem for Terrestrial planets over the orbital time period of each of the planet with $TOL = [10^{-12} - 10^{-4}]$. The eccentricity of Mercury, Venus, Earth, and Mars is approximately 0.21, 0.0068, 0.017, and 0.093, respectively. Whereas, for the Terrestrial planets, the corresponding time of the orbital motion to complete one vibration is approximately 28π, 72π, 117π, and 219π, respectively. Different time-steps are chosen to perform these numerical experiments. We evaluate the position and velocity on each of the time-step using ODE solvers. The values of positions, velocities, and time are stored in separate files. Then we obtain error in positions and velocities with respect to the true solutions obtained at the saved values of time.

Fig. 1, shows four sets of experiments with time-steps π/2, π/4, π/8, and π/16 to obtain the maximum global error in position using ODE23, ODE45, and ODE113 integrators applied to the Sun-Mercury system with $TOL = [10^{-12} - 10^{-4}]$. Fig. 1(a) shows four sets of numerical experiments with time step $\pi/2$. From Fig. 1(a), we observe that ODE45 gives the least error, which is approximately 1.9929×10^{-9} for the Sun-Mercury system at a combination of tolerance 10^{-12} and time-step $\pi/2$. The ODE23 integrator gives the 2nd least error, which is approximately $3.6778 \times 10^{-9}$and the 3rd least error, which is approximately 5.8392×10^{-9} is attained by ODE113 integrator at the same combination of tolerance and time-step. Fig. 1(a) depicts a clear pattern. When the tolerance is increased, the maximum global error in position is also increased.

Fig. 1(b) shows the same set of experiments with time-step $\pi/4$. We see that at tolerances less than 10^{-10} all three integrators lose their accuracy by approximately two orders of magnitude as compared to the global errors at time-step $\pi/2$. When the time-step is $\pi/8$ and tolerance 10^{-12}, as shown in Fig. 1(c), all the three integrators have achieved almost the same accuracy, which is approximately 0.0272. Fig. 1(d) shows the experiments performed with time-step $\pi/16$. We observed that all three integrators lose their accuracy at tolerances less than 10^{-6} and give us straight lines which shows that there is no change in error when we further reduce the tolerance.

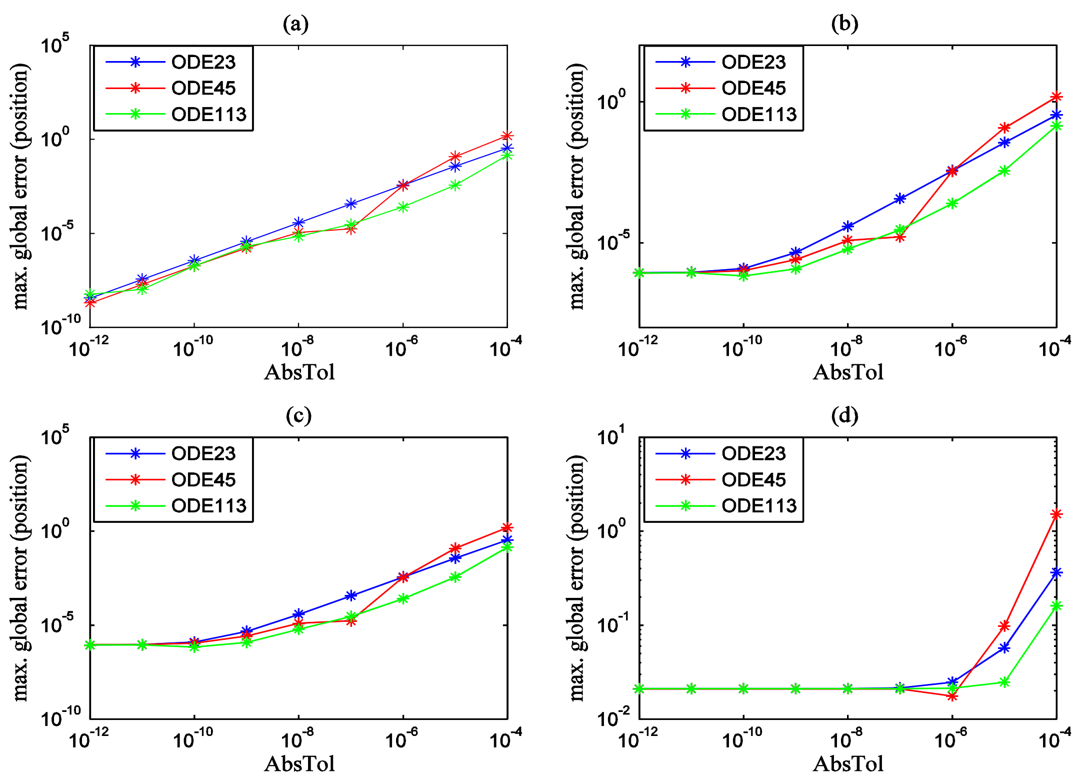


Fig. 1. The maximum global error in position using ODE23, ODE45, and ODE113 for the Sun-Mercury system against tolerances ranging from 10^{-12} to 10^{-4} and using time-steps from π/16 to π/2.

From all sets of experiments in Fig. 1, we observe that the accuracy of the given ODE solvers is improved if the time-step is increased at tolerance 10^{-12}. We conclude that a combination of tolerance $10^{-12}$and time-step π/2 gives better results in terms of maximum global error in position for all three integrators. We also observe that using ODE45 solver the least maximum global error in position is approximately 1.993×10^{-9}, which is the best observed accuracy. The ODE45 integrator achieved approximately 45% and 65% better accuracy than ODE23 and ODE113, respectively.

Fig. 2 shows four sets of experiments with time-steps π/2, π/4, π/8, and π/16 to obtain the maximum global error in position using ODE23, ODE45, and ODE113 integrators applied to the Sun-Venus system with $TOL = [10^{-12} - 10^{-4}]$. Fig. 2(a) shows four sets of numerical experiments with time step $\pi/2$. From Fig. 2(a), we observe that ODE113 gives the least error, which is approximately 5.0325×10^{-10} for the Sun-Venus system at a combination of tolerance 10^{-12} and time-step $\pi/2$. The ODE45 integrator gives the 2nd least error, which is approximately $1.0321\times 10^{-8}$and the 3rd least error, which is approximately 2.2642×10^{-8} is attained by ODE23 integrator at the same combination of tolerance and time-step. The ODE113 integrator achieved approximately 95% and 97% better accuracy than ODE45 and ODE23, respectively. Fig. 2(b) and Fig. 2(c), show the same sets of experiments with time-steps $\pi/4$ and $\pi/8$, respectively. We observed that at time-steps $\pi/4$ and $\pi/8$, the behavior of the global error obtained by using three integrators was very similar to that for the experiments performed with time-step $\pi/2$.

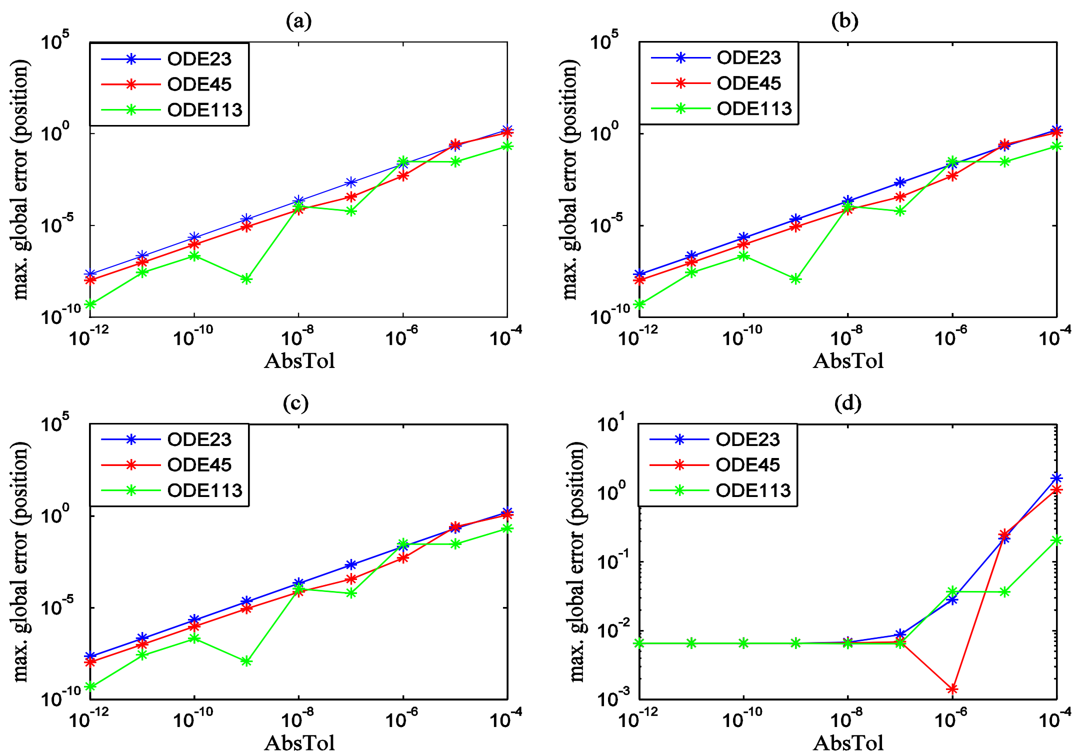


Fig. 2. The maximum global error in position using ODE23, ODE45, and ODE113 for the Sun-Venus system against tolerances ranging from 10^{-12} to 10^{-4} and using time-steps from π/16 to π/2.

Fig. 2(d) shows the experiments performed with time-step $\pi/16$. We observed that all three integrators lose their accuracy using tolerances less than 10^{-7} and give us a straight line which shows that there is no change in the error when we further reduce the tolerance. We observed from Fig. 2 that all three ODE solvers behave in a similar manner at time-steps $\pi/2$, $\pi/4$, and $\pi/8$ for the Sun-Venus system. However, for the efficiency reason, the time-step $\pi/2$ is recommended, because all the three integrators take least amount of CPU time with time-step $\pi/2$.

Fig. 3 shows four sets of experiments with time-steps π/2, π/4, π/8, and π/16 to obtain the maximum global error in position using ODE23, ODE45, and ODE113 integrators applied to the Sun-Earth system with $TOL = [10^{-12} - 10^{-4}]$. Fig. 3(a) shows four sets of numerical experiments with time step $\pi/2$. From Fig. 3(a) we observe that ODE23 gives the least error, which is approximately $6.3404\times 10^{-6}$for the

Sun-Earth system at a combination of tolerance 10^{-12} and time-step $\pi/2$. The ODE45 integrator gives the 2nd least error, which is approximately 6.3714×10^{-6} and the 3rd least error, which is approximately 6.3974×10^{-6} is attained by ODE113 integrator at the same combination of tolerance and time-step. The ODE23 integrator achieved approximately 0.49% and 0.89% better accuracy than ODE45 and ODE113, respectively.

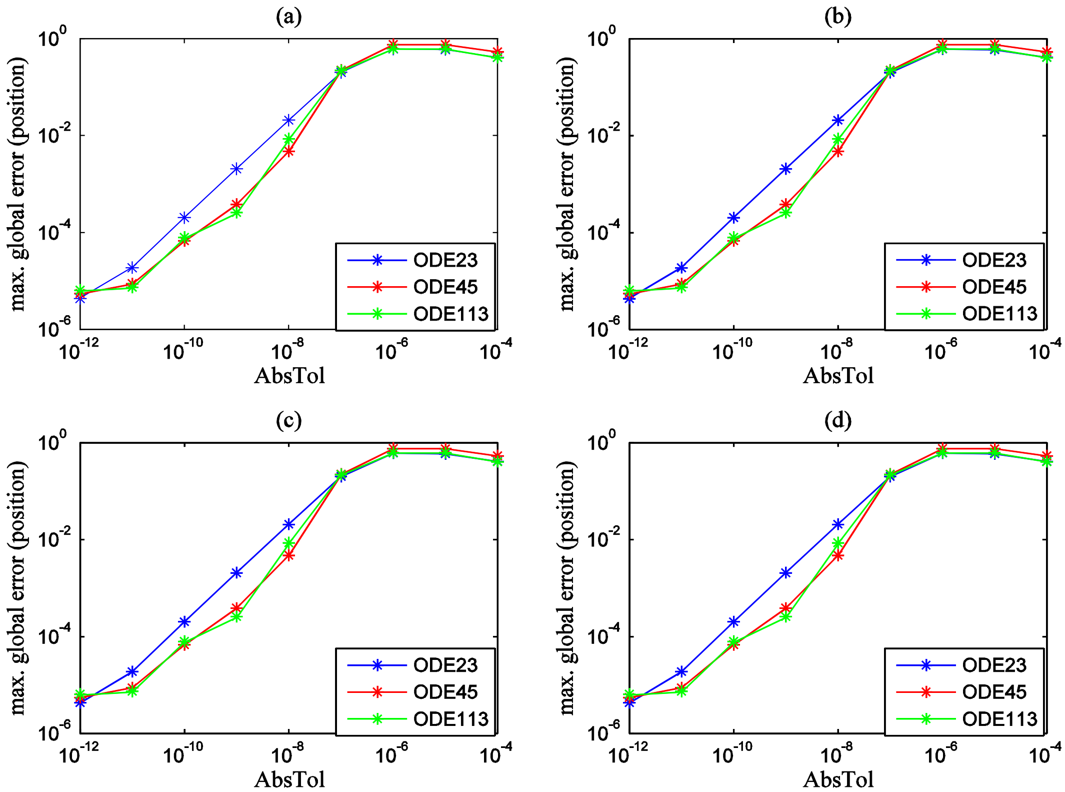


Fig. 3. The maximum global error in position using ODE23, ODE45, and ODE113 for the Sun-Earth system against tolerances ranging from 10^{-12} to 10^{-4} and using time-steps from π/16 to π/2.

Fig. 4 shows four sets of experiments with time-steps π/2, π/4, π/8, and π/16 to obtain the maximum global error in position using ODE23, ODE45, and ODE113 integrators applied to the Sun-Mars system with $TOL = [10^{-12} - 10^{-4}]$. Fig. 4(a) shows four sets of numerical experiments with time step $\pi/2$. From Fig. 4(a), we observe that ODE113 gives the least error, which is approximately 9.7285×10^{-8} for the Sun-Mars system at a combination of tolerance 10^{-12} and time-step $\pi/2$. The ODE45 integrator gives the 2nd least error, which is approximately 1.7289×10^{-7} and the 3rd least error, which is approximately 2.6425×10^{-7} attained by ODE23 integrator at the same combination of tolerance and time-step. The ODE113 integrator achieved approximately 43% and 63% better accuracy than ODE45 and ODE23, respectively. When the time-step is $\pi/4$, as shown in Fig. 4(b), the behavior of the errors obtained by using three integrators is very much similar to that for the experiments performed with time-step $\pi/2$. Fig. 4(c) and Fig. 4(d) show that the experiments were performed with time-step $\pi/8$ and $\pi/16$, respectively. We observe that given integrators lose their accuracy using tolerances less than 10^{-7} and give us a straight line which shows that there is no change in error when we further reduce the tolerance. Furthermore, we have performed the previous sets of numerical experiments to investigate the error growth in velocity. We observed almost the same trend as of the errors in position but with certain orders of magnitude difference in accuracy.

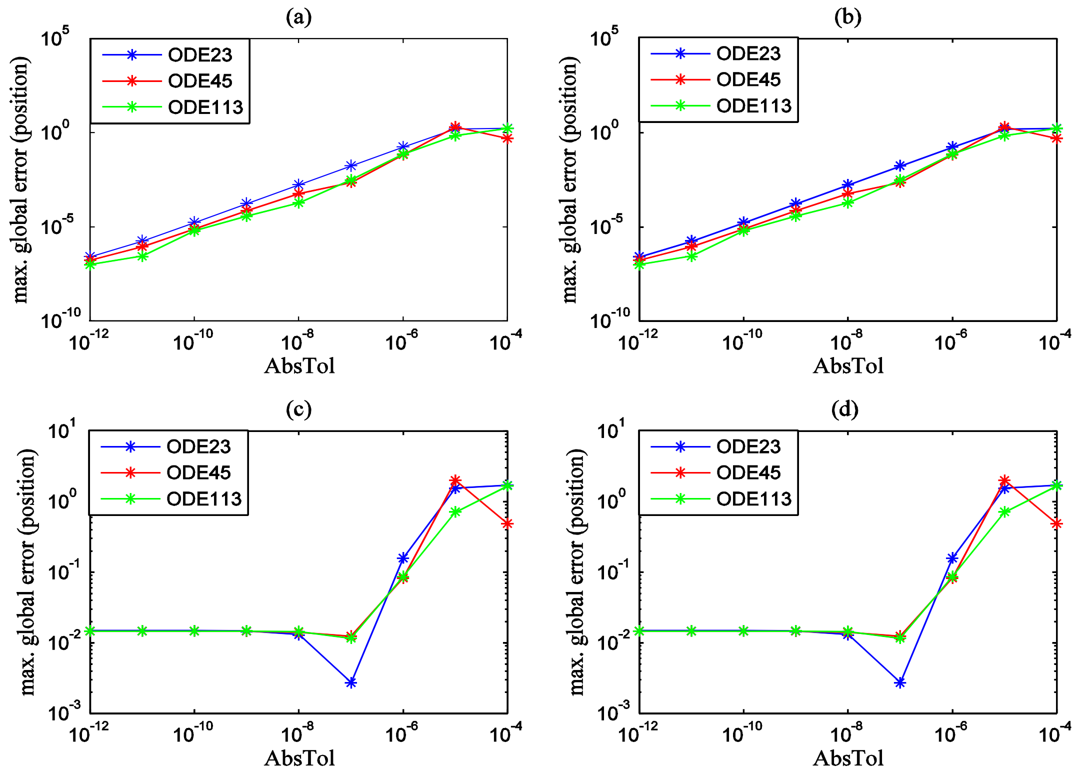


Fig. 4. The maximum global error in position using ODE23, ODE45, and ODE113 for the Sun-Mars system against tolerances ranging from 10^{-12} to 10^{-4} and using time-steps from $\pi/16$ to $\pi/2$.

Let us now consider the accuracy of the integrators ODE23, ODE45, and ODE113 using the combination of tolerance 10^{-12} and time-step $\pi/2$ over the orbital time period of each Terrestrial planet by estimating the relative error in energy and angular momentum.

Fig. 5 shows four sets of experiments with time-step of $\pi/2$ to observe the error behavior in total energy using three integrators ODE23, ODE45 and ODE113 applied to the Sun-Mercury, Sun-Venus, Sun-Earth, and Sun-Mars system over the orbital time period of each Terrestrial planet. The tolerance and time step is selected to give the smallest maximum global error. From Fig. 5(a), we observe that the best observed accuracy in terms of the relative error in energy is again achieved by the ODE45 integrator for the Sun-Mercury system. From Fig. 5(b), for the Sun-Venus system, we observe that the best observed accuracy in terms of relative error in energy is again achieved by the ODE113 integrator. From Fig. 5(c), we observe that the best observed accuracy in terms of relative error in energy is achieved by ODE113 rather than ODE23 for the Sun-Earth system. From Fig. 5(d), we observe for the Sun-Mars system that the best observed accuracy in terms of relative error in energy is again achieved by ODE113. As for Terretrial planets, we repeated the same sets of experiments and observe a similar behavior for the relative error in angular momentum.

Now we consider the efficiency of the ODE solvers discussed in this paper in terms of the amount of work done to attain a given accuracy. One way of observing the amount of work is to count the number of function evaluations against the maximum global error in position. Table 1 shows the number of function evaluations against the least maximum global error in position for the ODE45, ODE23 and ODE113 integrators with tolerance of 10^{-12} and time-step $\pi/2$.

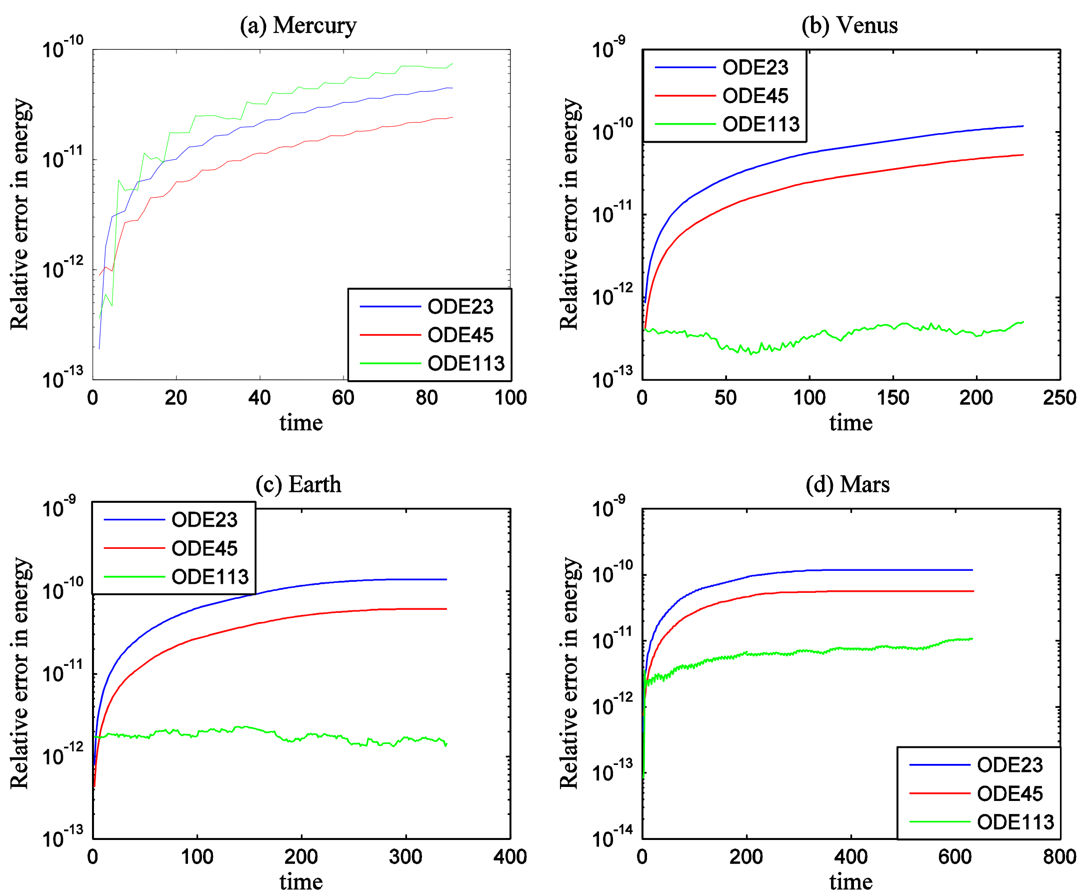


Fig. 5. The relative error in energy using ODE23, ODE45 and ODE113 integrators applied to the Sun-Mercury, Sun-Venus, Sun-Earth, and Sun-Mars system over the orbital time period of each Terrestrial planet.

For the Sun-Mercury system, the integrator ODE23 is approximately 25.4 times more expensive than ODE45 and the integrator ODE45 is approximately 9.13 times more expensive than ODE113. For the Sun-Venus system, the integrator ODE23 is approximately 22.5 times more expensive than ODE45 and the integrator ODE45 is approximately 14.2 times more expensive than ODE113 integrator. For the Sun-Earth system, the integrator ODE23 is approximately 22.6 times more expensive than ODE45 and the integrator ODE45 is approximately 13.9 times more expensive than ODE113. Whereas, the integrator ODE23 is approximately 23.6 times more expensive than ODE45 and the integrator ODE45 is approximately 11.4 times more expensive than ODE113 for the Sun-Mars system. Overall, we observe that the ODE113 integrator takes the least number of function evaluations for each Terrestrial planet and the ODE23 is the most expensive ODE solver, because it takes the most number of function evaluations.

Table 1. Number of function evaluations against least maximum global error in position for ODE23, ODE45 and ODE113 integrators at $h = \pi/2$ and tolerance $= 10^{-12}$over the orbital time period of each Terrestrial planet.

Solvers	**Mercury**	**Venus**	**Earth**	**Mars**
ODE23	995641	2258569	3688531	7180777
ODE45	39223	100375	163087	304303
ODE113	3966	7061	11679	26735

5. CONCLUSIONS

The main purpose of this paper was to observe the accuracy and efficiency of different ODE solvers applied to the real world problems involving the Sun and the Terrestrial planets. The simulations were performed over one orbital period of each of Terrestrial planet. For these simulations, we performed numerical experiments using three ODE integrators ODE23, ODE45, and ODE113 for the $TOL = [10^{-12} - 10^{-4}]$. We observed that when the tolerance is increased, the maximum global error in position is increased. We also observed that these ODE solvers are more accurate when the time-step is large at tolerance 10^{-12}. For the given range of tolerances from 10^{-12} to 10^{-4}, and time-steps from $\pi/16$ to $\pi/2$, we observed that for the Sun-Mercury system the integrator ODE45 achieves the best observed accuracy. The ODE45 integrator has achieved approximately 45% and 65% better accuracy than ODE23 and ODE113, respectively. For the Sun-Venus system, we observed that the integrator ODE113 achieves the best observed accuracy. The ODE113 integrator achieved approximately 95% and 97% better accuracy than ODE45 and ODE23, respectively. We observed that the integrator ODE23 achieves the best observed accuracy for the Sun-Earth system. The ODE23 integrator achieved approximately 0.49% and 0.89% better accuracy than ODE45 and ODE113, respectively. Finally we observed the results for the Sun-Mars system that the integrator ODE113 achieves the best observed accuracy. The ODE113 integrator achieved approximately 43% and 63% better accuracy than ODE45 and ODE23, respectively.

We also analyzed the efficiency of the ODE solvers discussed in this paper by counting the number of function evaluations against the least maximum global error in position for Terrestrial planets. We observed that the best observed accuracy attained by the integrator ODE45 for the Sun-Mercury system uses approximately 9.13 times more function evaluations than ODE113 and approximately 25.4 times less function evaluations than ODE23. The best observed accuracy attained by the integrator ODE113 for the Sun-Venus system uses approximately 14.2 times less function evaluations than ODE45 and approximately 319.9 times less function evaluations than ODE23. For the Sun-Earth system, the best observed accuracy attained by the integrator ODE23 uses approximately 22.6 times more function evaluations than ODE45 and approximately 315.8 times more function evaluations than ODE113. For the Sun-Mars system, the best accuracy attained by the integrator ODE113 uses approximately 11.5 times less function evaluations than ODE45 and approximately 268.6 times less function evaluations than ODE23.

6. REFERENCES

1. Grazier, K.R., W.I. Newman, W.M. Kaula & J.M. Hyman. Dynamical evolution of planetesimals in outer solar system. *Icarus* 140(2): 34-352 (1999).
2. Sharp, P.W. N-Body Simulations: The Performance of some integrators. *ACM Transactions on Mathematical Software* 32(3): 375-395 (2006).
3. Heun, K. Neue Methode zur approximativen integration der differentialgleichungen einer unabhängigen veränderlichen. *Mathematical Physics* 45: 23-38 (1900).
4. Butcher, J.C. *Numerical Methods for Ordinary Differential Equations.* John Wiley & Sons, England (2008).
5. Ashino, R., M. Nagase & R.Vaillancourt. Behind and beyond the MATLAB ODE suite. *Computers and Mathematics with Applications* 40: 491-512 (2000).
6. Bogacki, P. & L.F. Shampine. A 3(2) pair of Runge-Kutta formulas. *Applied Matematics Letters* 2(4): 1-9 (1989).
7. Dormand, J.R. & P.J. Prince. A family of embedded Rune-Kutta formulae. *Journal of Computational and Applied Mathematics* 6(1): 19-26 (1980).
8. Shampine, L.F., I. Gladwell & S. Thompson. *Solving ODEs with MATLAB.* Cambridge University Press, USA (2003).
9. Shampine, L.F. & M.K. Gordon. *Computer Solution of Ordinary Differential Equations: The Initial Value Problem*. W.H. Freeman, San Francisco, USA (1975).
10. Danka, D.L. & M. Varga. Simulation of the two-body problem in geogebra. *Acta Electrotechnica et Informatica* 12(3): 47–50 (2012).
11. Goldstein, H. *Classical Mechanics,* 3rd ed. Addison Wesley, USA (1980).
12. Hunten, D. Atmospheric evolution of the terrestial planets. *Science, New Series* 259(5097): 915-920 (1993).
13. Wieczorek, M.A. The gravity and topography of the terrestial planets. *Treatise on Geophysics* 5: 165-206 (2006).
14. Garlick, M.A. & D. Rothery. *The story of the Solar System.* Cambridge University Press, USA (2012).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 213–224 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Large Time Step Scheme Behaviour with Different Entropy Fix

Ihtram ul Haq[1]*, Mukkarum Hussain[2], Muhammad Jawed Iqbal[1], and Noor Fatima Siddiqui[3]

[1] Institute of Space and Planetary Astrophysics, University of Karachi, Karachi, Pakistan
[2] Institute of Space Technology, Karachi, Pakistan
[3] Department of Mathematics, University of Karachi, Karachi, Pakistan

Abstract: The progress of numerical techniques for scalar and one dimensional Euler equation has been a great interest of researchers in the field of CFD for decades. In 1986, Harten developed a high resolution and efficient large time step (LTS) explicit scheme for scalar problems. Computation of nonlinear wave equation depicts that Harten's LTS scheme is a high resolution and efficient scheme. However, computations of hyperbolic conservation laws show some spurious oscillations in the vicinities of discontinuities for larger values of CFL. Zhan Sen Qian investigated this issue and suggested to perform the inverse characteristic transformations by using the local right eigenvector matrix at each cell interface location to overcome these spurious oscillations. Harten and Qian both used Roe's approximate Riemann solver which has less artificial viscosity than exact method at sonic points. The reduced artificial viscosity reduces the accuracy of Roe's method at sonic points. Roe's approximate Riemann solver cannot capture the finite spread of expansion fans due to the inadequate artificial viscosity at expansive sonic points. As a consequence of this expansion shocks that are nonphysical may occur. The existence of the expansion shock is said to violate the entropy condition. A variety of entropy fix formulae for Roe scheme have been addressed in the literature. In present work large time step total variation diminishing (LTS TVD) scheme developed by Harten and improved by Qian have been tested with different entropy fix and its effect has been investigated. Computed results are analyzed for merits and shortcomings of different entropy fix with large time step schemes.

Keywords: CFL restriction, explicit scheme, inverse shock, Shock tube problem, SOD, TVD scheme, 1D Euler equation

1. INTRODUCTION

The transient 1D Euler equation is hyperbolic, no matter whether the flow is locally subsonic or supersonic. The marching direction for 1D Euler equation is the time direction. Methods to solve hyperbolic system of equations are primarily derived for non-linear wave equation and then implemented on hyperbolic system of equations. Lax in 1954, modified Euler's Forward Time Central Space (FTCS) method and presented first-order accurate method to solve nonlinear wave equation. Lax method is stable for Courant-Friedrichs-Lewy (CFL) condition less than 1 and predicts the location of moving discontinuity correctly [1-2]. This method is very dissipative and smears discontinuities over several mesh points and become worse as CFL decreases. Lax-Wendroff proposed a second-order accurate method for non-linear wave equation. His method sharply defined discontinuity and also stable for CFL less than 1 but produce undesirable oscillations when discontinuities are encountered. Similar to Lax method quality of results computed by Lax-Wendroff method degrade as CFL decrease.

Lax and Lax-Wendroff central finite difference schemes are stable and converge if flow field is sufficiently smooth but produce unwanted oscillations when discontinuities are met. It is due to the fact that series expansion for obtaining a difference approximation is only valid for continuous functions and

Received, February 2016; Accepted, May 2016
*Corresponding author: Ihtram ul Haq; Email: ihtram2010@hotmail.com

has continuous derivatives at least through the order of difference approximation [1] [3]. Godunov recognized this deficiency and proposed a finite volume scheme instead of a finite difference scheme to avoid the need of differentiability. He used exact Riemann problem solution for evaluating the flux term at the cell interface. Computation of nonlinear wave equation is easily accomplished by using Godunov method but this method is very inefficient and take long time when applied to system of equations [4-6]. To overcome this problem Roe suggested solving linear problem instead of actual nonlinear problem [7]. Roe's approximate Riemann solver is efficient but cannot distinguish between expansion shock and compression shock. This is due to the violation of entropy condition and hence expansion shocks that are nonphysical may occur in computed results [8-9]. A number of entropy fix have been recommended in literature to overcome this problem. Roe's upwind approximate Riemann solver capture physics in more appropriate way than Lax and Lax-Wendroff central schemes but is only first order accurate. Like second order central methods, higher order upwind methods have the same deficiencies and produce undesirable oscillations when discontinuities are encountered [9-10].

Harten introduced the concept of Total Variation Diminishing (TVD) scheme. TVD schemes are monotonicity preserving schemes and therefore it must not create local extrema and the value of an existing local minimum must be non-decreasing and that of a local maximum must be non-increasing [10-12]. He worked on non-oscillatory first order accurate scheme and modified its flux function to obtain a second order accurate TVD explicit difference schemes for scalar and system of hyperbolic conservation laws. Numerical dissipation terms in TVD methods are nonlinear. The quantity varies from one grid point to another and usually consists of automatic feedback mechanisms to control the amount of numerical dissipation. After this break through a number of TVD scheme have been proposed and discussed in literature [13-17].

Stability criteria for explicit formulation limits time stepping and thus increase computational cost. Similar to previously discussed schemes, explicit formulation of Harten and other TVD schemes are also stable only for CFL less than 1. It is a challenging task to develop an explicit scheme which is stable for higher values of CFL number. In literature this kind of schemes are known as large time step (LTS) schemes and an active field of research for last three decades. Leveque described a method for approximating nonlinear interactions linearly which allows Godunov's method to be applied with arbitrarily large time steps [18-19]. Harten extended Leveque work and proposed second-order accurate LTS TVD explicit schemes for the computation of hyperbolic conservation laws. Computation of nonlinear wave equation depicts that Harten's LTS scheme is a high resolution and efficient scheme [20]. However, computation of system of hyperbolic conservation laws show some spurious oscillations in the vicinities of discontinuities when CFL > 1. Zhan Sen Qian worked on Harten LTS TVD scheme and observed that these spurious oscillations are due to the numerical formulation of the characteristic transformation used by Harten for extending the method for hyperbolic conservation laws [21-23]. Zhan Sen Qian showed that if the inverse characteristic transformations are performed by using the local right eigenvector matrix at each cell interface location then these spurious oscillations are eliminated. His computations for shock tube problem confirm that the modified large time step total variation diminishing (MLTS TVD) scheme eliminate spurious oscillations for system of hyperbolic conservation laws without increasing the entropy fixing parameter.

Harten and Qian both used Roe's approximate Riemann solver which has less artificial viscosity than exact method at sonic points. Roe's approximate Riemann solver differs from Godunov's exact Riemann solver only at sonic points. Roe's method has less artificial viscosity than Godunov's method at sonic points. The reduced artificial viscosity reduces the accuracy of Roe's method at sonic points. Roe's approximate Riemann solver cannot capture the finite spread of expansion fans due to the inadequate artificial viscosity at expansive sonic points. As a consequence of this is that expansion shocks that are nonphysical may occur [7-8]. This nonphysical behavior is due to the fact that the scheme cannot distinguish between an expansion shock and a compression shock. Each is a valid solution for this formulation. The existence of the expansion shock is said to violate the entropy condition. In order to make this scheme to satisfy the entropy condition it must be properly modified, such a correction is usually designated as entropy fix. A variety of entropy fix formulae for the Roe scheme have been addressed in the literature. Most famous are due to Harten-Hyman and Hoffmann-Chiang [9-11].

In present work large time step scheme [22] behavior with different entropy fix has been investigated. Computed results are analyzed for merits and shortcomings of different entropy fix with large time step schemes. Shock tube problem is used for validation purpose. Reasons of attraction in this test case are availability of analytical solution and at the same time presence of complex flow features namely, expansion, shock wave, and contact discontinuities.

2. NUMERICAL METHOD

1D Euler equation in conservation form is used in present study. Detail about governing equations and numerical method is given below:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = 0 \tag{1}$$

$$\frac{\partial U}{\partial t} + A\frac{\partial U}{\partial x} = 0 \tag{2}$$

where; $U = \begin{bmatrix} \rho \\ \rho u \\ \rho E \end{bmatrix}$; $F = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ (\rho E + p)u \end{bmatrix}$ (3)

$$A = \frac{\partial F}{\partial U} = \begin{bmatrix} 0 & 1 & 0 \\ (\gamma-3)\frac{u^2}{2} & (3-\gamma) & (\gamma-1) \\ (\gamma-1)u^3 - \gamma u E & -\frac{3}{2}(\gamma-1)u^2 + \gamma E & \gamma u \end{bmatrix} \tag{4}$$

equation (1) in numerical flux form can be written as:

$$U_i^{n+1} = U_i^n - \lambda\left(f^n_{i+\frac{1}{2}} - f^n_{i-\frac{1}{2}}\right) \;; \quad \lambda = \frac{\Delta t}{\Delta x} \tag{5}$$

The numerical flux for Qian's modified LTS TVD is given by:

$$f_{i+\frac{1}{2}} = \frac{1}{2}\left[F_{i+1} + F_i\right] + \left[\frac{1}{2\lambda}\sum_{k=1}^{m} R^k_{i+\frac{1}{2}}\left(g^k_{i+1} + g^k_i\right) - \frac{1}{\lambda}\sum_{l=-K+1}^{K-1}\left\{\sum_{k=1}^{m} R^k_{i+l+\frac{1}{2}}\, C_l\left(v^k + \beta^k\right)_{i+l+\frac{1}{2}}\, \alpha^k_{i+l+\frac{1}{2}}\right\}\right] \tag{6}$$

where; $v^k_{i+\frac{1}{2}} = \lambda\, a^k_{i+\frac{1}{2}}$; $a = u,\ u+c,\ u-c$ (7)

$$\beta^k_{i+\frac{1}{2}} = \begin{cases} \dfrac{\left(g^k_{i+1} - g^k_i\right)}{\alpha^k_{i+\frac{1}{2}}}, & \text{for } \alpha^k_{i+\frac{1}{2}} \neq 0 \\ 0, & \text{for } \alpha^k_{i+\frac{1}{2}} = 0 \end{cases} \tag{8}$$

$$\alpha^k_{i+\frac{1}{2}} = R^{-1}\Delta_{i+\frac{1}{2}}U \tag{9}$$

$$\sigma^k_{i+\frac{1}{2}} = \frac{K}{2}\left[\psi\left(\frac{v^k_{i+\frac{1}{2}}}{K}\right)\left\{1 + \frac{K-1}{2}\psi\left(\frac{v^k_{i+\frac{1}{2}}}{K}\right)\right\} - \frac{K+1}{2}\left(\frac{v^k_{i+\frac{1}{2}}}{K}\right)^2\right] \tag{10}$$

$$C_{\pm k}(v) = \begin{cases} c_k\left(\mu_{\mp}(v)\right), & \text{for } 1 \le k \le K-1 \\ \frac{K}{2}\psi\left(\frac{v}{K}\right), & \text{for } k = 0 \end{cases} \tag{11}$$

$$\mu_{\pm}(v) = \frac{1}{2}\left[\psi\left(\frac{v}{K}\right) \pm \frac{v}{K}\right] \tag{12}$$

$$R = \begin{bmatrix} 1 & 1 & 1 \\ u & u+c & u-c \\ \frac{u^2}{2} & \frac{u^2}{2} + uc + \frac{c^2}{(\gamma-1)} & \frac{u^2}{2} - uc + \frac{c^2}{(\gamma-1)} \end{bmatrix} \tag{13}$$

$$R^{-1} = \begin{bmatrix} 1 - \frac{(\gamma-1)u^2}{2c^2} & (\gamma-1)\frac{u}{c^2} & -\frac{(\gamma-1)}{c^2} \\ -\frac{u}{2c} + \frac{(\gamma-1)u^2}{4c^2} & \frac{1}{2c} - \frac{(\gamma-1)u}{2c^2} & \frac{(\gamma-1)}{2c^2} \\ \frac{u}{2c} + \frac{(\gamma-1)u^2}{4c^2} & -\frac{1}{2c} - \frac{(\gamma-1)u}{2c^2} & \frac{(\gamma-1)}{2c^2} \end{bmatrix} \tag{14}$$

$$\tilde{g}^k_{i+\frac{1}{2}} = \sigma^k_{i+\frac{1}{2}} \alpha^k_{i+\frac{1}{2}} \tag{15}$$

$$g^k_i = \text{minmod}\left(\tilde{g}^k_{i+\frac{1}{2}}, \tilde{g}^k_{i-\frac{1}{2}}\right) \tag{16}$$

Mathematical expressions of coefficient of numerical viscosity proposed by Roe, Lax-Wendroff and Harten are given in equations 17 to 20 respectively.

$$\psi(v) = |v| \tag{17}$$

$$\psi(v) = v^2 \tag{18}$$

$$\psi(v) = \begin{cases} \frac{1}{2}\left(\frac{v^2}{\varepsilon} + \varepsilon\right) & |v| < \varepsilon \\ |v| & |v| \geq \varepsilon \end{cases} \tag{19}$$

$$\psi(v) = v^2 + \frac{1}{4} \tag{20}$$

Table 1 shows the expressions for coefficient functions $C_l(x)$ which were obtained after putting different values of K in equation 11.

Table 1. $C_l(x)$ at different K.

K	C_1	C_2	C_3
2	x^2		
3	$x^2(3-x)$	x^3	
4	$x^2(6-4x+x^2)$	$2x^3(2-x)$	x^4

3. TEST CASE DESCRIPTION

Shock tube problem is used for validation purpose. SOD and Inverse Shock boundary conditions are used (Table 2). The size of computational domain is $0 \leq x \leq 1$ and number of grid points are 1000. Entropy fix parameter ε is taken 0.1 while minmod limiter is used for all computations. Initial discontinuity centered on $x = x_o$ and $t = 0$ has following conditions:

$$U(x,t) = \begin{cases} U_L, & x < x_o \\ U_R, & x \geq x_o \end{cases} \qquad \text{where, } x_o=0.5$$

Table 2. Boundary conditions for SOD and Inverse Shock cases.

Boundary conditions	p_R	ρ_R	v_R	p_L	ρ_L	v_L
SOD	0.1	0.125	0.0	1.0	1.0	0.0
Inverse Shock	1.0	1.0	5.91608	29.0	5.0	1.183206

4. RESULTS AND DISCUSSION

In present work, Qian MLTS TVD scheme presented in equation (6) is studied with different entropy fix functions. Shock tube problem for SOD and Inverse Shock boundary conditions are solved to investigate the performance and behavior of MLTS TVD scheme with different entropy fix in the regions of expansion fan, discontinuities and strong shock waves.

The physical time for the flow processes are 0.15 and 0.05sec for SOD and Inverse Shock cases respectively. Simulations are carried out on Intel(R) Core(TM) 2, CPU @ 2.13 GHz, 2 GB RAM.

Fig. 1 to Fig. 5 show the density profile for SOD boundary conditions. Fig. 1 shows whole domain while Fig. 2 to Fig. 5 focus on start of expansion, end of expansion, contact, and shock regions respectively. Fig. 2 depicts that for lower values of CFL (0.9 and 1.9) the result with different entropy fix at the start of expansion fan is similar while for larger values of CFL (2.9 and 3.9) the result with entropy fix presented in equation 20 produces less dissipation. At the end of expansion fan minor oscillation is observed for equation 18 and equation 20, these oscillations grow up as K value increase, see Fig. 3(d). Results with entropy fix presented in equation 17 and 19 are more accurate as compare to other at the end of expansion fan location.

Investigation of Fig. 4 and Fig. 5 shows that the resolution for the contact discontinuity is captured worse than the shock wave for different entropy fix. The shock wave is a compressive wave and the characteristic lines are convergent so the dissipation near to the shock wave is controlled in a small level in the time marching steps. Near the contact wave the discrepancy of analytical and numerical results are worse for all entropy fix. Reason behind this is the fact that the contact discontinuity is a kind of linear wave in the theory of the hyperbolic conservation laws and the characteristic lines are parallel to each other so the dissipation can not be restrained. During the time marching step the dissipation is accumulated and the contact wave may span more and more grid points. Fig. 4 depicts that for lower values of CFL (0.9 and 1.9) the result with different entropy fix at contact is similar. For larger values of CFL (2.9 and 3.9) the result with entropy fix presented in equation 17 and equation 19 produces less dissipation and is non oscillatory while other two entropy fix produce oscillations.

Fig. 5 shows that for lower values of CFL (0.9 and 1.9) the result with different entropy fix at shock is similar. For larger values of CFL (2.9 and 3.9) the result with entropy fix presented in equation 18 and equation 20 produces less dissipation. Results near shock are oscillation free for all entropy fix used in present study. Fig. 6 to Fig. 10 show the density profile for Inverse Shock boundary conditions. Fig. 6 shows whole domain while Fig. 7 to Fig. 10 focus on start of expansion, end of expansion, contact, and shock regions, respectively. Fig. 6 depicts that for equations 17, 18 and 20 schemes does not have sufficient artificial viscosity for all CFL values and therefore nonphysical expansion shock is available in the result. This nonphysical behavior is due to the fact that the scheme with equations 17, 18 and 20 entropy fix cannot distinguish between an expansion shock and a compression shock. Both are valid result for this formulation.

Although scheme with entropy fix presented in equation 20 provide good result for smaller values of CFL however for larger CFL values (CFL = 3.9) this scheme become unstable and after some iteration it's blown up. Same is happened with scheme using entropy fix presented in equation 18. Schemes with entropy fix presented in equation 17 and equation 19 are found to be stable for all CFL values investigated in present studies.

It has been explored from Fig. 6 to Fig. 10 that scheme with entropy fix presented in equation 17 does not have sufficient artificial viscosity and produce nonphysical expansion shock for all CFL values investigated in present studies. Therefore scheme with entropy fix presented in equation 19 only produces real results without nonphysical expansion shock and stability issues.

It is summarized that for simple cases all entropy fix provide physical results with minor oscillation and dissipation issues for all CFL values investigated in present studies. However for complex test cases, like inverse shock boundary condition, only Harten's derived entropy fix (equation 19) is stable and depict physical behavior precisely for LTS TVD.

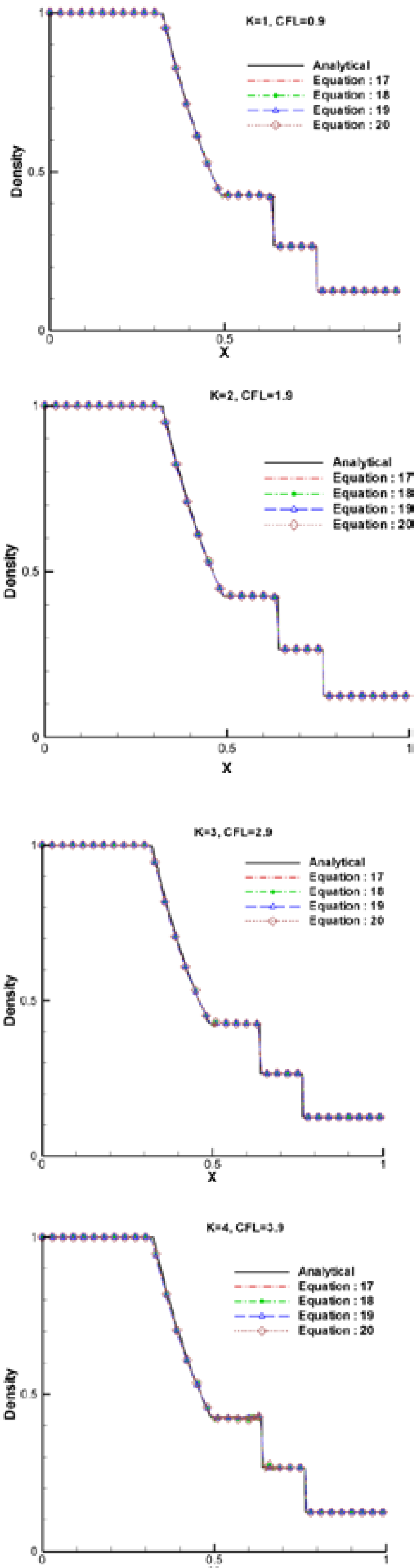


Fig. 1. Overall density profile, SOD case.

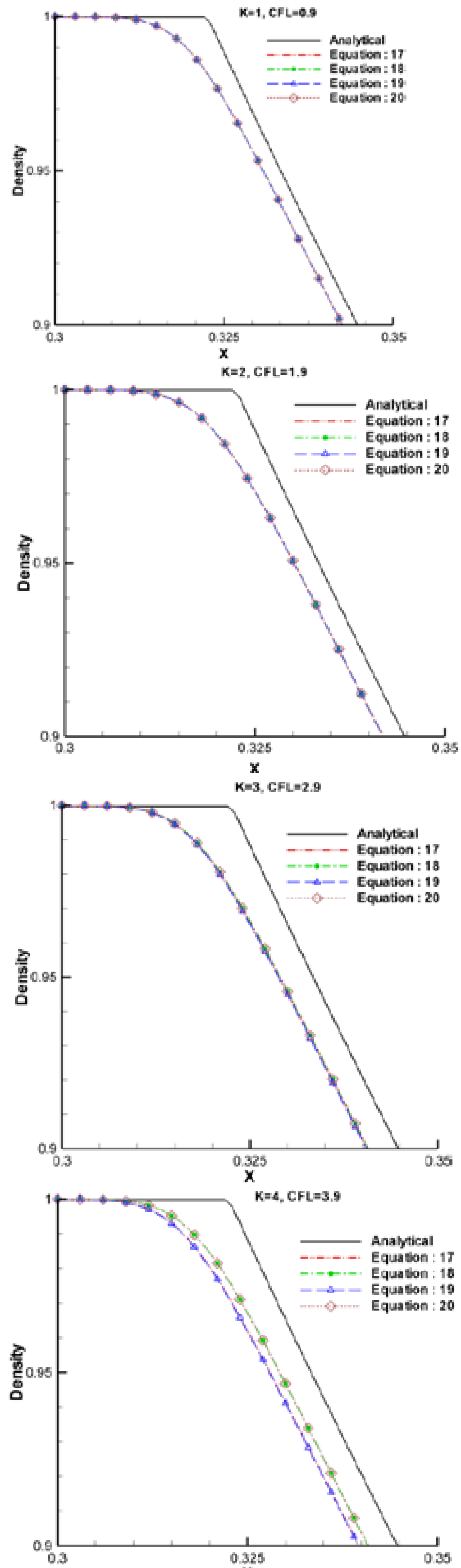


Fig. 2. Start of expansion, SOD case.

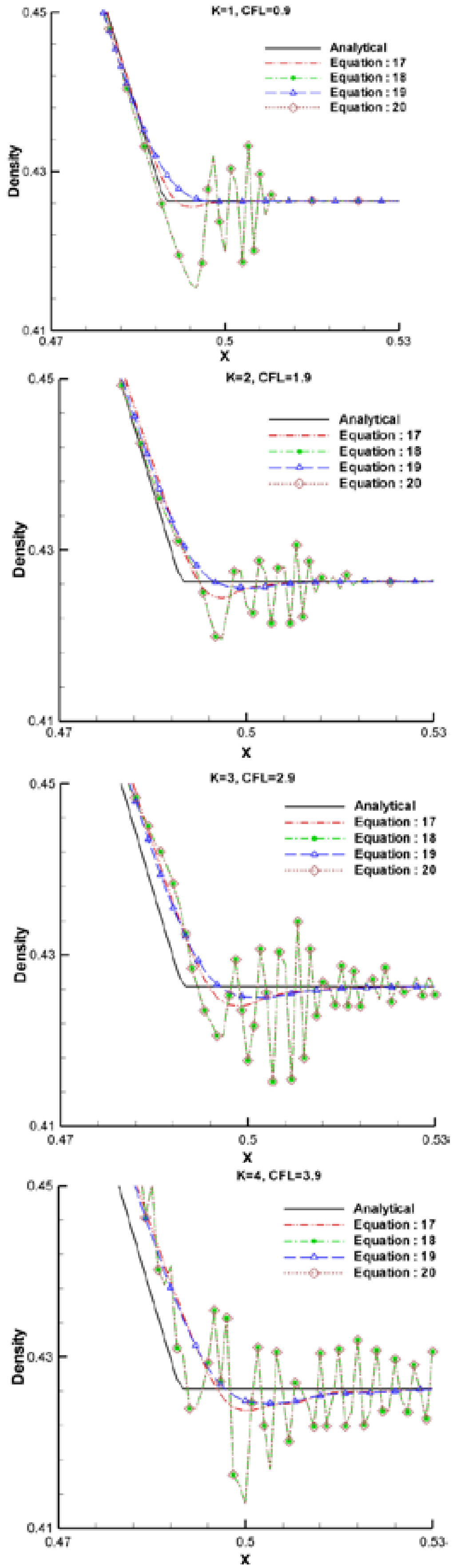


Fig. 3. End of expansion, SOD case.

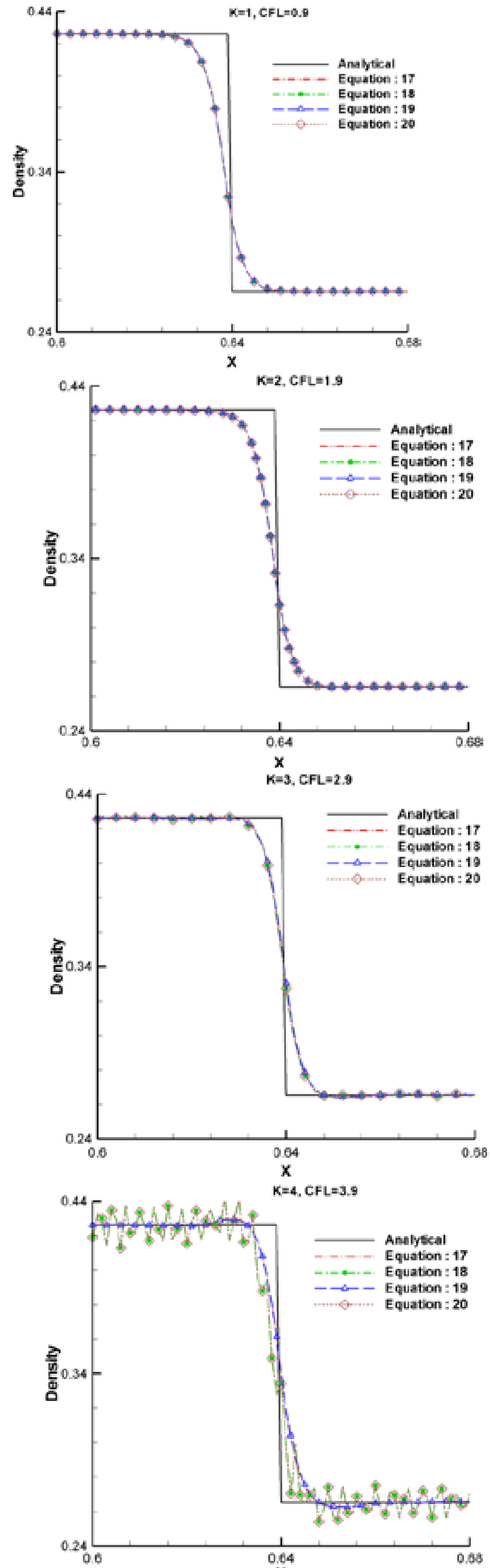


Fig. 4. Contact region, SOD case.

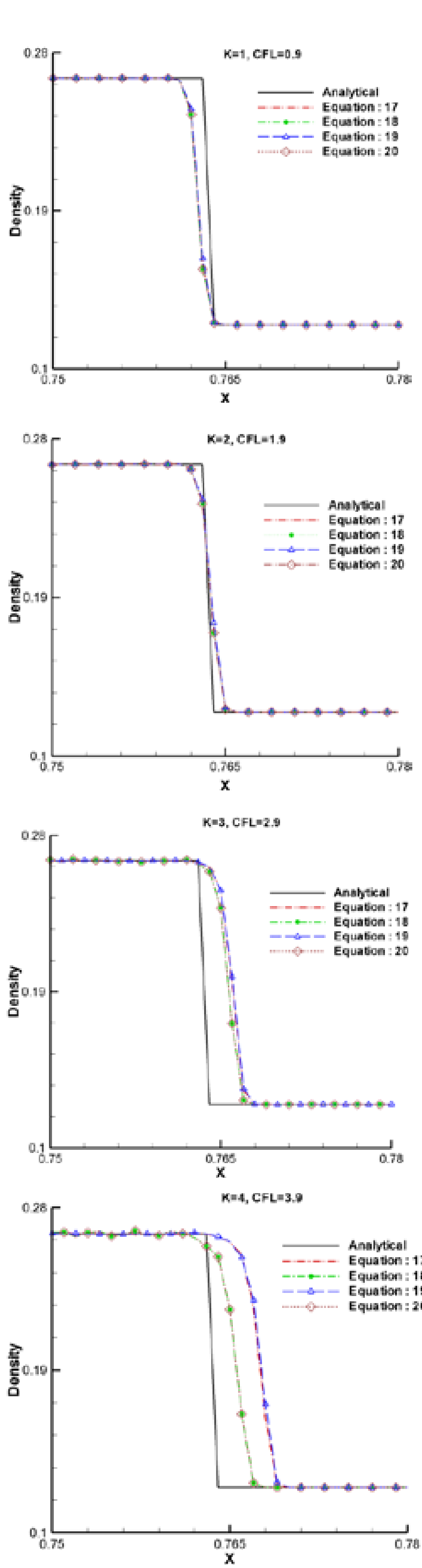


Fig. 5. Shock region, SOD case.

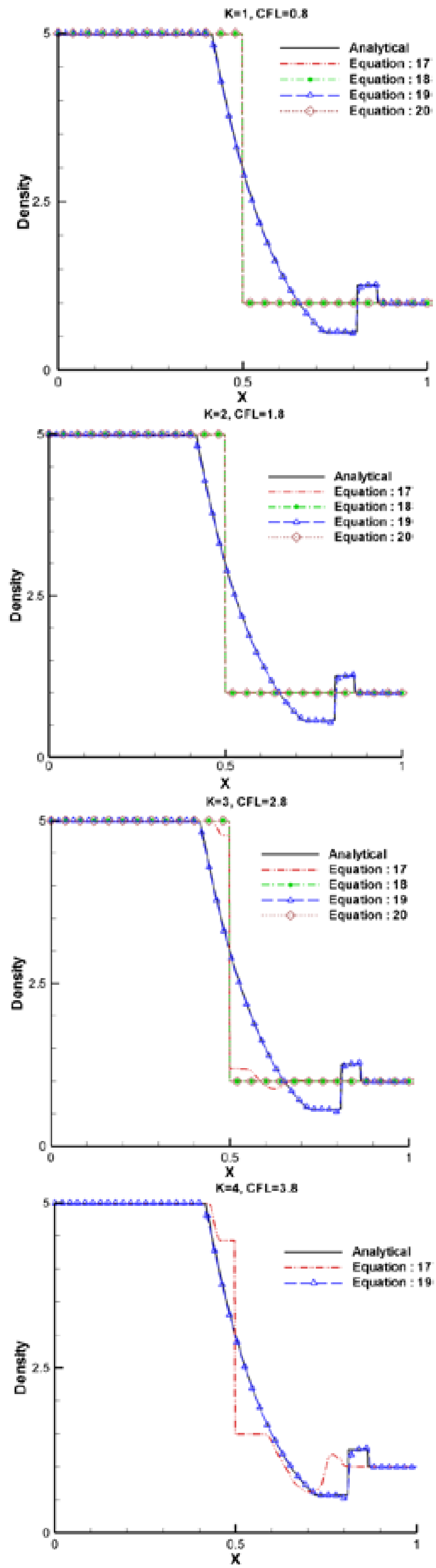


Fig. 6. Overall profiles, inverse shock. Case.

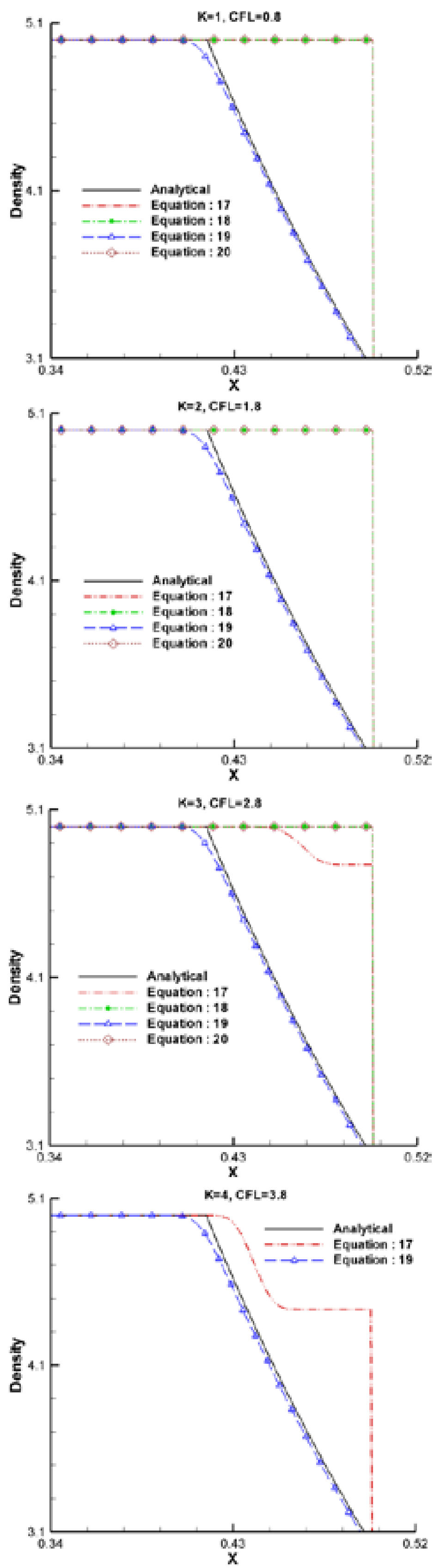


Fig. 7. Start of expansion, inverse shock case.

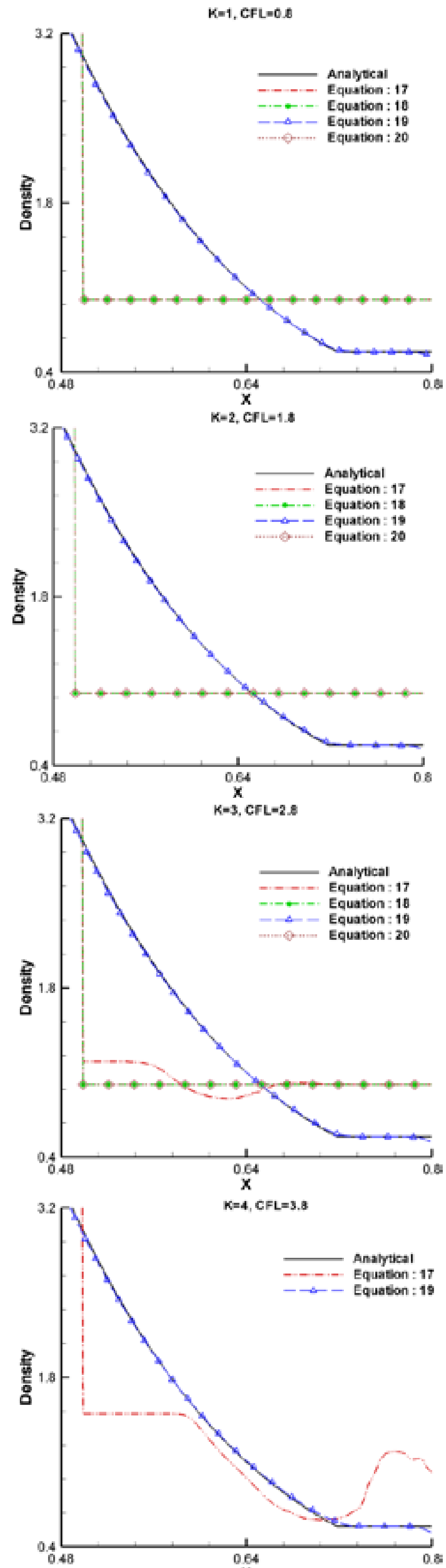


Fig 8: End of expansion, inverse shock case.

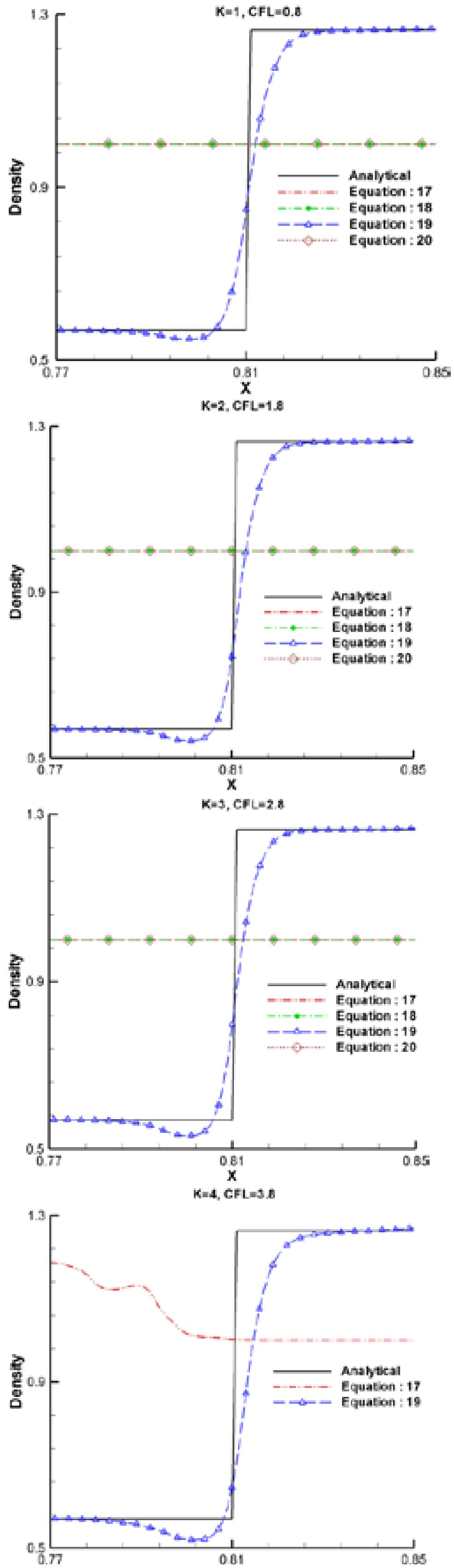


Fig. 9: Contact region, inverse shock case.

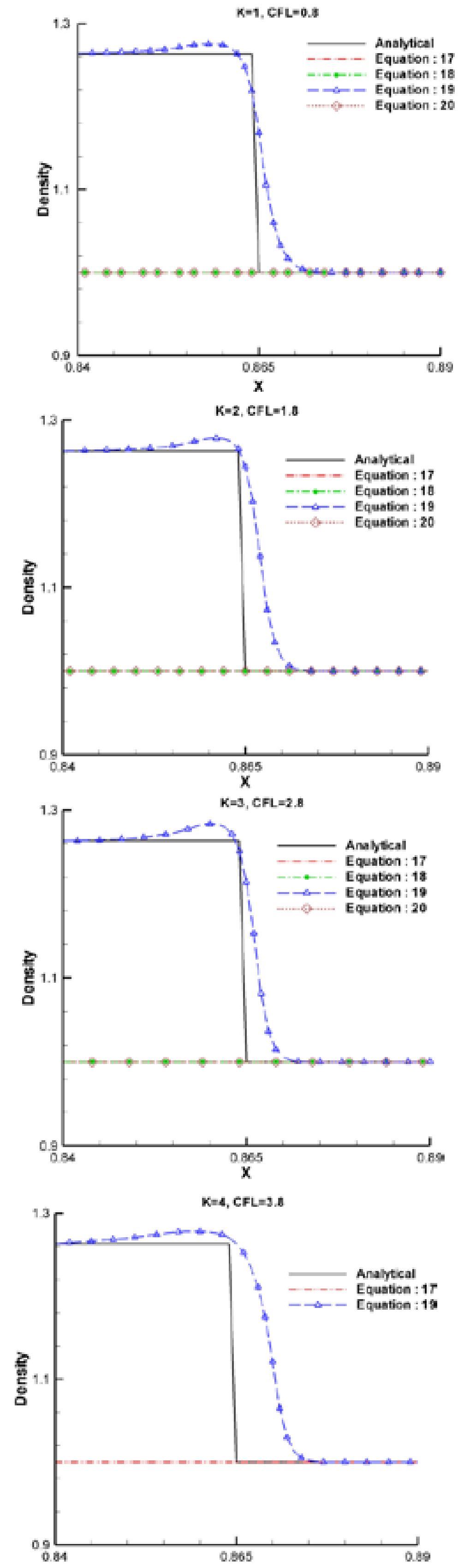


Fig. 10. Shock region, inverse shock case.

5. CONCLUSIONS

In present work LTS TVD scheme developed by Harten and improved by Qian have been investigated for different entropy fix conditions. Computed results are analyzed for merits and shortcomings of different entropy fix with large time step schemes. Shock tube problem is used for validation purpose. It is concluded that for simple cases all entropy fix provide physical results with minor oscillation and dissipation issues for all CFL values investigated in present studies. However for complex test cases, like inverse shock boundary condition, only Harten's derived entropy fix is stable and depicts physical behavior precisely for LTS TVD.

In the present work the numerical results were computed for 1D shock tube problem only. Similar studies must be pursued for more complex 2D and 3D flows in future. Large time step scheme behavior with different entropy fix for various Eigen vector matrix and flux limiter should also be investigated.

NOMENCLATURE

A	inviscid flux jacobion matrix	n	number of time steps taken
$C_{l(x)}$	coefficient functions	p	pressure
$C_{k(v)}$	entropy fixing function	t	time
E	total energy	Δt	time step
F	physical flux	u	velocity in x-direction
K	CFL restriction parameter	v	local CFL number
R	eigen vector matrix	x	axial distance
R^{-1}	inverse eigen vector matrix	Δx	grid spacing
U	conservative variable vector	α	characteristic variable
a	characteristic speed	β	numerical characteristic speed
c	speed of sound	ε	entropy fix parameter
f	numerical flux	Φ	numerical dissipation term
g	flux correction	γ	ratio of specific heat
$\breve{g}$	limiter function	λ	mesh ratio
i	grid points	μ	CFL parameter
k	characteristic direction	ρ	density
m	number of eigen values	σ	limiter function parameter
		ψ	entropy correction function

6. REFERENCES

1. Harten, A. The artificial compression method for computation of shocks and contact discontinuities: III. Self-adjusting hybrid schemes.*Mathematics of Computation* 32 (142): 363-389 (1978).
2. Harten, A. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics* 135: 260-278 (1997).
3. Harten, A. On a Large time-step high resolution scheme.*Mathematics of Computation* 46(174): 379-399 (1986).
4. Hoffmann, K., & S. Chiang. *Computational Fluid Dynamics. Engineering Education System (EES),* 4th ed., Vol I. Wichita, Kansas, USA (2000).
5. Hoffmann, K., & S. Chiang. *Computational Fluid Dynamics. Engineering Education System (EES)*, 4thed.,Volume II. Wichita,Kansas, USA (2000).
6. Huang, H., C. Leey, H. Dongz, & J. Zhang. Modification and applications of a large time-step high resolution TVD scheme. In: *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, AIAA*, p. 2013-2077 (2013).
7. John, D., & J. Anderson. *Computational Fluid Dynamics: The Basics with Applications.* McGraw-Hill, USA (1976).
8. Laney, C. *Computational Gasdynamics.* Cambridge University Press, New York, USA, (1998).
9. Lax, P.D. *Hyberbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves.* SIAM, New York (1973).
10. Mukkarum, H., U.H. Ihtram, & F. Noor. Efficient and accurate scheme for hyperbolic conservation laws.*International Journal of Mathematical Models and Methods in Applied Sciences,* NUAN 9: 504-511 (2015).
11. Noor, F., & H. Mukkarum. To study large time step high resolution low dissipative schemes for hyperbolic conservation laws. *Journal of Applied Fluid Mechanics,* vol. 9 (2016).

12. Roe, P. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics* 43: 357-372 (1981).
13. SOD, G.A. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics* 27 (1978).
14. Sweby, P.K. High resolution schemes using flux limiters for hyperbolic conservation laws. *Journal on Numerical Analysis,* SIAM 21(5): 995-1011 (1984).
15. Tannehill, J., D. Anderson, & H. Pletcher. *Computational Fluid Mechanics and Heat Transfer,* 2nd ed. Taylor & Francis, CA, USA (1984).
16. Toro, E.F. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction,* 3rd ed., Springer, Berlin (2009).
17. Versteeg, H., & W. Malalasekera. *An Introduction to Computational Fluid Dynamics: The Finite Volume Method,* 2nd ed. Pearson Prentice Hall, England, United Kingdom (2007).
18. Yee, H. *Upwind and Symmetric Shock-Capturing Schemes*. NASA Technical Memorandum 130: 89464 (1987).
19. Yee, H.C. *A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods*. Ames Research Center, NASA Technical Memorandum 101088 , 228., Moffett Field, CA, USA (February 1989).
20. Yee, H., G. Klopfer, & J. Montagn. *High-Resolution Shock-Capturing Schemes for lnviscid and Viscous Hypersonic Flows*. NASA Technical Memorandum 100097, 38. USA (April 1988).
21. Yee, H., N. Sandham, & M. Djomehri. Low dissipative high order shock capturing methods using characteristic based filter. *Journal of Computational Physics* 150: 199-238 (1999).
22. ZhanSen, Q.C.L. A class of large time step godunov scheme for hyperbolic conservation laws and applications. *Journal of Computational Physics* 230: 7418-7440 (2011).
23. ZhanSen, Q.C.-H.L. On large time step tvd scheme for hyperbolic conservation laws and its efficiency evaluation.*Journal of Computational Physics* 231: 7415-7430 (2012).

Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 225–235 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Application of Bayesian Monte Carlo Technique to Calculate Extreme Rainfall over Sindh Province in Comparison with Maximum Likelihood Method

Muhammad Ali[1, *], Muhammad Jawed Iqbal[2], and Zohaib Aziz[3]

[1] Department of Computer Science, Sindh Madrasatul Islam University, Aiwan-e-Tijarat Road, Karachi-74000, Pakistan
[2]Institute of Space and Planetary Astrophysics, University of Karachi, Karachi, Pakistan
[3]Department of Statistics, Federal Urdu University of Arts, Science and Technology, Karachi, Pakistan

Abstract: The consequences of statistical modeling of extreme rainfall are pivotal for civil engineering and planning division in order to instill the capability of structures of building that can withstand the extreme situation. Yearly maximum rainfall data of Karachi, Badin, Chhor from 1961-2010, and Rohri from 1971-2010 have been used in this study. The method of Maximum Likelihood (ML) and Bayesian techniques have been implemented to estimate the parameters of Generalized Extreme Value (GEV) distribution and also to compute return levels against sundry return periods. Non-informative priors are used to get the posterior densities. To gauge and compare the results of the above mentioned methods, acceptance rates and forecasting errors have been used as Goodness of Fit (GoF) test. Though both the methods are applicable, but the GoF test highlights that ML method is slightly better than Bayesian for observing the annual maximum rainfall in Sindh province of Pakistan.

Keywords: Yearly maximum rainfall, extreme value modeling, return period, goodness of fit test, return levels.

1. INTRODUCTION

The phenomenon of Earth's climate change with connected rainfall are mainly driven by the changes [1], caused by both natural and anthropogenic activities in the twentieth century, but their comparative roles and local influences are still under extreme debate [2]. The weather conditions which lie outside a locality's normal range are termed as extreme weather events. Extreme weather events, such as heat waves, cold waves, fog, thunderstorms, snowstorms, cyclones, hailstorms, floods and heavy rains etc., have a key importance. The economy and people's life generally hinge on these events [3]. Extreme value modeling of climatic processes is a customary practice for large scale construction [4], brought in generalized extreme value distribution. This distribution is repeatedly and extensively applied to model the extreme hydrological events, for example flood flow, extreme rainfall intensity [5], coastal water level [6] and extreme temperature [7]. The most climatic models in recent events suggest that the global warming is likely to augment the frequency of extreme weather events in many places [8]. The hot extremes, heat waves, and heavy precipitation events will persist to become more frequent [9].Maximum Likelihood(ML) method is used to embrace the result of covariates, for example the chance that one or more constraint of the distribution may contain a trend owing to climatic changes [10]. By means of historical data, estimates are calculated for design parameters, which are then used in construction to have a minimal failure probability [5]. These extreme events are frequently allied with climatic changes, which may be succeeded by natural calamities like landslides and flash floods. According to the report of the UNFCCC (United Nations Framework Convention on Climate Change), the Asian climatic change would influence the resources of water, food security and agriculture, biodiversity and ecosystems, public health and coastal zones. For any developing country similar to Pakistan, the natural disasters will

Received, March 2016; Accepted, May 2016
*Corresponding author: Muhammad Ali; Email: m.alishigri@gmail.com

assuredly affect the country's production. There is a conclusion from the analysis of climate change that it also influences negatively on paddy cultivation of rice [11]. It has been noted that such extreme climatic events are predictable at all. From statistical analysis of climate change, it is found that, the devastating effect of extreme rainfall may also be decreased by precautionary measures. In case of an extreme value modeling, one must consider other sources of knowledge like the known physical constraints, the maximum possible value or may be derived from an understanding of associated processes and possibly the same variable at a different location. Sometimes we may observe data, not to be completely representative for the whole period and there may be historical evidences, though not in the form of data, but in the shape of behavior , which is significantly more extreme than that which has been considered. As a result, there are a number of reasons that why it is to be expected that an expert with knowledge of the physical processes may have information that is pertinent to extreme behavior , which is also independent of the existing data. This leads obviously to the Bayesian inferential framework as a beginning for undertaking an extreme value analysis. Several attempts have been made to use Bayesian methodology in the extreme value analysis. For example, the most comprehensive analysis by Smith and Naylor [12] shows the effect of different prior assumptions on the posterior distributions of parameters of the Weibull distribution. The theoretical background of Bayesian estimation of extreme quantities, with partial regard for the significance of prior structure, has been explained by Pickands [13]. The advent of Markov chain Monte Carlo methodology estimates the marginal densities using Gibbs MCMC sampler [14]. In this way the current work is an attempt to calculate different return levels of extreme rainfall in Pakistan having best fitted model.

2. DATA DESCRIPTION AND METHODOLOGY

Peak Over Threshold (POT) and Annual Maximum Series (AMS) are commonly used in extreme rainfall analysis [8]. The AMS contains the maximum rainfall value occurring in a particular day of a year while POT involves all maximum rainfall values greater than the defined threshold value. POT is usually used to resolve the wastage of the data in AMS method. Due to the complex situation in choosing a suitable threshold in POT, the AMS is preferable method in extreme value analysis [15]. Therefore we use annual maximum series of rainfall of four meteorological stations of Sindh, from Jan, 1961 to Dec, 2010 for Karachi, Badin, Chhor and Jan, 1971 to Dec. 2010 for Rohri station. It is clear from Fig. 1, that there is no any particular trend in the data series. 40 years normal values of extreme rainfall of Karachi, Badin, and Chhor are 62.8 mm, 69.1 mm, and 69.8 mm respectively, while 30 years normal value of rainfall at Rohri is 41.2 mm. The maximum amount of extreme rainfall from 1961 to 2010 are 207.0mm that occurred in 1977, 241.0mm that occurred in 1979, 251.2mm that occurred in 1998 and 184.5mm in 1978 at Karachi, Badin, Chhor and Rohri, respectively.

Suitable probability distributions must be used to model extreme behavior of rainfall in order to get the best inference. Numerous literatures for example [16, 17] available to get the best-fitted distribution for annual maximum rainfall data; recommend GEV distribution as the best fitted one. It is found that for sample size $n \geq 30$, maximum likelihood method generally performs well over all values of shape parameter (k) in the range $-0.5 < k < 0.5$ [1]. Here we use probabilistic approach in order to get the return levels of extreme rainfall against different return periods. In this study we use two approaches (Bayesian approach and Maximum likelihood approach) and also compare their results to get a specific forecasting model for future extreme rainfall scenario of Sindh.

2.1. Bayesian Approach

Bayesian statistics investigate the uncertainty about the unknown parameters by using probability statements so that the unknown parameters are here regarded as random variables. These probability statements are conditional on observed values of rainfall. To compute the posterior distribution Markov Chain Monte Carlo (MCMC) simulation technique is repeatedly used for estimating parameters. To make use of MCMC, a parent distribution for producing simulated value of parameter is necessarily to be introduced. To select the anticipated distribution, it is very essential to check the appropriateness of that distribution, as the bad choice may significantly delay the convergence for equilibrium point. The appropriate acceptance rates in getting suitable proficiency of Morkov Chain Monte Carlo simulation are around 10 to 40% [6, 15]. So the likelihood function for $Z_1, Z_2, \ldots, Zn$ is given by:

$$L(\mu,\varphi,\xi:Z_1, \ldots ,Z_n) = \prod_{i=1}^{n} f(Zi|\mu, \varphi, \xi), \tag{1}$$

The density of posterior distribution is directly proportional to the product of Prior and Likelihood distribution as;

$$f(\mu,\varphi,\xi|Z_1, \ldots ,Z_n) \propto L(\mu,\varphi,\xi:Z_1,\ldots,Z_n) \times g(\mu,\varphi,\xi) \tag{2}$$

Where g, L and f are prior, likelihood and posterior distribution, respectively. The prior distribution shows the set of confidence about the required parameters. In this analysis 'g' is the non-informative prior distribution which shows that the significantly prior information about extreme rainfall over Sindh is not available at the moment. The location (μ) shape (ξ) and scale (φ) parameters are assumed to be normally distributed.

Generally the purpose of extreme rainfall analysis is to calculate the expected values of rainfall (return levels). If 'y' denotes the future rainfall values with pdf:

$$h(y|Z_1, \ldots ,Z_n) = \iiint f(y|\mu, \varphi, \xi) f(\mu, \varphi, \xi|Z1, \ldots , Zn) d\mu d\varphi d\xi \tag{3}$$

Then the estimates of probability of 'n' year return levels for sample $\theta_1, \theta_2, \ldots, \theta_R$ will be:

$$\Pr[y \leq q_n|z_1, \ldots , z_n] \approx \frac{1}{R}\sum_{i=1}^{R} \Pr(y \leq qn|\theta i) \tag{4}$$

Since Eq. (3) is very complicated to solve directly, so we use MCMC simulation technique to find the posterior distribution. For MCMC simulation we use R-statistical tool with loading two appropriate extreme value modeling (EVM) packages named as texmex and mvtnorm.

2.2. Probability Distribution and Maximum Likelihood Method

Let $Z_1, Z_2 \ldots, Z_n$ be the daily rainfall series where $M_n = \max\{ Z_1, Z_2 \ldots, Z_n\}$ with n=365, is the annual maximum rainfall data. Asymptotic considerations recommend that the distribution of M_n should be something like that of a member of the generalized extreme value (GEV) family having probability distribution function (PDF):

$$f(x) = \frac{1}{\varphi}[1 + \xi\left(\frac{Zi-\mu}{\varphi}\right)]^{-1-\frac{1}{\xi}} \times e^{-[1+\xi\left(\frac{Zi-\mu}{\varphi}\right)]^{-\frac{1}{\xi}}} \tag{5}$$

Where μ,φ and ξ are location, scale and shape parameter with parameter space $-\infty< \mu < \infty$, $\varphi> 0$ and $-\infty< \xi < \infty$. Eq. (5) can be integrated analytically to obtain cumulative density function (CDF) as;

$$F(x) = e^{-[1+\xi\left(\frac{Zi-\mu}{\varphi}\right)]^{-\frac{1}{\xi}}} \tag{6}$$

This equation can be reversed to find an appropriate formula for return levels (q_r) of rainfall corresponding to 1/r years return period [3].

$$q_r = \mu + \frac{\varphi}{\xi}\{1-[-\log(1-p)]^{-\xi} \} \tag{7}$$

The above mentioned parameters are then calculated by both ML method and Bayesian method.

3. GOODNESS OF FIT TEST

Now we will discuss about the best fitted model between Bayesian and maximum likelihood method for the estimation of GEV parameters and future predicted extreme rainfall of above said stations. In this work we use three Goodness of Fit (GoF) tests, named as Relative Absolute Squared Error (RASE), Relative Root Mean Squared Error (RRMSE) and Probability Plot Correlation Coefficient (PPCC). RRMSE and RASE include the measurement of inconsistency among observed and predicted values through the parent distribution (GEV), while PPCC measures the correlation between observed and forecasted values. Now we use annual maximum rainfall data from 1971 to 2000, and forecast the return levels from 2002 to 2010 through both Bayesian and Maximum likelihood method. A method is said to be better in forecasting, if it has less values of RRMSE and RASE and closest value of PPCC to 1 and vice versa. Now as suggested by Zin [17],we will also observe the above said three GOF tests by using the following formulas as;

$$\text{RRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\text{Xi,o}-\text{Q(xi)}}{\text{Xi,o}}\right)^2} \tag{8}$$

$$\text{RASE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\text{Xi,o}-\text{Q(Xi)}}{Xi,o}\right| \tag{9}$$

$$\text{PPCC} = \frac{\sum_{i=1}^{n}\left(\text{Xi,o}-\bar{\text{X}}\right)\left(\hat{\text{Q}}_{(\text{Xi})}-\bar{\text{Q}}_{(\text{Xi})}\right)}{\sqrt{\sum_{i=1}^{n}\left(\text{Xi,o}-\bar{\text{X}}\right)^2\sum_{i=1}^{n}\left(\hat{\text{Q}}_{(\text{Xi})}-\bar{\text{Q}}_{(\text{Xi})}\right)^2}} \tag{10}$$

Where Xi, o is observed values of i^{th} order statistics of yearly extreme rainfall from 2002 to 2010, and Q(xi) is the estimated values for this period. Now the summary of GoF tests is shown in table (2). The statistical analysis of table (2) indicates that, although there is no any big difference between ML and Bayesian method, showing that both the methods are suitable to forecast the extreme rainfall values, but a little bit difference in forecasting errors (i.e. less values of RASE and RRMSE), indicates that ML is more appropriate method for the observed data of Sindh. On the basis of AIC and MSE we can also compare the accuracy of our model for extreme yearly rainfall of observed meteorological stations of Sindh as shown in table (3).

Table 1. Comparison of Estimates of GEV parameters for different cities of Sindh with Standard Deviation (SD) in parenthesis.

GEV Parameters	Rohri		Chhor		Badin		Karachi	
	Bayesian	**ML**	**Bayesian**	**ML**	**Bayesian**	**ML**	**Bayesian**	**ML**
μ (location)	21.65 (3.56)	21.59 (3.4)	40.69 (5.28)	40.72 (5.15)	44.43 (6.03)	44.59 (5.9)	39.21 (5.6)	39.11 (5.42)
Φ (Scale)	2.96 (0.16)	2.9 (0.16)	3.48 (0.13)	3.45 (0.13)	3.6 (0.13)	3.57 (0.13)	3.51 (0.14)	3.46 (0.14)
ξ (Shape)	0.39 (0.166)	0.37 (0.164)	0.19 (0.13)	0.17 (0.13)	0.11 (0.13)	0.08 (0.14)	0.195 (0.15)	0.18 (0.15)

Table 2. Comparative study of forecasting errors between Bayesian and maximum likelihood estimates of Sindh.

Station Name	Method of Estimation	Goodness of Fit Test		
		RASE	RRMSE	PPCC
Rohri	Bayesian	2.36	7.08	0.945
	ML	2.13	6.39	0.956
Chhor	Bayesian	1.56	4.68	0.249
	ML	1.44	4.31	0.276
Badin	Bayesian	1.24	3.72	0.217
	ML	1.13	3.38	0.241
Karachi	Bayesian	0.73	2.19	0.266
	ML	0.70	1.09	0.273

Table 3. Comparative study of AIC, MSE and acceptance rate for observed meteorological stations of Sindh.

S.No	Sation's Name	WMO Number	Latitude	Longitude	AIC	MSE	Acceptance Rate
1	Rohri	41725	27^{o} 40'	68^{o} 54'	382	0.110	0.311
2	Chhor	41685	29^{o} 53'	69^{o} 43'	518	0.013	0.315
3	Badin	41785	24^{o} 38'	68^{o} 54'	526	0.082	0.309
4	Karachi	41780	24^{o} 54'	66^{o} 56'	521	0.327	0.312

Table 4. Comparison of return levels against different return period of Bayesian and ML methods for Sindh.

Station Name	Method of estimation	Return Level (mm)				
		10 year	25 year	50 year	75 year	100 year
Rohri (1971-2010)	ML	86	134	122	216	243
Chhor	ML	126	173	213	239	258
Badin	ML	133	175	209	230	245
Karachi	ML	128	178	222	250	271

4. RESULTS AND DISCUSSION

The amount of rainfall expected to occur at a station in future for a defined period of time is said to be return level of that station. The likelihood function can be assembled for complex situation of modeling for example non-stationary, effects of covariates and regression model etc. [18, 19].

MCMC technique is a way of simulation, for a complex distribution. The simulated values for all parameters of GEV distribution have been found together in a similar zone as revealed in Fig (2). The trace plots for scale and location parameters of GEV distribution for 40000 iterations are also shown in Fig (2). For non-informative priors, the variances have to be chosen large enough to get flat priors.

Fig (3) shows the posterior density plots for location, scale and shape parameters of parent distribution (i.e. GEV distribution), from which we can observe that all these figures are symmetrical in shape. The wide spread distributions in these figures indicate a big variance in non-informative prior distribution. In this study we have used Gibbs sampling in combination of Metropolis-Hasting scheme to get the desired posterior distribution. Our calculation also shows that the acceptance rate of Karachi, Badin, Chhor and Rohri are 31.2%, 30.9%, 31.5% and 31.1% respectively.

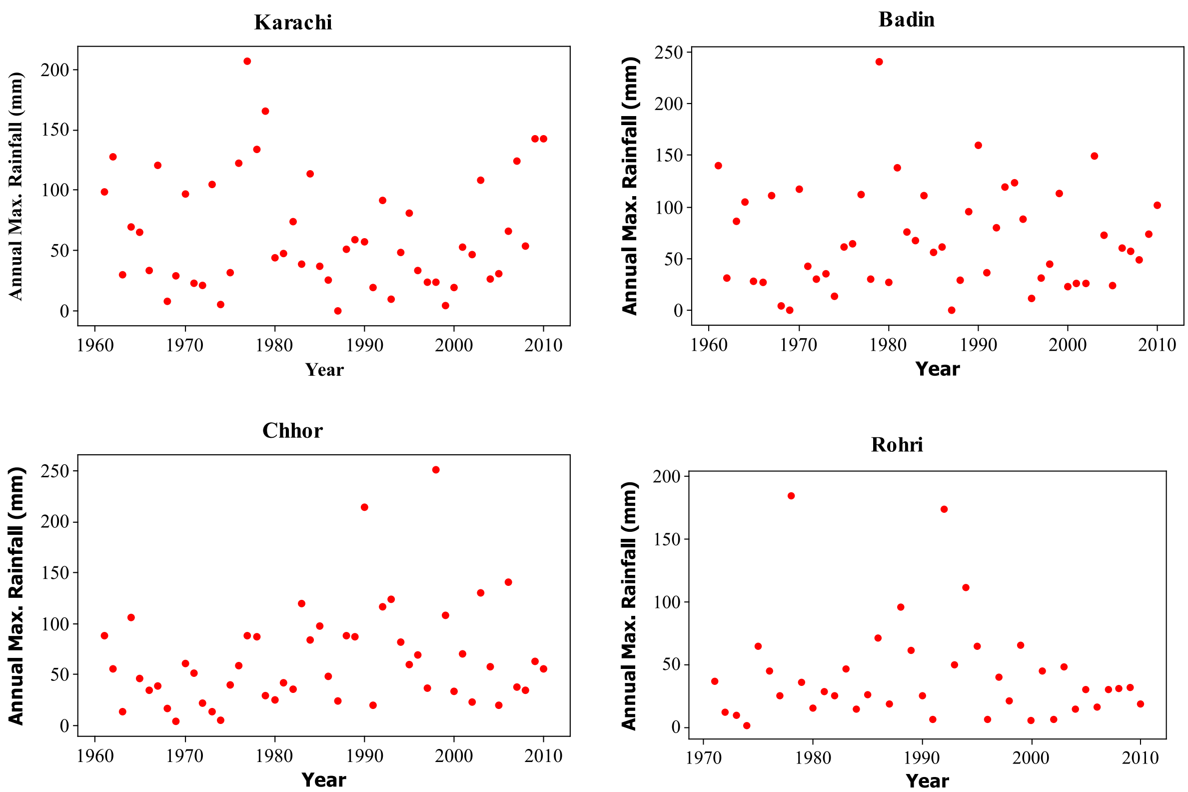
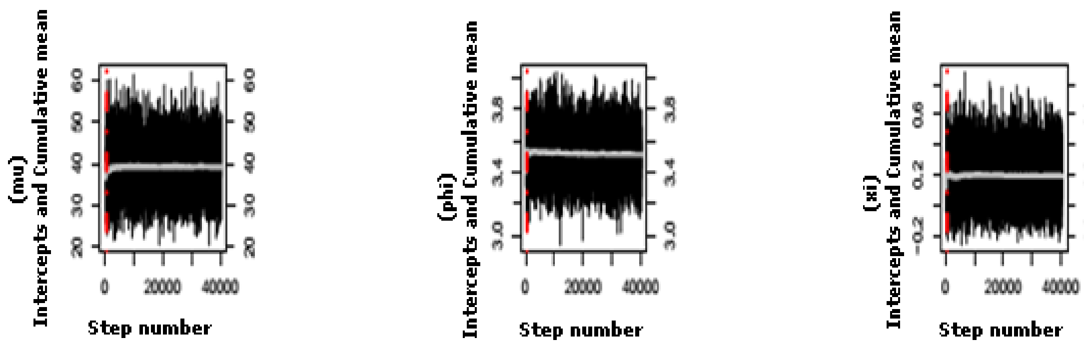


Fig. 1. Annual maximum rainfall over Sindh, 1960-2010.

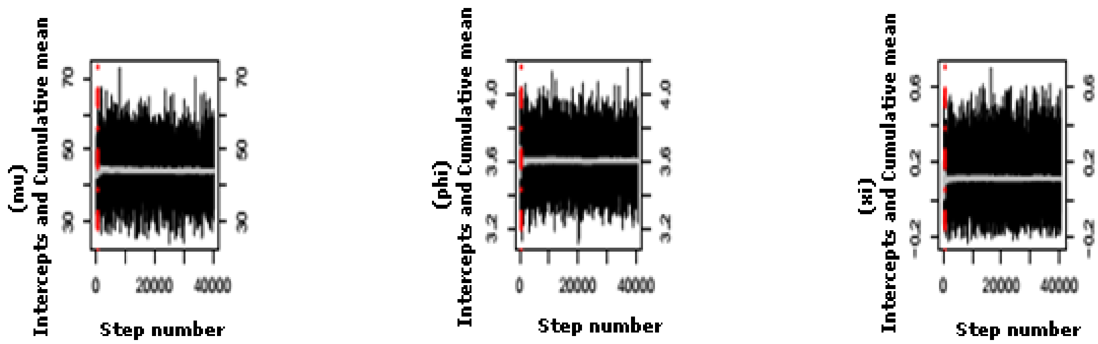
The ML method selects those parameters which maximize the likelihood of the given data. In ML method, parameters are supposed to be unidentified but fixed, and are calculated approximately. The suitability of this method for observed data points of given meteorological stations, can also be seen from Figs. (4).The model values and empirical values of Karachi, Badin, Chhor and Rohri approximately overlap the actual line in ML method. Therefore we can conclude that ML can explain more accurately the yearly maximum rainfall behavior of these stations. The values of parameters and their standard deviations shown in Table (1), also indicate that the ML method gives slightly better results than Bayesian.

Now with the help of R-statistical tool for extreme value modeling (EVM), we have first predicted the return levels of the above mentioned stations of Sindh from 2002 to 2010 designed for both methods Bayesian and ML. Comparing these forecasted values with observed values of yearly maximum rainfall from 2002 to 2010, we have calculated forecasting errors using eq (8), eq (9) and eq (10). Forecasting errors conclude that maximum likelihood is slightly better than Bayesian for yearly maximum rainfall of above mentioned stations of Sindh. Hence we have forecasted different return levels for 10, 25, 50, 75 and 100 years return periods as shown in Table 4. Table 4 depicts that the 50 years return levels estimated by ML method for Karachi, Badin, Chhor and Rohri are 222 mm, 209 mm, 213 mm and 122 mm respectively. While that of 100 years return levels are 271 mm, 245 mm, 258 mm and 243 mm yearly maximum rainfall in 24 hours, respectively. Comparing these predicted return levels of the above stations except Rohri, the amount of yearly maximum rainfall values of Karachi is the greatest while that of Badin is the least. Our analysis also suggests that, for short term forecast, the results obtained by Bayesian and ML are comparable to each other but for long range forecast they differ from one another. When these extreme rainfalls remain continuous for few days, it may lead floods, which will be very destructive and devastating for that region. Therefore, the purpose of calculating these return levels is to inform engineers and higher authorities, about the incoming situation of extreme rainfall over Sindh, in order to take precautionary steps to save our country from different kinds of losses.

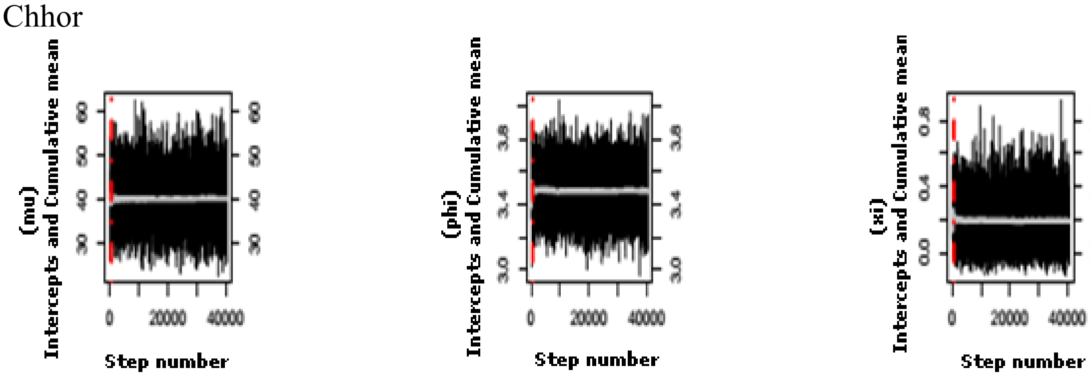
Karachi



Badin



Chhor



Rohri

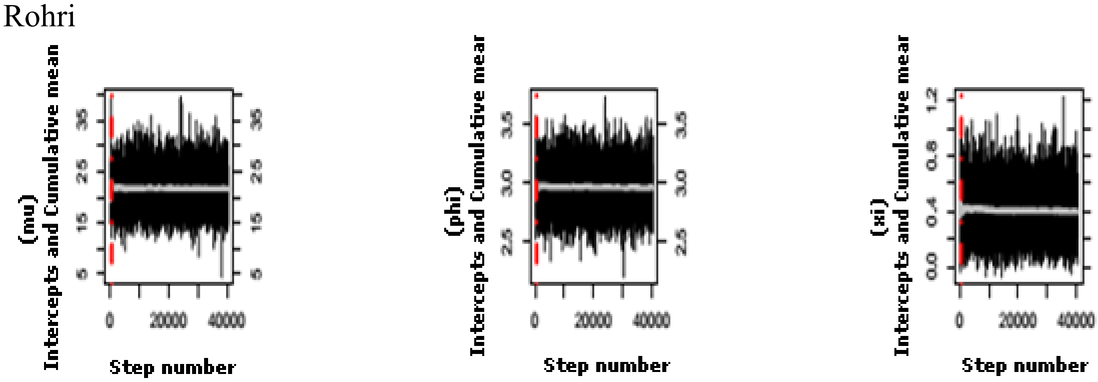


Fig. 2. Trace plots for location, scale and shape parameters of GEV distribution for 40000 iterations for Sindh.

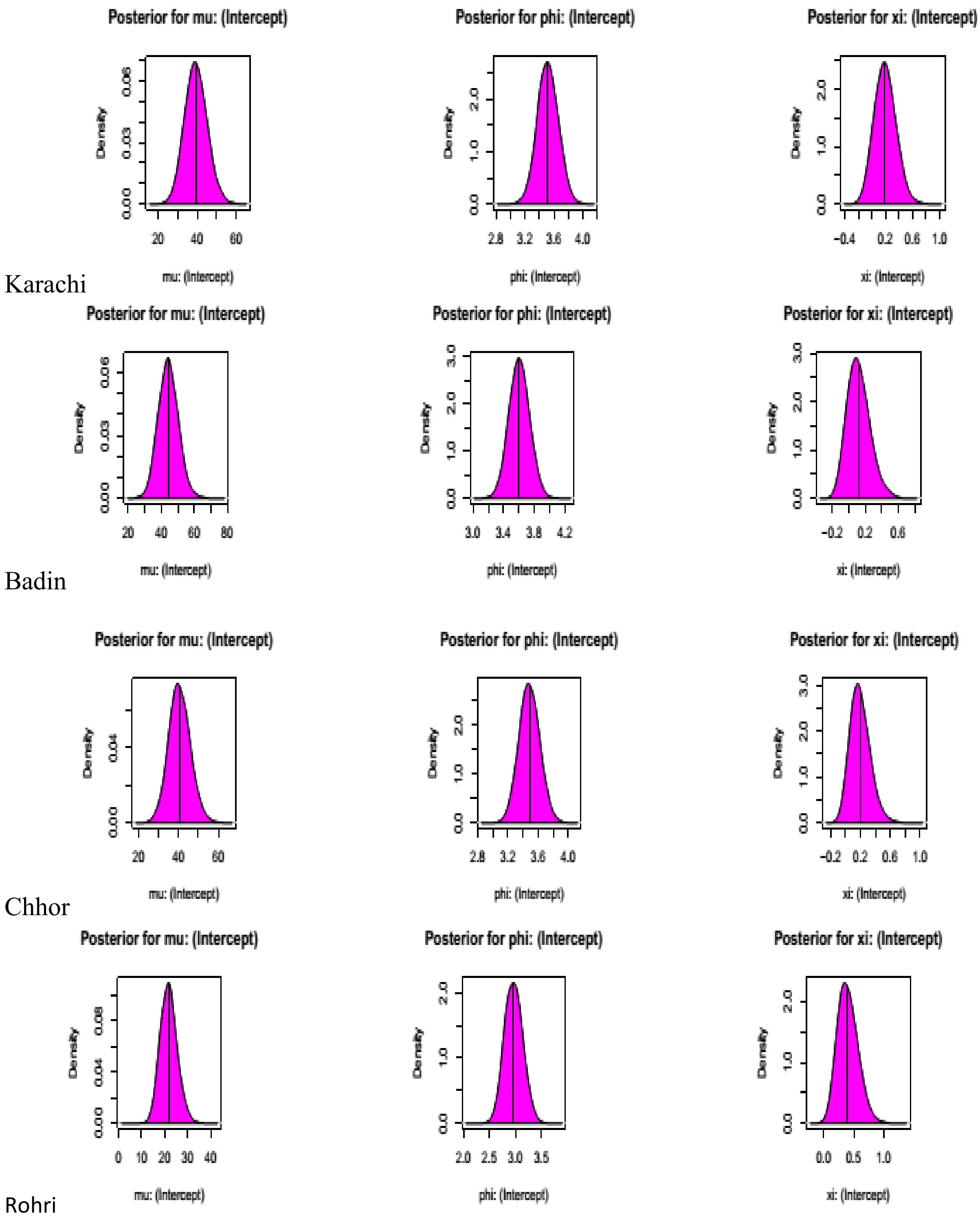


Fig. 3. Posterior density plots for different cities of Sindh.

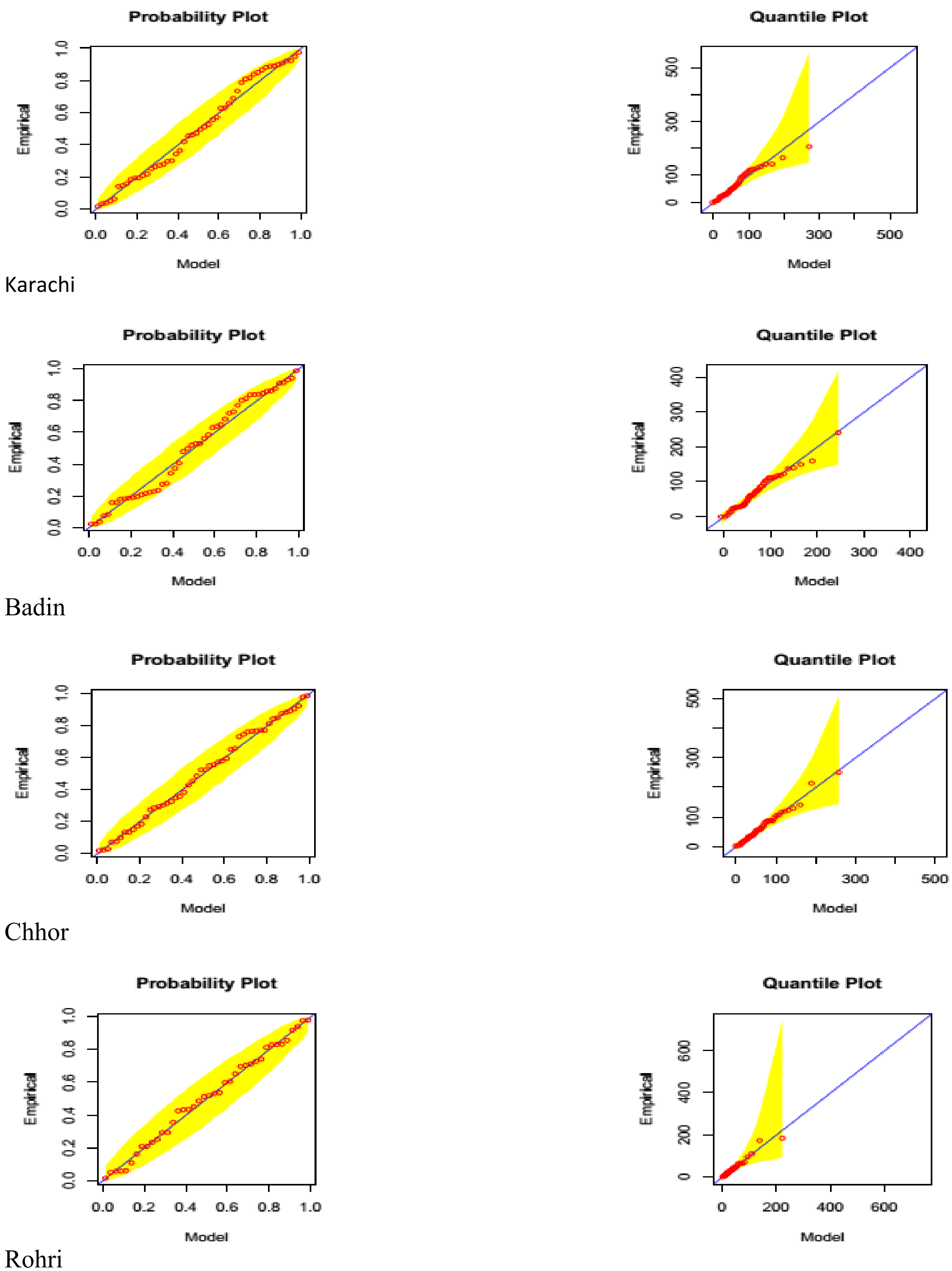


Fig. 4. Probability plot and Quantile plot of maximum likelihood method for Sindh.

5. CONCLUSIONS

The main purpose of carrying out this analysis is to apply, and compare the results of Bayesian approach and Maximum Likelihood method, and also to estimate the return levels against different return periods. We utilize yearly maximum rainfall values of Karachi, Badin, Chhor and Rohri, for the period of January 1961 to December 2010. It was observed that maximum likelihood and Bayesian MCMC technique are strongly related because both methods involve likelihood function in their initial steps. In this study we used Gibbs sampling in combination of Metropolis-Hasting scheme to get the desired posterior distribution. Our calculation also shows that The acceptance rates in getting high proficiency of MCMC simulation for Karachi, Badin, Chhor and Rohriare 31.2%, 30.9%, 31.5% and 31.1%, respectively. Because of non-informative priors, no any big difference has been found between Bayesian and Maximum Likelihood method in calculating the values of GEV parameters. But the deeply analysis of forecasting errors i.e. Relative Root Mean Squared Error (RRMSE), Relative Absolute Squared Error (RASE)and Probability Plot Correlation Coefficient(PPCC),shows that ML method is slightly better than Bayesian method for the calculation of return levels of observed meteorological stations. Furthermore, Fig (4) also showed that the model values and empirical values of the above said meteorological stations are approximately overlapping the actual line in ML method. Therefore we can conclude that ML can explain more accurately the yearly maximum rainfall behavior of these stations. The values of parameters and their standard deviations shown in Table (1), also indicated that the ML method gives slightly better results than Bayesian, so we can use it as the best fitted model. We also observe that for short term forecast, the results obtained by Bayesian and ML methods are very close to each other but for long range forecast they differ from one another.

Hence the estimated return levels for 50 years return period by ML method are 222 mm, 209 mm, 213 mm and 122 mm while those for 100 years return period are 271 mm, 245 mm, 258 mm and 243 mm for above mentioned meteorological stations in 24 hours, respectively. These return levels are very useful for planning division, civil engineering, forecasting sections of Pakistan Meteorological Department, ministry of climate change, etc. So our work suggests to upgrade the flood forecasting system by using modern technology i.e. GIS based technology, and to get remarkable improvement in the river structures of Sindh province of Pakistan.

6. ACKNOWLEDGEMENT

We are very thankful to Pakistan Meteorological Department (PMD) for providing the required daily rainfall data.

7. REFRENCES

1. Bansod, S.D., H.N. Singh, S.D. Patil, & N. Singh. Recent changes in the circulation parameters and their association with Indian summer monsoon rainfall. *Journal of Atmospheric and Solar-Terrestrial Physics* 77: 248-253 (2012).
2. Zin, W.Z.W., A.A. Jemain, & K. Ibrahim. The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment. *Theoretical and Applied Climatology* 96(3-4): 337-344 (2009).
3. Ali, M., & M.J. Iqbal. A probabilistic approach for estimating return periods of extreme annual rainfall in different cities of Khyber Pakhtunkhwa (KP), Pakistan. *Nucleus* 49(2): 107-114 (2012).
4. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society* 81(348): 158-171 (1955).
5. Coles, S.G., & J.A. Tawn. A Bayesian analysis of extreme rainfall data. *Applied Statistics* 45: 463-478 (1996).
6. Huang, Wenrui, S. Xu, & S. Nnaji. Evaluation of GEV model for frequency analysis of annual maximum water levels in the coast of United States. *Ocean Engineering* 35(11): 1132-1147 (2008).

7. Siliverstovs, B., R. Ötsch, C. Kemfert, C.C. Jaeger, A. Haas, & H. Kremers. Climate change and modeling of extreme temperatures in Switzerland. *Stochastic Environmental Research and Risk Assessment* 24(2): 311-326 (2009).
8. Cotton, J., & C. Ferro. *Bayesian Priors and Estimation of Extreme Rainfall in Camborne*. MAS3029 Mathematical Sciences Project 1, University of Exeter, Exeter EX4, Devon, UK (2008).
9. IPCC. IPCC Report "*Climate Change 2007: The Physical Science Basis, Summary for Policymakers"*. Contribution of Working Group I to the Fourth Assessment Report, Paris (2007).
10. Katz, R.W., G.S. Brush, & M.B. Parlange. Statistics of extremes: Modeling ecological disturbances. *Ecology* 86(5): 1124-1134 (2005).
11. Eli, A., M. Shaffie, & W.Z.W. Zin. Preliminary study on Bayesian extreme rainfall analysis: A case study of AlorSetar, Kedah, Malaysia. *Sains Malaysiana* 41(11): 1403-1410, (2012).
12. Smith, R.L., & J.C. Naylor. A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*: 358-369 (1987).
13. Pickands III, J. Bayes quantile estimation and threshold selection for the generalized Pareto family", In: *Extreme Value Theory and Applications*, Springer US, p. 123-138. (1994).
14. Gelfand, A.E., & A.F. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85(410): 398-409 (1990).
15. Madsen, H., F.F. Rasmussen, & D. Rosbjerg. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resources Research* 33(4): 747-757 (1997).
16. Zalina, M., M. Desa, V. Nguyen, & A. Kassim. Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science & Technology* 45(2): 63-68 (2002).
17. Zin, W.Z.W., A.A. Jemain, K. Ibrahim, S. Jamaludin, & S.M. Deni. A comparative study of extreme rainfall in Peninsular Malaysia: with reference to partial duration and annual extreme series. *Sains Malaysiana* 38: 751-760 (2009).
18. Coles, S. & M.J. Dixon. Likelihood-based inference for extreme value models. *Extremes* 2(1): 5-23 (1999).
19. Coles, S.G., & A. Jonathan. Tawn. A Bayesian analysis of extreme rainfall data. *Applied Statistics* 45(4): 463-478 (1996).
20. Ailliot, P., C. Thompson, & P. Thomson. Mixed methods for fitting the GEV distribution. *Water Resources Research*, doi: 10.1029/2010WR009417 (2011).

Proceedings of the Pakistan Academy of Sciences

Instructions for Authors

Aims and Scope: *Proceedings of the Pakistan Academy of Sciences* is official journal of the Academy, published quarterly, in English. This open access journal publishes research papers in *Engineering Sciences & Technology, Life Sciences, Medical Sciences***,** *and Physical Sciences*. State-of-the-art reviews (~20 pages, supported by recent references) summarizing R&D in a particular area of science, especially in the context of Pakistan, and suggesting further R&D are also considered. Manuscripts undergo double-blind review. Authors are not required to be Fellows or Members of the *Pakistan Academy of Sciences* or citizens of Pakistan.

Manuscript Format

Manuscript may contain **Abstract, Keywords, INTRODUCTION, MATERIALS AND METHODS, RESULTS, DISCUSSION (or RESULTS AND DISCUSSION), CONCLUSIONS, ACKNOWLEDGEMENTS and REFERENCES** *and any other information that the author(s) may consider necessary.* The Manuscript sections must be numbered, i.e., **1. INTRODUCTION, 2. MATERIALS AND METHODS,** and so on.

Manuscripts, in *Times New Roman*, 1.5-spaced (but single-space the Tables), with line numbering and one-inch margins on all sides on A-4 size paper, should not exceed 20 pages including Tables and Figures. Number manuscript pages throughout. The text (in **Font Size 11,** except for the sections mentioned in **Font Size 10)** must be typed in a single column across the paper width. All Tables and Figures must be placed after the text, i.e., after REFERENCES section.

Title Page: For Online submission of the manuscript, Title Page should not be included. However, for submitting as an email attachment, the title page must be as under:

(a) **Title** of the article **(**Capitalize initial letter of each main word; font size 16; **bold**), max 160 characters (no abbreviations or acronyms), depicting article's contents; (b) Author' first name, middle initial and last name (font size 12, **bold**), and professional affiliation (i.e., each author's Department, Institution, Mailing address and Email; but no position titles) (font size 12); (c) Indicate the corresponding author with *; (d) **Short running title**, max 50 characters (font size 10).The **next Page** should start with **Title** of the Article, followed by entire manuscript. **Headings and Subheadings** (font size 11): All flush left

LEVEL-1: **ALL CAPITAL LETTERS; bold**

Level-2: **Capitalize each main word; bold**

Level-3: Capitalize each main word; Bold, Italic

Level-4: *Run-in head*; *Italics*, *in the normal paragraph position. Capitalize the initial word only and end in a colon (i.e., **:**)*

Abstract (font size 10; max 250 words)**:** Must be self-explanatory**,** stating rationale, objective(s), methodology, main results and conclusions of the study. Abbreviations, if used, must be defined on first mention in the Abstract as well as in the main text. Abstract of review articles may have variable format.

Keywords (font size 10)**:** Three to eight keywords**,** depicting the article.

For Online submission:

- **Separate submission of Abstract along with Title is required**
- **Author(s)' names and affiliations should not be included in the manuscript or the Abstract**

INTRODUCTION: Provide a clear and concise statement of the problem, citing relevant recent literature, and objectives of the investigation.

MATERIALS AND METHODS: Provide an adequate account of the procedures or experimental details, including statistical tests (if any), in a concise manner but sufficient enough to replicate the study.

RESULTS: Be clear and concise with the help of appropriate Tables, Figures and other illustrations. Data should not be repeated in Tables and Figures, but must be supported with statistics.

DISCUSSION: Provide interpretation of the RESULTS in the light of previous relevant studies, citing published references.

ACKNOWLEDGEMENTS (font size 10)**:** In a brief statement, acknowledge financial support and other assistance.

REFERENCES (font size 10)**:** Cite references in the text **by number only** in **square brackets**, e.g. "Brown et al [2] reported ...” or “... as previously described [3, 6–8]”, and list them in REFERENCES section, in the order of citation in the text, Tables and Figures (not alphabetically). Only published (and accepted for publication) journal articles, books, and book chapters qualify for REFERENCES.

List of REFERENCES must be prepared as under:

a. **Journal Articles** *(Name of journals must be stated in full)*

1. Golding, I. Real time kinetics of gene activity in individual bacteria. *Cell* 123**:** 1025–1036 (2005).
2. Bialek, W. & S. Setayeshgar. Cooperative sensitivity and noise in biochemical signaling. *Physical Review Letters* 100**:** 258–263 (2008).
3. Kay, R.R. & C.R.L. Thompson. Forming patterns in development without morphogen gradients: differentiation and sorting. *Cold Spring Harbor Perspectives in Biology* 1: doi: 10.1101/cshperspect.a001503 (2009).

b. **Books**

4. Luellen, W.R. *Fine-Tuning Your Writing*. Wise Owl Publishing Company, Madison, WI, USA (2001).
5. Alon, U. & D.N. Wegner (Ed.). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC, Boca Raton, FL, USA (2006).

c. **Book Chapters**

6. Sarnthein, M.S. & J.D. Stanford. Basal sauropodomorpha: historical and recent phylogenetic developments. In: *The Northern North Atlantic: A Changing Environment.* Schafer, P.R. & W. Schluter (Ed.), Springer, Berlin, Germany, p. 365–410 (2000).
7. Smolen, J.E. & L.A. Boxer. Functions of Europhiles. In: *Hematology, 4^th^ ed*. Williams, W.J., E. Butler & M.A. Litchman (Ed.), McGraw Hill, New York, USA, p. 103–101 (1991).

Tables, with concise but self-explanatory headings must be numbered according to the order of citation (like **Table 1**, **Table 2**). Round off data to the nearest three significant digits. Provide essential explanatory footnotes, with superscript letters or symbols keyed to the data. Do not use vertical or horizontal lines, except for separating column heads from the data and at end of the Table.

Figures may be printed in two sizes: column width of 8.0 cm or page width of 16.5 cm; number them as **Fig. 1**, **Fig. 2**, … in the order of citation in the text. Captions to Figures must be concise but self-explanatory. Laser printed line drawings are acceptable. Do not use lettering smaller than 9 points or unnecessarily large. Photographs must be of high quality. A scale bar should be provided on all photomicrographs.

Declaration: Provide a declaration that: (i) the results are original; (ii) the same material is neither published nor under consideration elsewhere; (iii) approval of all authors has been obtained; and, (iv) in case the article is accepted for publication, its copyright will be assigned to *Pakistan Academy of Sciences*. Authors must obtain permission to reproduce, where needed, copyrighted material from other sources and ensure that no copyrights are infringed upon.

Reviewers: Authors may suggest four relevant reviewers, two local and two from scientifically advanced countries.

Online Manuscript Submission:

Proceedings of the Pakistan Academy of Sciences is using a web-enabled online manuscript submission and review system, which offers authors the option to track the progress of manuscripts in real time.

For online submission, authors are required to register at website of the Academy (www.paspk.org) by clicking the “Journal and Publications” tab.

Manuscripts must be submitted in Microsoft Word (.doc or .docx format; **pdf** files not acceptable). Figures can be submitted in Word format, TIFF, GIF, JPEG, EPS, PPT.

NOTE: In case of online submission, authors are not required to submit the manuscript as email attachment.

In case of any difficulty while submitting your manuscript, please get in touch with:

Editor-in-Chief
Pakistan Academy of Sciences
3-Constitution Avenue, G-5/2, Islamabad, Pakistan
Email: editor@paspk.org
Tel: +92-51-920 7140
Website: www.paspk.org

PROCEEDINGS
OF THE PAKISTAN ACADEMY OF SCIENCES:
A. Physical and Computational Sciences

CONTENTS

Volume 53, No. 2, June 2016 **Page**

Research Articles

PAKISTAN ACADEMY OF SCIENCES, ISLAMABAD, PAKISTAN

HEC Recognized, Category X; PM&DC Recognized

Website: www.paspk.org