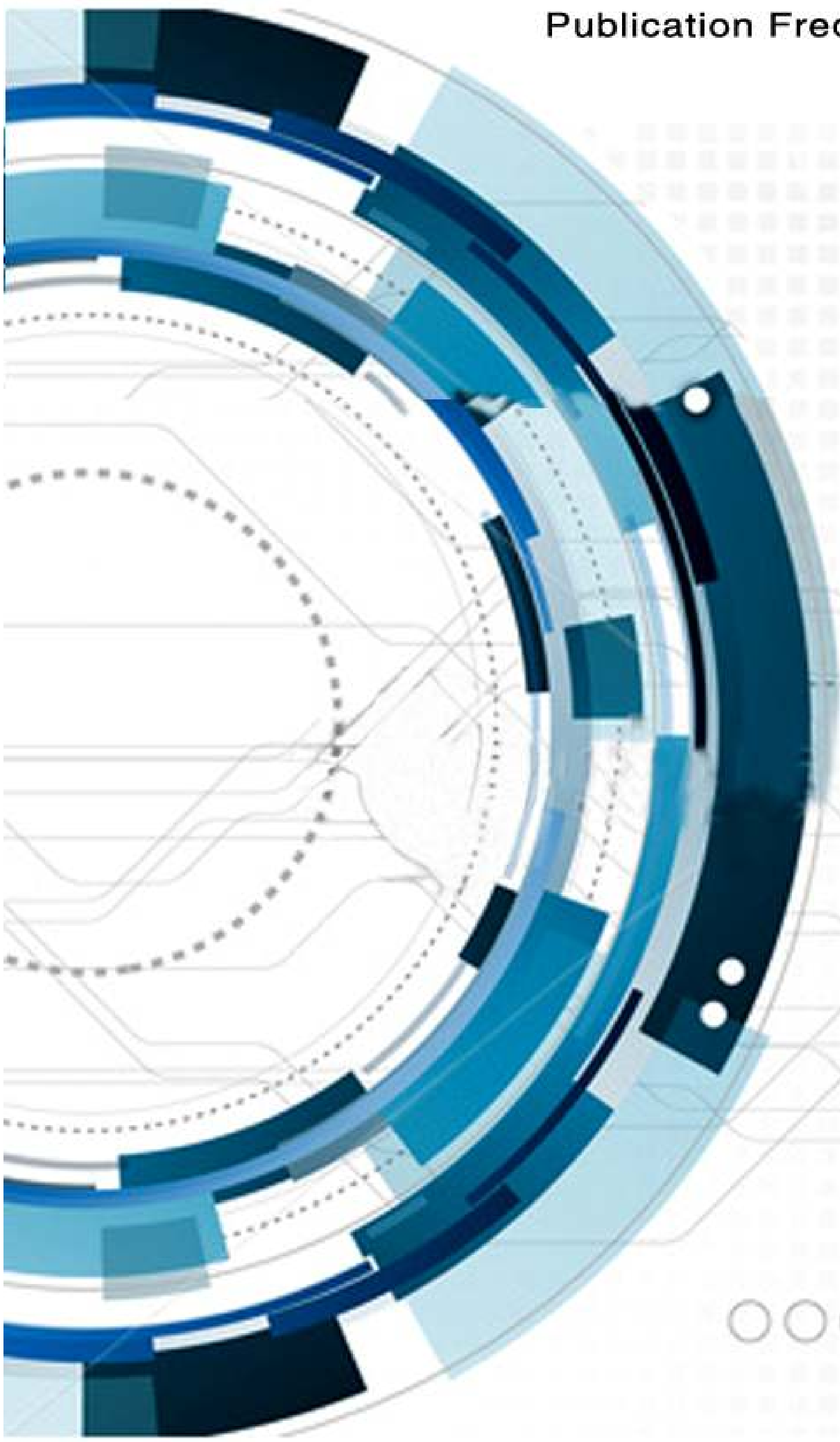
Editor-in-Chief
Dr. Kouroush Jenab

INTERNATIONAL JOURNAL OF ENGINEERING (IJE)

ISSN : 1985-2312

Volume 5 ▪ Issue 3 ▪ August 2011
Publication Frequency: 6 Issues / Year



CSC PUBLISHERS
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF ENGINEERING (IJE)

VOLUME 5, ISSUE 3, 2011

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-2312

International Journal of Engineering is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJE Journal is a part of CSC Publishers
Computer Science Journals
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF ENGINEERING (IJE)

Book: Volume 5, Issue 3, August 2011
Publishing Date: 31-08-2011
ISSN (Online): 1985-2312

IJE Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2011

EDITORIAL PREFACE

This is the third issue of volume five of International Journal of Engineering (IJE). The Journal is published bi-monthly, with papers being peer reviewed to high international standards. The International Journal of Engineering is not limited to a specific aspect of engineering but it is devoted to the publication of high quality papers on all division of engineering in general. IJE intends to disseminate knowledge in the various disciplines of the engineering field from theoretical, practical and analytical research to physical implications and theoretical or quantitative discussion intended for academic and industrial progress. In order to position IJE as one of the good journal on engineering sciences, a group of highly valuable scholars are serving on the editorial board. The International Editorial Board ensures that significant developments in engineering from around the world are reflected in the Journal. Some important topics covers by journal are nuclear engineering, mechanical engineering, computer engineering, electrical engineering, civil & structural engineering etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 5, 2011, IJE appears in more focused issues. Besides normal publications, IJE intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

The coverage of the journal includes all new theoretical and experimental findings in the fields of engineering which enhance the knowledge of scientist, industrials, researchers and all those persons who are coupled with engineering field. IJE objective is to publish articles that are not only technically proficient but also contains information and ideas of fresh interest for International readership. IJE aims to handle submissions courteously and promptly. IJE objectives are to promote and extend the use of all methods in the principal disciplines of Engineering.

IJE editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJE. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJE provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Engineering (IJE)

TABLE OF CONTENTS

Volume 5, Issue 3, August 2011

Pages

Analysis of Hand Anthropometric Dimensions of Male Industrial Workers of Haryana State

Arunesh Chandra chandra_arunesh@yahoo.co.in
Faculty of Engineering/Mechanical
Engineering Department/
Krishna Institute of Engineering & Technology
Ghaziabad, UttarPradesh, 201206, India

Pankaj Chandna pchandna08@gmail.com
Faculty of Engineering/Mechanical
Engineering Department/
National Institute of Technology
Kurushetra, Haryana, 136119, India

Surinder Deswal sdeswal@nitkkr.ac.in
Faculty of Engineering/Civil
Engineering Department/
National Institute of Technology
Kurushetra, Haryana, 136119, India

Abstract

The purpose of this paper is to analyse the thirty-seven hand anthropometric characteristics of the industrial worker of the Haryana state. A survey of convenience sample of eight hundred and seventy eight male industrial workers was conducted in the year 2009. Paper contains data from all the four divisions of Haryana state of India and from the five age groups. Minimum, maximum, mean, standard deviation, skewness, coefficient of variation, 5th, 50th, and 95th percentile for each hand anthropometric dimension were calculated for the entire state. The normality assumption was evaluated for each hand dimension, separately. It was found that in most hand dimensions there were differences between five age groups. Additionally, the statistical analysis was carried out to correlate various hand dimensions and to obtain prediction equation between different variables. It has been found that most of the hand dimensions are correlated significantly with each other. The data gathered may be used for the design of hand tools, gloves, machine access spaces and hand-held devices and for selection of hand tools for use by Industrial worker working in the Haryana state of India.

Keywords: Hand Anthropometric Measurements, Industrial Worker, Hand Tools, Prediction Equation

1. INTRODUCTION

The economic growth and technological improvements have lead to greater demand and development of machines and devices used in industrial settings. With these dramatic changes there has also been greater interaction between man and machines. Anthropometric data are one of essential factors in designing machines and devices as described by [1 & 2]. Incorporating such information would yield more effective designs, ones that are more user friendly, safer, and enable higher performance and productivity. According to [3 & 4] the lack of properly designed machines and equipments may lead to lower work performance and higher incidence to work related injuries [5] have discussed that for years, anthropometry has been used in national sizing surveys as an indicator of health status. Anthropometric measurement of human limbs plays an important role in design of workplace, clothes, hand tools, manual tasks or access spaces for the hand and many products for human use.

Many studies have been conducted in the past to study the hand anthropometry. The depth and breadth of each segment of the hand were measured at points that were spaced at equal distance between the joints of the hand by [6]. Data on the mean length of the proximal and middle phalangeal segments for the fingers was published by [7]. Also [8] described that the interaction of handle size and shape with the kinematics and anthropometry of the hand have a great effect on hand posture and grip strength. Anthropometric survey measuring 18 dimensions of the right hand female workers living in Western Nigeria was conducted by [9] and the means of the collected data were compared with those females from USA, UK and Hongkong. Grip tasks for six subjects were studied using the hand measurement system by [10] the result showed that the flexion angle for the five fingers decreased with increasing grip span. [11 & 12] have stressed the importance of interplay of hand anthropometry and handle size or shape in influencing hand posture, grip span or grip strength. [13] estimated internal biomechanical loads of the hand from external loads and finger lengths that were themselves estimated from measured hand length and breadth; and found that hand anthropometric measurements, especially palm width, are better predictors of hand strength than stature and body weight. The effects on hand grip forces by relatively small changes in hand or handle size have also been demonstrated by [14] for torquing on cylinders [15] for gripping cylinders and [16] for gripping and squeezing on parallel handles of a standard handgrip dynamometer. Hence, measurement of small difference in hand size is important in understanding gripping forces. An important implication of the above discussion is that the anthropometry of the hand must be known for any target population for whom hand tools and other manual devices are to be designed. [17] stated that today, there is a growing demand among professional hand tool users to have ergonomically designed product. Further [18, 19 & 20] have discussed that poor ergonomic hand tools design is a well known factor contributing to biomechanical stresses and increasing the risk of cumulative trauma and carpal tunnel syndrome disorders of users. According to [21]) hand anthropometry is useful for determining various aspects of industrial machineries so as to design the equipment and machines for better efficiency and more human comfort. [22] discussed the potentially harmful effects of ignoring anthropometric differences between populations may be manifested when a developing nation, for example, imports equipment from a developed nation since the latter tends to design their equipment based on the anthropometric data of their own population. Reliable data on the association between hand injuries or disorders and hand anthropometry are almost absent in the developing countries. According to [23 & 24] the continued reliance on muscular power in tool use, in developing countries, and the widespread use of hand tools that do not fit the hands properly results in problems of health, safety and task performance. Further data on relevant anthropometric dimensions of the populations of the importing countries for equipment design may help alleviate the problems. Only a limited work has been reported in connection of hand anthropometry data for the populations of developing countries by [25, 26, 27, 28, 29 & 30].

Keeping the above-mentioned factors in consideration, the present analysis is an attempt to study the impact of collected hand anthropometric data of male industrial worker of Haryana state. As Haryana state of India has total geographical area of 44212 sq. meters. As per Census data 2001 male population of state is 11364000 with about total 498656 (5%) of male population working in almost about 72643 registered industrial units with output @ 6430 Crores with major SME (small manufacturing enterprises) clusters and SEZ (small economic zone) in the Haryana State of India. These movements and others provide incentives for foreign suppliers and investors to open factories and service sectors in Haryana state of India. Many of the industries being developed, therefore, would depend heavily on tools and equipment imported from IC (Industralized Countries) with the negative consequences as described above, if no attempt will be made to match equipment design with human characteristics. The present study thus represents an effort for analysing hand anthropometry data of male industrial worker. The data from this study will also help to understand the anatomical relationships among the various segments of the hand within the Haryana Industrial worker population.

2. METHODS

2.1 Subjects and Apparatus

Sets of thirty-seven hand dimensions were measured for each industrial worker. Selection of these dimensions were made on the basis of their relevance to the design of industrial tools, machine guarding and other manual equipments, and also because they have been measured in previous research studies in different populations. The figures of the hand dimensions are provided in figures 1(a) and 1(b). A total of 878-convenience sample of participants were measured from thirty-eight small and medium scale industries located in different divisions of the state. The range included companies from the automobile, tools and instruments, railway workshop, agricultural and metal sectors, among other, mainly located in the four different divisions (Ambala, Rohtak, Gurgaon and Hisar) of the Haryana state of the India. Subjects were selected according to their availability and willingness to participate without payment or any other kind of reward they were informed with the objectives of the study, anthropometric dimensions, clothing requirements, measurements procedures and freedom to withdraw. Age of the subjects varied between 18 and 62 years old with an average age of 37.91 years, whereas average stature height and body weight of the subjects was found out to be 1653.23 mm and 65.14 kg respectively. The sample comprised essentially individuals from industry. Underlying the choice of subjects from industry is the fact that this account for approximately 5% of active adult male population of Haryana state (Census, 2001). The methods of hand anthropometric measurements were same as stated by [31 & 32]. Regular measurement tools are used such as Hardenpen anthropometer for stature measurement and arm length measurement, small anthropometer for elbow length measurement, digital vernier caliper for length, breadth and depth measurement of hand, measuring tape for circumferential measurements, a wooden cone designed locally and specially to measure internal grip diameter, inner caliper for measurement of grip span and the body weight was measured by portable weighing digital scale. Table 1 describes the age distribution of the sample of the subjects measured.

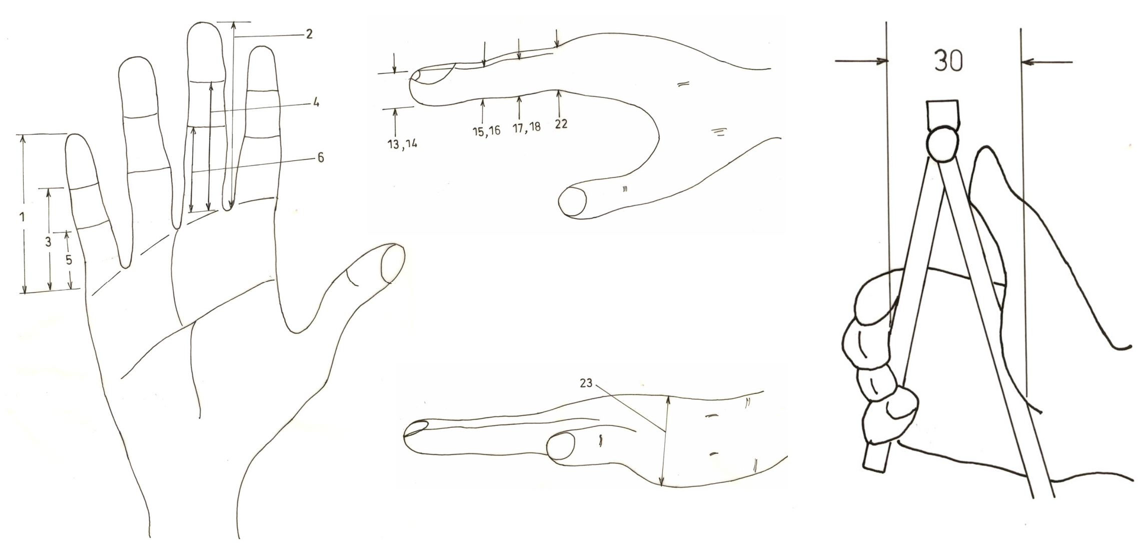


FIGURE 1(A): Selected right hand Anthropometric dimensions of Male Industrial worker Defined In Table 3

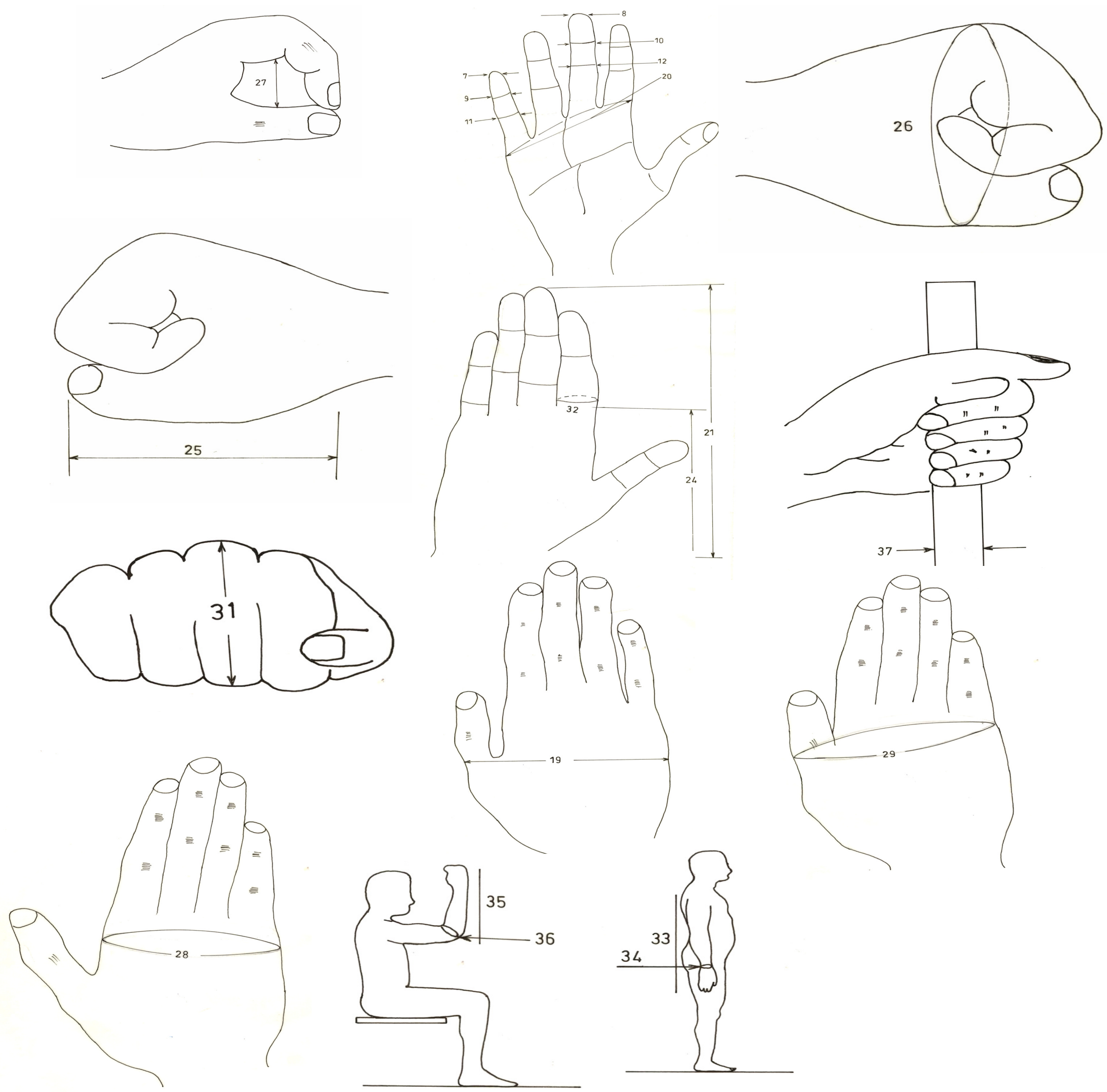


FIGURE 1(B): Selected right hand Anthropometric dimensions of Male Industrial worker Defined In Table 3

Age Group (Years)	Male Industrial Worker	
	Number	Percentage
18 – 25	133	15.15
26 – 35	253	28.82
36 – 45	221	25.17
46 – 55	218	24.83
56-Above	53	6.04

TABLE 1: Age distribution of Subjects

3. RESULTS

According to [33 & 34] there are many factors in human measurements that intervene as sources of error and results can be systematically different in spite of the measures being highly trained.

In anthropometric research the measurer cannot perceive the anomalous measures, as the norm has a very wide range and the size differences among the subjects of a sample are much higher than the accuracy of experimental devices, sometimes a factor of 10 or higher. Thus the data collected was further analyzed using SPSS statistical package (version 16.0) for normality distribution of each hand dimension, using the Kolmogorov-Smirnov and using the Shapiro-Wilk test at the 5% level of significance, the results of the tests are shown in table 2. Outputs are also obtained from box-plots generated from the explore command and the extreme outliers that is 1.77% of the collected readings are rejected for further analysis as they are not following the normal distribution curve as these may be systematic or bias errors which are possible which may not be clearly noticeable and occasionally these may be systematic errors in the measurement processes which could have a significant effect on both mean values of experimental variable and their standard deviation could cause mistaken conclusions over considered population.

S.No.	Measured Parameter	Kolmogorov Smirnov		Shapiro-Wilk	
		Statistic	Significance	Statistic	Significance
1	Age	0.072	0.000	0.975	0.000
2	Stature height	0.069	0.000	0.971	0.000
3	Weight	0.052	0.004	0.993	0.030
4	Finger tip to root digit 5	0.041	0.056	0.994	0.042
5	First joint to root digit 5	0.039	0.092	0.994	0.063
6	Second joint to root digit 5	0.049	0.009	0.991	0.008
7	Finger tip to root digit 3	0.063	0.000	0.982	0.000
8	First joint to root digit 3	0.039	0.083	0.991	0.004
9	Second joint to root digit 3	0.048	0.011	0.991	0.006
10	Breadth at tip digit 5	0.035	0.200	0.992	0.013
11	Breadth at first joint digit 5	0.044	0.032	0.993	0.035
12	Breadth at second joint digit 5	0.045	0.025	0.991	0.005
13	Breadth at tip digit 3	0.032	0.200	0.995	0.124
14	Breadth at first joint digit 3	0.050	0.007	0.990	0.003
15	Breadth at second joint digit 3	0.045	0.025	0.991	0.005
16	Depth at tip digit 5	0.055	0.002	0.987	0.000
17	Depth at first joint digit 5	0.042	0.044	0.994	0.075
18	Depth at second joint digit 5	0.070	0.000	0.977	0.000
19	Depth at tip digit 3	0.051	0.006	0.993	0.026
20	Depth at first joint digit 3	0.044	0.031	0.989	0.001
21	Depth at second joint digit 3	0.044	0.032	0.990	0.002
22	Grip span	0.051	0.006	0.994	0.066
23	Max. breadth of the hand	0.092	0.000	0.978	0.000
24	Breadth of the knuckles	0.079	0.000	0.984	0.000
25	Hand length	0.084	0.000	0.986	0.000
26	Palm length	0.077	0.000	0.989	0.002
27	Depth of the knuckles	0.136	0.000	0.963	0.000
28	Max. depth of the hand	0.070	0.000	0.989	0.002
29	Fist length	0.105	0.000	0.987	0.000
30	First phalanx digit 3 length	0.105	0.000	0.967	0.000
31	Fist circumference	0.064	0.000	0.991	0.007
32	Hand circumference	0.072	0.000	0.988	0.000
33	Max. hand circumference	0.057	0.001	0.990	0.003
34	Index finger circumference	0.127	0.000	0.972	0.000
35	Wrist circumference	0.087	0.000	0.990	0.003
36	Arm length	0.077	0.000	0.982	0.000
37	Elbow length	0.050	0.006	0.986	0.000
38	Elbow flexed	0.070	0.000	0.989	0.001
39	Max. internal grip diameter	0.175	0.000	0.934	0.000

40	Middle finger palm grip diameter	0.196	0.000	0.913	0.000

TABLE 2: Comparison of the empirical distribution of the sample vs. the theoretical (Normal) distribution for Male Industrial Worker

With consideration of normal distribution table 3 provides the minimum, maximum, mean, standard deviation, coefficient of variation, skewness of each hand dimension and the values of each hand dimension at the 5th, 50th, and 95th percentile.

S. No	Hand dimensions	Min.	Max.	Mean	SD	CV	Skew ness	Percentile		
								5th	50th	95th
1	Finger tip to root digit 5	49.79	68.10	59.13	3.39	5.73	-0.117	52.97	59.95	66.89
2	First joint to root digit 5	27.31	41.58	34.23	2.76	8.06	-0.040	28.16	34.37	39.33
3	Second joint to root digit 5	12.93	22.55	17.52	1.96	11.19	0.046	14.12	17.45	21.53
4	Finger tip to root digit 3	69.79	90.80	79.05	4.31	5.45	0.384	71.44	79.18	88.41
5	First joint to root digit 3	43.76	60.51	52.06	3.54	6.80	0.091	45.13	52.51	59.36
6	Second joint to root digit 3	19.46	32.41	25.53	2.71	10.61	0.139	21.24	25.72	30.52
7	Breadth at tip digit 5	10.62	15.84	12.97	1.04	8.02	0.021	11.22	13.13	15.28
8	Breadth at first joint digit 5	12.72	17.60	15.10	0.93	6.16	-0.024	13.53	15.27	17.14
9	Breadth at second joint digit 5	14.73	19.79	17.06	0.99	5.80	0.276	15.50	17.17	19.26
10	Breadth at tip digit 3	12.85	18.56	15.79	1.12	7.09	-0.082	13.54	16.02	18.07
11	Breadth at first joint digit 3	15.07	19.64	17.35	0.90	5.19	0.191	15.76	17.49	19.53
12	Breadth at second joint digit 3	17.90	22.45	20.21	0.94	4.65	0.187	18.37	20.27	22.23
13	Depth at tip digit 5	9.46	13.86	11.37	0.86	7.56	0.293	10.02	11.53	13.45
14	Depth at first joint digit 5	11.22	16.31	13.70	0.99	7.23	0.078	12.06	13.78	15.84
15	Depth at second joint digit 5	13.84	19.97	16.50	1.24	7.51	0.388	14.57	16.55	19.32
16	Depth at tip digit 3	10.32	15.35	12.99	0.98	7.54	-0.103	11.39	13.17	14.99
17	Depth at first joint digit 3	12.83	17.85	15.51	1.13	7.29	-0.036	13.60	15.69	17.84
18	Depth at second joint digit 3	16.53	22.30	19.08	1.13	5.92	0.310	17.40	19.18	21.47
19	Grip span	82.32	114.66	98.07	6.30	6.42	-0.019	86.71	99.15	109.56
20	Max. breadth of the hand	95.00	110.00	101.83	3.38	3.32	0.278	95.00	102.00	110.00
21	Breadth of the knuckles	78.00	92.00	84.85	2.82	3.32	0.082	80.00	85.00	92.00
22	Hand length	170.00	202.00	185.77	6.32	3.40	0.216	175.00	187.00	201.00
23	Palm length	94.00	118.00	105.59	4.57	4.33	0.188	97.00	106.00	115.00
24	Depth of the	24.00	32.00	28.04	1.68	5.99	0.010	25.00	28.00	31.00

	knuckles									
25	Max. depth of the hand	35.00	54.00	44.62	3.41	7.64	0.071	40.00	45.00	51.00
26	Fist length	89.00	113.00	100.05	4.99	4.99	0.009	92.00	101.00	110.00
27	First phalanx digit 3 length	60.00	74.00	65.85	2.92	4.43	0.442	62.00	66.00	72.00
28	Fist circumference	252.00	305.00	277.65	10.57	3.81	-0.093	259.00	280.00	305.00
29	Hand circumference	225.00	265.00	243.82	8.52	3.49	-0.100	228.00	245.00	262.00
30	Max. hand circumference	310.00	379.00	344.50	12.87	3.74	-0.251	319.00	346.00	373.00
31	Index finger circumference	60.00	77.00	67.28	3.76	5.59	-0.075	61.00	68.00	74.00
32	Wrist circumference	149.00	185.00	164.54	6.92	4.21	0.153	152.00	165.00	180.00
33	Arm length	692.00	847.00	771.16	27.36	3.55	-0.025	727.00	776.00	821.00
34	Elbow length	423.00	501.00	459.91	15.70	3.41	0.260	434.00	462.00	493.00
35	Elbow flexed	223.00	320.00	263.72	18.11	6.87	0.113	234.00	266.00	295.00
36	Max. internal grip diameter	35.00	52.00	42.68	4.05	9.49	0.163	35.00	44.00	50.00
37	Middle finger palm grip diameter	12.00	22.50	16.33	2.47	15.12	0.188	12.50	17.50	21.00

TABLE 3: Hand Anthropometric data of sample (N=878, All measurements are in Millimeter)

In addition to the above analysis the male industrial worker groups were divided further into five age groups of 18-25, 26-35, 36-45, 46-55, and above 56 years, for which mean and standard deviations, were calculated separately as shown in table 4. Based on these values, the 5th, 50th and 95th percentiles can be calculated separately.

S. No	Hand dimensions	18-25 (n=133)		26-35 (n=253)		36-45 (n=221)		46-55 (n= 218)		56-Above (n= 53)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	Finger tip to root digit 5	60.71	3.79	60.12	4.00	59.75	4.51	59.37	3.47	58.16	4.06
2	First joint to root digit 5	35.49	3.05	34.78	2.48	34.13	3.37	34.06	3.47	33.35	2.78
3	Second joint to root digit 5	18.14	2.23	17.81	1.85	17.64	2.31	17.36	1.96	16.82	2.00
4	Finger tip to root digit 3	79.95	4.58	80.06	5.31	79.16	4.81	79.77	4.98	77.19	5.49
5	First joint to root digit 3	52.92	3.51	53.48	3.97	51.98	4.06	52.05	3.67	49.62	4.84
6	Second joint to root digit 3	26.14	2.62	26.25	3.07	25.52	2.38	25.93	2.63	23.49	3.24
7	Breadth at tip digit 5	12.82	1.19	12.87	1.26	13.20	1.20	13.69	1.09	13.28	1.25
8	Breadth at first joint digit 5	15.03	1.01	15.03	1.02	15.34	1.03	15.64	1.13	15.36	1.11
9	Breadth at second joint digit 5	16.74	0.96	16.80	1.04	17.28	1.07	17.75	1.04	17.81	1.23
10	Breadth at tip digit 3	15.36	1.38	15.76	1.49	16.12	1.18	16.50	1.26	15.94	1.14
11	Breadth at first joint digit 3	17.13	0.98	17.19	1.14	17.58	1.08	18.02	1.09	17.71	1.13
12	Breadth at second joint	19.80	1.21	20.08	1.17	20.44	1.18	20.62	1.10	20.08	1.18

	digit 3										
13	Depth at tip digit 5	11.24	0.82	11.40	1.01	11.56	1.13	11.92	0.99	11.98	0.87
14	Depth at first joint digit 5	13.40	0.93	13.61	1.00	13.84	1.42	14.41	1.04	14.01	0.89
15	Depth at second joint digit 5	16.11	1.14	16.47	1.58	16.84	1.63	17.15	1.21	16.91	1.22
16	Depth at tip digit 3	12.76	0.97	12.88	1.13	13.25	1.03	13.52	1.08	13.38	1.09
17	Depth at first joint digit 3	15.30	0.97	15.31	1.30	15.84	1.33	16.10	1.25	16.32	1.07
18	Depth at second joint digit 3	19.00	1.24	18.92	1.19	19.31	1.23	19.75	1.35	19.66	0.92
19	Grip span	99.93	6.58	98.08	7.08	98.54	6.60	99.38	6.72	94.64	5.04
20	Max. breadth of the hand	101.41	4.51	101.85	4.11	102.68	3.80	103.54	4.26	102.26	4.30
21	Breadth of the knuckles	84.56	3.24	85.26	3.19	85.61	3.48	86.21	3.84	85.08	3.95
22	Hand length	186.41	8.32	187.25	8.28	188.30	7.96	188.10	7.88	182.82	7.52
23	Palm length	107.40	5.12	105.39	5.25	106.40	5.19	106.21	5.61	102.46	5.46
24	Depth of the knuckles	27.43	1.71	27.64	1.68	28.59	1.82	28.46	2.01	27.88	2.00
25	Max. depth of the hand	43.96	3.12	44.08	3.17	45.76	3.73	46.24	4.01	44.46	2.66
26	Fist length	99.79	5.87	100.39	6.03	101.13	5.18	101.52	4.92	99.25	4.67
27	First phalanx digit 3 length	66.48	3.15	66.56	3.62	66.13	2.98	66.69	2.99	65.23	3.07
28	Fist circumference	275.41	11.14	277.65	13.33	281.00	13.08	284.39	13.42	278.98	15.97
29	Hand circumference	242.08	10.22	243.44	10.84	247.99	9.29	248.37	11.52	239.94	10.21
30	Max. hand circumference	342.91	13.59	344.92	17.97	348.32	13.63	347.83	18.22	345.77	17.27
31	Index finger circumference	65.52	3.90	66.27	3.74	68.82	3.34	69.59	3.92	68.24	3.34
32	Wrist circumference	161.59	7.06	163.65	7.50	167.14	7.56	168.72	8.84	165.18	9.87
33	Arm length	772.81	30.49	777.05	33.50	773.12	31.06	773.60	27.53	768.04	27.73
34	Elbow length	463.21	17.58	462.43	20.93	462.61	16.20	463.50	16.87	456.50	13.91
35	Elbow flexed	262.79	19.78	263.30	18.24	268.10	16.29	267.75	20.13	267.64	19.92
36	Max. internal grip diameter	43.82	3.78	43.80	4.13	42.16	4.84	43.44	4.46	42.16	5.03
37	Middle finger palm grip diameter	17.09	2.69	16.87	2.67	16.21	2.79	16.17	2.65	15.51	2.26

TABLE 4: Hand Anthropometric data of sample classified by Age (Mean values and standard deviation) all measurements are in millimeter

Table 5 shows the correlation coefficients between different hand anthropometric dimensions. These coefficients were calculated to see to what extent these dimensions are related to each other and to what extent equipment design decisions could be based on such correlation. The simple and multiple regression analyses were done between hand length, hand circumference and other hand dimensions in order to find out the best set of predictors related to hand length and hand circumference and are provided in Table 6(a) and 6(b).

4. DISCUSSIONS

From 32486 measured hand variables, 578 measured readings are rejected using stem-and-leaf plots, histograms and box plots on SPSS software, based on the modifications of the Kolmogorov-Smirnov and Shapiro-Wilk test as it is suitable for continuous distribution to examine the test of normality distribution of data. Thus rejecting 1.77% (578) sample data which may due to certain type of error while measuring the hand dimensions the result obtains indicates that the hand variable have statistical distribution that can fit closely to normal distribution curve, as usual in from the result of the normality test given in Table 2. These test indicates that the thirty five out of thirty seven hand variables were normal with some deviation in other two variables, these two variable maximum internal grip diameter and middle finger palm grip diameter are also approximately normal ($p < 0.05$) knowing that a dimension is normal makes it possible to easily derive percentiles in the distribution using the standard normal (Z) table. Otherwise, the cumulative distribution may be used. The frequency distribution would look like a symmetrical bell-shaped or normal curve, with most subjects having values in the mid range and with a smaller number of subjects with high and low scores. As all the hand anthropometric dimensions follow a normal distribution curve and errors made in using the normal distribution are either not significant, statistically or are of little practical importance thus the probability density function of the underlying distribution is estimated based on a sample from the population without any prior knowledge of the mean, variance etc. of the population

Table 3 presents the summary data obtained for mean and standard deviation, as well as other important statistical information namely minimum, maximum, and coefficient of variation, skewness, and important percentile values for all the hand measurements of the male industrial worker. Coefficient of variation (the ratio of standard deviation to mean) among the thirty seven hand dimensions ranged from 3.32 to 15.12 % with 34 of them below 10% far lower than we can assume or suggested by [35]. As the skewness of all the thirty-seven hand dimension is less than plus or minus one (<+/- 1.0); thus hand dimension is atleast approximately normal and skewness is not significantly different from normal, and hence we can use the mean, standard deviation and different percentile values to easily determine the proportion of the population who fall within a specific range of value for a given hand dimension. These values may also be used for comparison with those published for other population.

The values of mean and standard deviation (SD) for five age groups of male industrial workers surveyed, namely 18-25, 26-35, 36-45, 46-55, and > 56 years; pertaining to thirty seven hand anthropometric dimensions were calculated and are presented in table 4. The data show an increase in most hand dimensions in the middle age before declining with an increasing age. This classification revealed that there are clear differences between the five groups. Moreover young and middle aged worker are smaller than 56 and above age industrial worker in breadth at second joint digit 5, depth at tip digit 5 and depth at first joint digit 3. However in other hand dimensions, the 56 and above age industrial worker are generally smaller than both the young and the middle aged. Figure 2 illustrate the average values obtained of hand length and hand circumference for five different age groups. This shows that, hand length and hand circumference vary significantly with age. These differences are very important and should be taken into consideration in designing the hand tools or equipment that should be controlled by hands of different age groups. [36] and many others researchers support these findings that anthropometric data have indicated difference among age groups. It will be interesting to find out whether these are significant difference between different age groups most of the hand dimensions with significant differences with were not related to vertebral compression. The exact reason for the significant differences remain unknown we could not identify them in this study. The differences found in the hand anthropometric dimensions of the different age groups emphasize the usefulness of this study and of the results presented herein.

Correlations among measured hand segments were performed among hand length and hand circumference. Testing the significance of correlation revealed that almost all values are significant and positively correlated between the hand length and hand circumference, suggest that it is possible to predict hand dimensions with 95% confidence, by measuring the hand length and hand circumference alone. Linear regression equations are provided in Table 6(a) and 6(b) respectively. The statistically significant correlation between the hand lengths (L) related variables are coded by Y_1 to Y_{28} and the hand circumference (C) related variables are coded by Y_{29} to Y_{34}.

Code	Variable	Coefficient of Correlation	Prediction Equation
L	Hand length	-	-
Y_1	Finger tip to root digit 5	0.602**	$Y_1 = 0.4346L - 20.736$
Y_2	First joint to root digit 5	0.486**	$Y_2 = 0.3866L - 35.296$
Y_3	Second joint to root digit 5	0.299**	$Y_3 = 0.2098L - 21.258$
Y_4	Finger tip to root digit 3	0.697**	$Y_4 = 0.5322L - 18.892$
Y_5	First joint to root digit 3	0.610**	$Y_5 = 0.4082L - 23.512$
Y_6	Second joint to root digit 3	0.470**	$Y_6 = 0.2922L - 28.622$
Y_7	Breadth at tip digit 5	0.110*	$Y_7 = 0.12L - 8.9$
Y_8	Breadth at first joint digit 5	0.139**	$Y_8 = 0.1276L - 7.886$
Y_9	Breadth at second joint digit 5	0.181**	$Y_9 = 0.12L - 5.09$
Y_{10}	Breadth at tip digit 3	0.038	$Y_{10} = 0.1478L - 11.988$
Y_{11}	Breadth at first joint digit 3	0.168**	$Y_{11} = 0.125L - 6.03$
Y_{12}	Breadth at second joint digit 3	0.272**	$Y_{12} = 0.1518L - 8.458$
Y_{13}	Depth at tip digit 5	0.060	$Y_{13} = 0.0952L - 5.792$
Y_{14}	Depth at first joint 5	0.163**	$Y_{14} = 0.1292L - 9.452$
Y_{15}	Depth at second joint 5	0.152**	$Y_{15} = 0.1746L - 14.786$
Y_{16}	Depth at tip digit 3	0.022	$Y_{16} = 0.1188L - 8.688$
Y_{17}	Depth at first joint digit 3	0.141**	$Y_{17} = 0.1332L - 8.642$
Y_{18}	Depth at second joint digit 3	0.243**	$Y_{18} = 0.1386L - 6.046$
Y_{19}	Grip span	0.419**	$Y_{19} = 0.6674L - 24.464$
Y_{20}	Maximum breadth of the hand	0.466**	$Y_{20} = 0.46L + 17.4$
Y_{21}	Breadth of the knuckles	0.415**	$Y_{21} = 0.38L + 15.2$
Y_{22}	Palm length	0.290**	$Y_{22} = 0.6L - 6.0$
Y_{23}	Depth of knuckles	0.411**	$Y_{23} = 0.18L - 4.8$
Y_{24}	Maximum depth of hand	0.254**	$Y_{24} = 0.38L - 25.8$
Y_{25}	Fist length	0.306**	$Y_{25} = 0.58L - 5.8$
Y_{26}	First phalanx digit 3 length	0.455**	$Y_{26} = 0.34L + 4.6$
Y_{27}	Elbow length	0.607**	$Y_{27} = 1.9796L + 101.2653$
Y_{28}	Arm length	0.582**	$Y_{28} = 3.44L + 141.6$

TABLE 6(a): Coefficient of Correlation between Hand Length and related variables for Haryana State Industrial Workers and the corresponding prediction equation

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
1	1																																				
2	.833**	1																																			
3	.619**	.752**	1																																		
4	.649**	.516**	.312**	1																																	
5	.553**	.550**	.366**	.852**	1																																
6	.378**	.384**	.428**	.640**	.746**	1																															
7	.118**	-.043	.083	.089	-.009	-.006	1																														
8	.157**	-.043	.116*	.057	-.037	.040	.623**	1																													
9	.161**	.042	.143**	.174**	.066	.071	.511**	.684**	1																												
10	.029	-.091*	.055	-.014	-.171**	-.046	.556**	.562**	.404**	1																											
11	.086	.005	.098*	.098*	.025	.111*	.435**	.621**	.619**	.595**	1																										
12	.183**	.105*	.173**	.276**	.233**	.166**	.342**	.468**	.634**	.343**	.592**	1																									
13	.006	-.066	.002	.036	-.077	-.009	.516**	.513**	.534**	.441**	.480**	.507**	1																								
14	.148**	.014	.132**	.170**	.048	.188**	.416**	.513**	.543**	.294**	.417**	.501**	.545**	1																							
15	.219**	.069	.049	.211**	.072	.141**	.356**	.397**	.461**	.350**	.330**	.341**	.447**	.605**	1																						
16	-.050	-.174**	-.057	.000	-.093*	.017	.444**	.478**	.449**	.483**	.460**	.348**	.631**	.515**	.487**	1																					
17	.051	-.061	.025	.101*	.019	.157**	.309**	.500**	.488**	.318**	.535**	.490**	.451**	.596**	.528**	.553**	1																				
18	.160**	.073	.091*	.240**	.186**	.218**	.234**	.362**	.443**	.268**	.517**	.572**	.442**	.555**	.502**	.489**	.662**	1																			
19	.179**	.165**	.027	.171**	.126**	.026	.061	-.085	-.032	-.036	.031	.093*	.173**	-.046	.104*	.058	.009	.083	1																		
20	.304**	.166**	.124**	.299**	.180**	.058	.339**	.362**	.355**	.335**	.315**	.273**	.291**	.224**	.286**	.251**	.257**	.273**	.262**	1																	
21	.250**	.135**	.145**	.389**	.265**	.262**	.326**	.366**	.368**	.345**	.297**	.338**	.371**	.303**	.364**	.302**	.304**	.272**	.181**	.668**	1																
22	.602**	.486**	.299**	.697**	.610**	.470**	.110*	.139**	.181**	.038	.168**	.272**	.060	.163**	.152**	.022	.141**	.243**	.419**	.466**	.415**	1															
23	.393**	.305**	.173**	.329**	.256**	.229**	.077	.077	.072	.000	.086	.149**	.053	.115*	.095*	.017	.164**	.158**	.512**	.377**	.290**	.772**	1														
24	.192**	.073	.118**	.232**	.127**	.135**	.294**	.375**	.434**	.292**	.419**	.460**	.301**	.318**	.308**	.336**	.406**	.374**	.031	.398**	.411**	.254**	.204**	1													
25	.127**	.012	.021	.146**	.092*	.051	.273**	.301**	.384**	.370**	.392**	.398**	.271**	.269**	.212**	.216**	.289**	.346**	.069	.397**	.254**	.205**	.099*	.411**	1												
26	.378**	.294**	.226**	.390**	.413**	.365**	.201**	.180**	.343**	.045	.282**	.391**	.153**	.351**	.383**	.157**	.375**	.408**	.215**	.341**	.306**	.569**	.512**	.362**	.371**	1											
27	.578**	.436**	.278**	.735**	.632**	.423**	.180**	.166**	.300**	.084	.220**	.488**	.230**	.311**	.266**	.118**	.270**	.388**	.191**	.393**	.455**	.717**	.466**	.397**	.284**	.515**	1										
28	.144**	-.056	.010	.319**	.161**	.113*	.370**	.370**	.378**	.234**	.225**	.290**	.233**	.320**	.258**	.245**	.285**	.238**	.104*	.619**	.555**	.390**	.274**	.449**	.388**	.251**	.346**	1									
29	.265**	.082	.114*	.269**	.101*	.079	.432**	.427**	.452**	.463**	.375**	.383**	.371**	.379**	.315**	.314**	.309**	.307**	.125**	.770**	.649**	.388**	.278**	.525**	.596**	.372**	.394**	.682**	1								
30	.388**	.202**	.137**	.327**	.219**	.142**	.322**	.222**	.238**	.252**	.224**	.246**	.227**	.244**	.199**	.167**	.143**	.182**	.269**	.584**	.381**	.503**	.400**	.386**	.435**	.398**	.408**	.510**	.697**	1							
31	.152**	.031	.095*	.207**	.060	.046	.369**	.448**	.567**	.366**	.416**	.494**	.422**	.513**	.423**	.352**	.455**	.406**	.083	.467**	.463**	.242**	.172**	.549**	.547**	.391**	.374**	.496**	.631**	.354**	1						
32	.256**	.085	.156**	.288**	.135**	.124**	.315**	.364**	.513**	.328**	.330**	.456**	.277**	.409**	.328**	.224**	.290**	.359**	-.043	.561**	.506**	.284**	.093*	.473**	.486**	.301**	.373**	.592**	.709**	.450**	.603**	1					
33	.471**	.298**	.160**	.545**	.434**	.298**	.105*	.101*	.114*	.035	.107*	.138**	.043	.096*	.159**	.056	.103*	.194**	.303**	.382**	.365**	.607**	.440**	.168**	.202**	.295**	.514**	.309**	.391**	.436**	.100*	.294**	1				
34	.255**	.133**	.124**	.238**	.117*	-.028	.214**	.269**	.413**	.197**	.277**	.395**	.242**	.264**	.156**	.242**	.216**	.257**	.122**	.533**	.402**	.286**	.223**	.363**	.418**	.251**	.343**	.439**	.579**	.331**	.395**	.579**	.366**	1			
35	.408**	.390**	.191**	.517**	.475**	.270**	-.010	-.056	-.044	-.131**	-.021	.074	-.069	-.098*	-.008	-.175**	-.167**	.016	.383**	.318**	.153**	.585**	.365**	.013	.041	.216**	.359**	.185**	.121**	.334**	-.047	.133**	.446**	.052	1		
36	.290**	.227**	.034	.371**	.306**	.160**	-.004	-.051	-.019	-.111*	-.103*	.080	.024	.018	.099*	-.017	-.002	.032	.399**	0.181**	.120**	.400**	.233**	-.031	.052	.112*	.275**	0.103*	.046	0.231**	.039	.080	.248**	.017	.681**	1	
37	.481**	.297**	.179**	.549**	.442**	.351**	.110*	.071	.185**	-.025	.088	.237**	.048	.143**	.219**	.014	.153**	.251**	.264**	.347**	.320**	.582**	.463**	.226**	.206**	.432**	.580**	.279**	.382**	.424**	.158**	.280**	.799**	.336**	.369**	.220**	1

TABLE 5: Matrix of the Pearson Correlation Coefficients obtained between the different Hand Anthropometric dimensions as per order provided In Figure 1 and Table 3

**Correlation is significant at the 0.01 level (2- tailed) *Correlation is significant at the 0.05 level (2- tailed)

Code	Variable	Coefficient of Correlation	Prediction Equation
C	Hand circumference	-	-
Y_{29}	Maximum hand circumference	0.510**	$Y_{29} = 1.6552C – 65.1724$
Y_{30}	Index finger circumference	0.496**	$Y_{30} = 0.3621C – 19.569$
Y_{31}	Wrist circumference	0.509**	$Y_{31} = 0.8448C – 38.3276$
Y_{32}	Elbow flexed	0.391**	$Y_{32} = 1.7414C – 158.8793$
Y_{33}	Maximum internal grip diameter	0.121**	$Y_{33} = 0.431C – 63.5345$
Y_{34}	Middle finger grip diameter	0.046	$Y_{34} = 0.181C – 27.2845$

TABLE 6(b): Coefficient of Correlation between Hand Circumference and related variables for Haryana State Industrial Workers and the corresponding prediction equation

** Significant at α = 0.01 * Significant at α = 0.05
Note all dimensions in mm

The tests of hypothesis that the intercepts or the slopes are zero were rejected for the level of significance shown in Table 6(a) and 6(b) Predictions should be confined to the ranges of hand length and hand circumference as prescribed by the regression analysis. The minimum and maximum values for hand length were 170 mm and 202 mm respectively and the counter values for the hand circumference were 225 mm and 244 mm respectively. Although this hand anthropometric data will be of great value in practical application it should be noticed that these are static anthropometric measurers. Therefore, the use of such data in design of equipment, tools, and workstation in which functional hand anthropometric data is needed, must be done considering the differences between the two referred types of hand anthropometric data

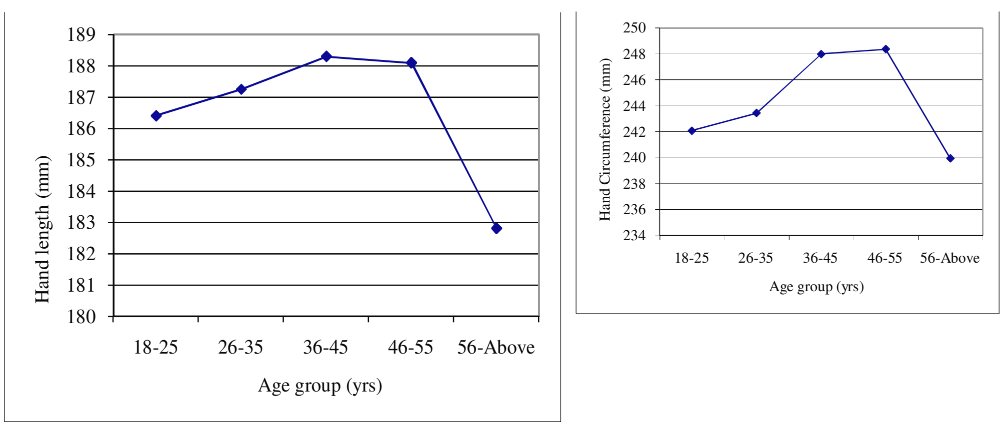


FIGURE 2: Variation of Hand Length and Hand Circumference (Mean Values in mm) for Age Groups (Yrs.) defined

5. CONCLUSIONS

Thirty-seven hand dimensions of eight hundred and seventy eight male industrial workers of Haryana state belonging to thirty-eight industries of Haryana state of India have been analysed in this work. This will be useful for the new designs/design modifications for hand tools, workstations, hand apparel, tools and protective equipment and other practical applications. Mean and standard deviation of the sample of different age groups shows that values of most of the hand anthropometric dimensions are higher in the middle age groups and lower with higher and lower age groups. With respect to the above analysis there are a few important remarks, which need to be emphasized.

- This study investigated assumptions of normality commonly made by designers in establishing workplace, equipment, as well as tool design recommendations and the objective of this analysis is to check precision in anthropometric measures. It was observed that 98.23% of collected reading of 37 hand variables of hand anthropometric dimensions fit closely to a normal distribution curve.

- The correlation coefficients among different hand dimensions were calculated to see to what extent these dimensions are related to each other. It was observed that 77% of correlation coefficients are significant at the 1% level, 5% of the correlation coefficients are significant at 5% level, and 18% of the remaining values are insignificant. Correlation among measured hand segments was performed among hand length and hand circumference and almost all values are significant and positively correlated.
- The sample size used (878) was satisfactory for all variables. Therefore designers for industrial worker of Haryana state can utilize the statistics presented and prediction equations present in this study to set specifications for the system used, such as hand tools and other hand held devices. These prediction equations can be used to predict 34 hand variable dimensions with 95% confidence by measuring the hand length and hand circumference alone.
- There is a need to enlarge the sample size, not only in terms of age range, namely to compensate for low frequency observed below 25 and above 56 years, but also to encompass other occupational groups such as agricultural worker, household worker, constructional workers and of female workers as their numbers are increasing day to day in the state.

6. REFERENCES

[1] B. Mebarki and B.T. Davies. “Anthropometry of Algerian Woman”. International Journal of Industrial Ergonomics 33(12): 1537-1547, 1990.

[2] W.G. Lewis and C.V. Narayan. “Design and sizing of ergonomic handles for hand tools”. Applied Ergonomics 24: 351-356, 1993.

[3] W.E. Botha and R.S. Bridger. “Anthropometric variability, equipment, usability and musculoskeletal pain in a group of nurses in the Western Cape”. International Journal of Applied Ergonomics 29(6): 481-490, 1998.

[4] P. Chandna, S. Deswal and A.Chandra. “An Anthropometric Survey of Industrial Workers of the Northern Region of India” International Journal of Industrial and Systems Engineering 6(1): 110-128, 2010.

[5] G.C. Marks, J.P. Habicht and W.H. Mueller. “Reliability, dependability and precision of anthropometric measurement”. American Journal of Epidemiology 130(3): 578-587, 1989.

[6] B. Buchholz and T.J. Armstrong. “An ellipsoidal representation of human hand anthropometry”. Human Factors 33: 429-441, 1991.

[7] K.N. An, E.Y. Chao, W.P. Cooney and R.L. Linscheid. “Normative model of human hand for biomechanical analysis”. Journal of Biomechanics 12: 775-788, 1979.

[8] B. Buchholz, T.J. Armstrong and S.A. Goldstein. “Anthropometric data for describing the kinematics of human hand”. Ergonomics 35: 261-273, 1992.

[9] O.O. Okunribido and K.A. Olajire. “A survey of hand anthropometry of female rural workers in Ibadan, Western Nigeria”. Ergonomics 11: 2-6, 1999.

[10] M.H. Yun. “Designing for diversity”. In Proceedings of the human factors and ergonomics society. 37th Annual Meeting, Seattle, Washington, 1993.

[11] C. Frannson and J. Winkel. “Hand strength: the influence of grip spans and grip type”. Ergonomics 34: 881-892, 1991.

[12] J.R. Blackwell, K.W. Kornatz and E.M. Heath. “Effect of grip span on maximal grip force and fatigue of flexor digitorum superficial”. Applied Ergonomics 30: 401-405, 1999.

[13] C.B. Irwin and R.G. Radwin. “A new method for estimating hand internal loads from external force measurements”. Ergonomics 51(2): 156-167, 2008.

[14] S.N. Imrhan and K. Farahmand. “Male torque strength in simulated oil rig tasks and the effects of grease-smeared gloves and handle length, diameter and orientation”. Applied Ergonomics 30(5): 455-462, 1999.

[15] K.A. Grant, D.J. Habes and L.L. Steward. “An analysis of hand designs for reducing manual effort: The influence of grip diameter”. International Journal of Industrial Ergonomics 10(3): 199-206, 1992.

[16] M. Eksioglu, J.E. Fernandez and J.M. Twomey. “Predicting peak pinch strength: artificial neural networks vs. regression”. International Journal of Industrial Ergonomics 18: 431-441, 2004.

[17] N.A. Snow and T.J. Newby. “Ergonomically designed job aids”. Performance and Instructional Journal 28: 26-30, 1984.

[18] A. Chandra, P. Chandna, S. Deswal and R. Kumar. “Ergonomics in the Office Environment: A review”. In Proceedings of the International Conference on Energy and Environment. Chandigarh, Haryana, India, 2009.

[19] L. Claudon. “Ergonomics hand tools design: Interview of users. Ergonomics and safety for global business quality and productivity”. In Proceeding of the second international conference ERGON-AXIA 2000. Warsaw, Poland. Central Institute for Labour Protection; Warsaw. (2000),

[20] P. Loslever and A. Ranaivosoa. “Biomechanical and epideminological investigation of carpal tunnel syndrome at work-places with high risk factors”. Ergonomics 36(5): 537-555, 1993.

[21] S.N. Imrhan, M. Nguyen and N. Nguyen. “Hand anthropometry of Americans of Vietnamese origin”. International Journal of Industrial Ergonomics 12: 281-287, 1993.

[22] J.D.A. Abeysekera and H. Shahnavaz. “Body size variability between people in developed and developing countries and its impact on the use of imported goods”. International Journal of Industrial Ergonomics 4: 139-149, 1989.

[23] O. Okunribido. “A survey of hand anthropometry female rural farm workers in Ibadan, Western Nigeria”. Ergonomics 43: 282-292, 2000.

[24] S.K. Kar, S. Ghosh, I. Manna, S. Banerjee and P. Dhara. “An investigation of hand anthropometry of agricultural workers”. Journal of Human Ergology 14(1): 57-62, 2003.

[25] L.T. Gite and B. G. Yadav. “Anthropometric survey for agricultural machinery design: An Indian Case Study”. Applied Ergonomics 20: 191-196, 1989.

[26] A. Nag, P.K. Nag and H. Desai. “Hand anthropometry of Indian women”. Indian Journal Medical Research 117: 260-269, 2003.

[27] S.N. Imrhan and M.G. Contreras. “Hand anthropometry in a sample of Mexicans in the US-Mexico border region”. In Proceedings of the XIX annual Occupational ergonomics and safety conference, Las Vegas, NV: 589-593, 2005.

[28] N. Mandahawi, S. Imrhan, S. Al-Shobaki and B. Sarder. “Hand anthropometry survey for the Jordanian population”. International Journal of Industrial Ergonomics 38: 966-976, 2008.

[29] S.N. Imrhan, M.D. Sarder and N. Mandaharu. “Hand anthropometry in Bangladeshis living in America and comparisons with other populations”. Ergonomics 52: 987-998, 2009.

[30] A. Chandra, P. Chandna and S. Deswal. “Hand Anthropometric Survey of Male Industrial Workers of Haryana State (India)”. International Journal of Industrial and Systems Engineering (In Press 2011).

[31] B.T. Davies. “Female hand dimensions and guarding of machines”. Ergonomics 23(1): 79-84, 1980.

[32] A.J. Courtney and M.K. Ng. “Hong Kong female hand dimensions and machine guarding”. Ergonomics 27(2): 187-193, 1984.

[33] H. Kemper and J. Pieters. “Comparative study of anthropometric measurements of the same subjects in two different institutes”. American Journal of Physical Anthropology 40: 340-344, 1974.

[34] E. Panchon, R. Lobato, F. Sanchez and A. Panchon. “Index for quality control in anthropometric surveys”. International Journal of Industrial Ergonomics 34: 479-482, 2004.

[35] S. Pheasant. “Bodyspace: Anthropometry, Ergonomics and the Design of Work”, Taylor & Francis, (1998).

[36] L.G. Martin and B.J. Soldo. “Racial and Ethnic Differences in the Health of Older Americans”, National Academy Press, Washington, DC, 285-296, 1997.

Averaging Method for PWM DC-DC Converters Operating in Discontinuous Conduction Mode With Feedback

Mohammed S. Al-Numay *alnumay@ksu.edu.sa*
Department of Electrical Engineering
KingSaudUniversity,
Riyadh11421, Saudi Arabia.

N.M. Adamali Shah *anoormuhamed@ksu.edu.sa*
Department of Electrical Engineering
KingSaudUniversity,
Riyadh11421, Saudi Arabia.

Abstract

In this paper, a one-cycle-average (OCA) discrete-time model for PWM dc-dc converter based onclosed-loop control method for discontinuous conduction mode (DCM) is presented. It leads toexact discrete-time mathematical representation of the OCA values of the output signal even atlow frequency. It also provides the exact discrete-time mathematical representation of theaverage values of other internal signals with little increase in simulation time. A comparison of thismodel to other existing models is presented through a numerical example of boost converter.Detailed simulation results confirm the better accuracy and speed of the proposed model.

Keywords: Switched Systems, Pulse-width Modulation (PWM), Power Converter, Discrete TimeModeling, one-cycle-averaging, Sampled Data Model.

1. INTRODUCTION

PWM converters are widely used for operating switch controlled systems. These systems areusually operated in two modes of operation, namely: continuous and discontinuous conductionmodes [5]. The discontinuous conduction mode (DCM) of operation typically occurs in dc/dcconverters at light load.The boundary between the continuous conduction mode (CCM) andDCM depends on the ripple current in the inductor or the ripple voltage in the capacitor. Forlowpower applications, many designers prefer to operate in the DCM in order to avoid the reverserecovery problem of the diode. DCM operation has also been considered a possible solution tothe right-half plane (RHP) zero problem encountered in buck-boost and boost derived topologies. In single-phase ac/dc converters with active power factor correction (PFC), the input inductorcurrent becomes discontinuous in the vicinity of the voltage zero crossing; some PFC circuits areeven purposely designed to operate in DCM over the entire line cycle in order to simplify thecontrol. Proper analytical models for DCM operation of PWM converters are therefore essentialfor the analysis and design of converters in a variety of applications (see [11] and referencestherein). These modes of operation are also very much useful for efficiently extracting maximumpower from the photovoltaic panel (PV) which is another main application [12]. These powerconverters are connected between the PV and load or bus. Due to the variety of applications ofPWM converters operating in DCM, there is a need for an accurate model for the analysis anddesign of such converters. Many efforts have been taken in this view for past three decades [8,11].

Most power electronic design procedures rely on averaging techniques. Averaging techniquesprovide the analytical foundation for most power electronic design procedures of the system level. These techniques are widely used in the analysis and control design of PWM power electronicsystem. In fact classical averaging theory is not applicable when there are state discontinuities.This is significant because all feedback controller converters are state discontinuous systems.

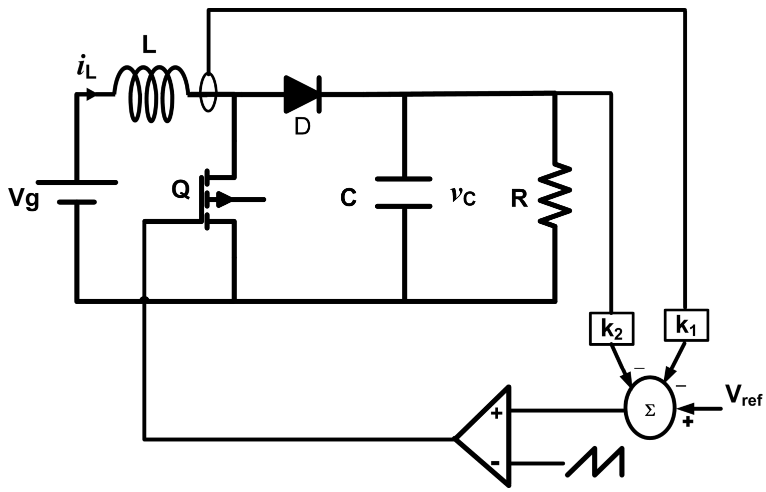


FIGURE 1: Closed loop Boost converter

Theperiodic solution of PWM converters is averaged to equilibrium. Although the periodic solutionat high switching operations has small amplitude (ripple), and averaging seems justified, theinherent dynamics for a periodic solution and at equilibrium are completely different. This issueis generally neglected in most power electronics literature. It is known that the averaged modelsare approximate by design [4, 9]. Moreover, it has been found that the directly obtained averagedmodels are inaccurate for converters operated in discontinuous conduction mode (DCM). Thisleads to the need of more efforts to obtain more accurate averaging methods.Averaging methods are sometimes used to produce approximate continuous-time models forPWM systems by neglecting the switching period of the switches and the sampling period ofthe controller [1, 2]. In averaging process the ripple in the current or in the voltage is also notconsidered. To overcome the above disadvantage, the sampled-data modeling techniques areadapted. This provides the most accurate result, which replicates the actual behavior of PWMsystems and is also suitable for digital control process. Sampled-data models allow us to focuson cycle-to-cycle behavior, ignoring intra cycle ripples. This makes them effective in generalsimulation, analysis and design. These models predict the values of signals at the beginningof each switching period, which most of the times represent peaks or valleys of the signals ratherthan average values. To better understand the average behavior of the system, a discrete-timemodel for the OCA signals was presented in [1]. Averaged discrete-time models for continuousconduction mode (CCM) and discontinuous conduction mode (DCM) without PWM feedbackcontrol are available. They are shown to be more accurate than continuous-time models andfaster [1, 2].

In this paper, a sampled-data model for PWM converters operating in DCM with feedback isformulated. This gives the exact discrete-time mathematical representation of the values of theoutput and internal signals with feedback loop at low frequency. A discrete-time model to providethe one-cycle-average (OCA) signals of PWM converters operating in DCM with feedback is proposed.This model provides the exact discrete-time mathematical representation of the averagedvalues of the output signal. It also provides the average values of other internal signals with littleincrease in simulation time. The main motivation for the new model is based on the fact that, inmany power electronic applications, it is the average values of the voltage and current rather thantheir instantaneous values that are of greatest interest. Numerical simulations show the accuracyof the propose model.

2. EXISTING AVERAGE MODELS

Different averaging methods for PWM converters are used for comparison, analysis and design.The mathematical models of the boost converter with PWM feedback control shown in Figure 1are presented in this section and will be used in Numerical example.

2.1 Switched Model

The DCM PWM converter can be described by

$$\dot{x}(t) = \begin{cases} A_1 x(t) + B_1 u(t), & t \in \tau_1 \\ A_2 x(t) + B_2 u(t), & t \in \tau_2 \\ A_3 x(t) + B_3 u(t), & t \in \tau_3 \end{cases} \quad (1)$$

$$y(t) = \begin{cases} C_1 x(t), & t \in \tau_1 \\ C_2 x(t), & t \in \tau_2 \\ C_3 x(t), & t \in \tau_3 \end{cases} \quad (2)$$

Where $u \in R^m$ is the input vector, $x \in R^n$ is the state vector, and $y \in R^p$ is the output vector. Thesystem switches between three topologies (A_1, B_1, C_1), (A_2, B_2, C_2), and (A_3, B_3, C_3), withswitching intervals determined by

$$\tau_1 := kT \le t < kT + d_k^1 T$$
$$\tau_2 := kT + d_k^1 T \le t < kT + (d_k^1 + d_k^2)T$$
$$\tau_3 := kT + (d_k^1 + d_k^2)T \le t < kT + T$$

Where T is the switch period, $(d_k^1 + d_k^2) \in [0, 1]$ are the switch duty ratios, and k is the discrete-time index. All auxiliary inputs will be assumed to be piecewise constants, i.e. $u(t) = u_k$ for all $t \in [kT, (k+1)T]$. This assumption is not necessary and is made for convenience only; more general cases would only require more complex notations. This is the exact switching model which will be used as the base model for comparison of different methods.

The state space matrices A_1,A_2,A_3,B_1,B_2,B_3,C_1,C_2, and $C_3$of boost converter shown in Figure1 are defined as

$$A_1 = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{RC} \end{bmatrix}; \; B_1 = \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix}; \; C_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & -\frac{1}{L} \\ \frac{1}{C} & -\frac{1}{RC} \end{bmatrix}; \; B_2 = \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix}; \; C_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{RC} \end{bmatrix}; \; B_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \; C_3 = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

The control scheme given in is applied, where the modulation signal is $m(t)=V_{ref}- k_1\, i(t) - k_2\, v(t)$ with $V_{ref}=0.13$, $k_1 =0.174$, and $k_2 = -0.0435$ as in [1, 7].

2.2 DCM State-space Average Model (SSA)

The SSA mathematical model of the boost converter is given as [11]

$$\dot{x}_1(t) = \frac{2x_1}{dT}\left(1 - \frac{x_2}{u}\right) + \frac{dx_2}{L} \quad (3)$$

$$\dot{x}_2(t) = \frac{x_1}{C} - \frac{d^2Tu}{2LC} - \frac{x_2}{RC} \tag{4}$$

2.3 Conventional Discrete-Time Model (CDTM)

The conventional discrete-time mode is given by

$$x_{k+1} = A(d_k^1, d_k^2)x_k + B(d_k^1, d_k^2)u_k \tag{5}$$

Where the input nonlinearities $A\ (d^1, d^2)$ and $B\ (d^1, d^2)$ are given by

$$A(d^1, d^2) := \Phi_3\Phi_2\Phi_1$$
$$B(d^1, d^2) := \Phi_3(\Phi_2\Gamma_1 + \Gamma_2) + \Gamma_3$$

The arguments d^1T, d^2T, and $(1-d^1-d^2)T$ for ($\Phi_1$andΓ_1), ($\Phi_2$andΓ_2) and ($\Phi_3$andΓ_3) respectively are omitted from the above equations for notation simplicity. Where

$$\Phi_i(t) := e^{A_i t}$$
$$\Gamma_i(t) := \int_0^t e^{A_i \tau} b_i d\tau$$

3. PROPOSED MODEL

This section introduces the new average discrete-time model for PWM converter operating inDCM with feedback loop. Description of the original system and derivation of the proposed modelare discussed here. The one-cycle average (OCA) representation of the output signal [1] is givenby

$$y^*(t) := \frac{1}{T}\int_{t-T}^{t} y(\tau)d\tau \tag{6}$$

The signal, $y^*(t)$ is used to develop a new discrete-time model for PWM converters operating inDCM. This model provides the basis for discrete-time simulation of the averaged value of anystate in the DCM PWM system, even during transient non-periodic operating conditions.

3.1 OCA Discrete-Time Model

It is desired to compute, without approximation, the evolution of all system variables at thesampling instants, $t = kT$ assuming three different topologies for the system. Since the state andoutput equations (1) - (2) are piecewise-linear with respect to time t, the desired discrete-time model can be obtained symbolically. Using the notation, $x_k := x(kT)$ and $y_k^* := y^*(kT)$, the result is the OCA large signal model

$$x_{k+1} = A(d_k^1,\ d_k^2)x_k + B(d_k^1,\ d_k^2)u_k \tag{7}$$
$$y_{k+1}^* = C(d_k^1,\ d_k^2)x_k + D(d_k^1,\ d_k^2)u_k \tag{8}$$

Where the input nonlinearities $A(d^1, d^2)$, $B(d^1, d^2)$, $C(d^1, d^2)$ and $D(d^1, d^2)$ are given by

$$A(d^1, d^2) := \Phi_3\Phi_2\Phi_1$$
$$B(d^1, d^2) := \Phi_3(\Phi_2\Gamma_1 + \Gamma_2) + \Gamma_3$$
$$C(d^1, d^2) := C_1\Phi_1^* + C_2\Phi_2^*\Phi_1 + C_3 \quad \Phi_3^*\Phi_2\Phi_1$$
$$D(d^1, d^2) := C_1\Gamma_1^* + C_2(\Phi_2^*\Gamma_1 + \Gamma_2^*) + C_3(\Phi_3^*(\Phi_2\Gamma_1 + \Gamma_2) + \Gamma_3^*)$$

The arguments d^1T, d^2T, and $(1 - d^1 - d^2)T$ for (Φ_1, Φ_1^*, Γ_1, and Γ_1^*), (Φ_2, Φ_2^*, Γ_2, and Γ_2^*) and (Φ_3, Φ_3^*, Γ_3, and Γ_3^*) respectively are omitted from the above equations for notation simplicity. Where

$$\Phi_i(\mathrm{t}) := e^{A_i t}$$
$$\Gamma_i(\mathrm{t}) := \int_0^t e^{A_i \tau} b_i d\tau$$
$$\Phi_i^*(\mathrm{t}) := \frac{1}{T}\int_0^t \Phi_i(\tau) d\tau$$
$$\Gamma_i^*(\mathrm{t}) := \frac{1}{T}\int_0^t \Gamma_i(\tau) d\tau$$

Note that the averaging operation adds "sensor" dynamics to the system;as a consequence, thelarge-signal model equations (7) and (8) is not in standard state-space form. By defining theaugmented state vector $x^* \in R^{n+p}$such that

$$x_{k+1}^* := \begin{bmatrix} x_{k+1} \\ C(d_k^1, d_k^2)x_k + D(d_k^1, d_k^2)u_k \end{bmatrix}$$

An equivalent (but standard form) representation of the OCA large-signal model is given by:

$$x_{k+1}^* = A^*(d_k^1, d_k^2)x + B^*(d_k^1, d_k^2)u_k$$
$$y_k^* = C^* x_k^*$$

Where

$$A^*(d_k^1, d_k^2) := \begin{bmatrix} A(d^1, d^2) & 0_{n\times p} \\ C(d^1, d^2) & 0_{p\times p} \end{bmatrix}$$
$$B^*(d_k^1, d_k^2) := \begin{bmatrix} B(d^1, d^2) \\ D(d^1, d^2) \end{bmatrix}$$
$$C^*(d_k^1, d_k^2) := [0_{p\times n} \quad I_{n\times p}]$$

Note that not only the OCA values of output signal will be available but also the values of thesignals (without averaging) at the beginning of every switching period as well.

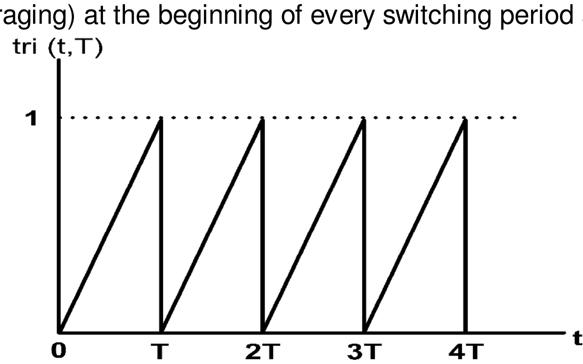


FIGURE 2: PWM sawtooth function

3.2 Feedback Computation

The modulation signal for feedback control is $m(t)=V_{ref}-k_1\, i(t) - k_2\, v(t) = V_{ref} - K\, x(t)$ and the duty ratio at each switching period is $d_k=t^*/T$. The time instant t^* at which the modulation signal crosses the sawtooth is computed by solving the nonlinear equation

$$tri(t^*,T) = V_{ref} - Kx(kT + t^*) \\ = V_{ref} - K\{\Phi_1(t^*)x_k + \Gamma_1(t^*)u_k\} \quad (9)$$

at each time instant *k*, where the sawtooth function is shown in Fig. 2 and mathematicallyrepresented by *tri(t, T)= (t/T)-floor(t/T)*. For reasonably high switching frequency, the value of$x(kT+t^*)$ can be approximated by neglecting the higher order terms in the Taylor expansion of the nonlinear functions $\Phi_1$andΓ_1. That is

$$\Phi_1(t^*) = I + A_1 t^* + \frac{A_1^2}{2!} t^{*^2} + \cdots \approx I + A_1 t^*$$

$$\Gamma_1(t^*) = (It^* + A_1 \frac{t^{*^2}}{2!} + \cdots)B_1 \approx It^* B_1$$

And hence, a good approximation of (9) becomes

$$tri(t^*,T) \approx V_{ref} - K\{(I + A_1 t^*)x_k + B_1 t^* u_k\}$$

Noting that $tri(t^*,T)$ equals to t^*/T for $t^* \in [kT, (k+1)T]$, we get

$$\frac{t^*}{T} \approx V_{ref} - K\{(I + A_1 t^*)x_k + B_1 t^* u_k\}$$

Or

$$d_k = \frac{t^*}{T} \approx \frac{V_{ref} - Kx_k}{KT(A_1 x_k + B_1 u_k) + 1}$$

Which provides a closed from solution for d_k. The duty ratio d_k can be computed without approximation by solving the nonlinear equation (9) for t^*.

4. NUMERICAL EXAMPLE

Since all of the aforementioned averaged models have been controlled with the same controllerdesign, a comparative study is carried out to investigate the accuracy and speed of the proposedmodel as compared to the existing averaged model. It should be noted that no approximation ismade in deriving the new discrete-time model, and all simulations were performed using Matlab.The results of all models are computed using built-in Matlab nonlinear equation solver. The state variables are $x_1 = i_L$and $x_2 = v_C$.

4.1 Ideal Condition

Ideal boost converter with PWM feedback control shown in Figure 1 is simulated using existingaveraged models and proposed model. The simulated voltage and current waveforms are shownin Figures 3 - 4. The steady-state average values predicted by the proposed model are moreaccurate than the ones obtained by the SSA models for the parameters R = 45 Ω, L = 100 μH,C = 4.4 μF, V_g= 5 V, and T_S= 100 μs.

The steady-state average values of the output voltage are v_C= 8.1125 V for SSA model andv_C= 8.3174 V for the proposed model. It should be noted that the accuracy of the SSA methoddecreases as the switching frequency decreases, while the proposed model does not depend onthe switching frequency as discussed below.

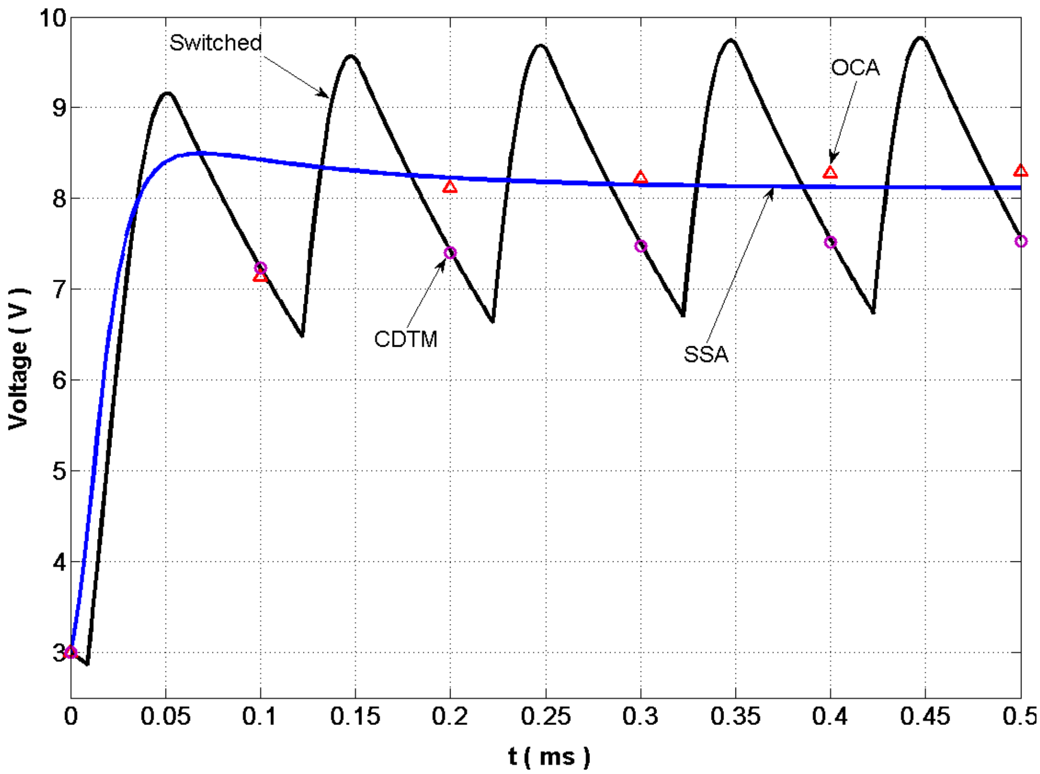


FIGURE 3: Simulation comparison of voltage waveform

4.2 Effect of Switching Frequency

To demonstrate the effect of switching frequency, consider operating the converter in a Higherfrequency, for example TS = 80µs i.e. f_S= 12.5 KHz. Figures 5 - 6 shows the effect ofincreasing the switching frequency on the simulation results of the output waveform. The steadystateaverage values of the output voltage are v_C= 8.055 V for SSA model and v_C= 8.475V for the proposed model. The accuracy of SSA averaged model decreases and move awayfrom the actual average as the switching frequency increases. The proposed model captures thecycle-to-cycle behavior of the system with no approximation regardless of changes in switchingfrequency.

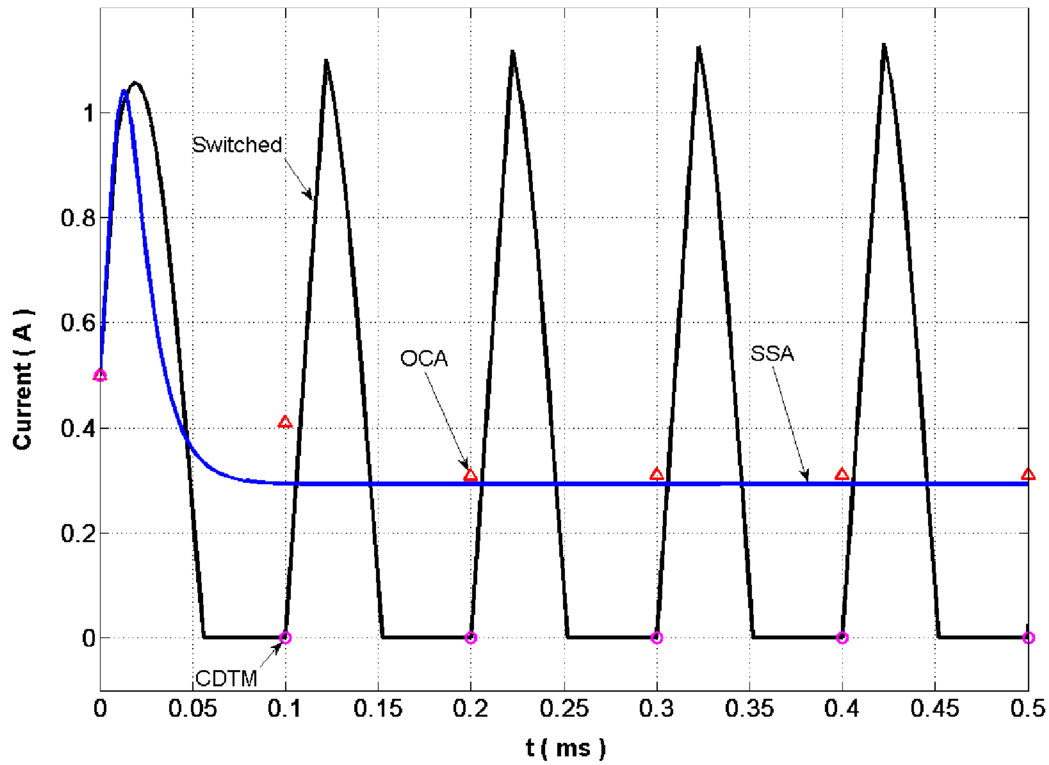


FIGURE 4: Simulation comparison of current waveform

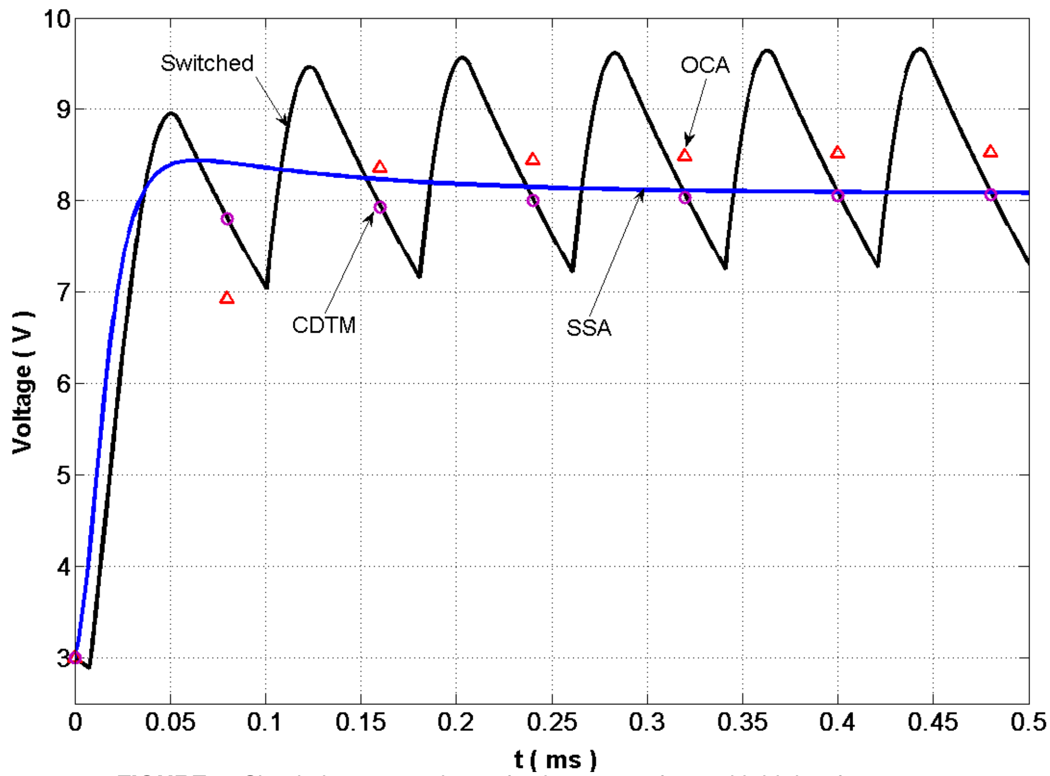


FIGURE 5: Simulation comparison of voltage waveform with higher frequency

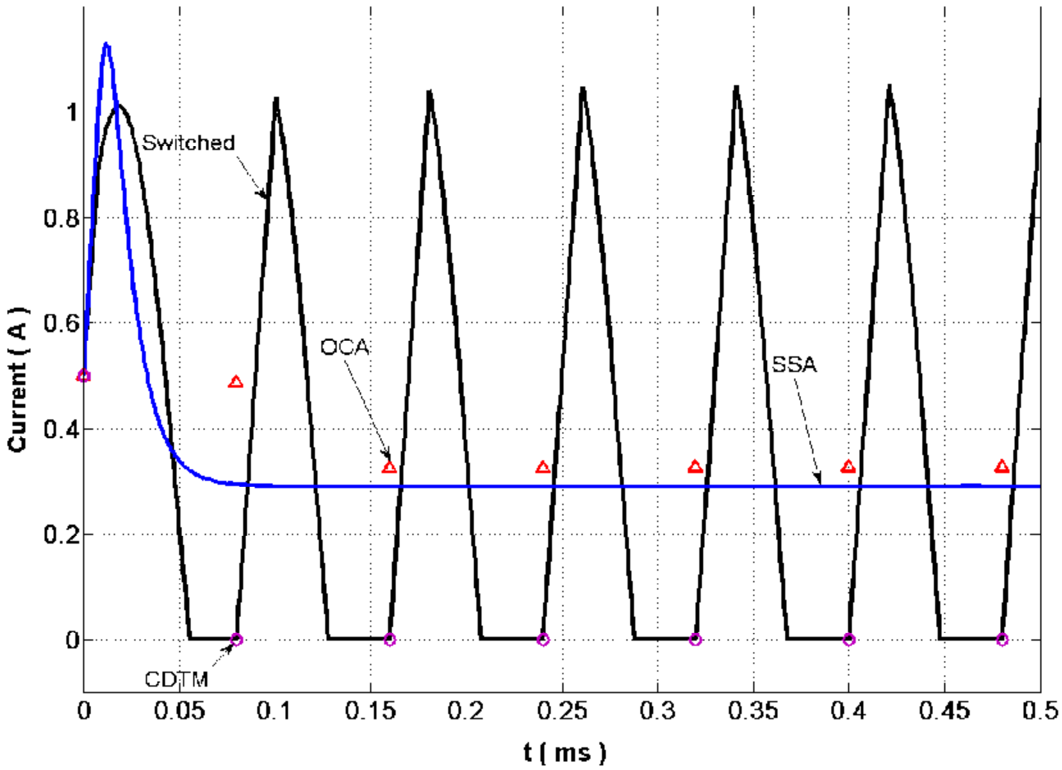


FIGURE 6: Simulation comparison of current waveform with higher frequency

4.3 Effect of Change in load

To study the effect of load resistance on the simulation results, a step change on the loadresistance R from 45 Ωto 55 Ωat time instant, t = 0.6 ms has been simulated and the resultsare shown in Figure 7. The steady-state average values of the output voltage are v_C= 8.1125V and v_C= 9.1125 V for SSA model and v_C= 8.3174 V and v_C= 9.625 V for the proposedmodel respectively. It can be observed that the average values produce by SSA model deviatemore from the exact waveforms as the load resistance increased. On the other hand the proposedmodel provides the same accuracy of waveforms regardless of the change in load resistance.

Table 1 summarizes the normalized simulation times for ideal boost converter with feedback fordifferent simulation methods.

Method	Normalized Simulation Time
Switched	42
SSA	1
CDTM	8.4
DCM OCA	8.81

TABLE 1: Simulation time for boost converter in DCM with feedback

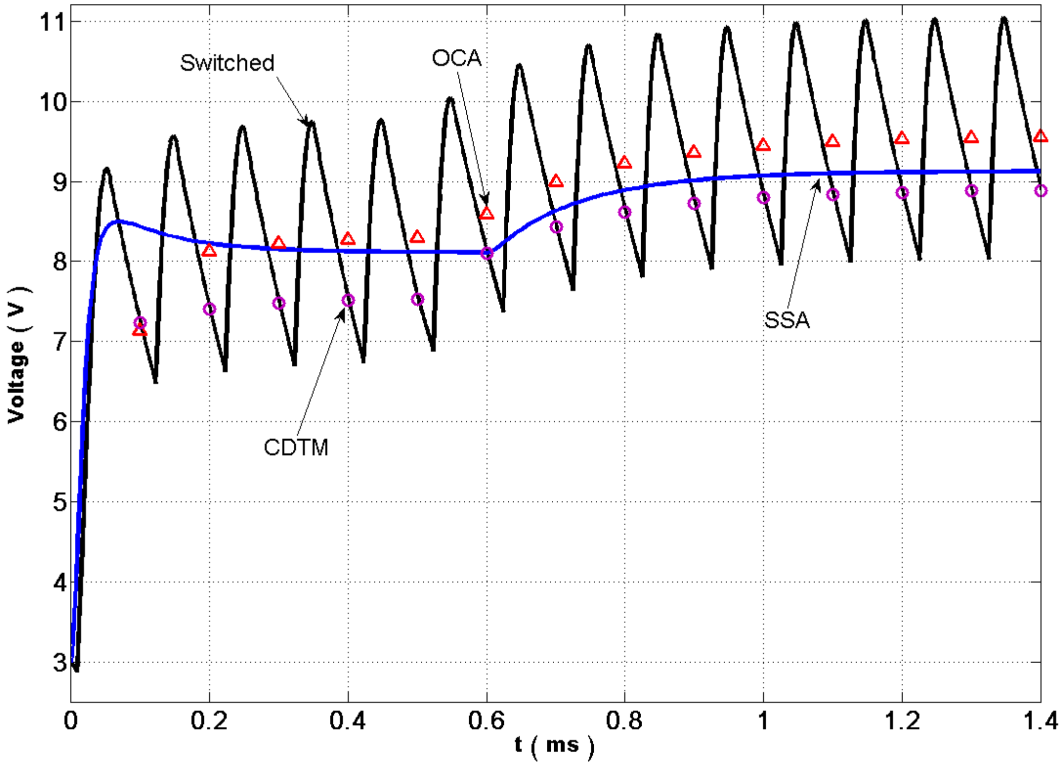


FIGURE 7: Simulation comparison of voltage waveform with variable load resistance

5. CONCLUSION

This paper proposed a new model which provides a discrete-time response of the one-cycle-average(OCA) value of the output signal for PWM dc-dc converters operating in the DCM withfeedback. It is compared to existing models through a numerical example of boost converter. Asa result of variations in the circuit parameters such as switching frequency and load resistance,a significant deviation in the average values of the converter's signals is predicted by the existingaveraging method. On the other hand, the proposed model is fast and can accurately simulatethe average behavior of the output voltage even though there is a large variation in the circuitparameters without any approximation in the design.

6. REFERENCES

[1] M. S. Al-Numay. "Discrete-Time Averaging of PWM DC-DC converters with feedback", *The 2006International Conference on Computer Engineering and Systems*, pp. 45–48,November 2006.

[2] M. S. Al-Numay. "Averaged Discrete-Time Model for Fast simulation of PWM Converters in DCM",*International Journal of Modeling and Simulation*, vol. 29, no. 1, pp. 46–51, 2009.

[3] V. A. Caliskan, G. C. Verghese, and A. M.Stankovic."Multifrequency averaging of DC/DC converters",*IEEE Transactions on Power Electronics*, vol. 14, no. 1, pp. 124–133, 1999.

[4] A. Davoudi, J. Jatskevich, and T. D. Rybel."Numerical state-space average-values modeling OfPWM DC-DC Converters Operating in DCM and CCM", *IEEE Transactions on Power Electronics*,vol. 21, no. 4, pp. 1003-1012, July 2006.

[5] N. Femia and V. Tucci."On the Modeling of PWM Converters for Large Signal Analysis in Discontinuous Conduction Mode", *IEEE Transactions on Power Electronics,* vol. 9, no. 5, pp. 487–496, September 1994.

[6] J. G. Kassakian, M. F. Schlecht, and G. C. Verghese: *Principles of Power Electronics Reading*,MA: Addison Wesley, 1991.

[7] B. Lehman, and R. M. Bass. "Switching Frequency Dependent Averaged Models for PWM DCDCConverters", *IEEE Transactions on Power Electronics*, vol. 11, no. 1, pp. 89-98,January1996.

[8] D. Maksimovic and S. Cuk. "A Unified Analysis of PWM Converters in Discontinuous Modes", *IEEE Transactions on Power Electronics,* vol. 6, no. 3, pp. 476–490, July 1991.

[9] R. D. Middlebrook and S. Cuk. "A general unified approach to modeling switching-converter Powerstages", *International Journal of Electronics*, vol. 42, pp. 521–550, 1977.

[10] S. R. Sanders, J. M. Noworolski, X. Z. Liu, and G. C. Verghese. "Generalized Averaging Methodfor Power Conversion Circuits", *IEEE Transactions on Power Electronics*, vol. 6, no. 2, pp. 251–259, April 1991.

[11] J. Sun, D. M. Mitchell, M. F. Greuel, P. T. Krein and R. M. Bass. "Averaged Modeling of PWMConverters Operating in Discontinuous Conduction Mode", *IEEE Transactions on Power Electronics*,vol. 16, pp. 482–492, April 2001.

[12] K. K. Tse, M. T. Ho, Henry S. H. Chung, and S. Y. (Ron) Hui. "A Novel Maximum Power Point Tracker for PV Panels sing Switching Frequency Modulation", *IEEE Transactions on PowerElectronics*, vol. 17, no. 6, pp. 980–989, November 2002.

[13] G. C. Verghese, M. E. Elbuluk and J. G. Kassakian, "A general approach to sampled-data modeling for power electronics circuits ", *IEEE Transactions on Power Electronics*, vol. 1(2), pp.76–89, 1986.

[14] R. C. Wong, H. A. Owen and T. G. Wilson,"An efficient algorithm for the time-domain Simulationof regulated energy-storage dc-to-dc converters", *IEEE Transactions on Power Electronics*, vol.2(2), pp. 154–168, 1987.

Semantic Web Mining of Un-structured Data: Challenges and Opportunities

Manoj Manuja manoj_manuja@infosys.com
Principal, Education & Research Dept,
Infosys Technologies Ltd.,
Chandigarh, India

Deepak Garg dgarg@thapar.edu
Thapar University,
Patiala, India

Abstract

The management of unstructured data is acknowledged as one of the most critical unsolved problems in data management and business intelligence fields in current times. The major reason for this unresolved problem is primarily because of the actuality that the methods, systems and related tools that have established themselves so successfully converting structured information into business intelligence, simply are ineffective when we try to implement the same on unstructured information. New methods and approaches are very much necessary. It is a known realism that huge amount of information is shared by the organizations across the world over the web. It is, however, significant to observe that this information explosion across the globe has resulted in opening a lot of new avenues to create tools for data management and business intelligence primarily focusing on unstructured data. In this paper, we explore the challenges being faced by information system developers during mining of unstructured data in the context of semantic web and web mining. Opportunities in the wake of these challenges are discussed towards the end of the paper.

Keywords: Semantic Web, Web Mining, Unstructured Data.

1. INTRODUCTION

The last few years have seen growing recognition of information as a key business tool for the success of the organizations across the world. The organizations which effectively identify, accumulate, study, scrutinize and thereafter act upon the information are definite winners in this new “information age”. Further to this, the realization of “web” has critically changed the perspective of how the organizations extract information from the available data in today’s world of dynamic business.

Therefore, the most important differentiator between a successful and an unsuccessful business is how an organization manages its data. The critical aspect in today’s business scenario is how data is converted into Information and subsequently how information is converted into knowledge.

2. BUSINESS DILEMMA IN A LARGE ORGANIZATION

It is very crucial to extract knowledge from un-structured data which is available in various formats and generated by heterogeneous sources across a big organization.

According to projections from Gartner, professionals will spend anywhere between 30 to 40 % of their office time in managing various documents, which is 20% more than what they used to spend on similar activities 10-15 years ago. Similarly, Merrill Lynch has anticipated that data which is unstructured will amount to more than 85 percent of all information available in a company.

It is very easy to extract useful knowledge from structured data using proven algorithms and patterns. But the problem comes when we have unstructured data to work with. It becomes very difficult to extract knowledge from the un-structured data because of non-availability of proven algorithms, schemas, patterns and information systems. Through this paper, we shall explore various challenges in the field of unstructured data mining using semantic web techniques and also available opportunities in the wake of these challenges.

3. DATA, INFORMATION AND KNOWLEDGE

In a practical real life scenario, data is available in three forms:

- **Unprocessed data** which is gathered in real time,
- **Extracted data** which gives us information and
- **Processed data** which provides us useful knowledge [1].

Knowledge is being used to determine and analyze the specifics of a given situation. Knowledge is a credence that is factual, vindicated, and relies on no false theories.

3.1 Un-Structured Data

It refers to computerized information that is either not having any data model or cannot be directly used by a computer program [1]. In other words, data with some form of structure may be characterized as unstructured if its structure does not reflect a useful schema to get a desired processing task.

Most of the business information exists as unstructured data – commonly appearing in e-mails, blogs, discussion forums, wikis, official memos, news, user groups, chatting scripts on social networking sites, project reports, business proposals, public surveys, research and white papers, marketing material, official and business presentations and most of the web pages on WWW.

3.2 Different forms of un-structured data

We have different types of un-structured data generated by different users across the globe:

- **Business Data**: This type of data is primarily generated in a business organization. Although a small part of it is structured data like employee information, salary details, company's balance sheet etc, but a large part of it is simply unstructured like customer communication and feedback, client presentations, minutes of project / team meeting, official memos and many more.
- **Social networking data**: This type of data is purely un-structured. Users basically use SMS type language which is not easily understandable by even human beings. Product reviews, and feedbacks are another important part of this database. Chat scripts are also constituents of such data.
- **General communication data**: This type of data mainly constitutes emails, blogs, wikis, news, discussion forums etc. Although templates to capture this data are structured but the contents inside those text blocks are mainly un-structured.
- **Audio-Visual data**: The data in the form of audio and video files is available in huge quantity across the world. There is no defined pattern available while we browse these files.

4. CURRENT SCENARIO OF DATA MINING OVER WEB

Keeping in view of the potential opportunity to extract business focused knowledge from the colossal amount of data available on www, a structured approach is being followed by data administrators and managers across the world when we talk about data mining over web. Below are the major fields which are being explored in terms of finding knowledge from un-structured data:

4.1 Data mining with a focus towards mining unstructured data

A lot of unstructured data is noisy text [2]. Spontaneous communication (such as e-mails, discussion forums, SMS, blogs, and collaborative web portals) contains noisy text and processing noise. We can define "noise" in text as the difference of any type found between the original and received text.

In the context of unstructured content [3], there is no conceptual and data type definitions available in textual documents, and we find it very tricky to extract information from the content [4]. Therefore, proficient algorithms duly supported by human intercessions are necessary to make the unstructured data smoothly readable and understandable by a computer machine [5]. A vast proportion of this unstructured data contains informal and semi-formal, internal and external communications of a given organization [6]. Usually humans can understand such text straightaway. However, with enormous quantity of such data content being available nowadays, both online and inside the enterprise, it becomes critical to mine such text using computers as it becomes very difficult and complex for a human being to mine huge data manually.
We can think of using available data mining generalized models to represent unstructured data also but with very less efficiency and proper outcome. There are a few algorithms available to extract useful information from unstructured data including Opinion mining [7] from noisy text data, but a generalized, rugged approach is still missing.

4.2 Semantic Web Mining
The semantic web is based on the visualization of Tim Berners-Lee [8], the inventor of the World-Wide-Web (WWW). According to him, "The semantic web is not at all visualized as a separate web but it is an expansion of the existing one, in which information is given well-defined sense and significance, better enabling PCs and people to work in cooperation."

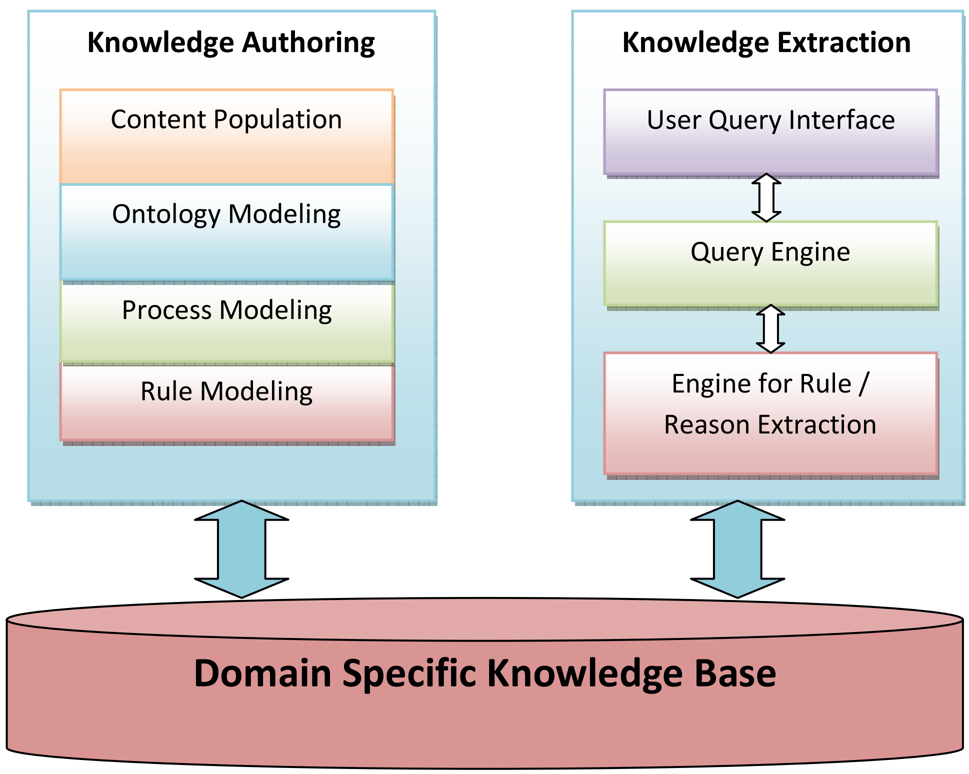


FIGURE 1: Semantic Web Solution Architecture

Semantic web mining intends at two emergent research areas of semantic web and web mining [9,10]. The idea is to improvise the results of web mining by taking advantage of the new semantic structures on the Web; and also, making use of web mining, for building up the semantic web by extracting similar meanings, useful patterns, structures, and semantic relations from existing web resources.
Figure 1 shows proposed solution architecture for semantic web mining. The architecture is primarily divided into three logical modules, namely knowledge extraction block, knowledge

authoring block and domain specific knowledge base. The availability of prevailing search engines has to a great extent improved our ability to carry-out a meaningful data search on the web. But, such search option is still primarily restricted to structured data. In semantic technology, the focus is generally to formulate flexible data model (called Triples) from the user friendly domain query. Semantic search engines are yet to prove themselves in the huge periphery of web search.

4.3 Ontologies Extraction

Extracting ontology from the web is a challenging task [10]. Ontology extraction and modeling use a lot of existing resources, like text, thesauri, dictionaries, databases and similar resources. Techniques from several related research areas e. g., machine learning, information retrieval [11], etc, are combined, and are applied together to discover the `semantics' in the data and to make them plain and clear.

A few systems have already been developed by research community across the world to extract ontology [12]. Several standards have been developed to implement the layered structure of the Semantic Web, such as the Resource Description Framework (RDF) [13] and Web Ontology Language (OWL) [14]. Resource Description Framework (RDF) is being used by people to represent metadata of web pages which can be processed by a machine. It describes a data model to represent all relations between different resources. Still, this “similar meaning and relation extraction” work is yet to mature on the global information retrieval platform simply because of the non-availability of proven processes, standards and systems.

5. CHALLENGES IN SEMANTIC WEB MINING OF UNSTRUCTURED DATA

There are many challenges when it comes to mining of the unstructured data in the context of semantic web:

i. **Structured-data mining focused search engines:** The emergence of some great search engines has significantly improved our skillful capability to search for data on the web; however, such search tool is vastly restricted to structured data only. Not many search engines are available in public domain which specifically addresses the requirement of mining / searching unstructured data flavored with semantic search. This is the biggest challenge being faced by industry and academia alike.

ii. **No standardized web form structure:** We can search for and extract information available as HTML, but till date, we are not proficient to gain easy access to the hidden web. It is very difficult to get to the accurate web form, and even harder to find a suitable truthful web application and related service. When we find the accurate web form or web service, then there is a supplementary step to understand its schema and reformulate the user's query to fit that schema. While human beings do this on a regular basis, one form at a time, it is very complex and cumbersome to automate the process of query reformulation, and therefore we cannot leverage the wealth of information residing behind web forms and services for the masses.

iii. **Non-availability of standard Semantics:** We cannot apply the techniques for exploiting corpora of documents directly for searching unstructured data. The main reason is that searching unstructured data requires an understanding of its underlying semantics. This structure is normally specified by the schema. However, in specifying these semantics, the actual words used and the information clustering purely depends more on the developer's whim, and little variations may result in a very different semantics altogether. Thus, it is a big challenge to have standard semantics available for general usage.

iv. **Lack of global Standards:** Very less international standards are available on Semantic Web Mining. Some big organizations and universities are working on it with little success. Non-availability of broad, rugged and internationally recognized set of standards addressing amalgamated mix of semantic web, web mining and unstructured data mining is the crucial challenge being faced by researchers across the world.

v. **Lack of proven frameworks:** There are some challenges involved in large scale integration on the web [15] namely the realization of the mining framework, the robustness of mining techniques, and the exploration of holistic insight.

vi. **Non-standard implementation:** There is a huge implementation challenge to develop a database management system for administering the entire process of information extraction in an efficient and effective manner [16]. Some industries and academic establishments have come up with their own KDIS (knowledge driven information systems), but all these systems are proprietary to individual organizations or universities. Because of this, the research in this area is not getting opened up for wider coverage and business implementation.

vii. **Non-standard Information Systems:** If we are to design such a system which can help the users to mine unstructured data, how should it look for? What will be the system capabilities? The key challenges include data model and representational issues; need for newer index structures; standardization for Information Extraction (IE); data cleaning and its fusion; accurate relationships in the context of IE and probabilistic databases; and lastly the role of knowledge which user possess and the iterative nature of user interaction. These are a few challenges researchers face during information system development.

viii. **No support for audio / video data mining:** There is any standard support available which can help the end-users to extract valuable and handy information from audio and video files available in huge quantity across the world.

ix. **Less-explored Ontology framework:** Developing the ontology for the large scale databases is a great challenge in itself. There are so many industry verticals and domains available across the organizations. Developing Ontologies for these verticals and domains is a major challenge.

x. **Lack of availability of best practices:** Data vocabulary complemented by content relevance is the operational challenges during the development of semantic web mining tools. There are no best practices available for this development as the technology area is not yet matured and lot of new developments are happening across the world in this field.

6. OPPORTUNITIES FOR SEMANTIC WEB DATA MINERS IN THE FIELD OF UN-STRUCTURED DATA

The unprecedented success of www has unfolded the true potential of two fast-emerging research areas of semantic web and web mining. Both of these areas complement each other and open up new opportunities for the researchers across the world. As discussed in preceding sections, the majority of the available data on web is totally un-structured which can be understood by human-beings only. But the amount of data suggests that the same can be processed by machines efficiently. Hence, there is a good opportunity for semantic and web miners to explore this situation to provide next level of mining paradigm to the world.

The intact opportunity in the field of semantic web mining can be elaborated and split into two unique parts as “semantic” – “web mining” or “semantic web” – “mining”. In the past few years, there have been many attempts at “breaking the syntax barrier” on the web [17].

Analysis of challenges mentioned in previous section gives us an ample opportunity to explore semantic web mining of un-structured data and extract huge amount of knowledge available un-tapped at www. A few opportunities are suggested below:

- To develop web mining techniques that will enable the power of www to be realized. These constitute development of web metrics and measurements, process mining, temporal evolution of the Web, web services optimization, fraud and threat analysis, and web mining and privacy.
- To design and develop search engines specifically focused towards mining un-structured databases. This is the need of the hour as the success and failure of semantic web mining of unstructured data will primarily depend on the availability of suitable and relevant search engines.
- To design and develop information systems for exploring unstructured data available in bulk on web primarily extracting content, structure and usage mining. An enterprise system

integrating these three spheres of web mining is very critical to the success of this field of research [15].

- To design and develop knowledge extraction and un-structured text processing algorithms which are either available as proprietary algorithms with some big organizations or not available at all for general usage and exploration.
- To design and develop models for concept and ontology extraction from unstructured data. This extraction is very important after the enormous explosion of social networking.
- To design and develop an ontology modeling algorithm which also addresses rule and process modeling in a particular industry vertical or domain area.
- There are not many KDIS available across the organizations which can help them feel the pulse of their customers, employees and vendors. Mining the opinion from a huge unstructured data available on www is one of the hottest research areas in current time. Extracting sentiment and opinion of customers' feedback is an exciting problem to work on.
- A few enterprises and research groups are working to make standards for semantic web, web mining and semantic web mining. This is one of the most critical fields in today's world which will provide directions to the research community on semantic web mining.
- To develop a user-friendly database management system to manage the entire process of information extraction [16]. To develop a data model which addresses representational issues of un-structured data with newer index structures is an important unexplored field. An end-to-end solution which may provide knowledge extraction and retrieval will be a big opportunity for the developers to develop.
- To standardize process for Information Extraction (IE); design efficient algorithms for data cleaning and fusion; design mechanism to find out relationships between uncertainty management in the context of IE and probabilistic databases.

7. CRITICAL ANALYSIS

Research in the field of semantic search engines is focused on various approaches and classification theories. Miller et al. [18] talked about Navigational Searches which points to the classification of documents based upon the intention of the user. Mangold [19] focused on architecture, coupling, user context, query modification, transparency, structure of ontology and relevant technology as parameters to realize semantic search. In another critical research on semantic search engine [20], it is pointed out that augmenting traditional keyword search with semantic techniques is considered as the important parameters to implement the semantic search engine.

Hildebrand et al. [21] suggested a search system based upon query construction in section with custom search algorithms. Dietze and Schroeder [22] suggest a new classification approach based on 9 criteria which include structured/unstructured file, text mining type, type of documents, number of documents, Ontologies, clustering, result type, highlighting, scientifically evaluated. Dong et al. [23] present a extended classification with semantic search algorithm based on the Graph, methodology on distributed hash tables and logics-based Information retrieval.

8. COMPARISON OF SEMANTIC SEARCH ENGINES

In the current search scenario, there's no denial about the super power and unquestionable popularity of the Google search engine, where results are based on page rankings and proprietary algorithms. But there are some very innovative ways available to search the web, using semantic search engines. A semantic search engine will definitely ensure more closely suggested relevant results based on the ability to understand the definition and user-specific meaning of the word or term that is being searched for. Semantic search engines are able to better understand the context in which the words are being used, resulting in smart, relevant results with more user satisfaction.

A comparison of semantic search engines is shown in Table – 1.

Semantic Search Engine	**URL**	**Main Approach**	**Features**
Kngine	http://kngine.com/	Based upon "Concepts"	It contains more than 8 million concepts
Yebol	http://www.yebol.com/	Based upon patented algorithms paired with human knowledge	Yebol automatically clusters and categorizes search terms, Web sites, pages and contents, instead of the common "listing" of Web search queries.
Hakia	http://hakia.com/	Based upon "Credible"	It divides the results into Web, News, Blogs, Twitter, Image and Video, and can be re-listed according to relevance.
Duckduckgo	http://duckduckgo.com/	Based upon classic search, information search	If we search for a term that has more than one meaning, it will give us the chance to choose what you were originally looking for, with its disambiguation results.
EVRI	http://www.evri.com/	Based upon Information search	Search results can be filtered into *Articles, Quotes, Images* and *Tweets*.
Truevert	http://www.truevert.com/	Based upon "Green Search Engine"	All results are filtered and organized from one specific perspective – with the topic of environmental awareness in mind.

TABLE 1: Comparison of Semantic Search Engines

9. SOME SEMANTIC WEB BASED WEB SITES

Although, there are many web pages available on www which are using semantic web as the base technology, we are sharing a few popular web sites as shown in Table 2.

Web Pages supported by semantic web	**URLs**
Brickipedia	http://lego.wikia.com/wiki/LEGO_Wiki
Familypedia	http://familypedia.wikia.com/wiki/Family_History_and_Genealogy_Wiki
Semantic MediaWiki	http://smwtest.wikia.com/wiki/Semantic_MediaWiki_Test_Wiki
Books Wiki	http://bookswiki.wikia.com/wiki/Books_Wiki
SuperWikia	http://super.wikia.com/wiki/Main_Page
Governance Wiki	http://government.wikia.com/wiki/Giki
Yellowikis	http://yellowikis.wikia.com/wiki/Main_Page
MyWikiBiz	http://mywikibiz.com/Main_Page
Common Sense Wiki	http://commonsense.wikia.com/wiki/Common_Sense_Wiki
Animepedia	http://anime.wikia.com/wiki/Animepedia
Creative Commons Wiki	http://wiki.creativecommons.org/Main_Page
semanticweb.org	http://semanticweb.org/wiki/Main_Page

TABLE 2: Semantic web based web sites

10. CONCLUSION

Semantic web mining is relatively new sub-field of data mining. It has a vast scope for investigation keeping in view of the availability of tons of unstructured data on WWW. Lack of available global standards on this subject opens up a enormous prospect for the research community to focus on this area in a big way. Non-availability of a rugged database management system to manage semantic web mining opens up new avenues for the researchers to develop KIMS (Knowledge extraction management system) for unstructured data available on the web. A user-oriented semantic search engine is the need of the day. These fields if explored in a right manner will provide unlimited opportunities to extract knowledge from the goldmine of unstructured data available across the globe.

11. REFERENCES

[1] J. Han, M. Kamber. (2001) “Data Mining Concepts and Techniques”. Academic Press, Morgan Kaufmarm Publishers. ISBN 1-55860-489-8.

[2] W. Fan, L. Wallace, S. Rich, Z. Zhang. (2006, September). “Tapping the power of text mining”. Communications of the ACM. Volume 49, Issue 9. pp. 76 – 82

[3] D. Bitton, F. Faerber, L. Haas, J. Shanmugasundaram. (2006). “One platform for mining structured and unstructured data: dream or reality?”. Proceedings of the 32nd international conference on Very large data bases. pp. 1261 – 1262

[4] R. J. Mooney, R. Bunescu. (2005). “Mining knowledge from text using information extraction”; ACM SIGKDD Explorations Newsletter. pp. 3 – 10

[5] R. Ghani and Carlos. (2006, December). “Data mining for business applications”. KDD-2006 workshop. Volume 8, Issue 2. pp. 79 – 81

[6] M. Rajman, R. Besancon. (1997). “Text Mining - Knowledge extraction from unstructured textual data”. In Proceedings of the 7th IFIP Working Conference on Database Semantics (DS-7). pp. 7-10

[7] L. Dey , S. K. M. Haque. (2009, July). “Studying the effects of noisy text on text mining applications”. Proceedings of The Third Workshop on Analytics for Noisy Unstructured Text Data. Barcelona, Spain

[8] Tim Berners-Lee. “Semantic Web Roadmap”. http://www.W3.org/

[9] B. Berendt, A. Hotho, and G. Stumme. (2002). “Semantic Web Mining and the Representation, Analysis, and Evolution of Web Space”. Proceedings of the First International Semantic Web Conference on The Semantic Web. pp. 264 – 278

[10] B. Berendt, A. Hotho, G. Stumme. (2002). “Towards Semantic Web Mining”; ISWC '02: Proceedings of the First International Semantic Web Conference on The Semantic Web; Publisher: Springer-Verlag

[11] A Maedche. (2002). “Ontology Learning for the Semantic Web”; Kluwer. ISBN: 0792376560

[12] M. Niepert, C. Buckner, J. Murdock, C. Allen. (2008). “InPhO: a system for collaboratively populating and extending a dynamic ontology”. Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, Pittsburgh PA, PA, USA. pp. 429-429

[13] Resource Description Framework (RDF) Schema Specification. (2000) In W3C Recommendation.

[14] Web Ontology Language (OWL), http://www.w3.org/2004/OWL/.

[15] K. Chang, B. He, Z. Zhang (2004, December). "Mining semantics for large scale integration on the web: evidences, insights, and challenges". ACM SIGKDD Explorations Newsletter, Volume 6 , Issue 2. pp. 67-74.

[16] A. Doan, R. Ramakrishnan, S. Vaithyanathan. (2006). "Managing information extraction: state of the art and research directions". Proceedings of the ACM SIGMOD international conference on Management of data. pp. 799 – 800

[17] G. Stummea, A. Hotho, B. Berendt. (2006). "Semantic Web Mining State of the art and future directions". Journal of Web Semantic. Web Semantics: Science, Services and Agents on the World Wide Web 4. pp. 124–143.

[18] Miller, Guha, R., McCool, R. (2003). "E. Semantic Search". Proceedings of the WWW'03, Budapest.

[19] C Mangold (2007). "A survey and classification of semantic search approaches". International Journal of Metadata, Semantics and Ontologies, pp. 23–34.

[20] D. Buscaldi, P. Rosso, E. S. Arnal (2005). "A wordnet-based query expansion method for geo-graphical information retrieval". Working Notes for the CLEF Workshop.

[21] M. Hildebrand, J. Ossenbruggen, and L. Van Hardman (2007). "An analysis of search-based user interaction on the semantic web". Report, CWI, Amsterdam, Holland.

[22] F. Figueira, J. Porto de Albuquerque, A. Resende, Geus, P. Lício de Geus, G. Olso (2009). "A visualization interface for interactive search refinement". 3rd Annual Workshop on Human-Computer Interaction and IR, Washington DC. pp. 46-49.

[23] H. Dong, FK Hussain, and E. Chang (2008). "A survey in semantic search technologies". 2nd IEEE International Conference on Digital Ecosystems and Technologies.

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Engineering (IJE)* is devoted in assimilating publications that document development and research results within the broad spectrum of subfields in the engineering sciences. The journal intends to disseminate knowledge in the various disciplines of the engineering field from theoretical, practical and analytical research to physical implications and theoretical or quantitative discussion intended for both academic and industrial progress.

Our intended audiences comprises of scientists, researchers, mathematicians, practicing engineers, among others working in Engineering and welcome them to exchange and share their expertise in their particular disciplines. We also encourage articles, interdisciplinary in nature. The realm of International Journal of Engineering (IJE) extends, but not limited, to the following:

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJE.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 5, 2011, IJE appears in more focused issues. Besides normal publications, IJE intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

IJE LIST OF TOPICS

The realm of International Journal of Engineering (IJE) extends, but not limited, to the following:

- Aerospace Engineering
- Agricultural Engineering
- Biomedical Engineering
- Chemical Engineering
- Civil & Structural Engineering
- Computer Engineering
- Control Systems Engineering
- Education Engineering
- Electrical Engineering
- Electronic Engineering
- Engineering Mathematics
- Engineering Science
- Environmental Engineering
- Fluid Engineering
- Geotechnical Engineering
- Industrial Engineering
- Manufacturing Engineering
- Materials & Technology Engineering
- Mechanical Engineering
- Mineral & Mining Engineering
- Nuclear Engineering
- Optical Engineering
- Petroleum Engineering
- Robotics & Automation Engineering
- Telecommunications Engineering

CALL FOR PAPERS

Volume: 6 - **Issue:** 1 - February 2012

i. Paper Submission: November 30, 2011 **ii. Author Notification:** January 01, 2012

iii. Issue Publication: January / February 2012

CONTACT INFORMATION

Computer Science Journals Sdn BhD
B-5-8 Plaza Mont Kiara, Mont Kiara
50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6207 1607
006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2011
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org