International Journal of Engineering (IJE)

Book: 2009 Volume 3, Issue 5
Publishing Date: 30-11-2009
Proceedings
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

Table of Contents

Volume 3, Issue 5, November 2009.

Pages

Reactivity Studies of Sludge and Biomass Combustion

Noorfidza Yub Harun Noorfidza.Yub_Harun@unb.ca
Graduate student
Faculty of Forestry and Environmental Management
University of New Brunswick, PO Box 4400
Fredericton, E3B 5A3, Canada

M.T. Afzal mafzal@unb.ca
Associate Professor
Department of Mechanical Engineering/FORUM
University of New Brunswick, PO Box 4400
Fredericton, E3B 5A3, Canada

Noorliza Shamsudin
Department of Chemical Engineering
Universiti Teknologi Petronas
Tronoh, 31750, Malaysia

Abstract

Sludge and biomass are wastes with energy value. Both can provide a renewable energy in the form of gaseous fuels through thermal conversion processes. Proper understanding of the thermal properties and reaction kinetic of sludge and biomass is important for efficient design, operation and modeling of the conversion process. This study was carried out to obtain the kinetics data of the sludge and biomass in pure oxygen atmosphere at 30 $mlmin^{-1}$ with the combustion temperature ranging from 50 to 900^{o}C. The effect of sample size and heating rate on thermal degradation was studied and kinetic parameters of sludge, bagasse and sawdust combustion are described using Arrhenius equation. Two distinct reaction zones were observed for sludge, bagasse and sawdust samples. Both the activation energy and pre-exponential factors were significantly higher in the first zone than that of in the second zone for sludge and bagasse where as an opposite trend was observed for sawdust.

Keywords: Activation Energy, Biomass, Combustion, Sludge.

1. INTRODUCTION

Present commercial energy is largely dependent on fossil fuels, which makes future sustainable development very difficult. There are drastic changes in the composition and behavior of our atmosphere due to the rapid release of polluting combustion products from fossil fuels. A significant amount of the NO_x and SO_x emissions from the energy sector is related to the use of fossil fuels for electricity generation. As the demand for electricity is growing rapidly, the emissions of these pollutants can be expected to increase unless alternatives are made available. Among the energy sources that can substitute fossil fuels, biomass wastes appear as one of the potential option [1].

Sludge from a petroleum refinery has several advantages. Calorific value of dry sludge corresponds to coal and energy content may be recovered through combustion [2]. The sludge from a petroleum refinery in Melaka has an annual production approximated 100,000 tonnes from drying and filters press alone. The limitation faced by land filling and recycling and the ban on sea disposal of sludge has lead to the expectation that the role of combustion of sludge will increase in future.

Various forms of biomass energy account for approximately 15% of the world energy supply [3]. Having an equatorial climate and fertile land, Malaysia is covered by large areas of tropical rainforest and agricultural vegetation. This provides a large resource base for biomass energy especially from plantation residues, mainly associated with the palm oil, forestry, wood, sugar cane plantation and rice milling industry [4]. It is estimated that there are about 20.8 million tones of potential biomass residues available in the country, which can generate about 700MW for energy consumption [11]. Even if only one third of the present annual output of biomass waste in the country is used to produce energy, an estimated RM2.0 billion a year can be saved in the country's import bill.

Thus, studies on the refinery sludge waste and abundant renewable forest and agriculture-based materials are necessary to see the potential of these materials to be used as solid fuel for energy production. The objective of this study is to determine the kinetic parameters of sludge, bagasse and sawdust.

2. THEORY

2.1 Determination of Kinetic Parameters

Based on Mansaray and Ghaly [18], kinetic parameters are determined by using typical curves of thermogravimetric data over an entire temperature range continuously. Whenever a peak is observed on a TGA trace as the sample is subjected to a controlled temperature ramp, it can be assumed that there has been a conversion, which can be represented as in Equation 1:

$$A \xrightarrow{k} B \tag{1}$$

Theoretically, for a solid fuel, the char reactivity towards a reactive gas is usually defined in terms of the conversion rate per remaining mass [6].

$$R = -\frac{1}{m}\frac{dm}{dt} = \frac{1}{1-X}\frac{dX}{dt} \tag{2}$$

where

$$m = m_o(1-X) \tag{3}$$

According to Pyris Software Kinetic Guide, for most system, the value of *n* is equal to 1. The degree of chemical conversion (*X*) is assumed to obey the following rate equation:

$$\frac{dX}{dt} = k(1-X)^n \tag{4}$$

The Arrhenius relationship is given by the following equation:

$$k = Z\exp(\frac{-E_a}{RT}) \tag{5}$$

Kinetics parameters are determined through experimental kinetic data. Coats and Redfern [7] have formulated equation from a single integral thermogravimetric analyzer (TGA) curve (weight loss vs. temperature) as in equation (6).

$$w(T) = \frac{dw}{dT} = \left(\frac{Z}{\beta}\right) e^{\frac{-E_a}{RT}} \tag{6}$$

Thus, the correlation between Equation (4), (5) and (6) was simplified into the following reaction rate expression as described in Equation (7).

$$[\frac{(m_0 - m)}{(m - m_f)}][\frac{1}{(T - T_0)}]\beta = Z \exp(\frac{-E_a}{RT}) \tag{7}$$

Equation (7) was reduced to a linear form by taking the natural logarithm of both sides of the equation to obtain the following equation:

$$\ln\left\{\left[\frac{(m_0 - m)}{(m - m_f)}\right]\left[\frac{1}{(T - T_0}\right]\beta\right\} = \ln Z + \left(\frac{-E_a}{RT}\right) \tag{8}$$

Change of variables was then introduced to Equation (8), which results in a linear relation between the new variables that is described in Equation (9) and then the unknown variables, *C* and *D* were solved [8].

$$Y = CX + D \tag{9}$$

Finally, the kinetic parameters that are the activation energy and pre-exponential factor were determined from the following equations respectively.

$$E_a = -CR \tag{10}$$

$$Z = \exp(D) \tag{11}$$

2.2 Thermal Degradation and Kinetic Parameters at Various Condition

The thermal degradation and kinetic parameters of selected materials were studied at two different operating conditions as follow:

i. Effect of particle size

The effect of particle size on the thermal degradation and kinetic parameters of sludge, bagasse and sawdust combustion were determined at two different sizes that are 425μm and 1.18 mm.

ii. Effect of heating rate

The experimental heating rates were set at 35°C and 50°C per min for each respective experimental run.

3. METHODOLOGY

Sludge obtained from a refinery and biomass; sawdust and bagasse, obtained in Perak, Malaysia were selected as the samples. The samples were oven-dried at 104°C to 110°C. After that, the samples were grounded and sieved to obtain homogeneous samples at size of 425μm and 1.18 mm, respectively.

3.1 Proximate and Elementary Analysis

Both proximate and elementary analyses are determined using Thermal Gravimetric Analyzer (TGA by Perkin Elmer) and CHNS-900 (by LECO) respectively. A known amount of sample was placed in a Thermal Gravimetric Analyzer (TGA by Perkin Elmer). The sample was held at 50 °C

for 1 minute before being heated from 50 °C to 110 °C at 60 °Cmin^{-1} with 99% purity nitrogen gas flow rate of 20 mlmin^{-1}. At 110 °C, the sample was on hold isothermally for 5 minutes and then heated up to 900 °C at 100 °Cmin^{-1}. The sample was then on hold at 900 °C for 15 minutes. The purge gas was then switched to 99.9% purity oxygen at a flowrate of 30 ml/min after 22 minutes of elapsed time. The thermograms produced were analyzed to determine the moisture, volatile matter, fixed carbon and ash. The elementary analysis was carried out using CHNS Analyzer and the calorific values were determined using the Oxygen Bomb Calorimeter (by IKA).

3.2 Thermal Analysis for Determination of Kinetic Parameters

The controlled parameters setting and experimental runs for the thermal analysis is given in Table 1. For each run, the samples were first held at 50 °C for 1 minute and then heated from 50 °C to 900 °C in oxygen atmosphere at a flowrate of 30 ml/min. The time taken to complete the thermal analysis is approximately 25 and 18 minutes for samples with heating rate of 35 °C and 50 °C respectively.

Raw material	Experiment No.	Particle size	Heating rate
Sludge	1	425 μm	35 ºCmin^{-1}
	2		50 ºCmin^{-1}
	3	1.18 mm	35 ºCmin^{-1}
	4		50 ºCmin^{-1}
Bagasse	5	425 μm	35 ºCmin^{-1}
	6		50 ºCmin^{-1}
	7	1.18 mm	35 ºCmin^{-1}
	8		50 ºCmin^{-1}
Sawdust	9	425 μm	35 ºCmin^{-1}
	10		50 ºCmin^{-1}
	11	1.18 mm	35 ºCmin^{-1}
	12		50 ºCmin^{-1}

Table 1: Controlled Parameters and Experimental Run for Thermal Analysis

4. RESULT AND DISCUSSION

4.1 Drying Process

The samples were dried in an oven at a constant temperature range of 104°C to 110 °C. The results showed that the moisture content (wet basis) of sludge, bagasse and sawdust was initially at 75.37%, 66.12 % and 16.12 % respectively in the beginning of the experiment (i.e. at 0 hour). Figure 1 illustrated the moisture content for the samples were observed to decrease gradually until the end of the experiment.

A rapid moisture loss during the initial state of drying was observed for the first 7 hours, 24 hours and 48 hours for sawdust, bagasse and sludge, respectively. This is attributed to the evaporation of moisture from the surface area of the sludge and biomass used [9]. According to Alves and Figueiredo [11], when heat is applied to wood particles, the particles begin to dry more intensely at the outer boundary, at which the temperature is high. Whereas the bound and free water tend to move outward by convection and diffusion although some may migrate towards the inner, colder parts of the solid, where recondensation occurs. However, as the drying process continued, the penetration of heat into the deeper part of the wood particle and the moisture movement to the surface become harder due to material resistance. Thus, resulting in slower rate of drying. This theory explains the results obtained for all drying curves of sawdust, bagasse and sludge.

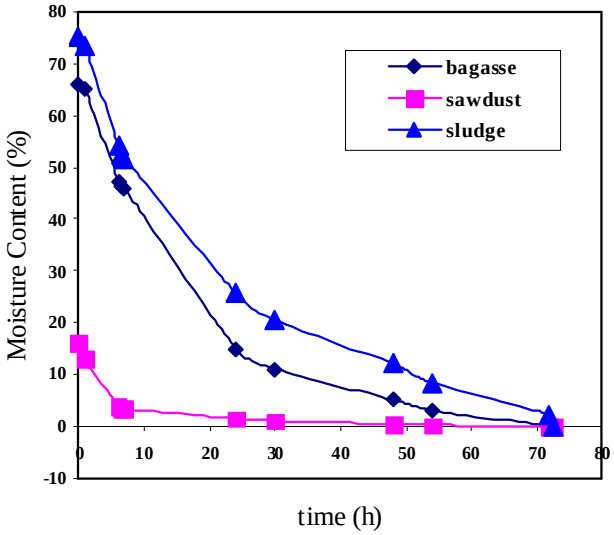


Figure 1: Drying Curve of Sludge, Bagasse and Sawdust

The duration of the rapid moisture loss for bagasse and sludge is longer than sawdust, which explained by the higher initial moisture content of bagasse and sludge, compare to sawdust. Longer time is required in order to release moisture content for material with higher initial moisture content when same heat or temperature applied.

4.2 Proximate and Elementary Analysis

Table 2 shows the results for both proximate and elementary analysis as well as calorific values of the selected materials. The moisture and volatile matters for sludge and bagasse are higher than sawdust. It was observed that some of the sludge and bagasse sample was not completely combusted, thus leaving some residue or ash of 25 and 4.6 weight percentage. The high value of ash in both sludge and bagasse compared to sawdust is contributed by the low value of carbon content. Charles [17] suggested that agricultural residues have lower carbon content as compared to wood residues. This is proven in the result obtained where the fixed carbon content of bagasse is less than sawdust. The values of proximate analysis for biomass samples were comparable with values obtained from literatures.

The fuel elements of biofuels are actually the volatile matters and the carbon content while moisture and ash contents contribute as the impurities [17]. The result obtained showed that the volatile matter of sludge and carbon content of sawdust is greater than bagasse. In contrast, the moisture as well as ash content of sawdust is less than sludge and bagasse. Furthermore, the energy content of sludge and biomass is measured by its calorific value. The calorific value for bagasse is 18.1MJ/kg which higher than that of sludge and sawdust, 16.9 and 17.0 MJ/kg

respectively. Lim [23] has reported the calorific value of sawdust and bagasse is around the same values, which is 18.86 $MJkg^{-1}$ and 17.33 $MJkg^{-1}$, respectively. Just sawdust is slightly higher than of bagasse.

The sulfur content of bagasse and sawdust were much higher than of sludge. The nitrogen content of sludge has much higher that that of biomass. This fact constitutes an advantage from a fertilizer point of view if considering the agricultural use of refinery sludge, could be an important drawback for combustion to be considered but other metal content in it may harmful the land and underwater stream. The NO_x emission could be expected during their combustion. Nevertheless, NO_x emission has been found decreased by combustion experiences. Albertson et al [2] reported that this could be related to the high volatile content in sludge, which accelerates the air consumed during combustion, diminishes the rate O/N in the flame and represses the thermal generation of NO_x.

Sample	Moisture (%) db	Volatile Matters (%) db	Ash (%) db	Carbon, C (%) db	Hydrogen, H (%) db	Nitrogen, N (%) db	Sulfur, S (%) db	High Heat Value, HHV (kJ/kg) db
Sludge	3.9	55.8	25	21.4	3.4	2.2	0.98	16880
Bagasse	2.77	48.13	4.6	44.5	5.32	0.23	6.05	18110
Sawdust	1.54	17.5	0.98	56.4	2.5	0.05	2.35	17029

db = dry basis

Table 2: Result of Proximate and Elementary Analysis for Sludge, Bagasse and Sawdust

However, the overall result showed that sawdust is having a better potential as fuel element as compared to sludge and bagasse in terms of higher C/O, C/H and C/N ratio.

4.3 Thermal Analysis

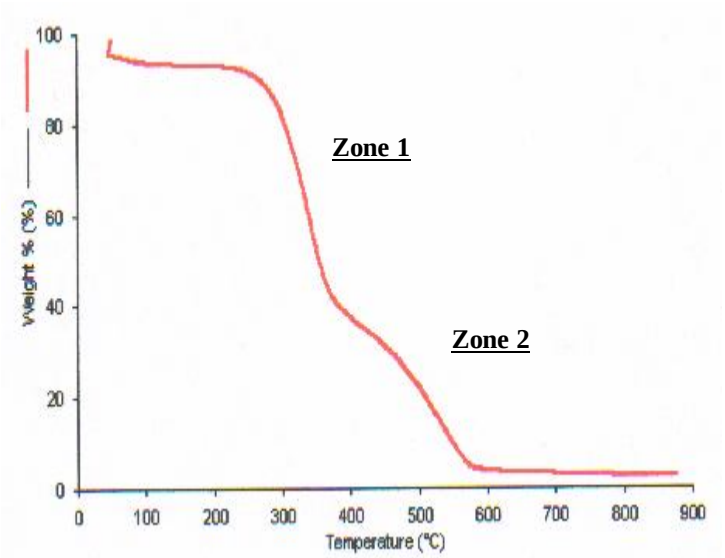


Figure 2: TGA thermograms by one of the experimentals runs

Figure 2 shows the TGA thermograms produced for thermal analysis by one of the experimental runs. The continuous records of weight loss and temperature were obtained to determine the thermal degradation characteristics as well as kinetic parameters of samples tested. It produces

two zones of degradation, which are decomposition, (Zone 1) and combustion reaction, (Zone 2) respectively. It was necessary to determine and use different kinetic parameters to describe the thermal degradation over the entire temperature range with higher accuracy [21].

Therefore, the thermal degradation characteristics and the kinetic parameters (activation energy and pre-exponential factor) of each sample were determined for the first and second reaction zones separately, and results obtained were depending on or subjected to the operating conditions.

4.3.1 Thermal Degradation

The results of thermal degradation of sludge, bagasse and sawdust in the first and second reaction zones are summarized in Table 3 and 4 respectively. The 2 zones region were plotted to see the effects of differences in particles size and heating rate as shown in Figure 3, 4 and 5.

It was observed from the thermograms that increasing the temperature from 50°C to 110°C resulted in weight losses. According to Mansaray and Ghaly [21], this is due to loss of water present in the samples and external water bounded by surface tension. In the first zone, the degradation temperature of sludge started from 120°C to 143°C and bagasse started approximately from 206 °C to 217 °C, whereas for the sawdust started from 259 °C to 270 °C. It was also observed that rapid degradation of sludge and bagasse took place in the first zone due to the rapid evolution of the volatile products.

The total degradation for sludge and bagasse were in the range of 51.3% to 56.7% and 42.0% to 48.6%, respectively, while sawdust has the total degradation range from 14.1% to 17.3%. The end of the first reaction zone was accepted as the beginning of the second reaction zone. The final temperature of the second reaction zone was approximately 397 °C to 565°C for sludge, 585 °C to 674 °C for bagasse, and 733 °C to 793 °C for sawdust. Result shows that for the second reaction zone the total degradation for sludge was in the range of 19.3% to 21.7 %, bagasse was in the range of 43.9 % to 45.3%, while sawdust was in the range of 76.4% to 79.5%. The residual weights of the sludge, bagasse and sawdust at the end of the experiment (at 900 °C) were in the range of 21.0 % to 24.25 %, 4.31 % to 4.71 % and 0.94% to 0.98%. These values were close to the initial ash contents of the original samples (Table 2), which indicate that, thermal degradation was completed at above the final degradation temperatures for all samples.

Sample	Exp. run	Water evolved (% db*)	Initial deg. Temp. (°C)	Final Temp. (°C)	Total deg. (%)
Sludge	1	3.1	120	213	54.7
	2	3.4	137	231	56.7
	3	3.5	115	243	51.3
	4	4.3	143	268	53.4
Bagasse	1	1.7	206	378	46.8
	2	2.1	217	381	48.6
	3	2.3	207	385	42.0
	4	2.7	211	389	43.4
Sawdust	1	1.1	259	415	16.8
	2	1.3	270	433	17.3
	3	1.5	259	422	14.1
	4	1.8	267	426	15.9

db = dry basis

Table 3: Thermal Degradation in First Reaction Zone

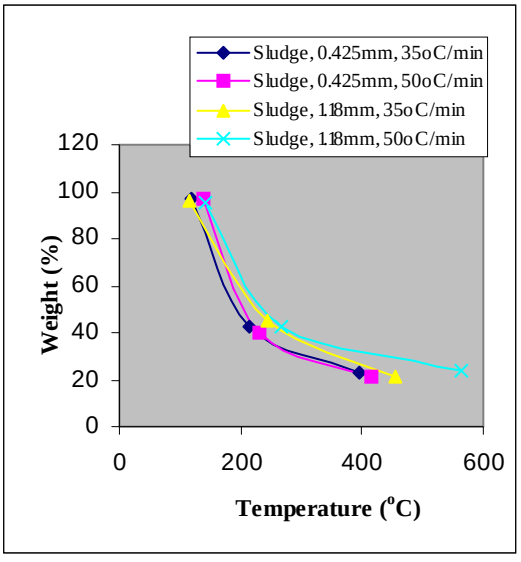


Figure 3: Thermal Degradation of Sludge

Sample	Exp. run	Initial deg. Temp. (°C)	Final Temp. (°C)	Total deg. (%)	Residual weight (%db)
Sludge	1	213	397	20.7	23.13
	2	231	415	21.7	21.00
	3	243	455	19.3	21.27
	4	268	565	20.4	24.25
Bagasse	1	378	585	44.4	4.31
	2	381	606	45.3	4.63
	3	385	667	43.9	4.71
	4	389	674	45.1	4.63
Sawdust	1	415	733	77.4	0.94
	2	433	793	79.5	0.98
	3	422	726	76.4	0.97
	4	426	752	78.2	0.97

Tabl

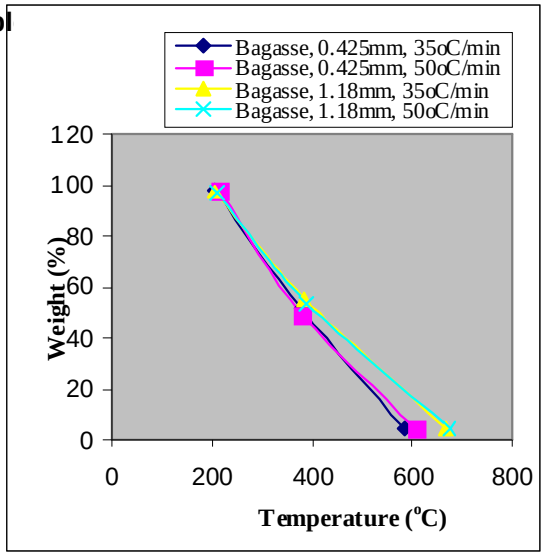


Figure 4: Thermal Degradation of Bagasse

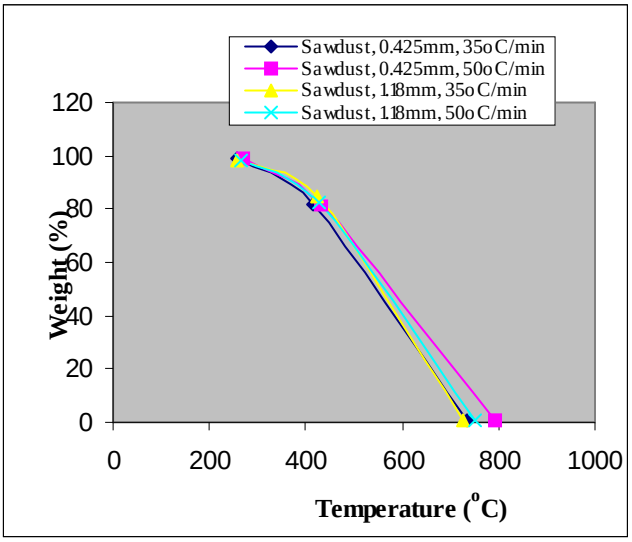


Figure 5: Thermal Degradation of Sawdust

The total degradation for sludge was lower in the second reaction zone as compared to the first reaction zone due to the fact that rapid volatilization in the first zone and left with low carbon content to degrade. The total degradation for bagasse is slightly lower in the second zone than in the first zone, may be due to the fact that lignin was decomposed to char [24]. However, sawdust total degradation in the second zone was higher than in the first zone. Hornof *et al.* [27] studied the effect of lignin content on the thermal degradation of wood pulp and found that the reactions above 330 °C are mostly dominated by the decomposition of lignin, which is the case mostly in the second reaction zone.

From the result obtained, it shows that the total degradation of sawdust is higher than of bagasse. This is because; sawdust is having higher carbon content and lower ash content as compared to bagasse as shown in Table 2.

4.3.2 Kinetic Parameters

The kinetic parameters (activation energy and pre-exponential factor) of each sample were determined for the first and second reaction zones separately by applying the least squares method to the thermogravimetric data. The kinetic parameters obtained for both first and second reaction zones are presented in Table 5.

It is observed that the activation energy is proportional to the pre-exponential factor as governed by the Arrhenius equation as in Equation 5. For the first zone, the activation energy of sludge falls in the range of 42.98 to 45.05 kJ/mol, bagasse falls in the range of 76.77 to 98.96 kJ/mol whereas for sawdust, the activation energy of sawdust falls in the range of 38.60 to 48.59 kJ/mol. The pre-exponential factor for the first reaction zone lies between 3.71×10^4 to 7.11×10^5 min^{-1} for sludge, 2.09×10^5 to 2.26×10^7 min^{-1} for bagasse and 1.77×10^2 to 3.97×101^2 min^{-1} for sawdust.

Sample	Exp. Run	Ea (kJ/mol)		Z (1/min)	
		$1^{st.}$ zone	$2^{nd.}$ zone	$1^{st.}$ zone	$2^{nd.}$ zone
Sludge	1	44.29	35.68	5.43×10^5	6.80×10^4
	2	45.05	37.71	7.11×10^5	7.11×10^4
	3	42.98	32.45	3.71×10^4	7.03×10^3
	4	43.97	34.97	3.94×10^4	7.23×10^3
Bagasse	1	97.32	72.72	1.45×10^7	3.80×10^4
	2	98.96	61.26	2.26×10^7	8.63×10^4
	3	76.77	69.29	2.09×10^5	5.52×10^3
	4	78.12	64.50	3.27×10^5	8.82×10^3
Sawdust	1	38.60	142.37	3.33×10^2	6.04×10^{10}
	2	48.59	152.14	3.97×10^2	7.10×10^{10}
	3	42.67	78.23	1.77×10^2	1.01×10^5
	4	43.34	116.70	1.87×10^2	1.22×10^5

Table 5: Kinetic Parameters of First and Second Reaction Zones

For the second reaction zone, the activation energy of sludge is between 32.45 to 37.71 kJ/mol, bagasse is between 61.26 to 72.72 kJ/mol while for sawdust, the value lies between 78.23 to 152.14 kJ/mol. Result obtained also shows that for the second reaction zone, the pre-exponential factor of sludge, bagasse and sawdust is between 7.03×10^3 to 7.11×10^4 min^{-1}, 5.52×10^3 to 8.63×10^4 and 1.01×10^5 to 7.10×10^{10} min^{-1}, respectively.

Figure 6 shows that the kinetic parameters for sludge and bagasse in the second reaction zone are lower than the first reaction zone. However, the kinetic parameters for sawdust in the second reaction zone are higher than in the first zone. Binning and Jenkins [10] and Ergudenler and Ghaly [19] also reported lower kinetic parameters in the second reaction zone compared to the first reaction zone for bagasses.

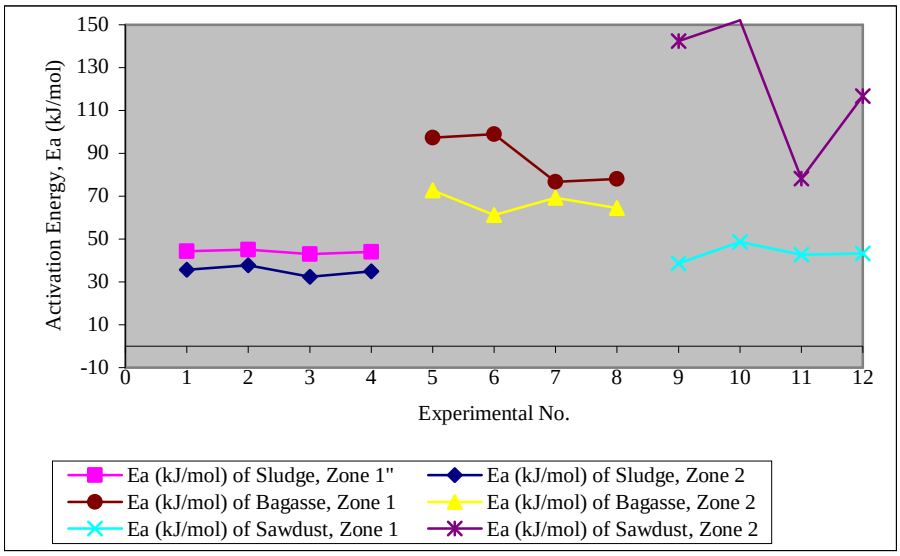


Figure 6: Activation Energy Analysis by Arrhenius Model

4.3.3 Effects of Operation Conditions on Thermal Degradation Characteristics and Kinetic Parameters

4.3.3(a) Effect of Particle Size

As the particle size was increased from 425 μm to 1.18 mm, the total degradation, activation energy and pre-exponential factor decreased in both first and second reaction zones. According to Shodor [31], one of the conditions for molecules to react, which is stated by the collision theory, is to have the right geometry. In conjunction with that, increasing the surface area of solids may increase the rate of a reaction. The reasons are, by increasing the surface area or in other words by decreasing the particle size, it will allow for more collisions of particles and give more molecules the right geometry to react. In contrast, by increasing the particle size will result in less collision and therefore slower reaction will take place thus causing the thermal degradation characteristics and kinetic parameters to decrease regardless of the reaction zones.

4.3.3(b) Effect of Heating Rate

The effect of heating rate on thermal degradation and kinetic parameters of the sample shows that as the heating rate was increased from 35°C to 50°C, the total degradation and kinetic parameters were also increased in the first and second reaction zone.

Basically, the total degradations value corresponds to the volatile evolution phenomenon. According to the research done by Sampath [32], during devolatilization, a substantial thermophysical and thermochemical heat requirement is associated with volatile evolution. Maloney [33] work suggests that the coal or biomass structure is initially constrained and, at high heating rates, a finite time is required for the structure to relax and respond to the thermal input. This induction period may be responsible for pushing volatile evolution to a higher temperature. Hence, as the heating rate increased, the total degradation was also increased regardless of the reaction zones.

The kinetic parameters also increased as the heating rate was increased. This is in accordance with to the collision theory, increase in temperature will give more molecules the right energy (so called the activation energy) thus, increases the rate of reaction. This is proven by the results obtained in this study.

5. CONCLUSION

Two distinct reaction zones were observed for sludge, bagasse and sawdust. Based on Arrhenius equation, the kinetic parameters; activation energy and pre-exponential factors, were found to be significantly higher in the first reaction zone than that of in the second zone for sludge and bagasse where as an opposite trend was observed for sawdust.

6. REFERENCES

1. E. Natarajan, A. Nordin, A.N. Rao. "*Overview of Combustion and Gasification of Rice Husk in Fluidized Bed Reactors*". Biomass and Bioenergy, 14(5): 533-546, 1998

2. O.E. Albertson, A. Baturay. "*Deodorization and Cleaning of Medium Temperature Wet Off-gases Derived from Burning of Wet Waste Sludge*". Environmental International, 17(2-3): 11-111, 1991

3. M. Hoogwijk, A. Faaij, R. van den Broak, G. Berndes, D. Gielen, W. Turkenburg. "*Exploration of the ranges of the global potential of biomass for energy"*. Biomass and Bioenergy, 25(2): 119-133, 2003

4. A.A. Tajuddin. *"Hydroelectricity and TNB's Other Renewable Energy Initiatives"*. In Kamaruzzaman, Mohd Yusoff Hj. Othman, Baharuddin Yatim of World Renewable Energy Congress '99 Malaysia. 1999

5. A. Ergundenler, A.E. Ghaly. "*Determination of reaction kinetics of wheat straw using thermogravimetric analysis*". Applied Biochemistry and Biotechnology, 34/35(1):75-91, 1992

6. A.F. Ismail, A.H. Shamsuddin, F.M.A. Mahdi. *"Reactivity Studies of Rice Husk Combustion Using TGA"*. In Kamaruzzaman, Mohd Yusoff Hj. Othman, Baharuddin Yatim of World Renewable Energy Congress '99 Malaysia. 1999

7. A.W. Coats, J.P. Redfern. "*Kinetic Parameters from Thermogravimetric Data"*. Nature 201: 68-69, 1964

8. John H. Mathews. "*NUMERICAL METHODS: for Mathematics, Science and Engineering"*. Prentice-Hall International Limited, London, UK. 1992

9. K.M. Sabil. "*Study on the Production of Charcoal by Carbonization Using Mangrove Wood in Malaysia*". In Proceedings of Second Engineering Congression Engineering Innovation and Sustainability: Global Challenges and Issues. Sarawak, Malaysia. 2002

10. A.S. Bining, B.M. Jenkins. "*Thermochemical reaction kinetics for rice strawfrom an approximation integral technique*". In ASAE Paper No. 92-6029. St. Joseph, MI, USA. 1992

11. S.S. Alves, Figuereido. "*A model for pyrolysis for wet wood*". Chemical Engineering Science, 44(12): 2861-2869, 1989

12. A. V. Bridgewater. *"Progress in Thermochemical Biomass Conversion"*, Blackwell Science Ltd., pp. 47-48, 61-63, 632-633 (2001)

13. A.V. Bridgewater. "*Developments in Thermochemical Biomass Conversion*", Blackie Academic & Proffesional, pp. 294-296 (1997)

14. *"Govt Agrees to Palm Oil Biomass for Power Project"*. In Business Times, 28/2/2001

15. *"Biomass future for by-products".* In Business Times, 3/9/2001

16. C.M. Earnest, R.I. Fyans. "*Recent Advances in Microcomputer Control Thermogravimetry of Coal and Coal Products*", ASTM STP 997, pp. 3 (1988)

17. Y. W. B. Charles, B.H. Essel. "*Biomass Conversion And Technology*", John Wiley & Sons, (1996)

18. E.C. Greencia, T. Aida, H. Niiyama, "*The heating rates of carbonization*". In Proceeding of Conference Advanced Catalytic Science and Technology. Tokyo Institute of Technology, Tokyo, Japan, 1997

19. A.E. Ghaly, A. Ergundenler. "*Thermal degradation of cereal straws in air and nitrogen*". Applied Biochemistry and Biotechnology, 27(1):111–126, 1991

20. I.J. Goldfarb, R. Guchan, A.C. Meeks. "*Kinetic analysis of thermogravimetry. Part II. Programmed temperatures*". In Report No. ARML-TR-68-181. Ohio: Air Force Laboratory, Wright-Patterson AFB, 1968

21. K.G. Mansaray, A.E. Ghaly. "*Determination of Kinetic Parameters of Rice Husk in Oxygen Using TGA*". Biomass and Bioenergy, 17(1): 19-31, 1999

22. K.O. Lim. "*The Energy Potential and Current Utilization of Agriculture and Logging Wastes in Malaysia*". The Planter, 57:182-187, 1986

23. K.O. Lim. "*Energy Productivity of Some Plantation Crops in Malaysia and the Status of Bioenergy Utilization*". In Kamaruzzaman, Mohd Yusoff Hj. Othman, Baharuddin Yatim of World Renewable Energy Congress '99 Malaysia. 1999

24. T.A. Milne, M.N. Soltys. "*Direct Mass Spectrometric Studies of the Pyrolysis of Carbonaceous Fuels: II. Qualitative Observations of Primary and Secondary Processes in Biomass*". Analytical and Applied Pyrolysis, 5(2): 111-131, 1983

25. T.A. Milne, M.N. Soltys. "*Direct Mass Spectrometric Studies of the Pyrolysis of Carbonaceous Fuels: I. A Flame Pyrolysis Molecular-beam Sampling Technique*". Analytical and Applied Pyrolysis, 5(2): 93-110, 1983

26. R.J. Evans, T.A. Milne, M.N. Soltys. "*Direct Mass Spectrometric Studies of the Pyrolysis of Carbonaceous Fuels: III. Primary Pyrolysis of Lignin* ". Analytical and Applied Pyrolysis, 9(3): 207-236, 1986

27. V. Hornof, B.V. Kokta, J.L. Valade, J.L. Fassen. "*Effect of Lignin Content of Thermal Degradation of Wood Pulp*". Thermochimica Acta, 19(1): 63-68, 1977

28. K. Raveendran, A. Ganesh, K.C. Khilar. "*Pyrolysis Characteristics of Biomass and Biomass Components*". Fuel, 75: 987-998, 1996

29. M. Rozainee, M. Rashid, S. Looi. "*Production of Renewable Energy From Biomass and waste Materials Using Fluidised Bed Technologies*". In Kamaruzzaman, Mohd Yusoff Hj. Othman, Baharuddin Yatim of World Renewable Energy Congress '99 Malaysia. 1999

30. M.S. Duvvuri, S.P. Muhlenkamp, K.Z. Iqbal, J.R. Weker. "*The Pyrolysis of Natural Fuels*". Fire and Flammability, 62: 467 – 468, 1975

31. http://www.Shodor.org/unchem/advanced/kin/index.html

32. R. Sampath, D.J. Maloney, W.M. Proscia, E.R. Monazam. "*Effect of Heating Rate on the Thermodynamic Properties of Pulverized Coal*". In Semi-Annual Progress Report. Department of Engineering, Clark Atlanta University, Atlanta, 1998

33. Maloney, D.J., Sampath, R., and Zondlo, J.W., "*Heat Capacity and Thermal Conductivity Considerations for Coal Particles During Early Stages of Rapid Heating, Combustion and Flame*". Combustion and Flame, 116: 94-104, 1999

34. S. Yusoff, "*Renewable Energy Form Palm Oil – Innovation on effective Utilization of Waste*". Cleaner Production, 14: 87-93, 2006

35. Z. Husain, Z.A. Zainal, M.Z. Abdullah. "*Analysis of Biomass-residue-based Cogeneration System in Palm Oil Mills*". Biomass and Energy, 24: 117-124, 2003

36. J. Werther, T.Ogada. "*Sewage Sludge Combustion"*. Energy and Combustion Science, 25(1): 55-116, 1999

Selection of the Best Module Design for Ultrafiltration (UF) Membrane in Dairy Industry: An Application of AHP and PROMETHEE

Mohsen Pirdashti pirdashti@shomal.ac.ir
Faculty of Engineering/Chemical Engineering Department/ Shomal University
Amol, 46134, Iran

Majid Behzadian behzadian_ie@yahoo.com
Faculty of Engineering/Industrial Engineering Department/ Shomal University
Amol, 46134, Iran

Abstract

Membrane with a type module has been expressed one of the key area of interest in dairy industry. Although recent publications have given a chance academics and practitioners to prove successful applications of membrane processes to the vast areas; a small number of publications have been devoted attention to the problem of capital equipment decision-making. To facilitate decision-making process in the membrane separation, this report focuses on the application of analytical hierarchy process (AHP) and Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE), from a group decision-making viewpoint; it uses the Delphi technique to evaluate available alternatives according the criteria elicited from expert's opinions. A real case study on the ultrafiltration membrane area is put forward to determine the best module design based on the five criteria expressed by decision-makers: sanitation design, clean–in-place, packing density, resistance to fouling and shear stress, and relative cost. Finally, the paper utilizes Expert Choice and DECISION LAB software's to facilitate calculations.

Keywords: Analytic Hierarchy Process (AHP), PROMETHEE, Delphi, multi-criteria decision making, group decision support systems, membrane separation, ultrafiltration.

1. INTRODUCTION

Dairy has been recognized as an important food industry which applies membranes in several sections (1). Since the discovery of asymmetric membrane by Loeb and Souriragin in the early 1960's, membrane processes as microfiltration (MF), ultrafiltration (UF) and reverse osmosis (RO) have widely been applied to the dairy, food and beverage industry (2). Compared to phase-inversion processes, membrane processes are inherently less energy intensive; they are able to save energy up to 30-50% of the current values. In addition, membrane processes are compact,

modular and easily amenable automation and scale-up; they are fast (due to fast mass transfer through the thin membrane layer), more efficient, unique in some applications (such as hemodialysis), consume no or less additives and chemical sand applicable to very dilute solutions (such as biotechnology products) with lower capital investments. Membrane processes are carried out in low temperature; hence they can be used for heat-sensitive materials such as food, medical and biotechnological products. Successful applications of membrane processes have been reported by a variety of industries around the world (3). Some of recent membrane applications in dairy industry are shown in table 1 (4).

Applications	**Membrane types**
Cheese whey concentration	UF-RO
Milk concentration	UF-RO
Desalting of whey	ED
Waste treatment	UF

TABLE 1: Some of membrane application in dairy industry.

UF is a pressure-driven membrane process widely considered as economical alternatives to conventional separation processes valuable products in the dairy industry; it has performed at the standard unit operations (5, 6). In addition to selection of the proper membrane martial and optimization of operating conditions (feed velocities, temperature, staging, etc.), bulk stream pretreatment (removal or stabilization of foulants), the selection of proper module configuration is one of the key stages in the industrial membrane system design (4). Module design for UF membrane in dairy application cases includes a considerable number of alternatives and criteria which should be analyzed and optimized (7). There has been a considerable demand for an appropriate management in determination of the best module as well as resolving the priority of other alternatives in comparison to the selected module. To cope with this demand, the AHP method and PROMETHEE from a group decision-making view of point has been put forward for a satisfactory solution. Flexibility of the results achieved in this study for AHP method, particularly in the sensitivity analysis stage, as well as taking into account the latest statistical information, and good agreement between the AHP result with PROMETHEE output has made this study a valuable contribution to selection of the best module in dairy application cases.

2. AVAILABLE MODULES

The term “module” is universally used; but the definition varies. Here, a module is the simplest membrane element that can be put into practice. Module design should deal with five major issues, plus a host of minor ones. First is economy of manufacture; Secondly, a module should mobilize a strong support and be able to properly maintain membrane integrity against damage and leaks; thirdly, it should deploy the feed stream so as to make intimate contact with the membrane, provide sufficient mass transfer to keep polarization in control, and do so with a minimum waste of energy; fourthly, the module should permit easy aggress of permeate; and fifthly it should permit the membrane to be cleaned when necessary. Despite the invention of various types of modules, a fair number of modules have commercially been operated; but the winning designs have varied depending on a few simple themes since 1996 (8). In order to membrane module is a unit assembly containing a combination of membranes and the membrane containment vessel (4). Commercially available modules include spiral wound, hollow fiber, tubular and plate-and-frame modules (5). These modules for membrane design were assessed in this study; a general description of the different module types is presented in the following sections.

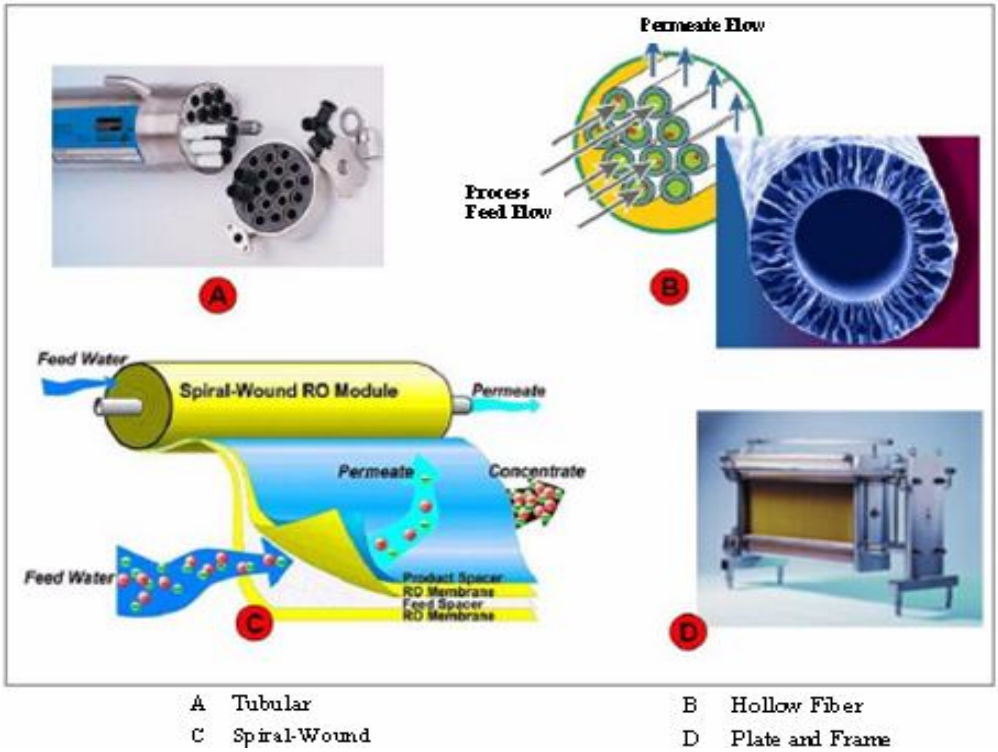


FIGURE 1: Membrane configurations for UF membrane.

Tubular Module (TM)

Tubular modules, which are among the earliest industrial designs, are often preferred for solutions containing suspended solids. The membranes are formed as tubes and are normally cast onto a supporting porous substrate. The tubes are not self-supporting and are normally inserted in a perforated tube. The tubes are typically housed in a shell and tube configuration. The membranes are normally sealed by means of elastomeric inserts at either end of the tube. The end caps determine if the flow is in series or parallel through the module. Permeate is collected in the shell side of the module (9).

Hollow-Fiber Module (HFM)

Hollow fiber modules were developed in the late 1960's for use in desalination and has subsequently been adopt for UF and MF applications. A hollow fiber module consists of a bundle of self-supporting hollow fibers set inside a cylindrical shell and potted at both ends to create a shell and tube arrangement. The fibers range in size from 100 µm to 2500 µm. (9).

Spiral-Wound Module (SWM)

The spiral wound membrane module consists of two flat sheet membranes wound around a central core. The membranes are oriented with the permeate site of the membranes facing. The membranes are separated by a spacer and then glued together to form an envelope. Two enveloped separated by a feed spacer are attached to the central core and wound Swiss role style around the core. The spacer on the feed side acts as a turbulence promoter and support. Once wound the ends by fixed by and anti-telescoping device. Modules are either tape bound and held in a steel housing or made as glass-reinforced modules. The feed runs axially through the unit. The permeate flows around the spiral and is removed via the central tube. This module is wrapped into a spiral and placed in a cylinder shell (9).

Plate and Frame Module (PFM)

The plate and frame configuration consists of membranes cut from flat sheets and bounded or clamped to a flat supporting frame is normally grooved or milled to permit flow of the permeate. The basic layout can be compared with that of a standard filter press (8). In Figure 1, the arrows show the upstream and Permeate paths. The upstream leaves as the retentive and is enriched in Non-permeate. Permeates is collected from channels in support plates and Leaves enriched in the most permeable component (10).

3. ANALYTICAL HIERARCHY PROCESS(AHP) APPLICATION

The Analytic Hierarchy Process (AHP) is a multi-criteria decision-making (MCDM) approach that simplifies complex, ill-structured problems by arranging the decision factors in a hierarchical structure. The AHP is a theory of measurement for dealing with quantifiable and intangible criteria that has been applied to numerous areas, such as decision theory and conflict resolution (11). One of the main characteristics of the AHP method is the possibility that AHP offers in group decision-making (12, 13,14). Since the evaluation of projects usually demands a remarkable team effort, AHP is an available method which provides decision-makers (DMs) with a systematic framework for group interaction and decision-making (15). The AHP method is selected for its specificity, which offers a certain freedom to a DM to express his preferences for particular criteria by using the original AHP measurement scale. The AHP method does not require such explicit quantification of criteria; but it needs specific hierarchical structuring of the MCDM problem. The method itself then generates the weights of the criteria by using the AHP measurement scale according to a specified procedure. Under such circumstances, a comparison of the results from such different methods applied to the same problem appears to be very interesting and challenging from both academic and practical perspectives (16). According to Saaty [17,18, 19], the following steps are intended to proceed the AHP method:

1. Define the problem and determine its goal,
2. Structure the hierarchy with the decision-maker's objective at the top with the intermediate levels capturing criteria on which subsequent levels depend and the bottom level containing the alternatives, and
3. Construct a set of $n \times n$ pair-wise comparison matrices for each of the lower levels with one matrix for each element in the level immediately above. The pairwise comparisons are made using the relative measurement scale in Table 2 (20, 21, 22). The pair-wise comparisons capture a decision maker's perception of which element dominates the other.
4. There are n (n-1)/2 judgments required to develop the set of matrices in step 3. Reciprocals are automatically assigned in each pair-wise comparison.
5. The hierarchy synthesis function is used to weight the eigenvectors by the weights of the criteria and the sum is taken over all weighted eigenvector entries corresponding to those in the next lower level of the hierarchy.
6. After all the pair-wise comparisons are completed, the consistency of the comparisons is assessed by using the eigenvalue, λ, to calculate a consistency index, CI:

$$CI = (\lambda - n)/ (n-1) \quad (1)$$

Where n is the matrix size. Judgment consistency can be checked by taking the consistency ratio (CR) of CI with the appropriate value in Table 3. Saaty [1980](21) suggests that the CR is acceptable if it does not exceed 0.10. If the CR is greater than 0.10, the judgment matrix should be considered inconsistent. To obtain a consistent matrix, the judgments should be reviewed and repeated.

7. Steps 3-6 are performed for all levels in the hierarchy (23).

Numerical rating	Verbal judgments of preferences
9	Extremely preferred
8	Very strongly to extremely
7	Very strongly preferred
6	Strongly to very strongly
5	Strongly preferred

4	Moderately to strongly
3	Moderately preferred
2	Equally to moderately
1	Equally preferred

TABLE 2: Pair-wise comparison scale for AHP preference.

Size of matrix	1	2	3	4	5	6	7	8	9	10
Random consistency	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

TABLE 3: Average random consistency.

Group AHP Method

While AHP can be used to capture the priorities of individual decision participants, it is necessary to combine these individual assessments into a consensus. To aggregate individual AHP judgments into a group decision, there are two perspectives as follows:

Aggregation of Individual Judgment

In this view, a group decision matrix is constructed from the unique matrix of each decision participant. An element of this matrix (a_{ij}^{G}) is calculated using a geometric average of the elements from each unique matrix,

$$a_{ij}^{G} = \left\{ \prod_{K=1}^{n} (a_{ijk})^{\beta_K} \right\}^{\frac{1}{\sum \beta_K}} = \left\{ \prod_{K=1}^{n} (a_{ijk})^{\beta_K} \right\} \quad ,i,j = 1,...,m \quad ,K = 1,...,n \tag{2}$$

Where β_k and a_{ijk} are the importance and efficiency of the K decision and the elements of the K matrix, respectively (24).

Aggregation of Individual Priorities (AIP)

In this approach, the order of the decision weights for each decision alternative for the *K* decision (W_i^{k}), *K*=1….*n*, where n is the number of decision-makers, is calculated and a group decision weight (W_i^{G}) for the alternative is constructed:

$$W^{G} = (W_i^{G}) \quad ;W_i^{G} = \prod_{K=1}^{n} (w_i^{K})^{\beta_K} \qquad i = 1,...,m \tag{3}$$

Where β_k indicates amount and importance of effectiveness of *K* decision and W^{G} matrix indicate aggregation weight of a single judgment in respect to each alternative.

In both approaches, each individual judgment affects the final judgment β_k. So that:

$$\sum_{K=1}^{n} \beta_k = 1 \tag{4}$$

After aggregating the individual judgments, matrices with the same dimensions as the unique individual matrices are constructed in which the local and final weights as well as the inconsistency of each matrix and total inconsistency are calculated with the same basic AHP method (24).

4. PROMETHEE II

The PROMETHEE II is used to provide a complete ranking on a finite set of feasible alternatives from the best to the worst. The central principle of PROMETHEE II is based on a pair-wise comparison of alternatives along each recognized criterion. Alternatives are evaluated according to different criteria, which have to be maximized or minimized. The implementation of PROMETHEE II requires relevant information on the weights and preference function of the criteria. PROMETHEE II assumes that the decision-maker is able to weigh the criteria appropriately, at least when the number of criteria is not too large (25). There are a number of papers in this regard combined PROMETHEE with AHP (26,27, 28). The final ranking of alternatives in this integration was done by PROMETHEE and the importance of criteria was determined by AHP. For each criterion, the preference function translates the difference between the evaluations obtained by two alternatives into a preference degree ranging from zero to one. In order to facilitate the selection of a specific preference function, Brans and Vincke (1985) (29)proposed six basic types, namely: Usual criterion, *U*-shape criterion, *V*-shape criterion, Level criterion, *V*-shape with indifference criterion, and Gaussian criterion. These six types are particularly easy to define. For each criterion, the value of an indifference threshold *q*, the value of a strict preference threshold *p*, or the value of *s* an intermediate value between *p* and *q* has to be fixed (30). The following paragraphs present stepwise procedure for the implementation of PROMETHEE II:

Step 1: Determination of deviations based on pair-wise comparisons

$$d_j(a,b) = g_j(a) - g_j(b) \tag{5}$$

Where $d_j(a,b)$ denotes the difference between the evaluations of *a* and *b* on each criterion.

Step 2: Application of the preference function

$$P_j(a,b) = F_j[d_j(a,b)] \qquad j = 1, \ldots, k \tag{6}$$

Where $P_j(a,b)$ denotes the preference of alternative *a* with regard to alternative *b* on each criterion, as a function of $d_j(a,b)$.

Step 3: Calculation of an overall or global preference index

$$\forall\ a, b \in A\ , \qquad \pi(a,b) = \sum_{j=1}^{k} P_j(a,b) w_j \tag{7}$$

Where $\pi(a,b)$ of *a* over *b* (from zero to one) is defined as the weighted sum $p(a,b)$ of for each criterion, and w_j is the weight associated with *j*th criterion.

Step 4: Calculation of outranking flows and partial ranking

$$\phi^+(a) = \frac{1}{n-1} \sum_{x \in A} \pi(a,x) \tag{8}$$

and

$$\phi^{-}(a) = \frac{1}{n-1}\sum_{x \in A} \pi(x, a) \quad (9)$$

Where $\phi^{+}(a)$ and $\phi^{-}(a)$ denote the positive outranking flow and the negative outranking flow for each alternative, respectively.

Step 5: Calculation of net outranking flow and complete ranking

$$\phi(a) = \phi^{+}(a) - \phi^{-}(a) \quad (10)$$

Where $\phi(a)$ denotes the net outranking flow for each alternative

5. THE PROCEDURE

Utilizing Delphi process

The Delphi method has been a popular tool in information systems research because the process increases the confidence with which researchers can use the results in subsequent studies and managers can make decisions based on the information gathered using this method. Researchers employ this technique primarily in cases where judgmental information is indispensable. The Delphi technique has been used in a number of real application cases: to predict technological developments (31), to identify issues affecting health care administration (32, 33), to assess interventions and policies in the mental health industry (34), to construct a model for project funding decisions at the National Cancer Institute (35), to evaluate the strategic importance of jobs in pay rate decisions (36), to evaluate emerging directions in information systems management (37), and to assess strategic responses to threats from global terrorism (38). A key advantage of Delphi is that it avoids direct confrontation between the participating experts (39). While real applications differ markedly according to the type of industry, the Delphi method consists of three essential processes that achieve information exchange among a group of DMs without introducing the potential biases of interpersonal interaction. The first process is to collect judgments, along with the underlying rationales, from individuals who are knowledgeable about an issue by questioning them individually. The second process is to collate and statistically summarize the individual judgments and rationales without revealing the identity of the individuals. The third process is to feed back the collated information to the DMs and ask them to reconsider their judgments. This sequence of collating, feedback and revision is repeated over several rounds until further repetitions produce practically no changes in individual judgments. Both Delphi and AHP assume knowledgeable DMs (40). Therefore, the careful selection of the participants for the DM group is important. This issue was discussed by Brockoff (1983) (40) and Preble (1984)(41). The experts who were selected to participate in this study include University Professors, Diary producers, Membrane inhibitors, who are familiar with the membrane process. While the experts in these groups are knowledgeable about the membrane process issues, they were provided with the most recent research results and statistical information from the UF separation. After identifying knowledgeable experts, the next step was sending questionnaires to elicit the experts' opinions about the factors affecting on selection module for UF membrane in dairy industry. The first questionnaire asked the participants to identify the factors they thought were most important in selecting a module. A comprehensive and mutually exclusive list was collated from the responses. A second questionnaire included the list generated from the initial responses and asked the experts to check those factors they considered to be important. A list of potential impacts was presented, and respondents indicated those impacts they considered most important. These were grouped under five objectives; these five factors were included as criteria:

Sanitation Design (SD)

Sanitation design plays an essential factor in the selection of a specific UF module. Sanitary operation is of paramount importance because the design allows for easy cleaning and sterilization.

Clean –in-Place (CIP)

Easy to clean or the ability to clean membranes when necessary and replace membranes is good. This is an important factor in comparing the different UF modules types.

Packing Density (PD)

Packing density is the ratio of Membrane surface to membrane module volume. That low packing density is a problem in high pressure where pressure vessel costs are significant.

Resistance to Fouling and Shear stress (RFS)

Fouling is a process resulting in loss of performance of a membrane due to the deposition of suspended or dissolved substances on its external surfaces, at its pore opening, or within its pores. It is one of the major considerations in the handling of milk products, particularly the UF of whey; however, resistance to fouling and self-cleaning is very important parameters for this application.

Relative Cost (RC)

The membrane module should satisfy a number of economic requirements. In order to optimization of initial module costs and operating life are very important in industrial membrane system design.

Application of expert choice software to facilitate the calculations

Expert Choice (43) provides strong support for DMs faced with solving complex problems, involving the evaluation of alternative courses of action on qualitative as well as quantitative criteria (44, 45). The software helps a DM to devise the structure of a complex problem as a hierarchy of criteria and alternatives; it guides the DM through a series of simple pairwise comparisons to solve the problem. While Expert Choice is powerful and intuitive, it is also easy to use; therefore, the solution is more likely to reflect the expertise of the DM while minimizing interference from the program and the computer. Figure 2 shows the hierarchical structure of the problem. The five criteria identified in the Delphi process and alternative modules are the second level and third level of structure, respectively.

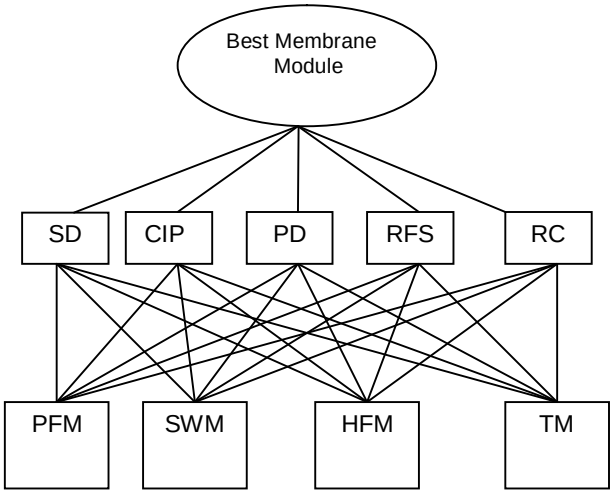


FIGURE 2: A hierarchal representation of problem with five criteria and four alternatives.

5. RESULTS

In the next step, each of the participants in the project used Expert Choice to assess the relative importance of the criteria and to evaluate the priority of alternative site locations. Pairwise comparisons form the core of the AHP technique. At each level of the hierarchy below the goal, a DM is asked to compare each possible pair of factors (*ci* and *cj*) and to provide judgments on the relative importance of each. As illustrated in Figure 3(a), each expert was asked to make pairwise comparisons between each possible pair of criteria.

	CIP	PD	RFS	RC
SD	3.0	5.0	5.0	7.0
CIP		2.0	2.0	3.0
PD			1.0	2.0
RFS				2.0

FIGURE 3(a): Compare the relative importance with respect to: GOAL.

These judgments provided inputs to Expert Choice. As described in the Appendix, once the pairwise comparison matrix at a given level of the hierarchy is complete, Expert Choice calculates the relative weights for the various factors at that level. Figure 3(b) shows the output for one of the experts. For this expert, SD considered the most important criterion while LE was assigned the least weight.

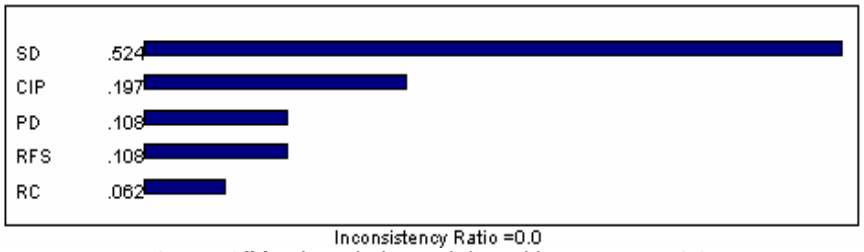


FIGURE 3(b): The relative weights with respect to: GOAL

In addition, Expert Choice computes an inconsistency ratio (IR) for each DM and encourages DMs whose IR exceeds 0.10 to reconsider their pairwise judgments. For example, if an expert rates CIP as two times more important than PD, and two times more important than RSF; then logically for that expert, PD and RSF should be equally important. Suppose, however, that in a pairwise comparison between these two criteria, the expert declares PD to be three times more important than RSF. In this case, a substantial inconsistency has occurred, and the calculated IR would be greater than 0.10. Expert Choice would encourage the DM to reconsider all of the underlying pairwise comparisons, and after a few trials, the DM should arrive at an acceptable level of consistency. Among the experts participating in this project, the largest IR for the initial comparison of the criteria was 0.02(<0.10). This level of inconsistency was very low, which indicated the meaningfulness of the criteria to the DMs. After assigning weights to the criteria, the next task for the experts was to evaluate the alternatives on these criteria. While there has been some criticism of AHP in the operations research literature, Harker and Vargas (1987) (46) and Saaty (1983) (47) have shown that AHP does have an axiomatic foundation. The cardinal measurement of preferences is fully represented by the eigenvector method, and the principles of hierarchical composition and rank reversal are valid. On the other hand, Dyer (1990a)(48) has questioned the theoretical basis underlying AHP, and argued that it can lead to preference reversals based on the alternative set being analyses. In response, Saaty (1990b) (49) explained how rank reversal is a positive feature when new reference points are introduced. In the module design decision, the geometric aggregation rule is used to avoid the controversies associated with rank reversal (49, 50, 51). When all the comparisons between criteria and alternatives were made by each expert, geometric averaging of the individual comparisons was used to synthesize

the criterion weights and alternative priorities for the expert group. The EC outputs are from all the AHP processes. The results reveal that SWM and HFM were the modules preferred by the experts are shown figure 4(a) to 4(c).

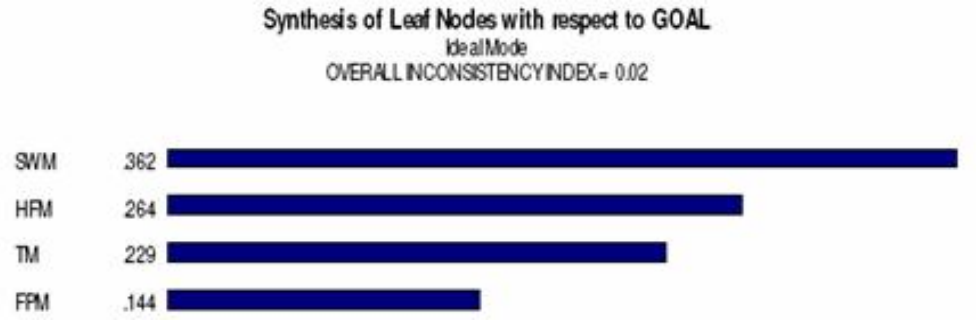


FIGURE 4(a): Synthesis of leaf nodes with respect to GOAL.

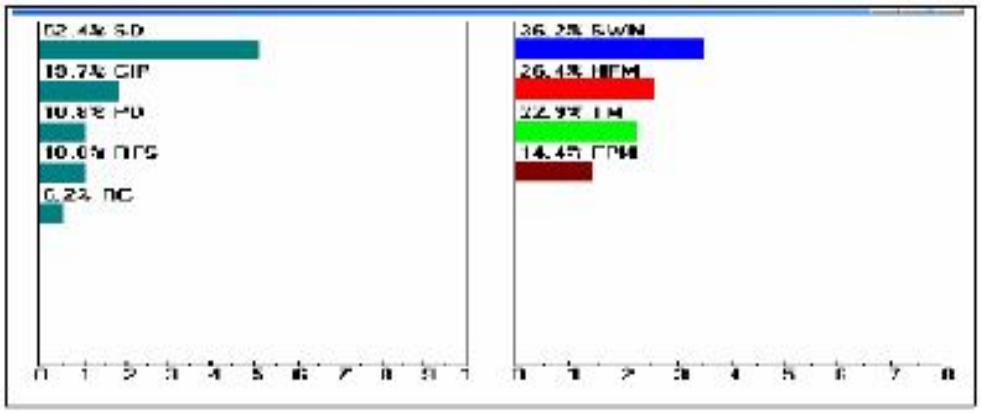


FIGURE 4(b): Dynamic sensitivity with respect to GOAL for nodes below GOAL.

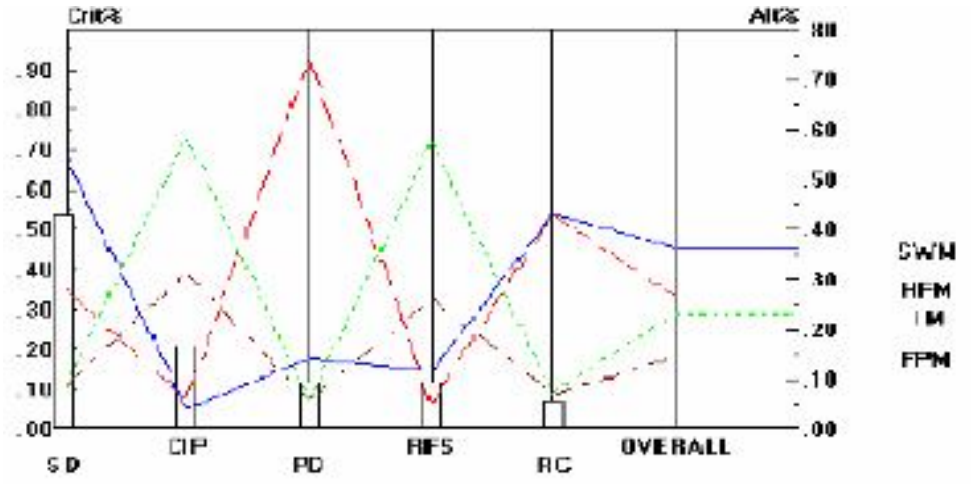


FIGURE 4(c): Performance sensitivity with respect to GOAL for nodes below GOAL.

The software is easy to use and understand; it provides visual representations of overall ranking on a computer screen.
In order to rank alternative technical requirements according to PROMETHEE II, the DECISION LAB software was run. The software, which was developed in collaboration with the Canadian company *Visual Decision*, is the current software implementation of the PROMETHEE & GAIA methods (52).

	RC	**CIP**	**RES**	**PD**	**SD**
Weight	0.062	0.107	0.108	0.108	0.524
Max/Min	Min	Max	Max	Max	Max
Type	5	5	5	5	5
q	1	1	1	1	1
p	3	3	3	3	3
PFM	3	5	5	4	3
SWM	7	3	4	5	7
HFM	6	3	2	8	5
TM	2	7	6	3	3

Table4: PROMETHEE II multi-criteria information.

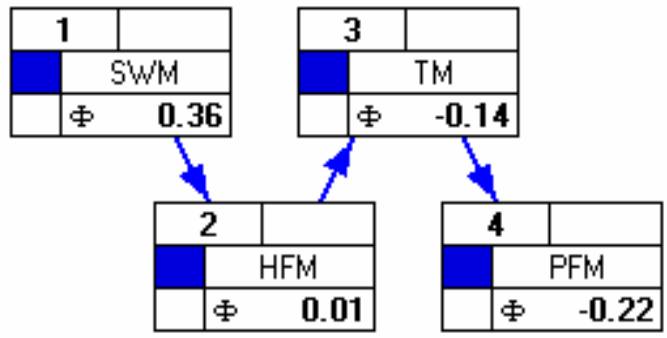


FIGURE 5: Ranking four alternatives by Decision lab software.

In this specific problem, the alternative *SWM* is preferred from the multi-criteria viewpoint, whereas *PFM* is selected the worst alternative, as shown in Table 4. Figure 5 shows the result of the complete ranking with PROMETHEE II for four modules, as an output of DECISION LAB and this is a good agreement between the AHP result with PROMETHEE output.

6. SENSITIVITY ANALYSIS

Sensitivity analysis was used to investigate the stability of the alternatives' assigned priorities to changes in the relative importance of the criteria. For example, RC was not reliably predictable, and the facilities provided at various industrial modules were likely to be improved in the future. Consequently, it would desirable to examine the impact of changes in these weights on the priorities of the alternative facility modules. For dairy application, the experts considered RC to be the most volatile criterion. Figure 6(a) to 6(d) show the potential impact of the changes in RC on the priorities of the alternative modules. A 650% decrease to 852% increase in the relative weight assigned to RC, from 0% to 55.4%, produced no change in the ranking of the modules in

comparison to Figure 6(a) and 6(d). Individually and as a group, the experts explored the impact of numerous scenarios on the weights and alternative priorities. They considered the solution presented in Figures 6(a) to 6(d) to be not only the most desirable, but also the most robust.

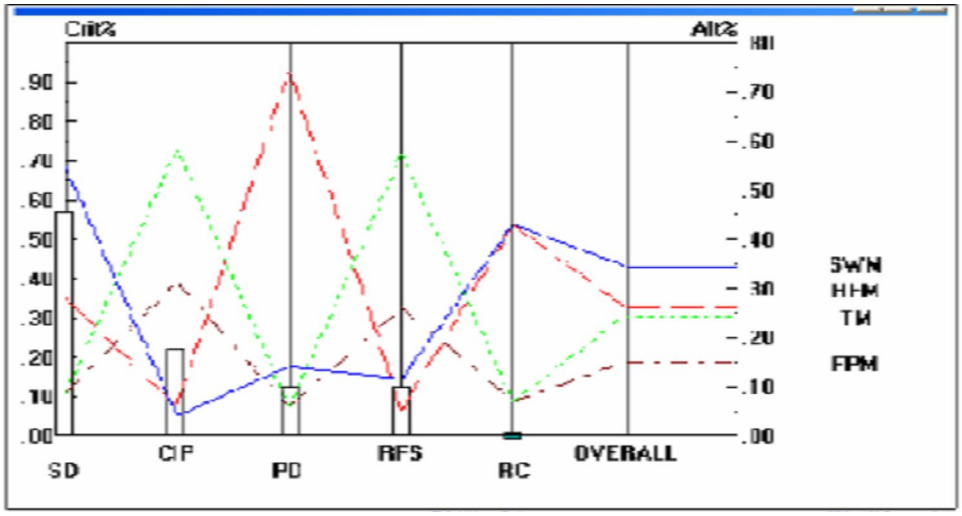


FIGURE 6(a): Performance sensitivity with respect to GOAL for nodes below GOAL.

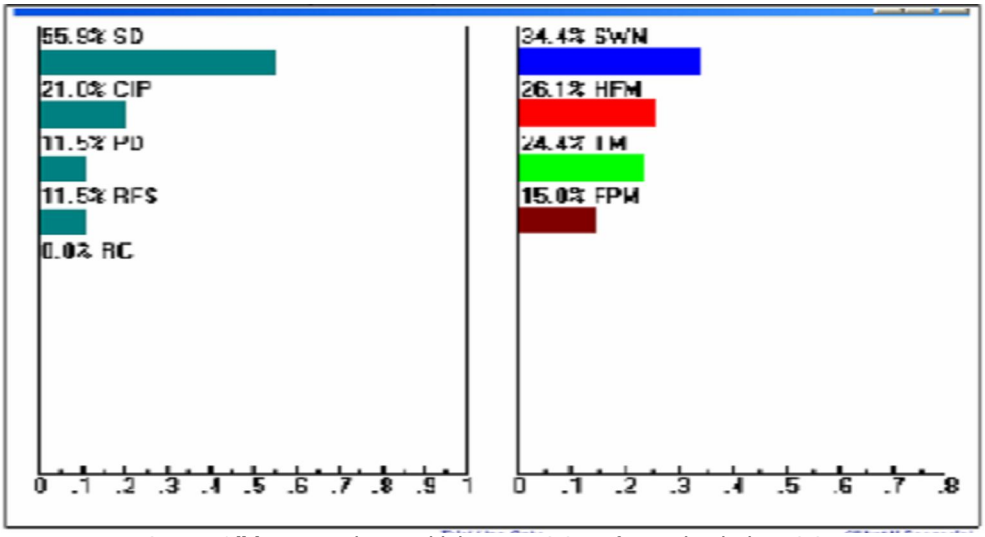


FIGURE 6(b): Dynamic sensitivity w.r.t GOAL for nodes below GOAL.

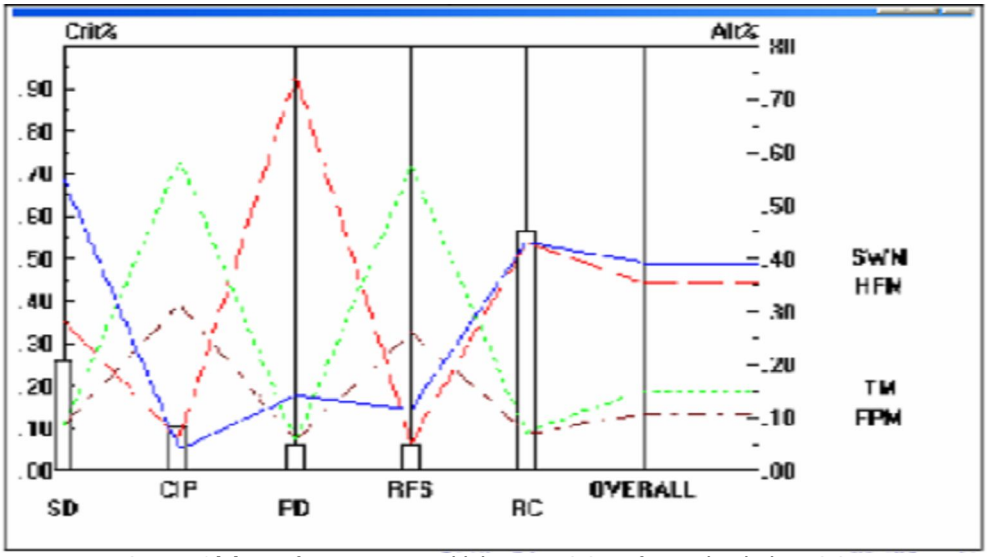


FIGURE 6(c): Performance sensitivity w.r.t GOAL for nodes below GOAL.

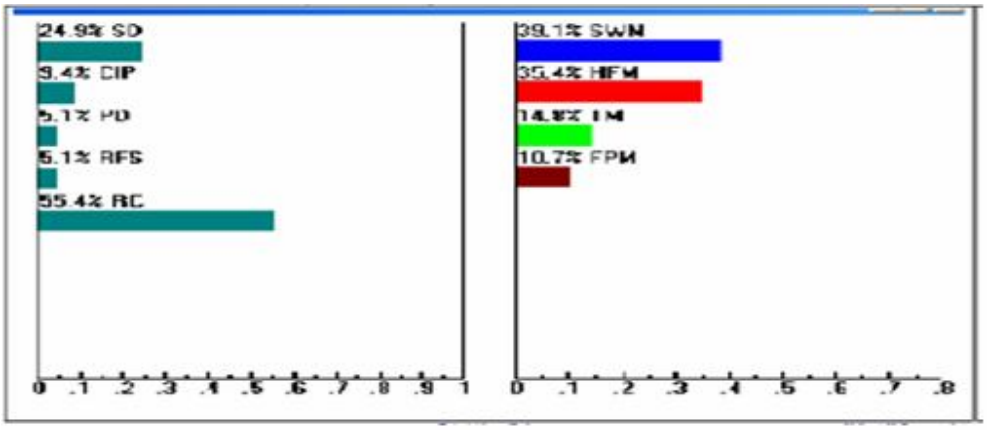


FIGURE 6(d): Dynamic sensitivity w.r.t GOAL for nodes below GOAL.

8. CONCLUTION

Despite the widespread interests for the application of membrane process to the dairy industry, arriving at a rational decision to select the best alternative from a group decision-making approach has been recognized as one of the main challenges among academics and practitioners. In this regard, the determination of the best Module design for UF membrane in dairy industry is a problem that involves both quantitative and qualitative criteria; however, the scale of a problem rises up due to complicated nature of the problem, presence of various alternatives and criteria, and a genuine difference of opinion between the experts. The major contribution of this paper is to use an AHP- Delphi multi-criteria model and a PROMETHEE-Delphi multi-criteria model to elicit, process, and synthesize these quantitative and qualitative expert opinions. For this purpose, the model applied Delphi technique to elicit expert opinions about criteria for the evaluation of four modules: SD and CIP, Sanitation design and Clean-in-place were considered important by the experts. The models also used AHP and Expert Choice and PROMETHEE and DECISION LAB to capture the priorities of individual decisions and to arrive at a common consensus among the DMs. As a result, the models developed in this paper provides a useful guideline as a structured and logical means of synthesizing judgments for

evaluating appropriate modules for an organization; it devised an elaborate structure in a difficult and often emotion-burdened decision. The second implication is the functionalities of the modules listed in the models. Thus, DMs can examine the strengths and weaknesses of each module. The authors believe that this study would be especially useful for academics and practitioners who are working in dairy industry and are interest in decision science.

9. REFERENCES

1. J.Wagner. *"Membrane Filtration Handbook Practical Tips and Hints"*. Second Edition, Revision 2 (2001)

2. B.Jiao, A.Cassano, E. Drioli. *"Recent Advances on Membrane Processes for the Concentration of Fruit Juices: A Review"*. Journal of Food Engineering, 63: 303-324, 2004

3. S.S.Madaeni, S.Zereshki. *"Membrane Technology as a Promising Alternative for Less Energy Consumption in Food Industry"*. In Proceedings of the Third International Conference on Energy Research and Development (ICERD- 3), 2006
4. Oak Ridge National Laboratory. *" Materials for Separation Technologies: Energy and Emission Reduction Opportunities"*. Energy Efficiency and Renewable Energy, U.S Department of Energy, Oak Ridge,TN 37830, (2005)

5. J.Schwinge, P.R.Neal, S.E.Wiley, D.F.Fltcher, A.G Fane . *"Spiral Wound Modules and Spacers: Review and Analysis"*. Journal of Membrane Science, 242: 129-153, 2004

6. S.Muthukumaran, S.E.Kenitish, M. Ashokkumar, G.W. Stevens. *"Mechanisms for the ultrasonic enhancement of dairy whey ultrafiltration"*. Journal of membrane science, 258: 106-114, 2005

7. P.Le-Clech. Z. Cao, P.Y.Wan, D.E.Wiley, A.G Fane. *"The application of constant temperature anemometry to membrane processes"*. Journal of membrane science, 284: 416-423, 2006

8. W.Eykamp. *"Membrane Separation Processes in Perry's Chemical Engineers Handbook"*.7th (Eds), Mc Graw-Hill,NY.22.39-22.39-22.41(1997)

9. Cousins, R.B. *" Membrane Technology"*. Department of Chemical and Process Engineering University of Strathclyde in Glasgow (2003-2004)

10. S. Ali, P.Boblak, E.Capili, S.Milidovich. *" Membrane Separation and Ultrafiltration",* CHE-396 Senior Design.,(1998).

11. L.Vargas. " *An overview of analytic hierarchy process: Its applications"*, European Journal of Operational Research 48 (1) 2–8, 1990

12. D. H. Byun, *"The AHP approach for selecting an automobile purchase model"*. Information & Management, 38: 289-297, 2001

13. M. T.Escobar, J.Aguaron, J. M. Moreno-Jimenes. *" A note on AHP group consistency for the row geometric mean priorization procedure"*. European Journal of Operational Research, 153: 318-322, 2004

14. V. S.Lai, B. K.Wong, W. Cheung. *"Group decision making in a multiple criteria environment: A case using the AHP in software selection"*. European Journal of Operational Research, 137: 134-144,2002

15. T.L. Saaty." *Decision making for leaders"*. Pittsburgh, PA: RWS Publications (1982)

16. M.Janic, A. Reggiani. *"An Application of the Multiple Criteria Decision Making (MCDM) Analysis to the Selection of a New Hub Airport"*. EJTIR, 2(2): 113-14, 2002

17. T. L. Saaty. " *Eigenvector and logarithmic least squares*", European Journal of Operational Research 48: 156-160, 1990

18. T.L. Saaty. *"Highlights and critical points in theory and application of the analytical hierarchy process"*. European Journal of Operational Research, 74: 426-447, 1994a

19. T.L. Saaty. *"How to make a decision: the analytic hierarchy process".* Interfaces, 24: 19-43, 1994b

20. T.L. Saaty. *"Modeling unstructured decision problems: a theory of analytical hierarchies"*. In Proceedings of the First International Conference on Mathematical Modeling, St. Louis, Missouri, 69–77, 1977

21. T.L. Saaty. " *The Analytical Hierarchy Process"*. Mc Graw-Hill. New York, (1980)

22. T.L. Saaty. *"How to make a decision: the analytical hierarchy process"*. European Journal of Operational Research, North-Holland, 48: 9-26,1990a

23. K. M. Al-Subhi Al-harbi. *"Application of the AHP in project management".* International Journal of Project Management, 19: 19-27, 2001

24. Z.Xu. *"On consistency of the weighted geometric mean complex judgment matrix in AHP"*. European Journal of Operational Research, 126, 683-687, 2000

25. C.Macharis, A.Verbeke, K. De Brucker. '*The strategic evaluation of new technologies through multi-criteria analysis: the advisors case'*, in Bekiaris E. and Y. Nakanishi (eds.) Economic impacts of Intelligent Transport Systems. Innovations and case studies, Elsevier, Amsterdam (2004)

26. Z.Babic, N.Plazibat. *"Ranking of enterprises based on multicriterial analysis".* International Journal of Production Economics, 29–35, 1998

27. R.U.Bilsel, G.Buyukozkan, D. Ruan. *"A fuzzy preference-ranking model for a quality evaluation of hospital web sites"*. International Journal of Intelligent Systems 21: 1181–1197, 2006

28. J.J.Wang, D.L. Yang. *"Using a hybrid multi-criteria decision aid method for information systems outsourcing"*. Computers & Operations Research 34: 3691– 3700, 2007

29. J.P. Brans, P.Vincke. *"A preference ranking organization method. The PROMETHEE method for MCDM"*. Management Science, 31: 641-656, 1985

30. J.P.Brans, B.Mareschal. *" PROMETHEE V–MCDM problems with segmentation constraints"*. INFOR, 30(2): 85-96, 1992

31. M.K. Malhotra, D.C. Steele, V. Grover.*"Important strategic and tactical manufacturing issues in the 1990s"*. Decision Sciences, 25(2): 189-214, 1994

32. R.P.Hudak, Jr. P.P. Brooke, K. Finstuen. P.Riley. *"Health care administration in the year 2000: practioners".* views of future issues and job requirements. Hospital and Health Services Administration, 38(2): 181-195, 1993

33. M.Tavana, D.T.Kennedy, P.Joglekar. *"A group decision support framework for consensus ranking of technical manager candidates"*. Omega, International Journal of Management Science, 24: 523-538. 1996

34. R. Bijl. *"Delphi in a future scenario study on mental health and mental health care".* Futures, 24(3): 232-250, 1992

35. N.G.Hall, J.C.Hershey, L.G. Kessler, R.C. Stotts. *"A model for making project funding decisions at the national cancer institute"*. Operations Research, 40(6): 1040-1052, 1992

36. T.E Weinberger. *" The strategic centrality of jobs: a measure of value"*. Compensation and Benefits Review, 24(1), 61-68, 1992

37. M.Tavana, D.T Kennedy. *" N-site: a distributed consensus building and negotiation support system"*. International Journal of Information Technology and Decision Making, 5(1): 123-154, 2006

38. F.Niederman, J.C. Brancheau, J.C Wetherbe. *"Information systems management issues for the 1990s"*. MIS Quarterly, 15(4): 474-500, 1991

39. Ch.Okoli, S. Pawlwoski . *"The Delphi method as a research tool: an example, design considerations and applications"*. Journal of Information and Management, 42: 15-29, 2004

40. Dalkey, C., Norman, L. R. Daniel, R.Lewis, D. Snyder," *Studies in the Quality of Life, Delphi and Decision-Making*", Lexington Books, PP. 20-21(1972)

41. K. Brockoff. *" Group process for forecasting"*. European Journal of Operational Research, 13(2),115-127, 1983

42. J.F. Preble. *"The selection of delphi panels for strategic planning purposes"*. Strategic Management Journal, 5(2): 157-170, 1984

43. www.Expertchoice.com

44. M.Tavana, D.T. Kennedy, B. Mohebbi. *"Total quality index: a benchmarking tool for total quality management"*. Benchmarking: An International Journal, 10(6): 507-527, 2003

45. M.Tavana, D.T. Kennedy , B.Leauby, J. Ugras. *"An investigation of the relative importance of the characteristics of information in governmental financial reporting"*. Business Journal, 19(1-2), 56-62, 2004

46. P.T.Harker, L.G. Vargas. *" The theory of ratio scale estimation: Saaty's analytic hierarchy process"*. Management Science, 33:1383-1403,1987

47. T.L. Saaty. "Axiomatic *foundations of the analytic hierarchy process"*. Management Science, 32: 841-855, 1983

48. J.S. Dyer. *" Remarks on the analytic hierarchy process"*. Management Science, 36: 249-258, 1990a

49. T.L. Saaty. *"An exposition of the AHP in reply to the paper 'Remarks on the analytic hierarchy process"*. Management Science, 36: 259-268, 1990b

50. J.S. Dyer . *"A clarification of Remarks on the analytic hierarchy process"*. Management Science, 36: 274-275, 1990b

51. P.T. Harker, L.G. Vargas. "Reply to 'remarks on the analytic hierarchy process by J.S. Dyer". Management Science, 36, 269-273, 1990

52. Decision Lab 2000-Getting started Guide. "www.visualdecision.com" "Visual Decision Inc. Montreal, Canada (1999)

Inventory Control In Supply Chain Through Lateral Transshipment – A Case Study In Indian Industry

Dharamvir Mangal mangaldharamvir1@rediffmail.com

Assistant Professor/ Mechanical Engineering Department
MD University Rohtak
TITS, Bhiwani, 127021
Haryana, India

Pankaj Chandna pchandna08@gmail.com

Associate Professor/ Mechanical Engineering department
NIT, Kurukshetra, 132119
Haryana, India

Abstract

Supply chain management emphasizes collaborative relationships between supply chain members. The purpose of this work is to examine the antecedents of retailer - retailer partnership and to explore its impact on the supply chain performance. We consider coordination among stocking locations through replenishment strategies that take explicitly into consideration transshipments, transfer of a product among locations at the same echelon level. A continuous review inventory system has been adopted, in which lateral transshipments are allowed. In general, if a demand occurs at a location and there is no stock on hand, the demand is assumed to be backordered or lost. Lateral transshipments serve as an emergency supply in case of stock out and the rule for lateral transshipments is to always transship when there is a shortage at one location and stock on hand at the other. The aim is to explore the role of lateral transshipment to control inventory and associated cost within supply chain and, from this, to develop an exploratory framework that assists understanding in the area. A simple and intuitive model is presented that enables us to characterize optimal inventory and transshipment policies for 'n' locations. The research is based on a case study of a bi-wheeler company in India by using its data and to strengthen its supply chain. This paper represents such an effort in that it integrates both inventory and transshipment components in the study of multi-location inventory systems. This work will enable the managers to overcome the uncertainties of demand and lead-time resulting into customer satisfaction and cost reduction.

Keywords: Inventory control 1, Supply chain management 2, Lateral transshipment 3

1. INTRODUCTION AND LITERATURE REVIEW

In the last couple of decades, the numbers of products offered to the market have generally exploded. At the same time, the product life-time has decreased

drastically. The combination of these two trends leads to increased inaccuracy of the demand forecasts, leading to firms facing an increased demand uncertainty resulting the increase in inventory levels. The role of inventory as a buffer against uncertainty has been established for a long time. However, more recently, the disadvantages of holding inventory have been increasingly recognised, particularly with regard to the adverse impact that this may have on supply chain responsiveness. Increasing globalisation has tended to lead to longer supply lead-times, which, by conventional inventory control theory, result in greater levels of inventory to provide the same service levels (Waters, 2002) [1]. In lean supply chain thinking, inventory is regarded as one of the seven "wastes" and, therefore, it is considered as something to be reduced as much as possible (Womack and Jones, 1996) [2]. Similarly, in agile supply chains, inventory is held at few echelons, with goods passing through supply chains quickly so that companies can respond rapidly to exploit changes in market demand (Christopher and Towill, 2001) [3]. There have been various supply chain taxonomies based on these concepts and most stress the need for inventory reduction within each of the classifications. Biju Kr. Thapalia et al., 2009 [4], covered the geographical risks for inventory strategies and their impact on supply chain with the help of a case study. F.T.S. Chan and H.K. Chan, 2009 [5], proposed an information sharing approach in multi echelon supply chains to convey exact information regarding inventory to the upstream level and simulation approach is used to test the effectiveness of proposed work. However, there has been some concern about the true costs of inventory and whether companies do in fact recognise these fully. For example, Christopher, (2005) [6] highlighted costs such as storage, obsolescence, damage, deterioration, shrinkage, insurance and management costs, as well as the more traditional cost of capital. With an incorrect assessment of inventory costs, there is the danger that companies may make inaccurate supply chain trade-offs in this respect and, therefore, hold too much inventory (Lee and Billington, 1992) [7].

Interest in the concept of supply chain management has steadily increased since the 1980s when companies saw the benefits of collaborative relationships within and beyond their own organization. The concept of supply chain is about managing coordinated information and material flows, plant operations, and logistics. It provides flexibility and agility in responding to consumer demand shifts without cost overlays in resource utilization. The fundamental premise of this philosophy is; synchronization among multiple autonomous business entities represented in it. That is, improved coordination within and between various supply-chain members. Increased coordination can lead to reduction in lead times and costs, alignment of interdependent decision-making processes, and improvement in the overall performance of each member as well as the supply chain. Kishore K. Pochampally et.al., 2009 [8] evaluated the performance of a reversed closed loop supply chain with numerical example. Supply chain management (SCM), which is also known as a logistics network (Simchi-Levi et al., 2003) [9] has been extensively studied in recent years. The logistical network consists of facilities and distribution options that perform the functions of procurement of materials, transformation of these materials into intermediate and finished products, and the distribution of these finished products to customers. SCM encompasses the management of all these (process) activities associated with moving goods from raw materials through to the end user. SCM coordinates and integrates all of these activities into a seamless process. It embraces and links all of the partners in the chain. For this reason, successful SCM

is the process of optimizing a company's internal practices, as well as the company's interaction with suppliers and customers, in order to bring products to market more efficiently.

As we know that firms can no longer effectively compete in isolation of their suppliers and other entities in the supply chain. A typical structure of a divergent inventory system is number of locations which replenish from a central supplier. Due to demand uncertainty inventory investments can be very high in such supply chain systems. A commonly used strategy to introduce flexibility in the system is to establish transshipment links between locations at the same echelon. This means that locations at the same echelon in some sense share inventory. Transshipments, the monitored movement of material between locations at the same echelon, provide an effective mechanism for correcting discrepancies between the locations' observed demand and their available inventory. As a result, transshipments may lead to cost reductions and improved service without increasing system-wide inventories. Lateral transshipments between stocking locations are used to enhance cost efficiency and improve customer service in different ways. There are basically two main approaches to capture the impact of transshipments between stocking locations. Within the first approach, transshipments are used after the demand is observed but before it is satisfied. If there is excess demand at some of the stocking locations while some have surplus inventory, lateral transshipments between stocking locations can work as a correction mechanism. Moreover, pooling the stocks can be viewed as a secondary source of supply for inventory shortages, especially when transshipments between stocking locations are faster and less costly than emergency shipments from a central depot or backlogging of excess demand. An alternative way of analyzing the impact of transshipments between stocking locations is to consider it as a tool to balance inventory levels of stocking locations during order cycles. To guarantee a certain level of customer service in all stocking locations, it is desirable to keep the inventory position at each location in balance relative to each other.

In our current research, we focus on collaborative planning and replenishment policy via lateral transshipment between retailers as a way to improve both cost and service.By allowing transshipments among locations better customer service can be achieved with retained inventory levels. We have recently witnessed an increasing use of transshipment, mainly as a result of better integration of the retailers participating in a distribution network. Normally, in divergent inventory systems, installations at some lower echelon replenish stock from a central warehouse at some upper echelon. The main benefit often associated with transshipments is balancing inventory levels at the various locations through emergency stock transfers. One of the first papers which mentioned the transshipment problem was Clark and Scarf, (1960) [10]. However, they ignored the problem due to the mathematical complexity. In another early paper, Krishnan and Rao, (1965) [11] developed a periodic review, single-echelon model in which they allow transshipments between the lower echelon stock facilities. Gross, (1963) [12] considered a transshipment model where it is assumed that both ordering and transshipments are made before demand is realized. Some papers which present results on optimality are Das, (1975) [13], Robinson, (1990) [14], Archibald et al., (1997) [15], Herer and Rashit, (1999) [16]. Das, (1975) [13] considered a two-location, single-echelon, and single-period problem with periodic review. The main

contribution of Das is the opportunity to transfer stock in the middle of a period. He derives optimal transfer and ordering rules by using stochastic dynamic programming, and shows that the so-called complete pooling policy is optimal. Robinson, (1990) [14] studied a multi-period and multi- location problem where transshipments between locations are possible. Under the assumptions of negligible transshipment and replenishment lead times, he demonstrates the optimality of the order-up-to policy. However, analytical results are only found in the two-location case. In the general case, Robinson suggests a heuristic solution method. Archibald et al., (1997) [15] examined an inventory system much related to Robinson, (1990). One difference is that in Archibald et al. Paper transshipments can be made at any time during the period. They formulate the problem as a Markov decision process, and allow emergency orders from the external supplier if transshipment from another location is not possible. Herer and Rashit, (1999) [16] studied transshipment with fixed and joint replenishment costs, but only with single period. Minner and Silver, (2005) [17] considered a distribution system with two identical locations, in which lateral transshipments are allowed. The rule for lateral transshipments is, however, not optimized. They assume that all unsatisfied demand after transshipments is lost, and develop heuristics in order to being able to evaluate costs.

In developing country like India, there has been a growing trend of realisation of supply chain optimisation. Rapid surge in global sourcing of auto components has also become a challenge for Indian manufacturers and suppliers although sourcing has reduced the cost of production substantially. Auto manufacturers in India and all tiers of the supply chain have immense opportunities to enhance their entire supply chain process with the successful implementation of SCM solution. By exploring Indian automobile sector, it has been found that uncertainties like demand and lead time have direct impact on managing inventories and managers are facing great difficulties while controlling these parameters. Customer satisfaction and cost reduction are again the key issues to be handled effectively and efficiently. Retailers in the supply chain face great uncertainty regarding consumer demands. If retailers that require long replenishment lead times from suppliers are located closer to each other, a lateral transshipment policy can be used as an effective alternative to minimizing total cost. Although transportation cost is increased, lateral transshipment is known as a better approach than a policy of no transshipments. The contribution of this paper is twofold. Firstly, we formulate a model that is simpler and more intuitive. Secondly, due to the simplicity of the model we are able to gain analytical insight into problems of higher dimensions than has been achieved earlier, i.e., problems with ‘n’ locations and general cost parameters. In this work, emergency lateral transshipment technique with variable transshipment cost is used to overcome these uncertainties. Variable transshipment cost, amongst retailers, is due to difference of distances between the retailers. In India as the retailers are apart by moderate distances so no additional transportation facility is required rather transshipment is done by road and by the vehicle itself. Model allows complete pooling between retailers. Each retailer faces normally distributed random demand pattern, demand at every retail outlet is independent of other’s demand. Lead-time is also normally distributed and independent of other retailer’s lead-time. All retailers are following periodic review policy. Based on this model, the programme has been formulated in C++. A case study has been done on an Indian automobile (bi-wheeler) manufacturing company. The formulated programme is being run on company’s actual data and the response is compared with existing results of the

company. Finally, it has been observed that comparison indicators like backorder quantity, surplus quantity and total associated cost have decreased drastically and it has been concluded that lateral transshipment is an effective technique to strengthen the whole supply chain, to achieve better customer satisfaction and in terms of improving the service levels.

2. MODEL FORMULATION

We consider an inventory system of a bi-wheeler manufacturing company in India that operates multiple retailers for the sale of bi-wheelers. The system uses a periodic-review, centralized inventory control policy. Each day, the retailers order replenishment bikes based on the total inventory available and reorder point. The external demand at each retailer comes from independent customers for the bi-wheelers.

1. Model Assumptions

The model we study assumes that central warehouse supplies a fixed quantity of units to the retailers and not allows transshipping more quantity, in case of stock-out in a particular period. If surplus quantity is remained after fulfilling the demand, retailer will hold it. Transshipment will be applied to respond to spot shortages so that if a retailer experiences a stock-out, the same will be transshipped from a nearest retailer with an adequate supply. During this practice the replenishment time is assumed to be zero or it is assumed that transshipment takes place instantaneously. Any demand that is not satisfied after transshipment is backlogged. The sequence of events in any period is delivery, demand, review, transshipment, and order. For computational tractability, the system is approximated by aggregating demand daily and allowing for a single transshipment opportunity at the end of the day right after the daily inventory review. We note this may be a limiting assumption since in practice transshipments might be used anytime during the review cycle as opportunities arise. Model allows complete pooling between retail outlets. The retailers have identical unit costs of shortage per period, holding and variable transshipment cost between any two retailers. This variable transshipment cost is due to the difference of distance between adjacent retailers.The model we study assumes that holding and shortage costs are based on a retailers' ending inventory level.

2. General Model with Emergency Transshipments

To minimize cost and inventory in a supply chain, a model has been formulated which considers a supply chain inventory system having one central warehouse or distribution centre with a very large capacity, serving to 'n' numbers of retailers with variable transshipment cost amongst the retailers.. Each retailer faces normally distributed random demand pattern, demand at every retail outlet is independent of other's demand. Lead-time is also normally distributed and independent of other retailer's lead-time. All retailers are following periodic review policy. Using this model an in-house programme has been formulated in C++ to study and to solve the existing problem of minimizing inventory and also the total associated cost.

Relationships of different costs and, different inventory policies are given as per following description. Inventory is checked at the end of every single period and if inventory is less than or equal to reorder level quantity then an order is placed. Maximum level of inventory is given as $I =$ *(Review Period + Mean Lead*

*Time)*Mean Demand or* $I = (R_p + M_l)M_d$ and reorder level is given as per following relation. $R_o =$ *Mean Lead Time*Mean Demand or* $R_o = M_l M_d$. When inventory reaches at reorder level or below this level, an order is placed. Here in transit inventory is also included, to calculate the ordered quantity by retailer *i*. Hence ordered quantity can be calculated as per following relation, $Q_{qi} =$ *Maximum Level of Inventory – (In transit Inventory + Surplus Inventory) or* $Q_{qi} = I - (Q_{in} + O_{ri})$. Surplus quantity of previous day is held by retailer. Thus total inventory for sale in particular period is given as, $I_{ti} =$ *Surplus inventory of previous day + Inventory reached that day to retailer i* or $I_{ti} = P_i + O_{ri}$. In this work applicable cost function includes holding, shortage and lateral transshipment costs, so expected cost for holding or surplus is given as, $R(CS) = \sum_{i=1}^{n}$ *Unit holding cost * surplus quantity of retailer i* or

$$R(CS) = \sum_{i=1}^{n} U_s S_i \quad (1)$$

Expected cost of shortage is given as $R(CB) = \sum_{i=1}^{n}$ *Unit penalty cost * Stock out quantity of retailer i or*

$$R(CB) = \sum_{i=1}^{n} U_p S_{oi} \quad (2)$$

and expected cost of lateral transshipment is given by $R(CL) = \sum_{i=1,j=1,i\neq j}^{i=n,j=n}$ *Unit transshipment cost*transshipment quantity from retailer i to j* or

$$R(CL) = \sum_{i=1,j=1,i\neq j}^{i=n,j=n} U_t O_{ri} \quad (3)$$

$U_t = A_{ij} b$, where A_{ij} is the distance between retailers i and j and b is unit distance travelling cost. In case of, without transshipment expected cost will be sum of expected holding cost and expected stock out cost. It can be written as following. $R_2(C)$ *= Expected Holding Cost + Expected Stock-Out Cost* or $R_2(C) = R(CS) + R(CB)$ or

$$R_2(C) = \sum_{i=1}^{n} U_s S_i + \sum_{i=1}^{n} U_p S_{oi} \quad (4)$$

Now expected cost per period, with transshipment, will be sum of expected holding cost, expected shortage cost, expected lateral transshipment cost. It can be given by following relationship. $R_1(C)$ *= Expected Holding Cost + Expected Shortage Cost + Expected Lateral Transshipment Cost* or $R_1(C) = R(CS) + R(CB) + R(CL)$ or

$$R_1(C) = \sum_{i=1}^{n} U_s S_i + \sum_{i=1}^{n} U_b S_{oi} + \sum_{i=1, j=1, i \neq j}^{i=n, j=n} U_t Q_{ij} \quad (5)$$

In this study, the ordering cost is not considered due to elimination of the effect of order batching. In this way, small batch sizes with high frequency ordering will be possible and, thus, effect of existence of large batch sizes arising from ordering costs is removed. The performance of system is measured by expected cost and service level. Service level can be shown in two ways. These are, demand service level and period service level. Demand service level (D_{SL}) gives better idea of satisfied customer. But when previous day's unsatisfied customer demand, does not affect next day's demand, then Period service level (P_{SL}) can be used to measure the performance.

Demand service level can be mathematically written as

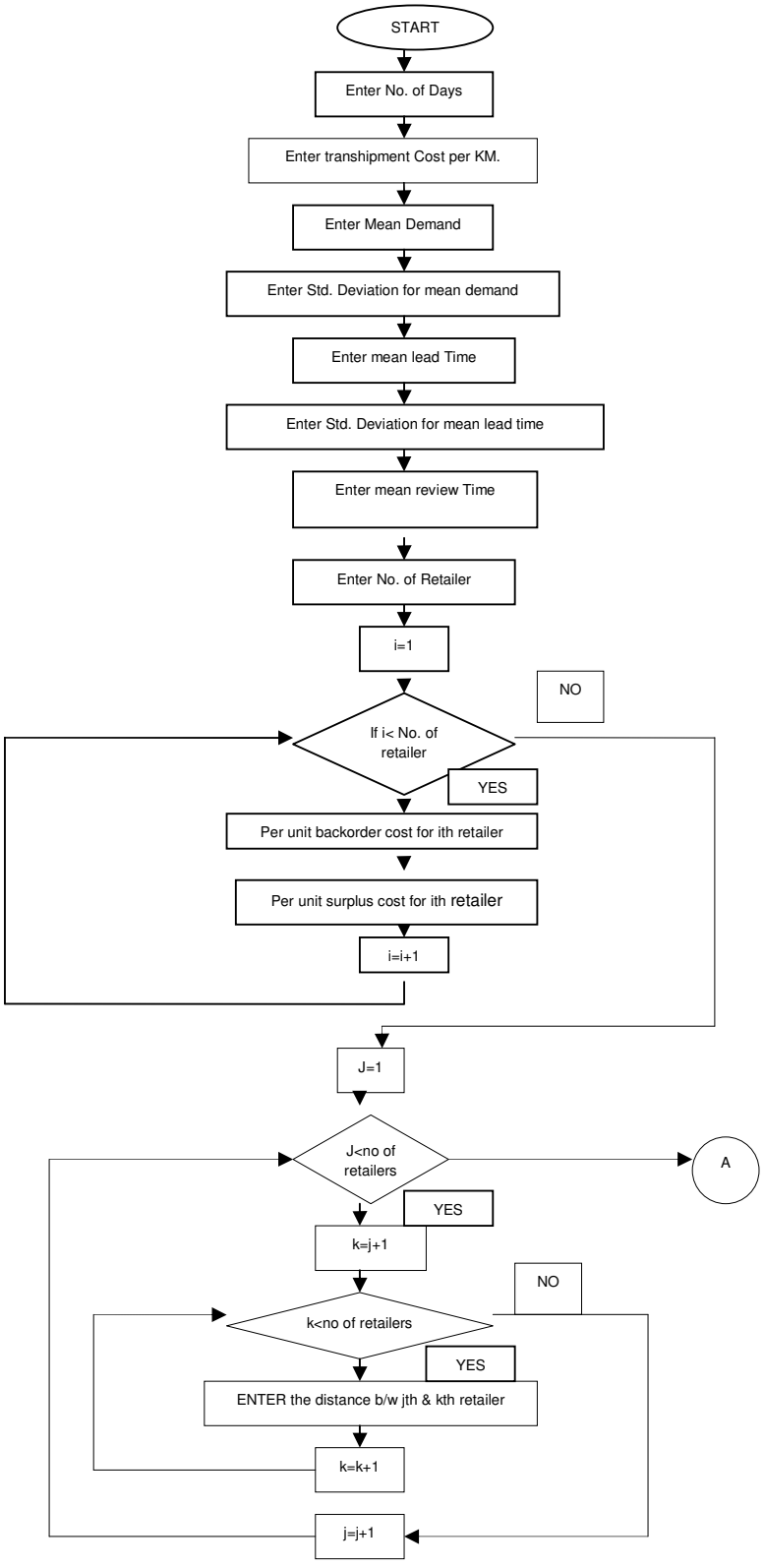
D_{SL} = 1 – Total stock out quantity/Total demand or

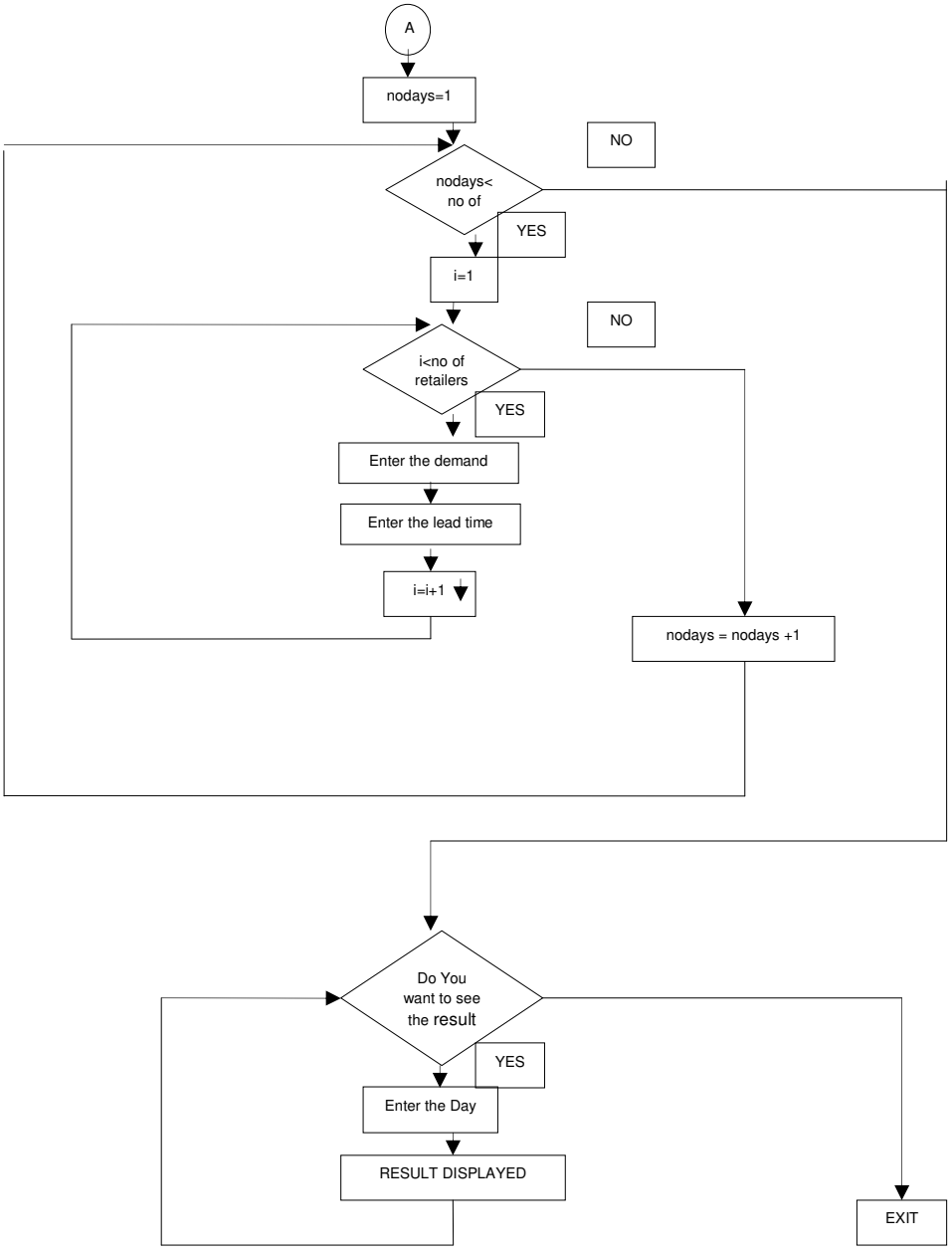
$$D_{SL} = 1 \quad \frac{\sum_{i=1}^{n} S_o}{\sum_{i=1}^{n} D_i} \quad (6)$$

Period service level can be written as follows

P_{SL} = 1 - Total no. of stock out periods / Total no. of periods or $P_{SL} = 1 - {}^{N_o}/_{N_t}$ (7)

We can use these relations to measure the service level of the system. Above mentioned formulated model is used to develop an in-house programme in C++. The formulated programme is now implemented on a case study as explained in section 3. Flow chart of the algorithm is as under.


START
Enter No. of Days
Enter transhipment Cost per KM.
Enter Mean Demand
Enter Std. Deviation for mean demand
Enter mean lead Time
Enter Std. Deviation for mean lead time
Enter mean review Time
Enter No. of Retailer
i=1
NO
If i< No. of retailer
YES
Per unit backorder cost for ith retailer
Per unit surplus cost for ith retailer
i=i+1
J=1
J<no of retailers
A
YES
k=j+1
NO
k<no of retailers
YES
ENTER the distance b/w jth & kth retailer
k=k+1
j=j+1



3. CASE STUDY

Automobile sector is one of the emerging markets in India and presently lateral transshipment is being less exercised at retailers' level in their supply chain structure. As inventory management and customers' satisfaction are the key areas in modern industrial sector, the same problems are also being faced by the supply chain managers.

We consider a company that manufactures bi-wheelers in India. The company keeps inventory of items at a distribution centre (DC) located in North India. It has a sales network of 498 outlets, and provides maintenance support to customers at 1500 workshops. Beside this company has 1400 rural outlets in towns with population 0f 25000 and below. As shown in Figure 1 the company opts supply chain strategies at different echelons but not on the lower side i.e. at retailers end. Retailers are not grouped with each other through any type of transshipment. In the pull type of inventory management system that is proposed in this work, where a demand is being generated by the retailer when having zero inventories, lateral transshipment between the retailers may be allowed. Also in the present distribution structure the retailers are apart by moderate distances, so transshipment may done by road and by the vehicle itself and it can be delivered on the same day.

This data collection has been made for the months of January, 2009 to June, 2009 by the authors. The data is collected for one brand for 900 demand periods (180 each) for 5 retailers. Collected data represents the day wise demand faced by the retailers, daily opening stock at retailers' end and following cost parameters. Holding cost per bike is Rs. 12/-per day, backorder cost is Rs. 900/- per bike and transportation cost per bike per kilometre is Rs. 0.80/-. Distance between retailers varies from 50 to 120 kilometres. Since retail outlets have variable demand and lead time, they face shortage or surplus. However with lateral transshipment both holding and shortage quantity decreases thereby reducing the total expected cost. If the retail outlets do not consider for lateral transshipment, they have to pay the holding cost for surplus inventory once the individual demand is satisfied and have to pay for shortage cost, if stock-out take place. Figure 2 represents the proposed structure in which retailers are grouped together for the sharing of information and material by adopting lateral transshipment. For each retailer constant holding cost, shortage cost and variable transshipment costs have been considered. The variability in transshipment cost is due to the differences of distances between retailers. Transshipment between retailers is made in such a way that if a retailer faces backorders he does transshipment from nearest retailer and onwards.

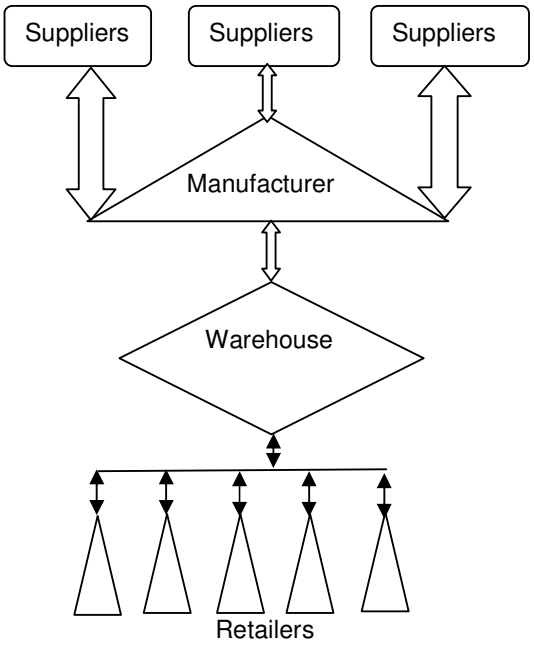


FIGURE 1: Investigated Supply Chain Structure

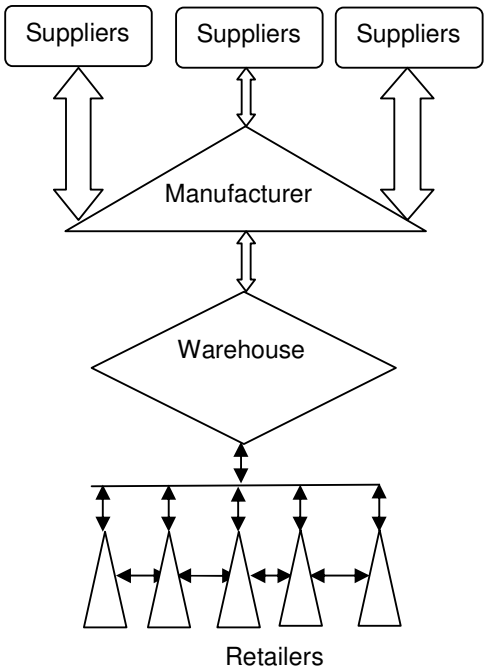


FIGURE 2: Proposed Supply Chain Structure

Ordering strategies and lead times are also taken into account. The developed programme is run on actual data of the bi-wheeler manufacturing company and the outcome has been compared with existing one.

4. RESULTS AND DISCUSSIONS

Emergency lateral transhipment technique, in considered supply chain structure is evaluated by making a comparison between the company's existing results and those which are obtained after the programme implementation. Figure 3 shows the comparison of surplus inventory at retailers' end without transshipment and with transshipment. It has been found that surplus stock without transshipment for 900 (Jan. 09 to Jun. 09) demand periods is 1545, 1312, 1374, 1624, 1488 and 1564 while with transshipment the stock is reduced to 788, 680, 808, 1152,1057 and 926 respectively.

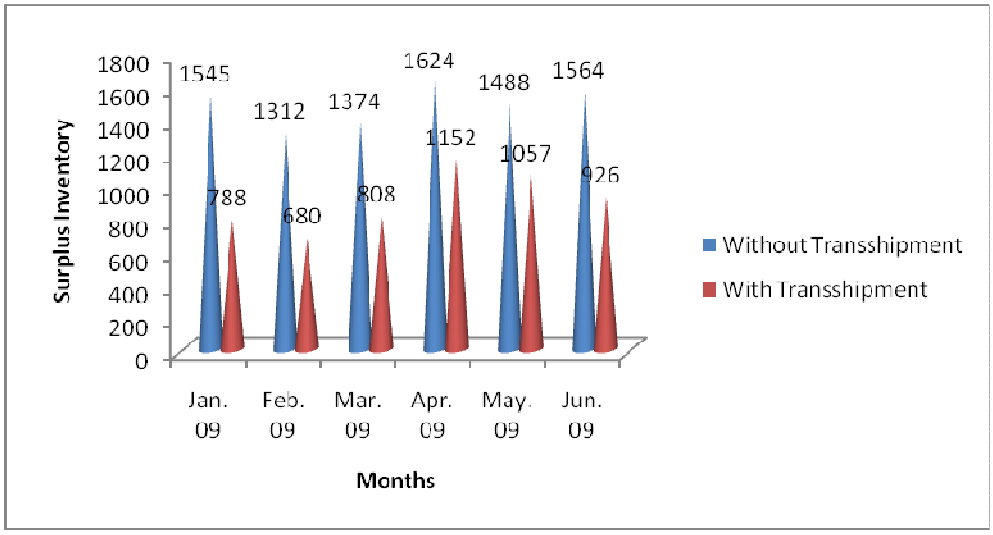


FIGURE 3: Comparison of surplus inventory with and without transshipment

Backorder quantities are also calculated in both the cases as shown in Figure 4 and it has been observed that with the use of lateral transshipment not even a single customer was backordered during the six months duration and without transshipment the numbers' of undelivered customers were 104 which reflects that by the use of lateral transshipment more customers are satisfied even with lesser inventories in hand.

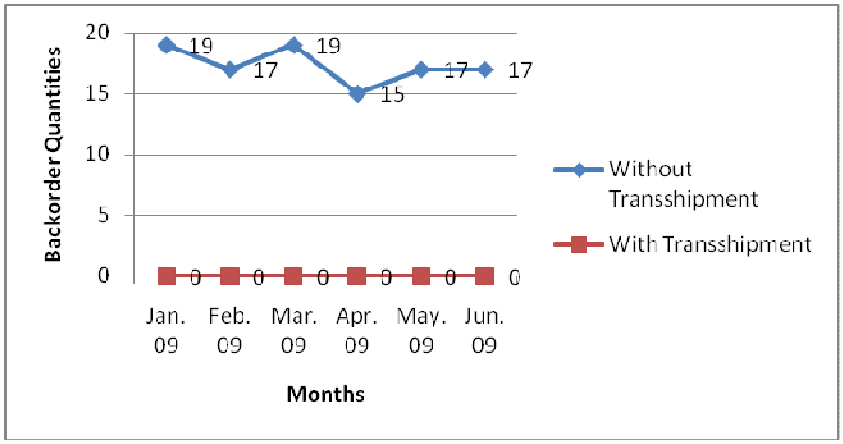


FIGURE 4: Comparison of backorder quantities with and without transshipment

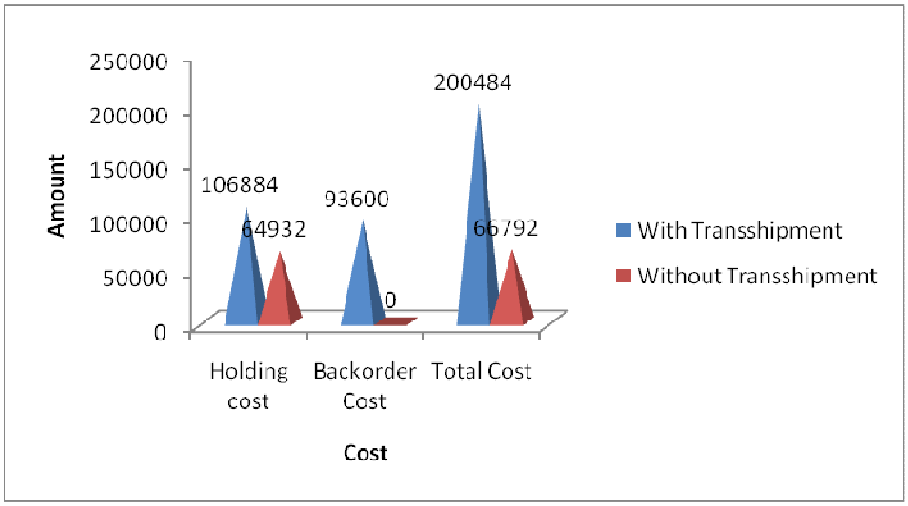


FIGURE 5: Comparison of holding, backorder and total costs with and without transshipments

As shown in Figure 5 similar effects have been obtained in regards to the holding and backorder costs. By using lateral transshipment holding cost has been reduced up to Rs. 64932/- and zero backorder cost as all the demands are fully satisfied. On the other hand these costs were found to be Rs. 106884/- and Rs. 93600/- when there is no lateral transshipment. Further total cost has been reduced to Rs. 66792/- from Rs. 200484/- even with the addition of transshipment cost of 104 bi-wheelers. Also with lateral transshipment demand service level and period service level has been enhanced from 0.9 and 0.75 to 1.0 and 1.0 respectively. Based on the results, this work has examined, from a cost-parametric perspective, the relative effectiveness of a transshipment approach in supply chain network, characterized by a single supply source at the higher echelon and multiple retail locations at the lower. Finally it has been observed that comparison indicators like backorder quantity, surplus quantity and total associated cost have decreased while using emergency lateral transshipment policy.

5. CONCLUSIONS

In this work, a multi-location inventory system is considered where transshipments are allowed as recourse actions in order to reduce the costs of surplus or shortage inventory after demands are realized. Based on the results, this work has examined, from a cost-parametric perspective, the relative effectiveness of a transshipment approach in supply chain network, characterized by a single supply source at the higher echelon and multiple retail locations at the lower. Our transportation network formulation has enabled us to gain analytical insight into problems of higher dimensions than has been achieved earlier.Furthermore, a number of simplifying assumptions such as zero lateral shipment lead-times, infinite supplier inventories, etc. are made here. In this paper a model has been formulated for one central warehouse serving to ‘n’ retailers. Emergency lateral transshipment technique is used for controlling inventories and associated costs for all the retailers and finally it has been found that surplus quantities and stock-out quantities are less in case of lateral transshipment, so holding cost and back order cost are decreased. For comparative purpose, a case study is solved by using their methodology and by

implementing the collected data in the programme developed. Our results indicate that, from a managerial standpoint, the notion of lateral transshipments appears to have substantial appeal. If the benefits of avoiding retail level shortages outweigh the additional delivery costs resulting from transshipments, customer service may be enhanced significantly, without the burden of additional safety stocks. Finally, we suggest that the issue of emergency shipments from one or more other supply sources be examined in future work in this area.

Acknowledgement
The authors would like to thank the officials of the automobile industry for their cooperation in data collection.

References

1. C.D.G. Waters. “Inventory Control and Management”. Wiley, Chichester, 2002
2. J.P. Womackand and D.T. Jones. “Lean Thinking: Banish Waste and Create Wealth in Your Corporation”. Simon & Schuster, London, 1996
3. M. Christopher and D. Towill. “An integrated model for the design of agile supply chains”. International Journal of Physical Distribution & Logistics Management, 31(4): 235-46, 2001
4. Biju Kr. Thapalia, Stein W. Wallace and Michal Kaut. “Using inventory to handle risks in the supply of oil to Nepal”. Int. J. Business performance and Supply Chain Modelling, 1(1): 41-60, 2009
5. F.T.S. Chan and H.K. Chan. “Effects of cascade information sharing in inventory and service level in multi echelon supply chain”. Int. J. Business performance and Supply Chain Modelling, 1(1): 1-7, 2009
6. M. Christopher. “Logistics and Supply Chain Management”. Prentice-Hall, Harlow, 2005
7. H.L. Lee and C. Billington. “Managing supply chain inventory: pitfalls and opportunities”. Sloan Management Review, Spring: 65-73, 1992
8. Kishore K. Pochampally, Surendra M. Gupta and Kannan Govindan. “Metrics for performance measurement of a reverse/closed – loop supply chain”. Int. J. Business performance and Supply Chain Modelling, 1(1): 8-32, 2009
9. Simchi-Levi, D. Kaminsky. P and Simchi-Levi. E. “Designing and Managing the Supply Chain: Concepts, Strategies, and Cases”. 2nd ed., McGraw-Hill, New York, NY, 2003
10. A.J. Clark and H. Scarf. “Optimal policy for a dynamic multi-echelon inventory model”. Management Science: 475–490, 1960
11. K. Krishnan and V. Rao. “Inventory control in N warehouses”. Journalof Industrial Engineering16: 212–215, 1965
12. D. Gross. “Centralized inventory control in multilocation supply systems”. In: Scarf, H. Gilford. D. and Shelly. M. (Eds.). “Multistage Inventory Models and Techniques”. Stanford University Press, Stan- ford, CA (Chapter 3), 1963
13. C. Das. “Supply and redistribution rules for two-location inventory systems: One-period analysis”. Management Science 21: 765–776, 1975
14. L.W. Robinson. “Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments”. Operations Research 38: 278–295, 1990

15. A.W. Archibald, S.A. Sassen and L.C. Thomas. “An optimal policy for a two depot inventory problem with stock transfer”. Management Science 43: 173–183, 1997
16. Y.T. Herer and A. Rashit. “Lateral stock transshipments in a two- location inventory system with fixed and joint replenishment costs”. Naval Research Logistics 46: 525–547, 1999
17. S. Minner and E.A. Silver. “Evaluation of two simple extreme transshipment strategies”. International Journal of Production Economics: 93–94, 1–11, 2005

Inclusion of the Memory Function in Describing the Flow of Shear-Thinning Fluids in Porous Media

M. Enamul Hossain menamul@kfupm.edu.sa
Department of Petroleum Engineering
King Fahd University of Petroleum & Minerals
Dhahran 31261, Saudi Arabia

L. Liu lei.liu@dal.ca
Department of Civil and Resource Engineering
Dalhousie University
Sexton Campus, 1360 Barrington Street, Halifax,
NS B3J 2X4, Canada

M. Rafiqul Islam Rafiq.islam@dal.ca
Department of Civil and Resource Engineering
Dalhousie University
Sexton Campus, 1360 Barrington Street,
Halifax, NS B3J 2X4, Canada

Abstract

This paper introduces mathematical models with memory to present the complex rheological phenomena combining the in situ shear rate and bulk rheology with fluid memory. The models are solved numerically and the results compared with established experimental data available in the literature. The new models showed excellent agreement with experimental results for the region where most existing models failed. The proposed models can be used in a number of applications, such as enhanced oil recovery (EOR), polymer manufacturing, etc.

Keywords: Non-Newtonian fluid; shear rate; complex fluid rheology; non-Darcy flow; non-linear dynamics; chaos in porous media.

1. INTRODUCTION

The flow of aqueous polymeric fluids through porous media has been studied in the past mainly due to its importance in the areas of EOR. Past investigations focus on the flow of inelastic shear-thinning fluids and more complex viscoelastic polymers. Even when the bulk rheology of the aqueous polymer solution is known, there is still an issue of how this relates to the in situ rheology of the fluid. Clearly, the internal geometry of the pore space in either a packed bed or a porous rock is far more complex than any regular rheumatic flow. In porous media, the apparent viscosity as an aggregate or "upscaled" measure of viscous, elastic and extensional flow effects, and the matter of how this should be done have been of concern. For polymeric solutions, the apparent viscosity is a function of flow rate through the porous medium, and the flow rate is further correlated with the in situ shear rate within the porous medium. This flow rate may also be interrelated with the fluid memory in the pore network where mineral precipitation or other history of movement may delay the response of the in situ fluid. This delay can lead to restriction of the

polymer flow through the pore throat. This paper attempts to eliminate these unsolved complexities of the fluid flow in porous media.
The Newtonian viscosity model is the most commonly used one for the study of fluid flow in various applications. In principle, the Newtonian model implies that the viscosity is independent of shear rate. Water, the most abundant fluid on earth, has been considered for centuries as an ideal example of a Newtonian fluid. Only recently, Li et al. [1] discovered that when water molecules are forced to move through a small gap (in nanoscale) of two solid surfaces (hydrophilic/wetting), its viscosity increases by a factor of 1000 to 10,000, resulting in a behavior similar to molasses. During their experiment on hydrophobic surfaces, they did not observe such an increase in viscosity. Their findings are in agreement with the molecular dynamics simulation results that show a dramatically decreased mobility for sub-nanometer thick water films under hydrophilic confinement. They concluded that water has viscous and solid-like properties at its molecular level and is organized into layers. At the nanometer scale, water is a viscous fluid and could be a much better lubricant.

This study received great attention from the general scientific community as well as the general public because of its potential applications to nano technology [2]. However, the fact that any fluid behaves differently under molecular constrains from larger scale has been known for some time. The slow-moving flow of a thin film of a liquid is a ubiquitous phenomenon. This flow exists in nature such as in lava flows, the linings of mammalian lungs, tear films in the eye, and in artificial instances such as microchip fabrication, tertiary oil recovery as well as in many coating processes [3]. Therefore, the natural phenomena, for which viscous fluid flow exists, are normally non-Newtonian type of flow [3-4]. If the most commonly accepted Newtonian fluid, water can exhibit viscous flow, it's a matter of research to say that all fluids are prone to behave like a non-Newtonian fluid. As time progresses and our ability to measure with greater accuracy improves, the true nature of fluids is likely to be revealed as non-Newtonian. Even though numerous models for simulating non-Newtonian fluid behavior exist, there is a need to develop a theoretical model that is valid for an entire range of stress-strain relationship.

If viscosity diminishes as shear is increased, the fluid is said to exhibit shear thinning. This occurs in simple liquids as well as in complex mixtures such as foams, micelles, slurries, pastes, gels, polymer solutions, and granular flows [5]. The flow and displacement of non-Newtonian and complex fluids (such as polymer gels and surfactants) in porous media is an important subject. These have a variety of industrial applications. It is well known that polymeric materials, emulsions, foams, and gels exhibit non-Newtonian behavior, even at meso scale. For various industrial applications, ranging from paints, manufacturing and food processing to petroleum production, numerous non-Newtonian models have been proposed.

Kondic et al. [6] studied the non-Newtonian fluid in a Hele-Shaw cell. They have used Darcy's law whose viscosity depends upon the squared pressure gradient to a fluid model with shear-rate dependent viscosity. However, their derivation does not allow strong shear-thinning which is related to the appearance of slip layers in the flow. The influence of shear effects on the adhesive performance of a non-Newtonian fluid under tension has been studied by Miranda [7]. He used the modified Darcy's law in developing a shear-rate model. His results show that, for a relatively small separation, the adhesion strength is considerably reduced if the fluid is shear thinning (thickening). Shear effects become negligible when large plate separation occurs. Afanasiev et al. [8] studied the drag-out problem for shear-thinning liquids at variable inclination angles. They used the power-law, the Ellis and Carreau model as their viscosity model. They considered steady state solution and the dependence on the rheological parameters. However, many applications of polymeric liquids and suspension show nonlinear stress-strain relationships. A distinct viscosity regimes show up during shear stress when most polymer solutes are used under shear-thinning conditions. The type of fluid (e.g. Newtonian or non-Newtonian) can be categorized based on the shear rate. It shows that polymeric fluid behaves as Newtonian at very low shear rates. As the shear rate increases the behavior starts to become nonlinear. When shear rate is further increased it moves into a regime where the viscosity can be modeled by a power-law relation. It is noted that the power-law model is only applicable at large shear rates,

Ellis model is for low shear rate [8]. Therefore, their observation further consolidates the need to develop a model that allows for very high shear rates.
Non-Newtonian fluids exhibit another particular phenomenon not observed in Newtonian fluids such as the negative wake behind a bubble are due to memory effects in stress relaxation [9]. Under shearing, a polymer solution may exhibit memory effects during a sequence of stretching of polymeric chains. Huang and Lin I [10] investigated the temporal memory and persistence time correlations of microstructural order fluctuations in quasi-two-dimensional dusty plasma liquids through direct monitoring of particle positions. They noticed that persistence length of the ordered and the disordered microstates both follow power-law distribution for the cold liquid. They also identified that increasing thermal noise level deteriorates the memory and leads to the less correlated stretched exponential distribution of the persistent length. Even though the existence has long been recognized (e.g., thixotropic fluid) [11-13], few models have been proposed to include the memory effect. Even fewer models have been developed that is continuous in nature, mostly settling for good agreement within a narrow range of constraints.

In the case of porous media flow, the problem is accentuated in the consideration of permeability variation with time. The fluid memory can be defined as the change of viscosity and other pressure dependent properties over time whereas the memory of rock can be characterized as the alteration of permeability with time. Therefore, if permeability diminishes with time, the effect of fluid pressure at the boundary on the flow of fluid through the medium is delayed. The information of this effect over time in the fluid flow will continue if the medium (i.e. flowing fluid) holds a memory [14-17]. However, most of the current flow models in porous media have been developed for purely Newtonian and Non-Newtonian fluids, no model being able to represent the fluid memory, applicable for the entire range of shear-thinning [18-29]. Most of the researchers have tried to relate the bulk properties of complex fluids to their behavior in a porous medium using a common approach, consisting of representing the medium by a bundle of parallel capillary tubes [27].

The majority of complex fluids used in oil field applications are non-Newtonian polymeric solutions demonstrating shear-thinning (pseudoplastic) behavior in a solution. The bulk macroscopic properties of these solutions, mainly their viscosity/shear rate dependency, are well understood and characterized using established models. Existing theoretical models as well as experimental findings are well established in the literature [18-29]. It is rather the numerical modeling that is incomplete [30]. Only recently, Hossain et al. attempted to model shear rate and viscosity of polymeric complex fluid as a function of time and other related bulk properties of fluid and media itself where memory has been incorporated to represent macroscopic and microscopic behavior of fluid and media in a more realistic way [16].

2. EXISTING MODELS

The two main types of polymers used in the oil industry for hydrocarbon recovery are synthetic polyacrylamide (in its partially hydrolyzed form, HPAM) and Xanthan biopolymer gum. Bulk properties measurements of polymeric solutions are a standard and reliable experimental procedure. Therefore, researcher's efforts have been made to extend the laws of motion for purely Newtonian fluids (Darcy's law) to rheologically complex ones using easily measurable properties such as the shear rate/viscosity behavior. A bundle of parallel capillary tubes approach has been used to measure the macroscopic and microscopic properties of porous media. This approach leads to the definition of an average radius which is dependent on macroscopic properties of the medium such as porosity, absolute permeability, and some measure of tortuosity. The available mathematical models (such as power law, Carreau, or Cross models) to describe the fluid rheology have been developed to define viscosity and apparent shear rate from the use of the Darcy velocity [19, 20, 23, 24, 31-35]. Experimental results show that the shape of the apparent viscosity curve is similar to that of bulk shear rate. Most of experimental works have been performed by Xanthan biopolymers, for which experimental results are available in the literature [19, 20, 24, 31, 33, 34, 35-39]. These investigations attempted to find the shape

factor, α_{SF}. For porous media, Chauveteru's form of the definition of porous media wall shear strain rate or in-situ shear rate is given by Eq. (1), listed below [18-23, 25, 26-29, 35, 37].

$$\dot{\gamma}_{pm} = \frac{\alpha_{SF}\, u}{\sqrt{k\,\varphi}} \tag{1}$$

In the above equation, α_{SF} represents the shape factor. In the context of polymer flooding (part of enhanced oil recovery schemes), in-situ rheology depends on polymer type and concentration, residual oil saturation, core material and other related properties [28, 33, 37, 39, 40]. A brief discussion of this topic has been presented by Lopez [35]. The existence of a slip phenomenon in the Newtonian region at ultra low flow rates is confirmed and the degree of shift (the α_{SF} factor) in the non-Newtonian region is quantified. It is shown that the adoption of rigorous and reproducible coreflood procedures is required to yield unambiguous data on in-situ polymer viscosity and polymer retention in real systems. Coombe et al. [41] analyzed the impact of the non-Newtonian flow characteristics of foam, polymers and emulsions. Balhoff and Thompson [42] modeled the flow of shear-thinning fluids, including power-law and Ellis fluids in packed bed using the network model. They pointed out that Eq. (1) has a generic form which depends on polymer type, medium structure and approach.

A number of constitutive equations have been developed in the past that capture the full bulk rheological behavior of pseudoplastic solutions [37]. To model the bulk rheology of the non-Newtonian fluid, Escudier et al. [24] performed a set of experiments on Xanthan gum. Lopez [27, 35] showed that Carreau-Yasuda, Cross and Truncated Power-law models behave almost same when they presented the bulk rheology (e.g. effective viscosity) of shear thinning fluid. Therefore, Carreau-Yasuda model is considered in this study and is used to develop and analyze the memory model in bulk rheology. Carreau-Yasuda model may be written as [20, 21, 24, 27, 35]:

$$\mu_{eff} = \mu_{\infty} + \frac{(\mu_0 - \mu_{\infty})}{\left[1 + (\lambda\dot{\gamma})^{a}\right]^{\frac{n}{a}}} \tag{2}$$

3. MATHEMATICAL MODEL WITH FLUID MEMORY

The exact form of the shear stress-shear rate (stress-stain) relationship depends on the nature of the polymeric solution. Therefore, recently, a question is always coming out about the effect of memory on rock/fluids in porous media when predicting oil flow outcomes. Hossain and Islam [30] have reviewed the existing complex fluid flow models with memory available in literature. None of them has focused the shear thinning fluid models which may couple with fluid memory. Recently Hossain et al. [16-17] have developed a model which represents a more realistic rheological behavior of fluid and media. They have developed a stress-strain relation coupling the macroscopic and microscopic properties with memory. They also did not consider the polymeric fluid properties in porous media. However, the conventional practice is to consider the Newtonian fluid flow equations as ideal models for making predictions. Even non-Newtonian models focus on what is immediately present and tangible in regard to fluid properties. This paper argues that the time function and time-dependent fluid and media properties can be coupled to capture the more complex behavior of shear thinning fluids in porous media.

In terms of practical point of view, several authors [29, 43, 44] have reported that the apparent viscosity of polymer solutions within various porous media are usually bead packs, sand packs and outcrop sandstone rocks. A detailed overview of the earlier works have been presented by the researchers [18, 19, 20, 26, 27] and summarized in the earlier section. Referring to them, the apparent viscosity for the polymeric solutions can be defined from Darcy's law as follows:

$$\mu_{app} = \frac{k\,A}{Q}\frac{\Delta p}{L} \tag{3}$$

For polymeric solutions, the apparent viscosity μ_{app} is a function of flow rate through the porous medium [29, 40], and the flow rate is further correlated with $\dot{\gamma}_{pm}$, the in situ shear rate within the porous medium, which is expressed as Eq. (1). In this simple equation, however, it is not obvious why $\dot{\gamma}_{pm}$ should vary linearly with the Darcy velocity, particularly when complex internal flows persist within the porous medium. The central theoretical problem of in situ rheology is, therefore, to establish clearly how the local (microscopic or pore-scale) rheology in a single pore relates to the aggregated or average bulk property as expressed by the apparent viscosity discussed above. Hence it is the effect of this local behavior, mediated through the interconnected network of pores in the porous medium to the macroscopic scale that must be found. Therefore, it is necessary to clarify how the local (microscopic or pore-scale) rheology relates to the average bulk rheology as expressed by the apparent viscosity. Some researchers [20, 27, 45] considered the local bulk rheology (e.g. Carreau model) for shear-thinning fluids and calculate the "upscaled" macroscopic apparent viscosity using a connected 2D network of cylindrical capillaries as an idealized porous media. Using these concept, Perrin et al. [29] showed that the average in situ shear rate in the network (which corresponds to $\dot{\gamma}_{pm}$) actually varied linearly with the flow velocity, u. We know that the complexity of the internal flows within the porous media leads to a non-linear behavior of shear rate with Darcy velocity [16, 17, 30].

To investigate the local phenomenology, we may introduce memory effects for the fluid complex rheological properties. In this regard, fluid precipitation of minerals and temperature may be considered. However, some fluids carry solid particles that may impede flow in some of the pores. The precipitation and obstruction may reduce the pore size and thus decrease the permeability in time. Some fluid may react chemically with the solid medium, altering pore size. The existence of these phenomena let the researcher consider local mineralization and permeability changes, which lead for a spatially variable pattern. Therefore, if permeability diminishes with time, the effect of fluid pressure at the boundary on the flow of fluid through the medium is delayed. The information of this effect over time in the fluid flow can be captured by introducing a memory function for the medium that holds the fluid.

Auriault and Boutin [46] pointed out that the macroscopic description is sensitive to the ratios between the different scales of the characteristic lengths of the pores, the fractures, and the macroscopic medium, respectively. They described the memory effect can exhibit in case of the ratio the characteristic lengths of the pores and the macroscopic medium. This memory is due to the seepage through the micropores. Under transient excitations, they concluded that the harmonic quasi-static excitations, with complex and frequency dependent effective coefficients in porous medium display memory effects. They referred to this mechanism as the 'viscoelastic property' of fluid saturated, cracked solids. Savins [43, 44] also acknowledged the effects of memory as a 'viscoelastic effects'. He mentioned the previous researchers who worked on the phenomenological theories of incompressible memory fluids.

Hossain et al. [16] initiated a rate equation model to study the flow of these fluids with memory. Caputo [14-15] modified Darcy's law by introducing the memory represented by a derivative of fractional order of differentiation which simulates the effect of a decrease of the permeability in time. If the fluid flows in x-direction, the mass flow rate equation may be written as (14-17, 30];

$$q_x = -\eta\,\rho_o\left[\frac{\partial^{\alpha}}{\partial t^{\alpha}}\left(\frac{\partial p}{\partial x}\right)\right] \tag{4}$$

where

$$\frac{\partial^{\alpha}}{\partial t^{\alpha}}[p(x,t)] = \frac{1}{\Gamma(1-\alpha)}\int_0^t (t-\xi)^{-\alpha}\,\frac{\partial}{\partial \xi}[p(x,t)]d\xi \quad , \quad \text{with}$$

$0 \leq \alpha < 1$

It is clear that the memory introduced in Eq. (4) to describe the flow of the fluid implies the use of two parameters, namely α and η . These two parameters are used instead of the permeability and viscosity in conventional Darcy's law. In Eq. (4), pressure gradient is negative and this has a decreasing slop [16-17]. Therefore, Eq. (4) can be written for fluid velocity which is related to pressure gradient as;

$$u_x = \eta\left[\frac{\partial^{\alpha}}{\partial t^{\alpha}}\left(\frac{\partial p}{\partial x}\right)\right] \tag{5}$$

where

$$\frac{\partial^{\alpha}}{\partial t^{\alpha}}[p(x,t)] = \frac{1}{\Gamma(1-\alpha)}\int_0^t (t-\xi)^{-\alpha}\,\frac{\partial}{\partial \xi}[p(x,t)]d\xi \quad , \text{ with } 0 \leq \alpha < 1$$

Eq. (5) may be written as,

$$u = \frac{\eta}{\Gamma(1-\alpha)}\int_0^t (t-\xi)^{-\alpha}\,\frac{\partial^2 p}{\partial \xi \partial x}d\xi \tag{6}$$

Substituting Eq. (6) in Eq. (1)

$$\dot{\gamma}_{pm} = \frac{\alpha_{SF}}{\sqrt{k\varphi}}\,\frac{\eta}{\Gamma(1-\alpha)}\int_0^t (t-\xi)^{-\alpha}\,\frac{\partial^2 p}{\partial \xi \partial x}d\xi \tag{7}$$

The above mathematical model provides the effects of the polymer fluid and formation properties in one dimensional fluid flow with memory and this model may be extended to a more general case of 3-Dimensional flow for a heterogeneous and anisotropic formation. It should be mentioned here that the first part of the Eq. (7) is the apparent core properties; second part is the effects of fluid memory with time and the pressure gradient. The second part is in a form of convolution integral that shows the effect of the fluid memory during the flow process. This integral has two variable functions of $(t-\xi)^{-\alpha}$ and $\partial^2 p/\partial \xi \partial x$ where the first one is a continuously changing function and second one is a fixed function. This means that $(t-\xi)^{-\alpha}$ is an overlapping function on the other function, $\partial^2 p/\partial \xi \partial x$ in the mathematical point of view. These two functions depend on the space, time, pressure, and a dummy variable. In the above, the parameter η captures the memory effect as a continuous function of time, whereas α defines the nature of the memory function. Even when α is set to zero, the memory effect is not eliminated. However, when namely η is set to zero, the governing equation reduces to a Newtonian fluid model.

Several workers [36, 47] used similar expressions to represent the apparent shear rate as an effective shear. Moreover, Savins [43, 44] refered to this problem as “scale up”. Sorbie et al. [47] reviewed some of the alternative approaches to this problem. In this paper, we ignore this shifting or scale up problems due to a constant factor involvement. This will not affect the big picture of the effective or apparent viscosity model. Therefore, to analyze the memory effect in the shear-thinning fluid viscosity, the model presented in Eq. (7) is applied in Eq. (2). Substituting Eq. (7) in Eq. (2), one obtains the following expression.

$$\mu_{eff} = \mu_{\infty} + \frac{\mu_0 - \mu_{\infty}}{\left[1 + \left(\frac{\lambda \eta \alpha_{SF}}{\sqrt{k\varphi}} \frac{\int_0^t (t-\xi)^{-\alpha} \frac{\partial^2 p}{\partial \xi \partial x} d\xi}{\Gamma(1-\alpha)}\right)^a\right]^{\frac{n}{a}}} \quad (8)$$

4. NUMERICAL SIMULATION

To solve the convolution integral of Eqs. (7) and (8), a reservoir model of length, L = 5000 m, width, W = 100 m, and height, H = 50 m is considered in this study. The porosity and permeability of the reservoir are 30% and $\mathbf{1 \times 10^{-14}}$ **to** $\mathbf{13.5 \times 10^{-14}}$ m^2 (10 md to 135 md), respectively. The reservoir is completely sealed and produces at a constant rate where the initial pressure is $p_i = 27579028\ pa$ (4000 psia). The fluid is assumed to be Xanthan gum, with the properties $c = 1.2473 \times 10^{-9}\ {}^{1}/_{pa}$, $\mu_0 = 13.2\ pa\ s$. The initial production rate is $q_i = 8.4 \times 10^{-9}\ {}^{m^3}/_{s}$ and the initial fluid velocity in the formation is $u_i = 1.2 \times 10^{-5}\ {}^{m}/_{s}$. The fractional order of differentiation, $\alpha = 0.2 - 0.8$, $\Delta x = 1000\ m$, and $\Delta t = 7.2 \times 10^{4}\ s$ have also been considered. The computations are carried out for $\mathbf{Time} = 40\ months$. The same input data has been used by Hossain et al. [16-17, 30] except the fluid properties. In solving this convolution integral with memory, trapezoidal method is used. All computation is carried out by using Matlab 6.5.

5. RESULTS AND DISCUSSION

The results of the in situ or apparent shear rate and effective viscosity are obtained by solving the governing Equations (7) and (8). These models are solved numerically. In this paper, we focused on the dependence of the shear rate on porosity, permeability, shape factor and flow velocity which is related to the effect of fluid memory and finding a numerical description for a sample reservoir.

To solve the Eq. (7), $\eta = 0.343249$ and $\alpha_{SF} = 1.25$ are used in numerical computation [16, 17, 20, 22, 27, 36, 45, 47]. In solving the proposed modified Carreau-Yasuda model presented in Eq. (8), the Xanthan gum fluid viscosities at low (μ_0) and high shear rate (μ_∞) are considered as 13.2 pa-s and 0.00212 pa-s respectively [24]. The power-law index, n and the time constant, λ are taken as 0.689 and 60.7 s, respectively [24]. Here, the diminishing permeability in the pore space and throat of the reservoir formation are varied within the range as mentioned before to considered fluid memory of the flowing shear thinning fluid.

5.1 Shear rate dependency on different parameters

5.1.1 Shear rate dependency on permeability for different α values

Figure 1 presents the variation of in situ or apparent shear rate with permeability for different α values. These are the nonlinear profiles for different α values. These curves show that as permeability increases, the in-situ shear rate increases slightly for a very tight reservoir. This trend doesn't persist when α value is increased. This indicates that when memory effect is dominant in a tight reservoir, it tries to restrict the fluid movement. However, as permeability increases, shear rate start to decrease up to a certain range of permeability values. The range is 45 mD – 55 mD. The fixed value depends on α values. After this transient permeability level, shear rate increases with the increase of permeability value. The overall trends of the curves are similar for all the α values. The memory of fluid has a great influence on in situ shear rate. For the same permeability value, the in situ shear rate decreases with the increase of α value at a range of up to 0.4 and after that it increases with the increase of α value. This behavior is a special characteristic of the fluid memory [16-17]. The magnitudes of the in situ shear rate variation are more dominant when the intensity of the fluid memory (α value) increases. Moreover, we already focused earlier that the variable permeability is one of the causes of fluid memory. The initial decrease and subsequent increase of shear rate variation with permeability indicates that fluid takes some time to move after pressure or force is exerted. This delay of fluid movement is nothing but a microscopic property of the viscoelastic fluid. This phenomenon is defined by fluid memory [14-17, 30]. Therefore the fluid movement can be characterized by the memory effects.

5.1.2 Shear rate dependency on shape factor for different permeability

The available literature [20, 22, 24, 27, 36] reported that the value of α_{SF} lies in the range of 1.0 – 14.1. This variation depends on the type of porous media. This may vary from 1.0 in regular unconsolidated packs up to 10.0 in consolidated sandstone [22]. The experimental results show that α_{SF} is in the range of 1.1 – 2.5 for ballotini beads and 1.9 – 9.1 for sandstone cores [20]. Therefore, to analyze the dependency of in situ shear rate with α_{SF}, it has been considered the value of α_{SF} in the range of 1.0 – 15.0.

Figure 2 shows the variation of apparent shear rate with shape factor, α_{SF} for $k = 70\ md, 80\ md$ **and** $\mathbf{90}\ md$ at a α value of 0.2. The in situ shear rate verses shape factor curve shows a linear relationship that has an increasing trend with an increase in the value of the shape factor. The effects of shape factor on apparent shear rate is more sensitive at higher permeability, for which the range of variation of $\dot{\gamma}_{apm}$ is wider. On the other hand, it is less sensitive at lower permeability.

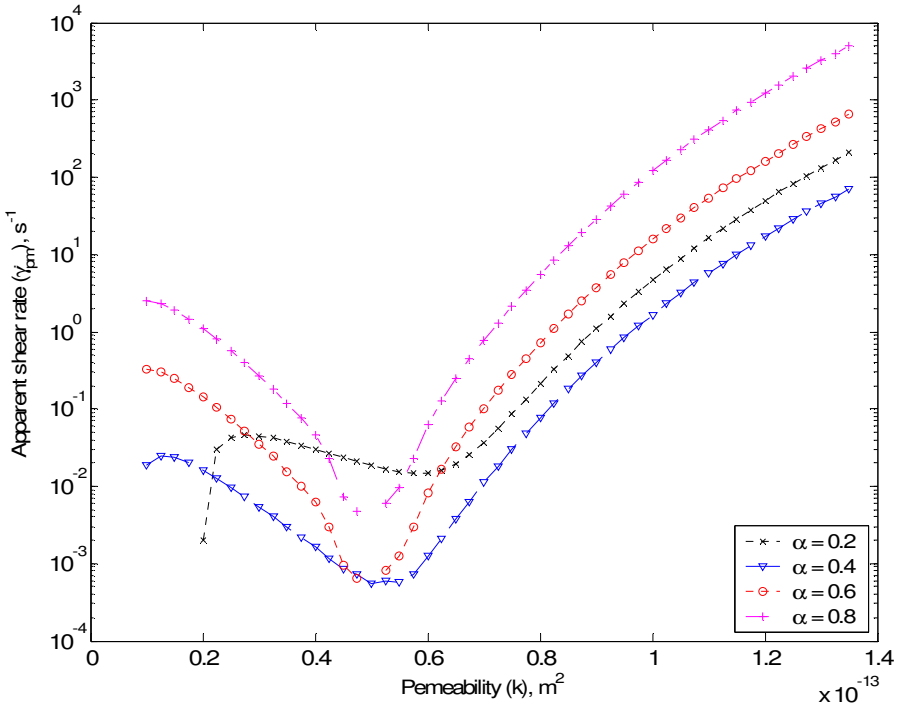


FIGURE 1: Variation of apparent shear rate with permeability of the porous media for $\alpha = 0.2$ to 0.8

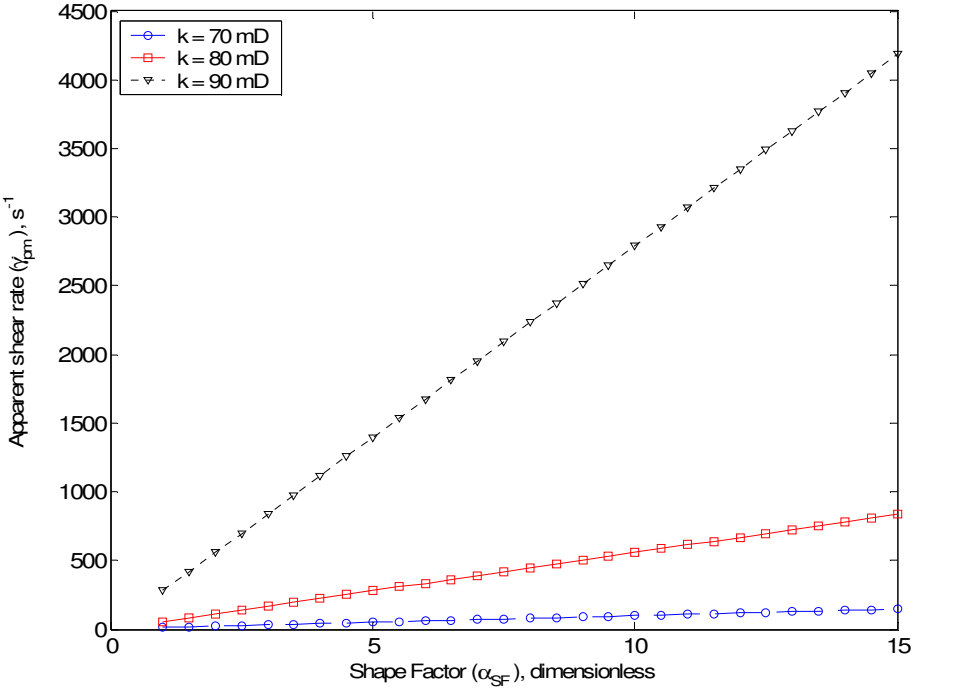


FIGURE 2: Variation of apparent shear rate with shape factor for $k = 20\ mD$ and $60\ mD$

5.1.3 Shear rate dependency on shape factor for different porosity

Figure 3 shows in a semi log plotting the variation of apparent shear rate with shape factor, α_{SF} for different porosity of the formation at a α value of 0.2. The present plotting trend of the curve shows a non-linear trend. However, this is a complete linear plotting in a Cartesian coordinate. For a tight reservoir (low porosity), shear rate is much higher compared to medium or highly porous media for the same shape factor. Shear rate increases with the increase of shape factor for all porosity values.

5.1.4 Shear rate dependency on flow velocity for different α values

To investigate the effects of flow rate on shear rate with memory, porosity and permeability are considered as 30% and 70 mD respectively. Flow velocity, *u* is varied in the range of 1.0×10^{-5} to 13.5×10^{-5} m/s.

Figures 4-7 present the variation of in situ shear rate with flow velocity for different α values. Shear rate decreases with the increase of flow rate at low velocity range. There is slightly faster decrease in the range of velocity increment of 1.5×10^{-5} to 2.0×10^{-5} m/s. After this velocity range, shear rate increases faster with the increase of flow velocity which is an asymptotic variation. This trend continues until the velocity of 5.5×10^{-5} m/s and again it starts to reduce after this velocity. The relationship between in situ shear rate and flow rate is nonlinear trend whereas Perrin et al. [29] showed that the average in situ shear rate varied linearly with the flow velocity. This non-linearity is only due to the memory effects on fluid flow behavior. Moreover, the shape and trend of the curves are similar for all the α values except the magnitude of the shear rate values which also leads to the existence of fluid memory on shear-thinning fluid. Here, shear rate increases with the increase of α values.

Figure 8 presents the conventional in situ shear rate model (Eq. (1)) available in the literature. The nonlinear trend of the $\dot{\gamma}_{pm}$ curves (Figs. 4-7) are due to the memory dependency of flowing fluid which did not captured by the Perrin et al. [29] and other researchers. So, it can be concluded that fluid memory has an influence on the shear thinning fluid and microscopic rheological behavior of the fluid may be characterized by this memory effect.

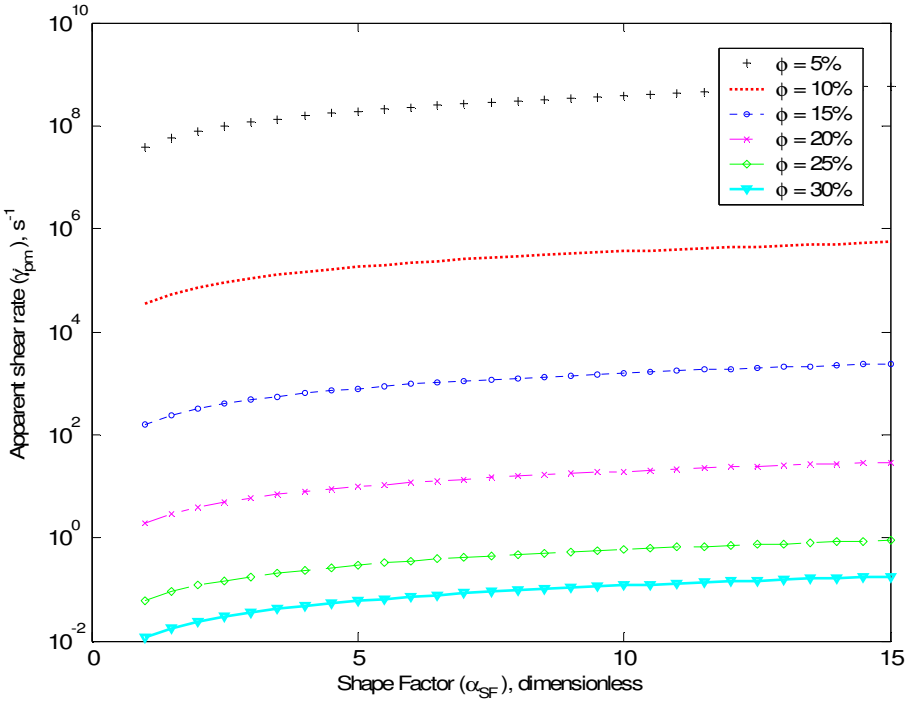


FIGURE 3: Variation of apparent shear rate with shape factor for $\varphi = 20\%$ and 40%

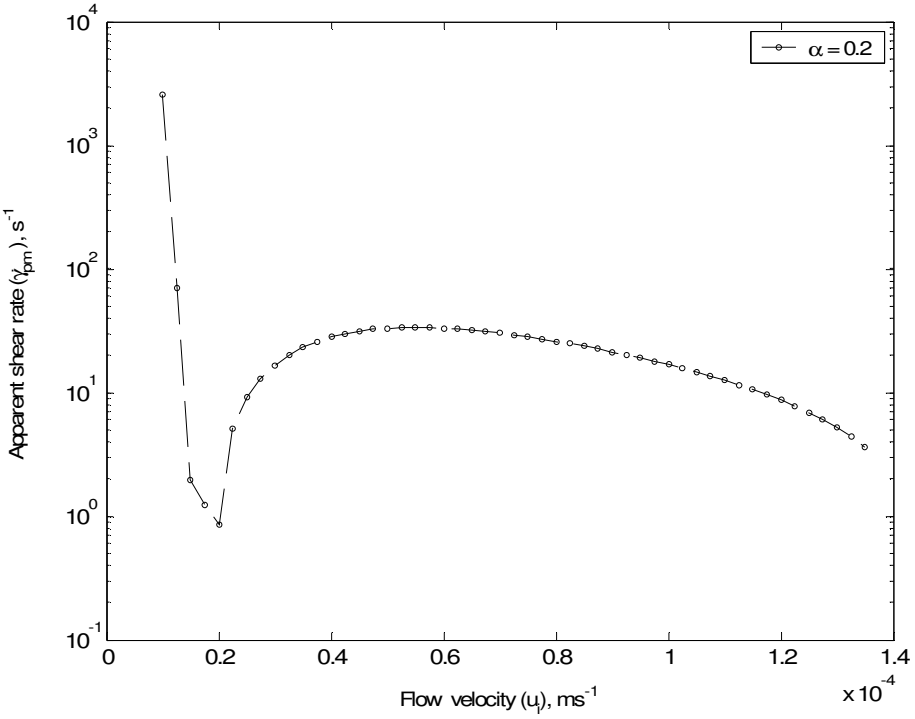


FIGURE 4: Variation of apparent shear rate with flow velocity for $\alpha = 0.2$

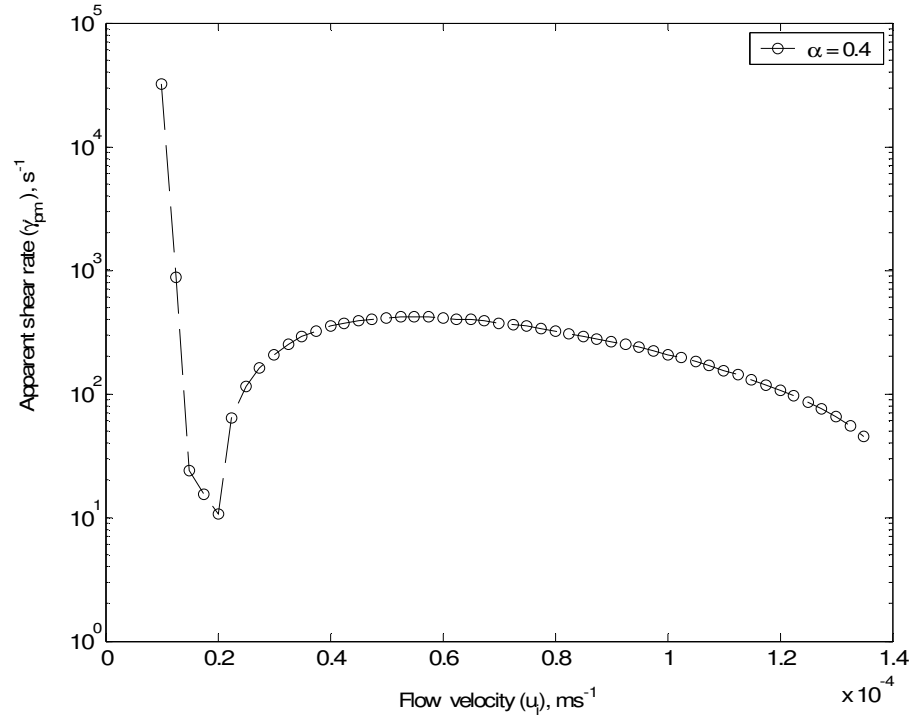


FIGURE 5: Variation of apparent shear rate with flow velocity for $\alpha = 0.4$

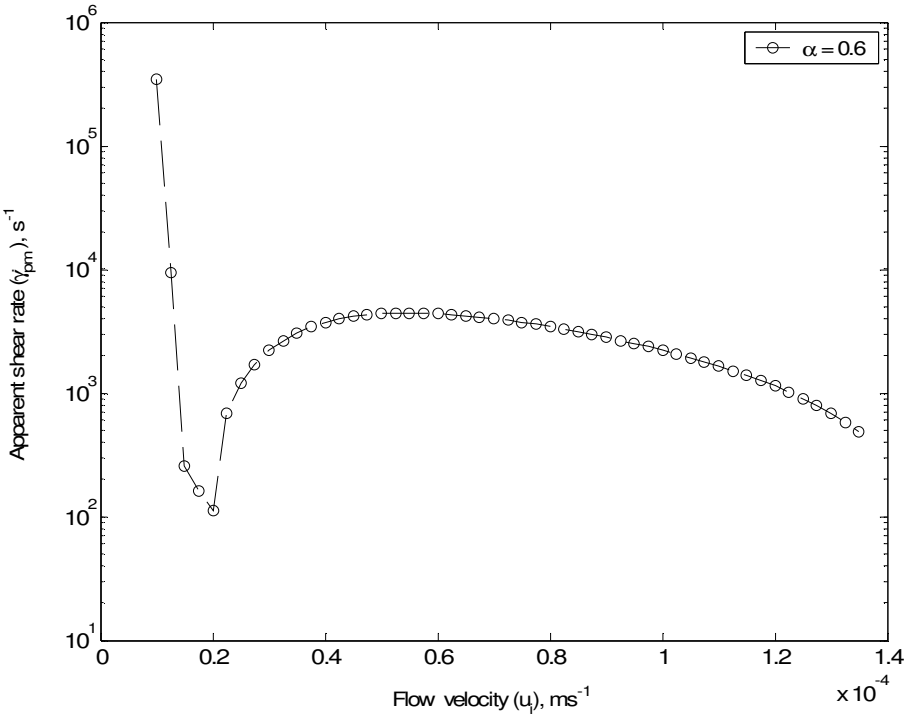


FIGURE 6: Variation of apparent shear rate with flow velocity for $\alpha = 0.6$

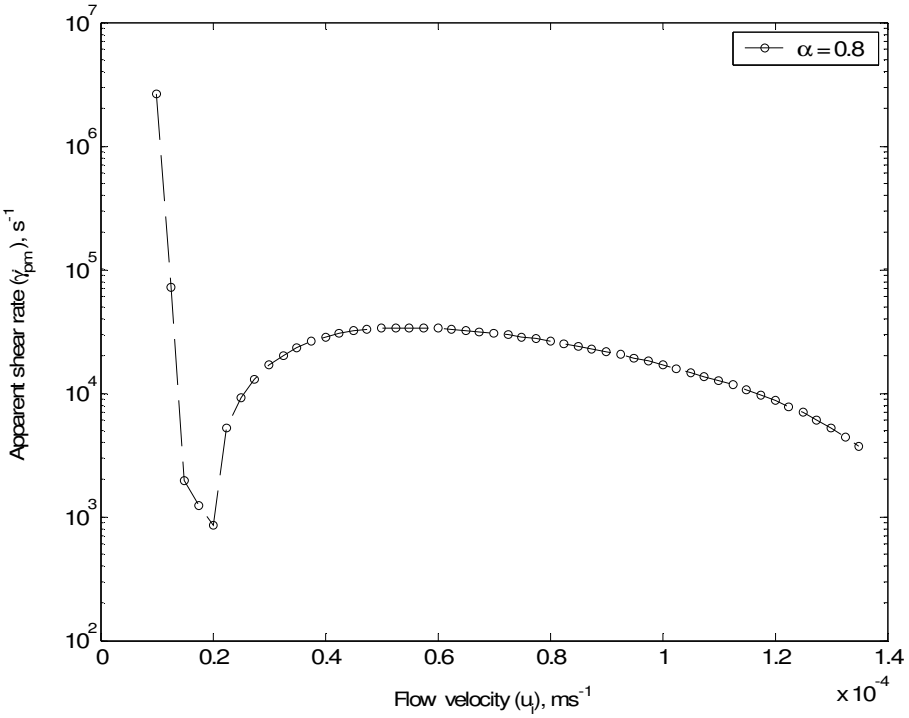


FIGURE 7: Variation of apparent shear rate with flow velocity for $\alpha = 0.8$

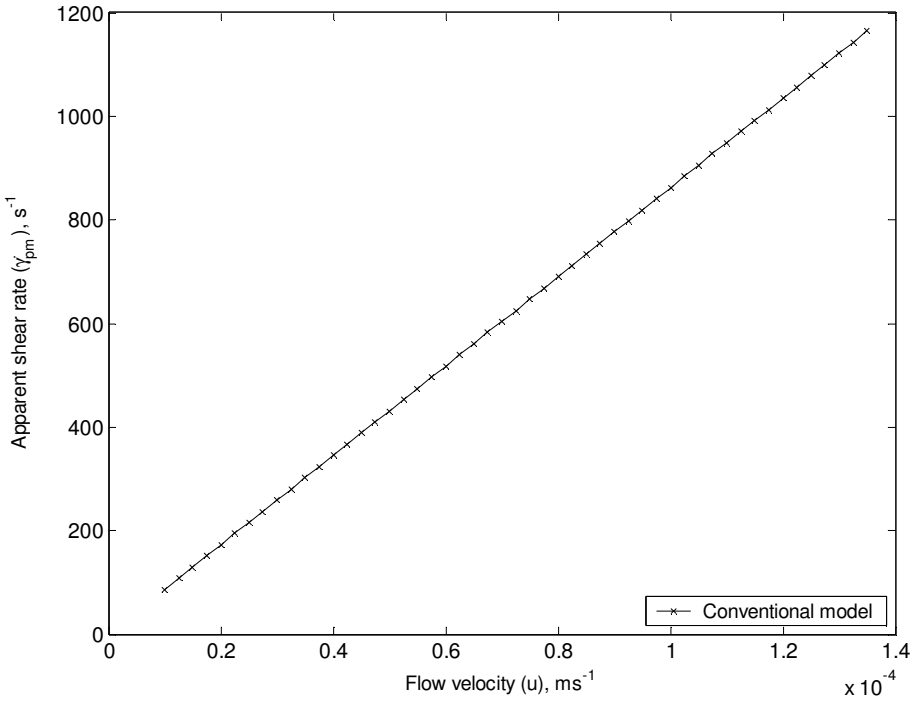


FIGURE 8: Variation of shear rate with flow velocity for the model presented in Eq. (1).

5.2 Comparison of proposed model with the conventional model based on flow velocity

Figure 9 presents a semi log plotting for the variation of in situ shear rate with flow velocity for different α values. This graph compares the proposed model results with those of conventional models. Here porosity and permeability are considered to be 30% and 70 mD, respectively for both the model. It is clear that memory has a potential influence on the shear thinning fluid which may be characterized by the microscopic rheological property of the media and fluid. Note that the memory effect, $\dot{\gamma}_{pm}$ increases with the increase of α values. This indicates that at a very low α value, the memory effect is not very significant at higher flow rates.

5.3 Effective viscosity dependency on shear rate for different α values

Figures 10 – 13 present in log-log plotting the variation of effective viscosity with shear rate for different α values. The trend and shape of the curves generated by the proposed model are similar. It is interesting to note that the original shape and trend of the Carreau-Yasuda model, as depicted in Eq. 2, is similar to that of the proposed modified model except the range and shifting of the data variation. This shifting may be characterized by adjusting the apparent and effective rheological properties or by "scaled up" as stated by Savins [43-44]. The data range and shifting range are different for different α values (Figs. 10 – 13). As α value is increased, the range of data expands in both for viscosity and shear rate values. This means that the zero shear region, transition region, power law region and infinite shear region as stated by Lopez [27, 35] for a viscosity-shear rate curve are more visible as α increases. The ranges of these different regions are dependent on fluid types used. Therefore, it may be concluded that the viscosity-shear rate curve regions are dependent on fluid memory and the regions are more dominant at higher fluid memory, denoted by higher α values. The curves show that the viscosity variation is low at low shear rate and it tends to be reduced with the increase of shear rate. As α increases, viscosity

reduction leads to the increase in shear rates. This is the behavior of a viscoelastic fluid, which is captured by fluid memory. Therefore, it may be concluded that the shear-thinning fluid has the memory when it tries to start a move in porous media.

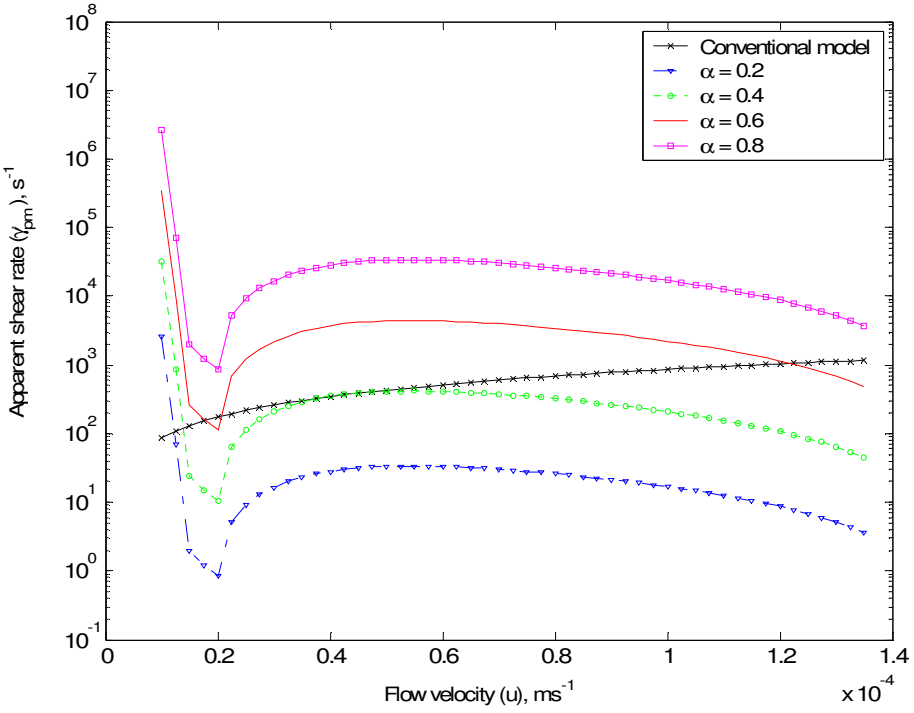


FIGURE 9: Variation of apparent shear rate with flow velocity for comparing the proposed model (Eq. (6)) and the model presented in (Eq. (1)).

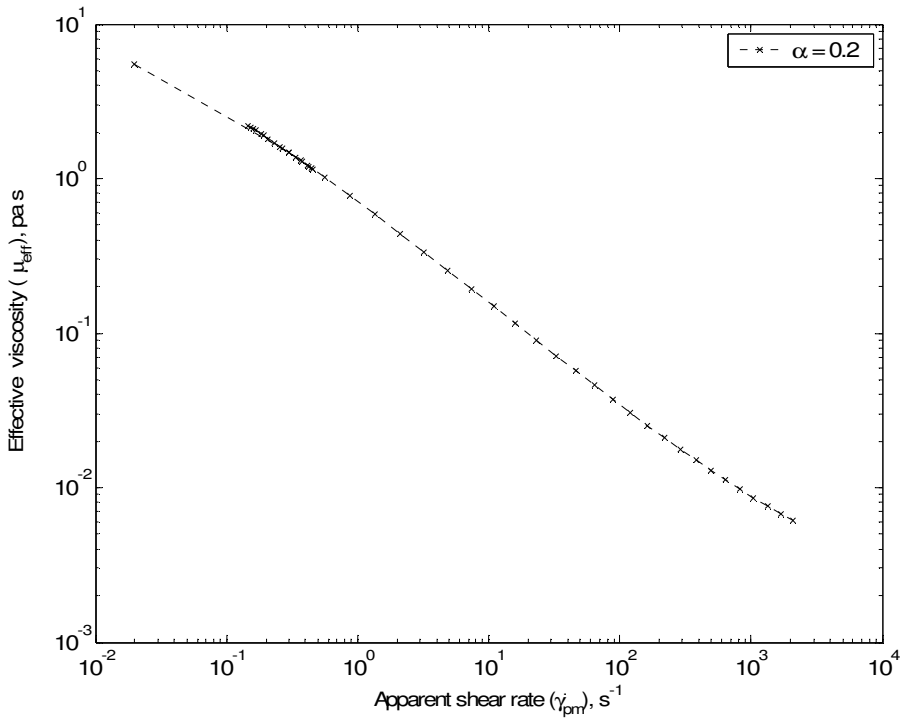


FIGURE 10: Variation of Effective viscosity with apparent shear rate for $\alpha = 0.2$

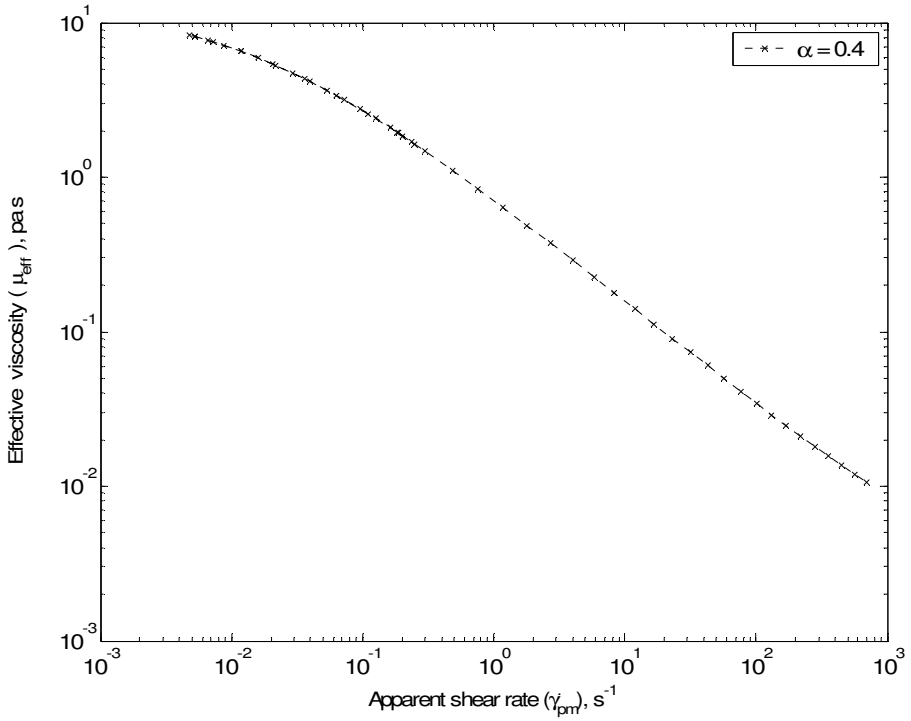


FIGURE 11: Variation of Effective viscosity with apparent shear rate for $\alpha = 0.4$

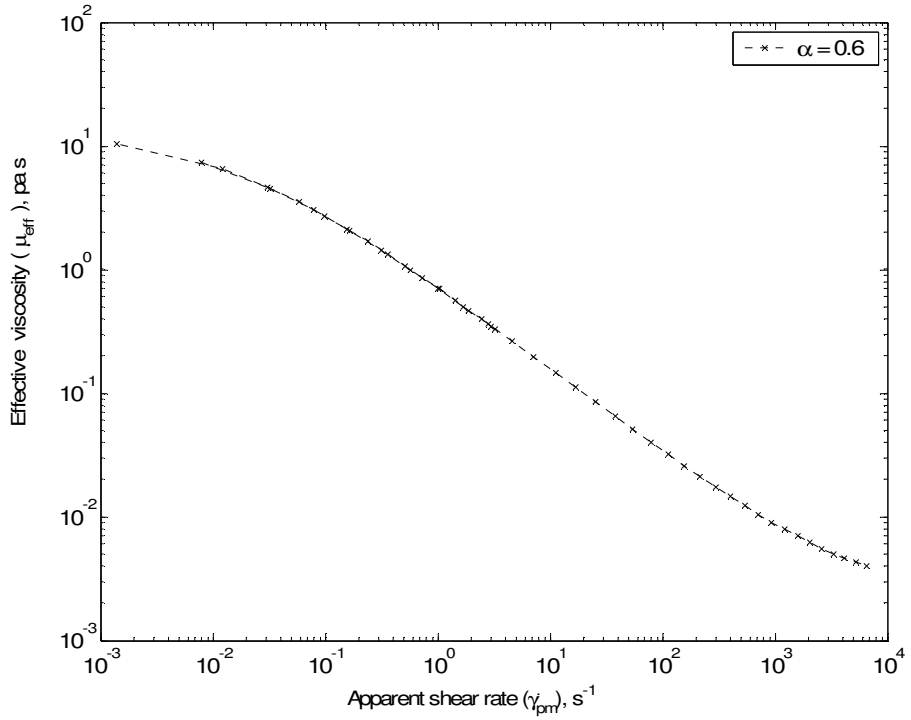


FIGURE 12: Variation of Effective viscosity with apparent shear rate for $\alpha = 0.6$

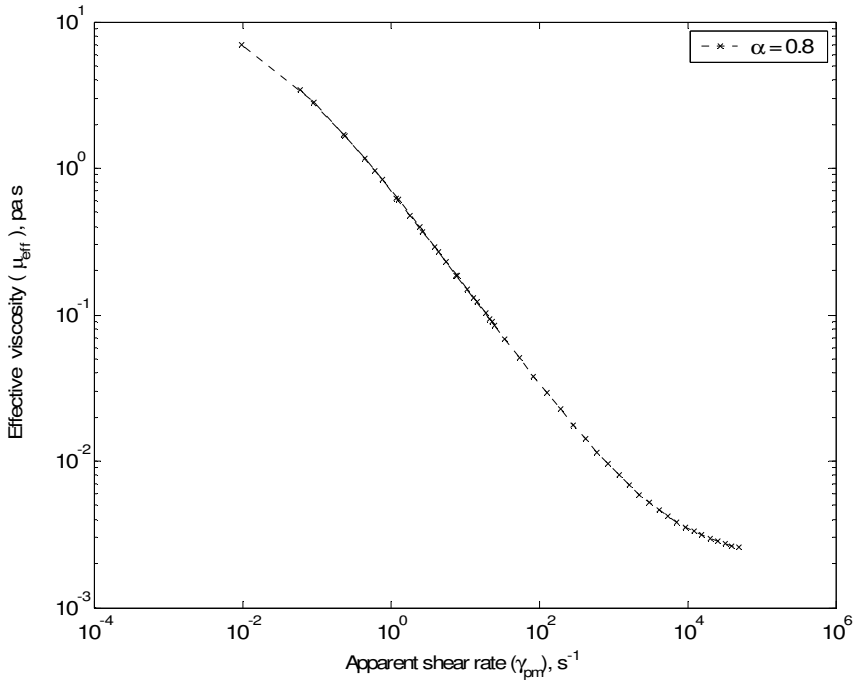


FIGURE 13: Variation of Effective viscosity with apparent shear rate for $\alpha = 0.8$

Figure 14 shows the variation of effective viscosity verses shear rate for different α values in a log-log plotting. In this figure, it is clear how the fluid memory plays a role in viscosity-shear rate relations. All data generated for different α values are overlapped. However, the ranges of data vary with the increase of α value. This conforms to what has been explained earlier.

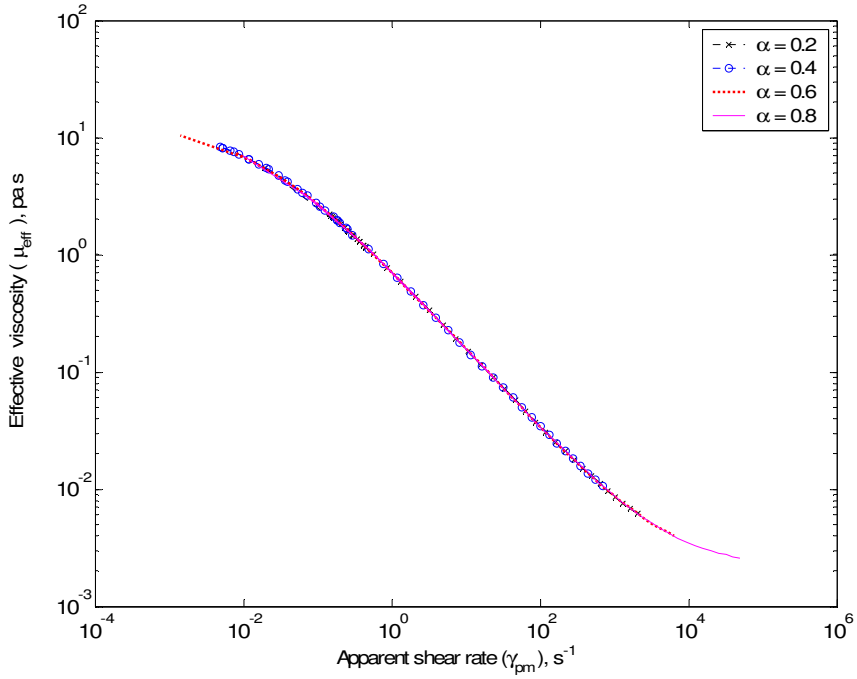


FIGURE 14: Variation of Effective viscosity with apparent shear rate for $\alpha = 0.2 \; to \; 0.8$

5.4 Comparison of proposed viscosity model with Carreau-Yasuda model

Figure 15 shows the variation of viscosity vs. shear rate of the proposed model (Eq. (8)) for different α values. Results of this model are compared with those of the Carreau-Yasuda model (Eq. (2)) in a log-log plotting. All the data generated by solving these two models are overlapped with each other except the range of data variation. For the same conditions and input data, the proposed model provides one with more information than Carreau-Yasuda model. The proposed model provides a wider range of data in both zero shear and infinite shear region. The existence of Carreau-Yasuda model is valid only in the power-law region if we compare it with proposed model. It is also noted that all the α values data lie within the transition and power-law region that are very difficult to capture and explain. If α increases, the data range extend to reach the other two regions, zero shear and infinite shear. Therefore, it may be concluded that the proposed model is more suitable for characterizing the rheological properties of the shear-thinning fluid flow in porous media.

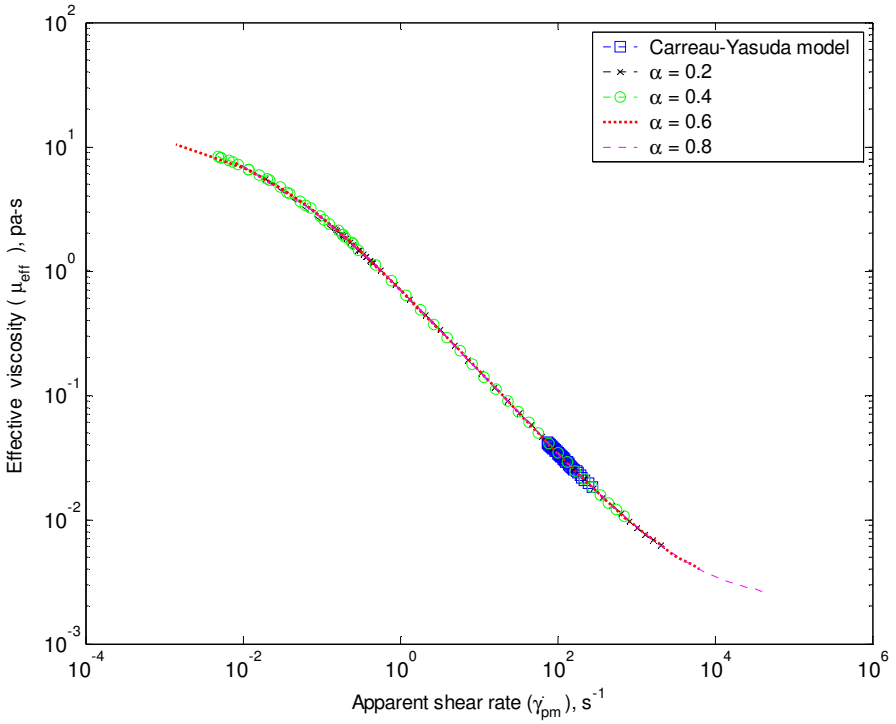


FIGURE 15: Comparison of proposed viscosity model with Carreau-Yasuda model.

6. CONSLUSIONS

In this study, two models are proposed to characterize the rheological behavior with memory for shear-thinning fluids. These models are validated with the available experimental data and also compared with currently used conventional mathematical models. The proposed models are effective in capturing physical phenomena. In addition, they are consistent with existing mathematical models. However, the proposed models capture a wider range of information, covering fluids that would not be tractable with existing models. The shear rate-flow velocity has a non-linear variation, which poses the intriguing question as to why conventional linear relationships do not hold. The answer lies within considerations of memory effects. In this paper, we focused on the dependence of the shear rate on porosity, permeability, shape factor and flow velocity which is related to the effect of fluid memory. Also considered is the viscosity

dependency on shear rate. This study concludes that fluid memory has a strong influence on shear-thinning fluid flow behavior during the propagation of a shear-thinning fluid in porous media and should be considered in various applications involving complex fluids.

7. ACKNOWLEDGEMENT

The authors would like to thank the Atlantic Canada Opportunities Agency (ACOA) for funding this project under the Atlantic Innovation Fund (AIF). The first author would also like to thank Natural Sciences and Engineering Research Council of Canada (NSERC) for funding.

8. NOMENCLATURE

a = parameter in Carreau–Yasuda model, dimensionless
A = cross sectional area of rock perpendicular to the flow of flowing fluid, m^2
k = initial reservoir permeability, m^2
L = length of a capillary or a core, m
n = power-law exponent for Carreau–Yasuda model, dimensionless
p = pressure of the system, N/m^2
Δp = differential pressure along a core or capillary of length L, N/m^2
Q = Au = initial volumetric flow rate, m^3/s
q_x = fluid mass flow rate per unit area in *x*-direction, $kg/m^2 s$
t = time, s
u = Darcy velocity ($= Q/A$), m/s
u_x = fluid velocity in porous media in the direction of x axis, m/s
α = fractional order of differentiation, dimensionless
α_{sp} = shape factor which is medium-dependent
$\dot{\gamma}_{sm}$ = apparent shear rate within the porous medium, $1/s$
ρ_0 = density at a reference pressure p_0, kg/m^3
φ = porosity of fluid media, m^3/m^3
μ = fluid dynamic viscosity, $pa\,s$
μ_0 = fluid dynamic viscosity at zero shear rate, $pa\,s$
μ_∞ = fluid dynamic viscosity at infinite shear rate, $pa\,s$
λ = time constant in Carreau–Yasuda model, s
η = ratio of the pseudopermeability of the medium with memory to fluid viscosity, $m^3\,s^{1+\alpha}/kg$
ξ = a dummy variable for time i.e. real part in the plane of the integral, s

9. REFERENCES

1. T. Li, J. Gao, R. Szoszkiewicz, U. Landman, and E. Riedo, "*Structured and viscous water in subnanometer gaps".* Physical Review B, 75(115415): 1-6, March 15, 2007
2. B. Mauk, Water Discovered to Flow Like Molasses. Special to LiverScience, May 11, (Link: http://www.livescience.com/technology/070511_small_water.html, accessed on June 10, 2007). 2007.

3. C.A. Perazzo, and J. Gratton, *"Thin film of non-Newtonian fluid on an incline"*. Physical Review E 67(016307): 1 – 4, 2003
4. P.E. Arratia, T. Shinbrot, M. M. Alvarez and F. J. Muzzio, *"Mixing of non-Newtonian fluids in steadily forced system"*. Physical Review Letters, PRL 94(084501): 1 – 4, 2005
5. Z. Donko´, J. Goree, P. Hartmann, and K. Kutasi, *"Shear viscosity and shear thinning in two-dimensional Yukawa liquids"*. Physical Review Letters, PRL 96 (145003): 1 – 4, 2006
6. L. Kondic, P. Palffy-Muhoray and M. J. Shelley, *"Models of non-Newtonian Hele-Shaw flow"*. Physical Review E, 54(5): R4536 – R4539, November, 1996
7. J. A. Miranda, *"Shear-induced effects in confined non-Newtonian fluids under tension"*. Physical Review E, 69(016311): 2004
8. K. Afanasiev, A. Münch, and B. Wagner, *"Landau-Levich problem for non-Newtonian liquids"*. Physical Review E 76(036307): 1 – 12, 2007
9. X. Frank and H. Z. Li, *"Complex flow around a bubble rising in a non-Newtonian fluid"*. Physical Review E, 71(036309): 1 – 5, 2005
10. Y. Huang, and I. Lin. *"Memory and persistence correlation of microstructural fluctuations in two-dimensional dusty plasma liquids"*. Physical Review E, 76(016403): 1 – 6, 2007
11. K. J. Lissant, *"Emulsions and Emulsion Technology: Part I"*. Marcel Dekker Inc., New York, U.S.A., 1974
12. K. J. Lissant, *"Emulsions and Emulsion Technology: Part II"*. Marcel Dekker Inc., New York, U.S.A., 1974
13. S. P. Parker, Mcgraw-Hill Encyclopedia of Science and Technology: 7th. Edition. McGraw-Hill Book Company Inc., McGraw-Hill Staff, Jonathan Weil, Betty Richman, 1992.
14. M. Caputo, Geothermics, *"Diffusion of fluids in porous media with memory"*. 23: 113 – 130, 1999
15. M. Caputo, Water Resources Research, *"Models of Flux in porous media with memory"*. 36(3): 693 – 705, 2000
16. M.E. Hossain, S.H. Mousavizadegan, C. Ketata and M.R. Islam, *"A Novel Memory Based Stress-Strain Model for Reservoir Characterization"*. Journal of Nature Science and Sustainable Technology, 1(4): 653 – 678, 2007
17. M.E. Hossain, S.H. Mousavizadegan and M.R. Islam, *"Rock and Fluid Temperature Changes during Thermal Operations in EOR Processes"*. Journal of Nature Science and Sustainable Technology, 2(3), in press
18. K. S. Sorbie, A. Parker, and P.J. Clifford, *"Experimental and theoretical study of polymer flow in porous media"*. Proceedings of SPE Annual Technical Conference and Exhibition, SPE 14231, Las Vegas, Nevada, 1987, September 22-25.
19. W.J. Cannella, C. Huh and R., Seright, *"Prediction of Xanthan Rheology in Porous Media"*. proceedings of the 63rd Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, SPE 18089, Houston, TX, USA, 1988.
20. K. S. Sorbie, *"Network modeling of Xanthan rheology in porous media in the presence of depleted layer effects"*. Proceedings of the SPE Annual Technical Conference and Exhibition, SPE 19651, San Antonio, Texas, 1989, October 8-11.
21. K. S. Sorbie, *"Depleted layer effects in polymer flow through porous media – 2 network calculations"*. Journal of Colloid and Interface Science 139(2): 315-323, 1990
22. K. S. Sorbie, *"Depleted layer effects in polymer flow through porous media – 1 single capillary calculations"*. Journal of Colloid and Interface Science 139(2): 299 – 314, 1990
23. K. S. Sorbie and Y. Huang, *"Rheological and transport Effects in Flow of Low-Concentration Xanthan Porous Media"*. Journal of Colloid and Interface Science, 145(1), 74 – 89, 1991
24. M.P. Escudier, I.W. Gouldson, A.S. Pereira, F.T. Pinho and R.J. Poole, *"On the reproducibility of the rheology of shear-thinning liquids"*. J. Non-Newtonian Fluid Mech, 97: 99–124, 2001
25. A. Fadili, P. M. J. Tardy and J. R. A. Pearson, *"A 3D filtration law for power-law fluids in heterogeneous porous media"*. J. of Non-Newtonian Fluid Mech., 106: 121-146, 2002
26. J. R. A. Pearson and P.M.J. Tardy, *"Models for flow of non-Newtonian and complex fluids through porous media"*. J. of Non-Newtonian Fluid Mech. 102: 447 – 473, 2002

27. X. Lopez, P.H. Valvatne and M.J. Blunt, *"Predictive network modeling of single-phase non-Newtonian flow in porous media"*. Journal of Colloid and Interface Science 264(1): 256–265, 2003
28. X. Lopez and M.J. Blunt, *"Predicting Impact of Non-Newtonian Rheology on Relative Permeability using Pore-Scale Modeling"*. Proceedings of SPE Annual Technical Conference and Exhibition, SPE 89981, Houston, U.S.A., 2004, September 26-29
29. C.L. Perrin, P.M.J. Tardy, K.S. Sorbie and J.C. Crawshaw, *"Experimental and modeling study of Newtonian and non-Newtonian fluid flow in pore network micromodels"*. Journal of Colloid and Interface Science 295: 542–550, 2006
30. M.E. Hossain and M.R. Islam, *"Fluid Properties with Memory – A Critical Review and Some Additions"*. Proc. 36th International Conference on Computers and Industrial Engineering, CIE – 00778, Taipei, Taiwan, 2006, June 20-23.
31. P. Vogel and G. A. Pusch, *"Some aspects of the injectivity of non-Newtonian fluids in porous media"*. Proceedings of the First European Symposium on EOR, Bournemouth, UK, 1981, September 21-22.
32. E. J. Kolodziej, *"Transport mechanisms of Xanthan biopolymer solutions in porous media"*. Proceedings of the 63rd SPE Annual Technical Conference and Exhibition, SPE 18090, Houston, Texas, 1988, October 2-5.
33. A.J.P. Fletcher, S.R.G. Flew, S.P. Lamb, T. Lund, E. Bjørnestad, A. Stavland, and N. B. Gjøvikli, *"Measurements of polysaccharide polymer properties in porous media"*. Proc. of the International Symposium on Oil field Chemistry, SPE 21018, Anaheim, California, 1991, February 20-22.
34. S. Hejri, G. P. Willhite and D.W. Green, *"Development of correlations to predict biopolymer mobility in porous media"*. SPE Reservoir Engineering. 6: 91-101, 1991
35. X. Lopez, "*Pore-Scale Modeling of Non-Newtonian Flow"*: PhD dissertation, Imperial College London, Department of Earth Science & Engineering, UK, 2004
36. G. Chauveteau and N. Kohler, *"Polymer flooding: The essential elements for laboratory evaluation"*. SPE 4745, Journal of Petroleum Technology April, 22-24, 1974
37. G. Chauveteau, *"Rod like Polymer Solutions Flown through Fine Pores: Influence of Pore size on Rheological Behavior"*. Journal of Rheology 26(2): 111-142, 1982
38. A.M. Sani, S.N. Shah, and L. Baldwin, *"Experimental investigation of Xanthan foam rheology"*. Proc. of the Production and Operations Symposium, SPE 67263, Oklahoma City, Oklahoma, March 24 – 27, 2001
39. C.D. Tsakiroglou, *"Correlation of the two-phase flow coefficients of porous media with the rheology of shear-thinning fluids"*. J. Non-Newtonian Fluid Mech. 117: 1–23, 2004
40. D.H. Fenwick and M.J. Blunt, *"Network Modeling of Three-Phase Flow in Porous Media"*. SPE Journal, March: 86 – 97, 1998
41. D.A. Coombe, V. Oballa and W.L. Buchanan, *"Scaling Up the Non-Newtonian Flow Characteristics of Polymers, Foams, and Emulsions"*. presented at the 12th SPE Symposium on Reservoir Simulation, SPE – 25237, Orleans, LA, U.S.A., February 28 – March 3, 1993
42. M.T. Balhoff and K.E. Thompson, *"A macroscopic model for shear-thinning flow in packed beds based on network modeling"*. Chemical Engineering Science 61: 698 – 719, 2006
43. J.G. Savins, *"Non-Newtonian Flow through Porous Media"*. Industrial and Engineering Chemistry, 61(10): 1- 17, 1969
44. J. G. Savins, *"Non-Newtonian flow through porous media"*. Industrial and Engineering Chemistry, 61(10): 18 – 47, 1969
45. G.Chauveteau and J.E. Glass, *"Water Soluble Polymers: Advances in Chemistry"*. American Chemical Society, Washington, DC, 213: 227–268, 1986
46. J.L. Auriault, and C. Boutin, *"Deformable Porous Media with Double Porosity Quasi-Statics. II: Memory Effects"*. Transport in Porous Media, 10: 153-169, 1993
47. K.S. Sorbie, P.J. Clifford and E.R.W. Jones, *"The rheology of pseudoplastic fluids in porous media using network modeling"*. Journal of Colloid and Interface Science 130(2): 508-534, 1989

Effect of Bend Curvature Ratio on Flow Pattern at a Mixing Tee after a 90 Degree Bend

Mohammadreza Nematollahi nema@shirazu.ac.ir
School of Mechanical Engineering
Shiraz University
Shiraz, 71348-51154, Iran

Mohammad Nazififard mnazifi@shirazu.ac.ir
School of Mechanical Engineering
Shiraz University
Shiraz, 71348-51154, Iran

Maziar Asmani maziarasmani@yahoo.com
School of Electrical and Computer Engineering
Iran Azad University of Qeshm Branch
Qeshm Free Area, 795151393, Iran

Hidetoshi Hashizume hidetoshi.hashizume@qse.tohoku.ac.jp
School of Engineering/Department of Quantum Science and Energy Engineering
Tohoku University
Tohoku, 980-8579, Japan

Abstract

Many nuclear power plants report high cycle thermal fatigue in their cooling system, caused by temperature fluctuation in a non-isothermal mixing area. One of these areas is the T-junction, in which fluids of various temperatures and velocities blend. The objective of this research is to classify turbulent jet mechanics in order to examine the flow-field structure under various operating conditions. Furthermore, this research discovers the optimum operating conditions of the mixing tee in this piping system. An experimental model, including the T-junction with a 90 degree bend upstream, is operated to analyze this mixing phenomenon based on the real operation design of the Phenix Reactor. The temperature and velocity data show that a 90 degree bend has a strong effect on the fluid mixing mechanism and the momentum ratio between the main velocity and the branch velocity of the T-junction, which could be an important parameter for the classification of the fluid mixing mechanism. By comparing their mean velocity distributions, velocity fluctuations and time-series data, the behavior of the branch jet is categorized into four types of turbulent jets; sorted from the highest to the lowest momentum ratios, the jets are categorized as follows: the wall jet, the re-attached jet, the turn jet, and the impinging jet. Ultimately, the momentum ration of the turn jet was selected as the optimum

operating condition as it has the lowest velocity and the lowest temperature fluctuations near the wall of the mixing tee.

Keywords: T-Junction, Mixing Phenomena, Secondary Flow, Phenix reactor.

1. INTRODUCTION

Thermal stress arising from temperature fluctuations in the cooling system of nuclear power plants is inevitable. When thermal stress is generated by a sudden change in temperature, the process is referred to as thermal shock; failure under repetitive application of thermal stress has been termed thermal fatigue. The severity of the failure is dependent on the shape of the component, the fluid mixing mechanism, and the temperature distribution and it's fluctuations. Studies of thermal fatigue in power plants were initially carried out for liquid-metal-cooled fast breeder reactors because of the high thermal conductivity of liquid-metal coolants. After many recent thermal fatigue events occurred in various nuclear power plants, such as the French PWR CIVAX in 1998, the Japanese PWR Tsuruga-2 in 1999, the Japanese PWR Tomari-2 in 2003, the focus of thermal striping studies shifted to not only fast breeder reactors, but also light water reactors. The T-junction was selected because it is a component common to the cooling systems of most nuclear power plants, having a high capability of thermal fatigue. Several types of numerical and experimental research have been performed based on the fluid mixing phenomena in the T-junction, such as the evaluation of thermal fatigue [1], the numerical simulation of the mixing phenomenon [2], and the analysis of the flow-field structure [3], all of which considered the T-junction as a single component. Although the T-junction has typically been considered a single component, the mixing tee is usually connected to another part of the complex piping system one of which is a 90 degree bend that typically exists upstream of the T-junction and has strong effects on the mixing mechanism [4,5]. The Particle Image Velocimetry (PIV) technique was used to experimentally investigate the fluid mixing phenomenon in the T-junction area with the 90 degree bend upstream [6–8]. This visualization displays origins of the velocity fluctuations in the T-junction area. The thermo-hydraulic characteristics of the turbulent jet are analyzed to better classify the fluid mixing mechanism. This classification is an important factor in order to categorize operating conditions and estimate the effects of thermal fatigue in each mixing condition in the piping system. The effects of velocity and pipe diameter on the fluid mixing mechanism were investigated to explain the behavior of turbulent jets and different operating conditions. Regarding the reported effects on the fluid mixing mechanism of a 90 degree smooth bend with a carving ratio of 1.41, in the present study attempts have been made to investigate the effects of the bend curvature ratio and the axial distance between the bend and T-junction as the target to find their effects. The results are presented here under the title of "Classification of Turbulent Jet in a T-junction area with a 90 degree Sharp Bend on its Upstream."

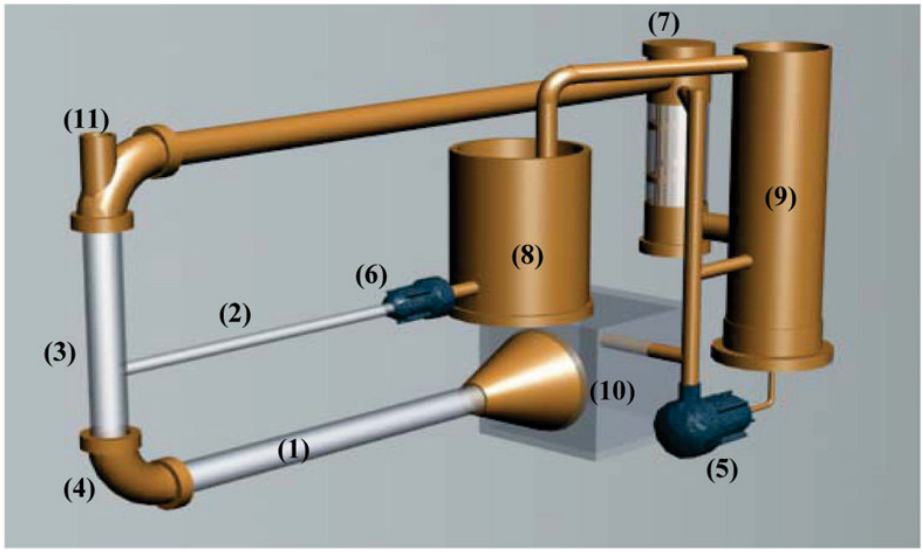
2. PRELIMINARIES

2.1. Experimental apparatuses

The fluid cycle system and T-junction area are shown in Fig. 1. Both the main and branch pipes are made of 3mm thick acrylic circular pipes, and the branch pipe is connected to the main pipe at the right angle to form the mixing tee with a square-edge. The internal main pipe diameter, D_m, is 108 mm and the branch pipe can have one diameter, as well, D_b = 21 mm. The main pipe runs vertically upward in the test section, which is connected to a measuring window downstream for visualizing the lateral flow-field.

A 90 degree sharp bend with a carving ratio of one is installed upstream of the T-junction. This bend is stainless steel with a 1.0 curvature ratio. The main flow is straightened by a reducer, a long main pipe before the 90 degree bend ($13D_m$), and a straightener tank made from acrylic

plates with 30 X 30 X 30 cm. The branch pipe is also long enough (60 D_b) to inject fully developed flow into the mixing area.
A heat-exchanger and a heating tank are used to control the inlet temperature of both the main flow and the branch flow in the mixing tee. They are respectively set behind the mixing tank and the branch pump. The heat exchanger has a secondary cooling system for decreasing the temperature with city water. While its flow rate is controlled by an inlet heat-exchanger valve, Whereas the heating tank uses gas and electric heaters to warm the branch flow.



(1) Main Pipe (2) Branch Pipe (3) Tee Junction (4) 90-Degree Bend
(5) Main Pump (6) Branch Pump (7) Heat Exchanger (8) Heating Tank
(9) Mixing Tank (10) Straightener Tank (11) Measuring Window

Figure 1. Experimental apparatuses.

The mixing tank is positioned in order to insert small tracers and remove bubbles from the main cycle. This tank is linked to other apparatuses by five pipe connections (the heat-exchanger pipe, emergency inlet pipe, heating tank pipe, main pump pipe and draying outlet pipe). There are two main and branch pumps with 600 and 75 L/min maximum flow rates. A flow rate control valve and an inverter individually adjust the mean velocity of these flows. There is a cube water jacket around the T-junction area which is made of the same acrylic plates as the piping material to decrease the effect of the laser light's refraction through the circular pipe wall during the visualization of the longitudinal sections.

2.2. Experimental conditions and analysis methods

Water was the only working fluid in the experimental loop throughout this research. The PIV system is used to visualize the flow characteristics in the T-junction area. Two conditions are selected to visualize the flow-field, the whole flow-field (long-shot) and the close-up flow-field with 150 mm*150 mm and 60 mm*60 mm view sections, respectively. The resolution of each section is 1018*1008 pixels. The interrogation cell is divided into 64 * 64 pixels at each segment for measuring the whole flow-field and 32 * 32 pixels for measuring a close-up area with a 50% area overlap and 100 time interval. The cross-section is visualized into the radial and axial velocity matrixes ($u_{r,i}$ and $u_{z,i}$). Both matrixes have 30 * 29 arrays when 64 * 64 pixels are cross-correlated, and have 62 * 59 arrays when 32 * 32 pixels cross-correlations are used. The laser sheet thickness is 1–3 mm, with a 200 mJ energy level; this width is selected based on three parameters, the visualization area, the amount of tracer in the fluid and the distance between the visualization cross-section and the camera lens. Two kinds of tracers with different diameters are

used, both made from nylon powder with 1.03 g/cm^3 density. The tracer with an 80 μm diameter is used for visualizing the whole flow-field and a 20 μm diameter for the close-up flow-field condition. The camera starts shooting with a 30 Hz frequency, which has a frame rate of 30 fps in triggered double exposure mode; each shot continually captures 99 images. Forty-nine velocity vector maps are obtained with only 0.03 s time gaps. Five shots are taken, and in total, 240 frames are used to evaluate both flow-fields. The average flow-field and intensity of velocity fluctuation is evaluated by the following equations:

$$U_{ave} = \frac{1}{n}\sum_{i}^{n}\sqrt{(u_{r,i}{}^2 + u_{z,i}{}^2)} \qquad (1)$$

$$s_j = \left(\frac{1}{n}\sum_{i=1}^{n}(u_{j,i} - \overline{u_j})^2\right)^{1/2} \qquad (2)$$

$$U_{mix} = \sqrt{(U_b^2 + U_m^2} \qquad (3)$$

$$I = \sqrt{(s_r + s_z)/2}/U \qquad (4)$$

Where $u_{r,i}$ and $u_{z,i}$ represent the instantaneous radial and axial velocities at the i frame, U_{ave} is the averaged absolute velocity, and n is the total number of the frames. U_b and U_m represent both the branch and main velocities. s_r and s_z are standard deviations of the velocity variation in the radial and axial directions. Lastly, the intensity of the velocity fluctuation represented. Here, both the main and branch flows are entirely in the turbulent regime. The temperature fluctuation intensity, ΔT_{rms}, is calculated at each measuring point from the time-series temperature data and is normalized by using the following equations:

$$\Delta T = T_{branch} - T_{main} \qquad (5)$$

$$\Delta T_{rms} = \left(\sqrt{\sum (Ti - Tm)^2 / m}\right)/\Delta T \qquad (6)$$

Where T_i and T_m represent instantaneous temperature and mean temperature, ΔT is the temperature difference between two fluids, and m is total sampling number of the temperature data. Another interesting parameter is fluctuation, which is introduced by: $\sqrt{(u_x - u_{x-1})^2 + (u_y - u_{y-1})^2}$. This parameters show the intensity secondary flow for each position.

3. Results & Discussion

3.1. Classification of turbulent jets

Turbulent jets in finite space show various behaviors. One of these finite spaces is the T-junction area in which two pipes with different diameters and flow are connected together at the right angle with a square-edge. Based on the velocity and momentum ratio of these pipes, the flow pattern in the mixing tee area has different mechanisms, thus many turbulent jets exist. The other parameters used to categorize the mixing mechanism are the Reynolds number of the branch pipe and the Dean number of the main pipe, which are more complicated than the momentum ratio, therefore, they do not show a clear classification of the jets.

Depending on the momentum/velocity ratio of the entering flows from branch pipe and the main pipe, the turbulent mixing patterns can be further divided into four branch jets such as the wall jet, re-attached jet, turn jet and impinging jet. The types of mixing flow are categorized by using the momentum ratio equation as follows:

$$M_R = \frac{\rho_m U_m^2 . (D_m \times D_b)}{\rho_b U_b^2 . \pi . (D_b/2)^2} \qquad (7)$$

M_R is the momentum ratio, ρ_m and ρ_b are the fluid densities for the main flow and branch flow, U_m and U_b are the mean velocity of the main and branch flow, D_m and D_b are the main and branch pipe diameters.

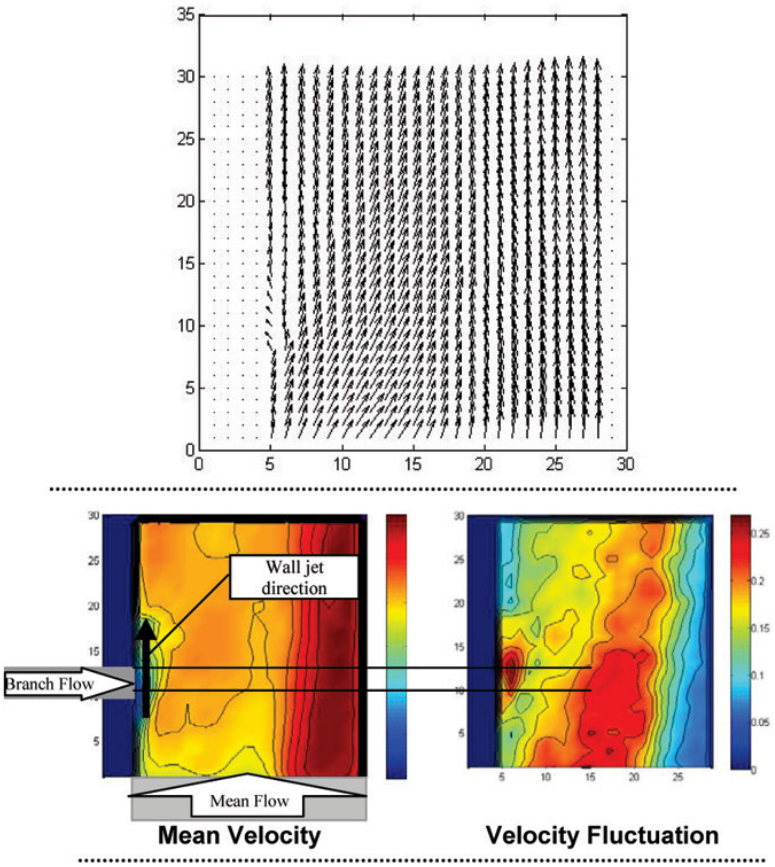


Fig. 2. Wall jet (Dm = 108 mm, Db = 21 mm, Um = 0.9 m/s, Ub = 0.3 m/s).

For distinguishing these four jets in the flow pattern, we need a basic introduction of each jet and some parameters to distinguish each group. Three parameters are used to investigate their structures and mechanisms: mean velocity distribution, velocity fluctuation and time-series data (Fig. 2).

Attending to these parameters, the branch jets are separated by mean velocity distribution and velocity fluctuation. The mean velocity distribution contains information about the average structure of the jet, while, the velocity fluctuations display the area with the highest variation. In the next step, the time-series data of each condition are investigated to explain the gradual change of the temporal flow structure.

For each condition, 240 frames are used to evaluate the mixing flow mechanism; the first three frames from each condition are shown in Fig.2. Finally, two jets are classified into the aforementioned categories based on the shape and behavior of the branch jets. For providing an exact investigation method for these jets, the structure and mechanism of each jet are described separately:

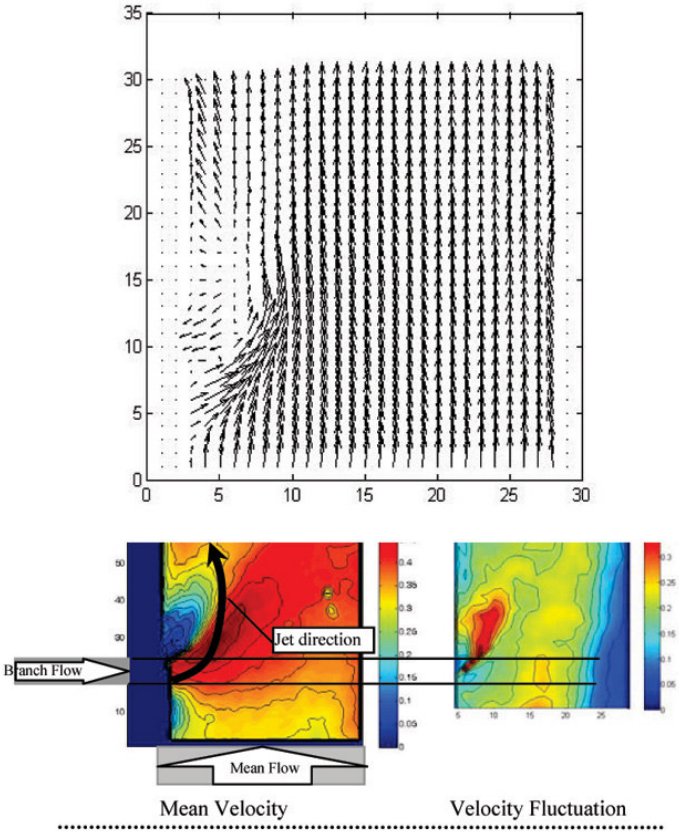


Fig. 3. Wall jet (Dm = 108 mm, Db = 21 mm, Um = 0.6 m/s, Ub = 0.2 m/s).

(a) WALL JET: The wall jet is characterized by a higher main flow and a lower branch flow. Here, the jet does not separate from the main pipe wall. Velocity fluctuations occur only near the wall of the main pipe as shown in Fig. 2.

(b) RE-ATTACHED JET: The re-attached jet is distinguished by an interaction between the secondary flow of the main pipe and the branch flow. The branch jet turns to the center axis of the main pipe and then turns again to the main pipe wall above the branch nozzle in Fig. 3.

(c) TURN JET: The turn jet exists when two inlet flows have comparable momentums, and the branch jet turns to the center axis of the main pipe in the same direction as the main flow provided in Fig. 4.

d) IMPINGING JET: The impinging jet occurs when the branch velocity is much higher than the main velocity, and as a result, the branch flow can touch the opposite wall of the main pipe as visualized in Fig. 5. According to the momentum/velocity ratios between the branch and the main flow, three branch pipe diameters were used to categorize branch jets. In order to categorize flow patterns, 205 conditions were selected.

Each condition was visualized at least five times (240 frames) with the same experimental conditions, such as temperature (18 ºC), 90 degree bend curvature ratio (C_R = 1.0), distance between the 90 degree bend and the branch pipe (d/D_m = 2), etc. By calculating the momentum ratios (M_R) of the main flow and the branch flow based on Eq. (7), a threshold for each flow pattern was defined, as shown in Table 1. Due to many large eddies in the T-junction area with high-mixing Reynolds numbers (Re = 33,000–150,000), it is clear that mainly the mixing phenomena are primarily controlled by the mechanism of these large eddies.

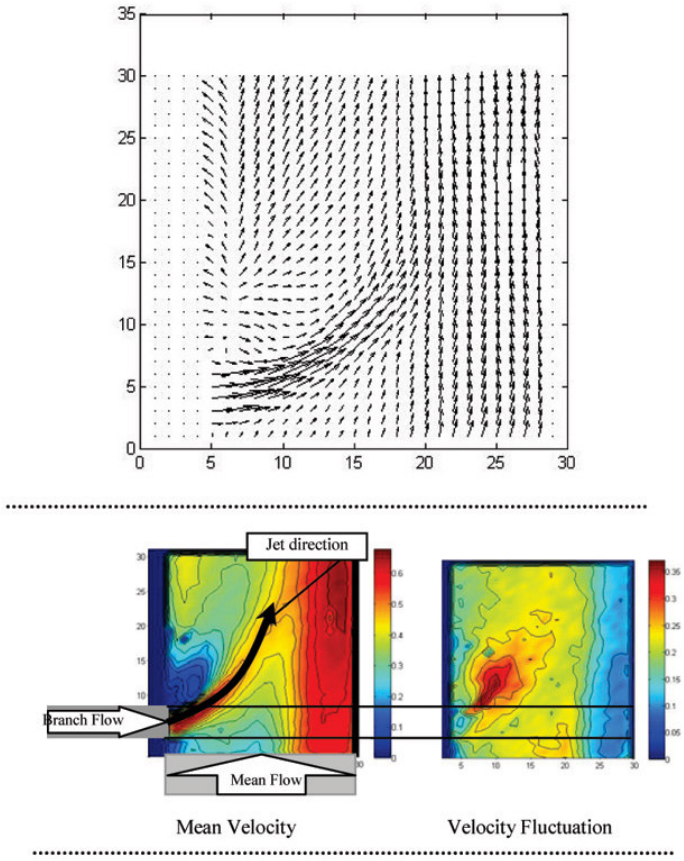


Fig. 4. Turn jet (D_m = 108 mm, D_b = 21 mm, U_m = 0.6 m/s, U_b = 0.6 m/s).

Table 1. Categorization of the branch jet in T-junction with sharp bend on its upstream

WALL JET	58 < M_R
Re-attached jet	9< M_R < 38
Turn jet	2< M_R < 9
Impinging jet	M_R < 1.5

3.2. Operating conditions

3.2.1. Momentum ratio effects

Most of these eddies are formed by pipe geometries and an interaction between the main and branch flows [4,6,9]. Fig. 6 shows the maximum intensity of the velocity fluctuation near the wall of the main pipe in the various momentum ratios.

Velocity fluctuation has minimum intensity when the momentum ratio is around M_R = 2.0. Only two main parameters, velocity ratio and branch pipe diameter, can change the momentum ratio. These two parameters provide different mechanisms for the mixing phenomena, and the effects of the velocity ratio and branch pipe diameter are evaluated to consider each mechanism separately. Fig. 7 shows velocity ratio effects on the fluctuation in the 21 mm branch pipe diameters which all categorized with regard to their type of jets. As can be seen in this figure the Phenix reactor Tee junction is located in the re-attached region with a bend by a curving ratio of 1.0.

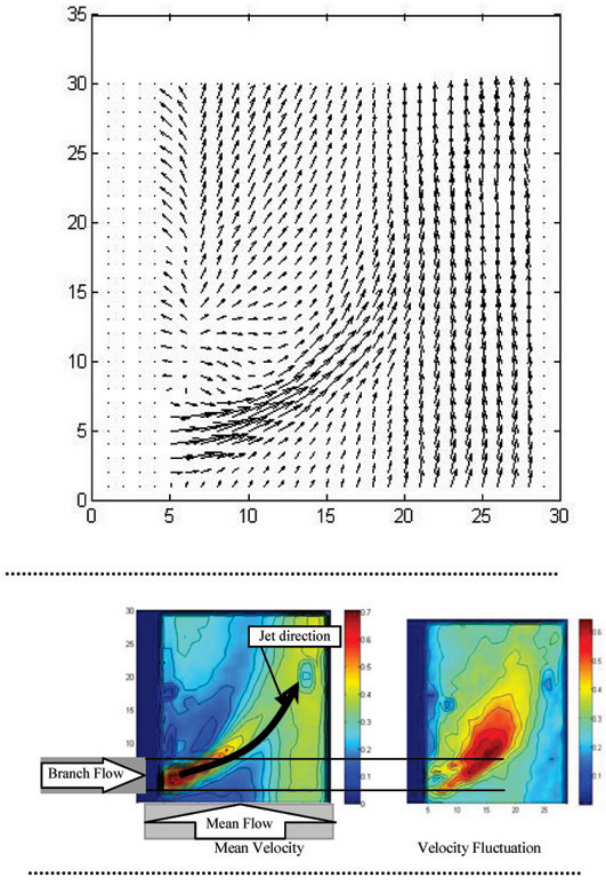


Fig. 5. Impinging jet (Dm = 108 mm, Db = 21 mm, Um = 0.3 m/s, Ub = 0.9 m/s).

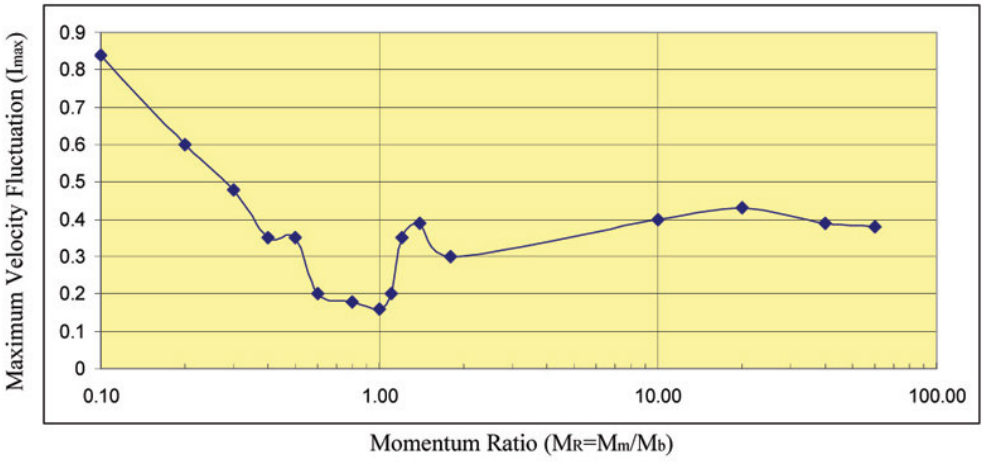


Fig. 6. Maximum intensity of the velocity fluctuation for curving ratio 1.41.

The result shows that for this configuration, by reducing M_R lower than 9, the jet type re-attached reform to the turn jet type. As it is known from equation (7), by reducing the velocity of the main pipe or increasing the velocity of the branch pipe, the lower value of M_R could be achieved. By

changing the bending ratio from 1.41 to 1.0 the results show that most of the data are in the turn jet region. Therefore with the sharpened bend, the re-attached region is compressed.
Since the experimental model is derived from the Phenix reactor structure design, the operating condition of Phenix is compared with other experimental conditions. It is clear that by decreasing the momentum ratio in the T-junction area of the Phenix reactor, the velocity fluctuation decreases. Two parameters have additional effects on the momentum ratio, the aforementioned pipe diameter and velocity.
The momentum is dependent on the diameter, as the following equations show:

$$M_b = C.\frac{1}{D_b^2} \qquad (8)$$

$$C = \frac{4Q_b^2}{\rho_b.\pi} \quad \& \quad Q_m \text{ and } Q_b = const. \qquad (9)$$

Q_m and Q_b are the constant flow rates in the main and branch pipes, respectively, and ρ_b is the branch flow density. It is important to know that decreasing the branch pipe diameter is helpful until the jet transfers to the optimum condition area. After that, the pipe diameter should no longer change.

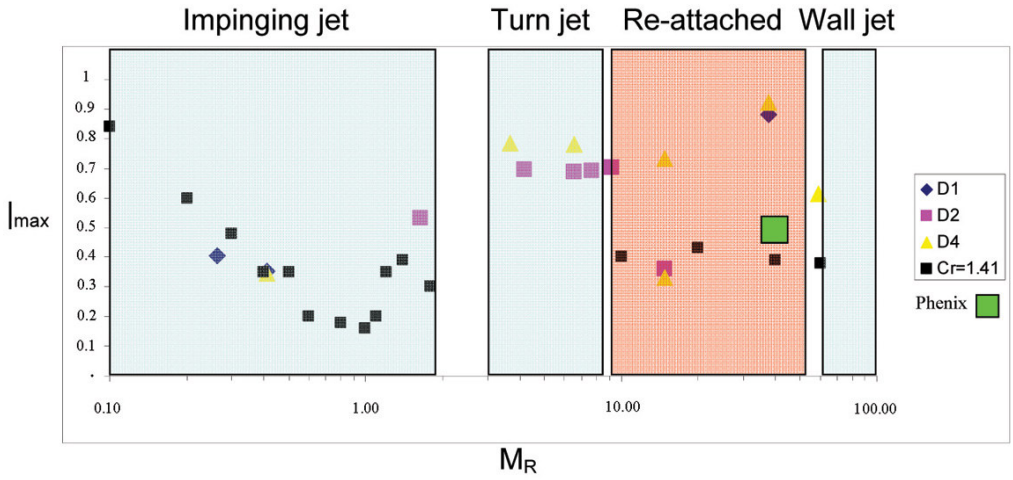


Fig. 7 shows velocity ratio effects on the fluctuation

3.3. Conclusion

The branch jet in the T-junction area acts as a turbulent jet in finite space. Various types of jets exist, dependent on the pipe's geometries and the physical properties of the working fluid. Momentum ratios between the main flow and the branch flow were selected to categorize different groups of branch jets based on the mechanisms and structures of each group. Finally, four groups of turbulent jets with different operating conditions in the T-junction area were introduced, such as the wall jet, re-attached jet, turn jet and impinging jet. The main results of the fluid mixing phenomena are listed as follows:

- By changing the bending ratio from 1.41 to 1.0 the results show that most of data are in the turn jet region. Therefore, with the sharpened bend, the re-attached region is compressed.
- The momentum ratios of the turn jet have the lowest velocity fluctuation.

- The flow rate of the turn jet gradually transfers to the flow rate of the re-attached jet with higher velocity fluctuations by increasing the momentum ratio.
- The flow rate of the turn jet sharply transfers to the flow rate of the impinging jet with a much higher velocity fluctuation, by decreasing the momentum ratio.
- The velocity and temperature fluctuations near the wall decrease by increasing the velocity ratio with a constant flow rate.
- By connecting the branch pipe closer to the bend, the optimum operating conditions transfers to the lower momentum ratio and becomes wider.
- The secondary flow of bend has strong effects on the fluid mixing in the T-junction and pushes the jet into a main pipe easily, which makes it act as under a higher momentum ratio without changing the flow rate.

Since many different types of T-junctions are used in nuclear power plants the temperature fluctuation in the mixing tee, which causes high cycle thermal fatigue, can be decreased more than 50% by changing simple geometries of the piping system, such as the branch pipe diameter, the distance between the bend and the branch pipe, the curvature ratio of the bend, etc. The present research is useful to assess the future design of piping systems for prolonged operation.

4. REFERENCES

[1] Faidy C., “EPRI-US NRC-OECD NEA”. In Third International Conference on Fatigue of reactor Components, Seville, Spain, 2004.

[2]Igarashi M. “Study on Fluid Mixing Phenomena for Evaluation of Thermal Striping in A Mixing Tee”. In 10th International Topical Meeting on Nuclear Reactor Thermal Hydraulic, NURETH-10, Seoul, Korea, 2003.

[3] Hosseini S.M. “ Experimental investigation of thermal-hydraulic characteristics at a mixing tee, in”.In International Heat Transfer Conference, FCV-17, Sydney, Australia, 2006.

[4] Hosseini S.M. “ The three-dimensional study of flow mixing phenomenon”. In International Conference Nuclear Energy for New Europe, ID: 037, Bled, Slovenia, September 2005.

[5] Hosseini S.M. “ Visualization of fluid mixing phenomenon”. Sixth International Congress on Advances in Nuclear Power Plants, ICAPP05, ID: 5006 (R006), Seoul, Korea, May 2005.

[6] Yuki K. “ vol. 2, 2001, pp. 1573–1578”. In Proceedings of the Fifth World Conference on Experimental Heat Transfer Fluid Mechanics and Thermodynamics, Japan, 2004.

[7] Yuki K. “ID: N6P082”.In Proceedings of the 15th International Conference on Nuclear Thermal Hydraulic, Operation and Safety (NUTHOS-6), Nara, Japan, 2004.

[8] H.C. Kao. “Some aspects of bifurcation structure of laminar flow in curved ducts”, J. Fluid Mech., pp. 519–539(1992).

[9] Carmine Golia, Bernardo Buonomo, Antonio Viviani, “Grid Free Lagrangian Blobs Vortex Method with Brinkman Layer Domain Embedding Approach for Heterogeneous Unsteady Thermo Fluid Dynamics Problems”. International Journal of Engineering (IJE), Volume 3, Issue 3, 313-329, May/June 2009.

System Level Power Management for Embedded RTOS: An Object Oriented Approach

Ankur Agarwal ankur@cse.fau.edu
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431, USA

Eduardo Fernandez ed@cse.fau.edu
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431, USA

Abstract

Power management systems for embedded devices can be developed in real-time operating system (RTOS) or in applications. If power management policies are applied in operating system (OS), then designers and developers will not have to worry about complex power management algorithms and techniques. They can rather concentrate on application development. The OS contains specific and accurate information about the various tasks being executed. An RTOS further has a comprehensive set of power management application programming interfaces (APIs) for both device drivers and applications within a power management component. Therefore, it is logical to place policies and algorithms in the OS that can place components not being used into lower power states. This can significantly reduce the system energy consumption. We present here an abstract model of a system power manager (PM), device power managers, and application power managers. We present relationship and interactions of these managers with each other using Unified Modeling Language (UML) class diagrams, sequence diagrams and state charts. We recommend that the PM must be implemented at the OS level in any embedded device. We also recommend the interfaces for interactions between PM and the devices power manager, as well as PM and application power manager. Device driver and application developers can easily use this object oriented approach to make the embedded system more power efficient, easy to maintain, and faster to develop.

Keywords: *Embedded Device, Object Oriented Design, Policy Manager, Real-Time Operating Systems, System Level Power management*

1. INTRODUCTION

Due to the nature of the use and blend of computationally extensive applications, power consumption is one of the major concerns in developing real-time OS devices [1, 2, 3]. There is always a need for longer battery life in order to avoid catastrophic data loss [4, 5]. The environmental impact of power consumption from electronic systems has raised serious concern [6]. Furthermore, excessive heat dissipation is a major obstacle in improving performance. Power management is widely employed to contain the energy consumption in power-constrained devices. The advancement in processor and display technology has far outpaced similar advancements in battery technology [4]. On the other hand, the battery capacity has

improved very slowly (a factor of two to four over the last 30 years), while the complexity of applications, the computational demands, and therefore the power needs have drastically increased over the same time frame. Power management techniques date back to 1989, when Intel shipped processors with the technology to allow the CPU to slow down, suspend, or shut down part or all of the system platform, or even the CPU itself, to preserve and extend the time between battery charges [7]. Since then, several power consumption strategies have been developed [8, 9, 10, 11, 12, 13, 14].

Figure 1 shows the embedded processor performance and power consumption evolution for the data in Table 1. It should be noted that higher MHz/mW implies higher efficiency. Thus, power management techniques have evolved along time. These techniques have been able to reduce (or keep constant) power consumption in portable devices.

Table 1: Embedded Processor performance and Power Consumption Data [15]

Processor	Intel StrongARM SA-1110	Intel Xscale PXA-250	Intel Xscale PXA-250	Intel Xscale PXA-250
Clock Speed (MHz)	206	200	300	400
Power Draw (mW)	800	256	411	563
MHz/mW	0.26	0.78	0.73	0.71

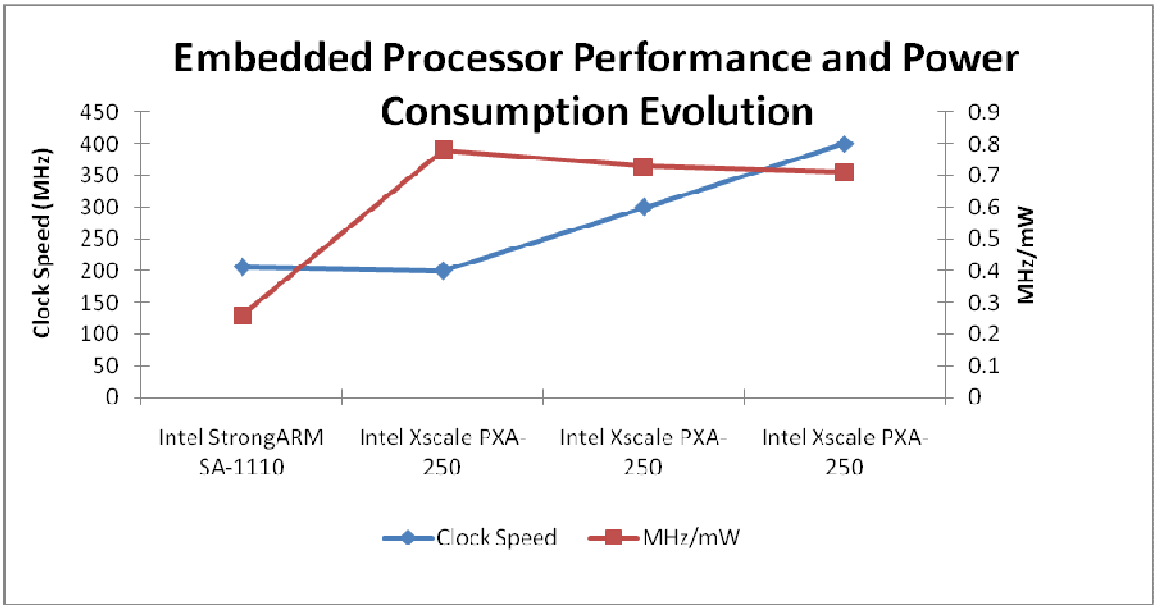


Figure 1: Embedded Processor performance and Power Consumption Evolution [15]

Power management has been a center of focus since the early 90's. Power management policies have been described at different levels of abstraction starting from the lowest level of abstraction: Transistor Level. Figure 2 shows the use of power management techniques at various levels of abstractions.

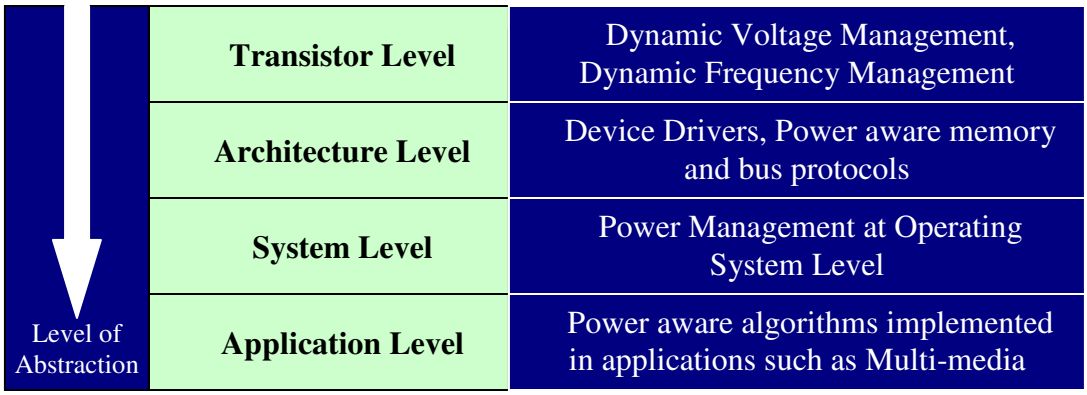


Figure 2: Power Management Schemes at Different Level of Abstractions

At the transistor, level dynamic voltage scaling (DVS) and dynamic frequency scaling have been exploited to reduce the power consumption in circuits at this level [16, 17]. Extensive research has been done on lower-levels such as transistor level, gate level and application level power management in the past few decades [18, 19]. Researchers have developed innovative bus protocols [20, 21], memory optimization algorithms [22, 23], and advancement in cache technology [24, 25] for managing power. Power management techniques in device drivers have resulted in power management techniques at architect's level of abstraction [26, 27, 28]. A system architect performs system level trade-offs by selecting an optimum architecture that satisfies power, area, cost, latency and other quality-of-service (QoS) and performance parameters [29, 30, 31]. Similarly, power management techniques have evolved at the application level [32, 33]. An application is usually given full control to choose the value of CPU power setting and parameters. The OS is then responsible for keeping track of all the power settings of different applications and applying them on different resources such as processors.

As per the Moore's law, the number of transistors on a chip doubles every eighteen months [34]. This has resulted in exponential increase in the complexity of embedded systems. Therefore, power control at lower levels, even though it is more accurate, becomes unfeasibly complex and is compounded by time-to-market pressures. Thus, with time the designs are now being described at higher levels of abstraction leading to the era of system level design, system-on-chip and networks-on-chip [30, 35]. Therefore, there is a strong need for specifying the power issues related to the design at the same level of abstraction, leading to the concept of system level power management. Moreover the application, semiconductor technology, cost, and time-to- market trends are causing a shift towards increased software content in embedded system and systems-on-chip. As a result, designers and users of embedded software must be increasingly aware of power issues. While power dissipation is inherently a property of the underlying system hardware, knowledge of embedded software that runs on the hardware is useful in order to analyze and improve system's power consumption characteristics. Modern OS not only contain precise information about the various tasks being executed but are also well developed with algorithms, that selectively place components into lower power states, thereby drastically reducing energy consumption [1]. However, the importance of reducing the power consumption in embedded OS has not been widely recognized and a large body of work has focused on estimating, managing, and reducing power consumption in various system components. An RTOS serves as an interface between the application software and the hardware. The embedded system design and its issues such as hardware resource management, memory management, process management and development of device drivers can be simplified by providing the designers with a well defined interface. With more features being supported by embedded systems, the applications and their development is becoming complex everyday. RTOS must provide a simple and encapsulated Application Programming Interface (API) so that the software remains portable across product users, product families and companies [36].

The motivation behind this paper is the need for the applications to provide well-defined interfaces between RTOS and device devices that can be used by the power manager to manage the overall power of the system. This paper also aims at providing a simple programming interface for the application developers to inform RTOS about applications' power and device requirements. Conversely, we also

recognize the need for RTOS to inform the application about the current battery status so that the application can keep user informed. Once the RTOS is aware of the power requirements, it should be able to bring the complete system into a lowest possible power state. For a simplified and well-encapsulated design, we provide an object-oriented representation for the power manager components that are embedded in the RTOS, device drivers and applications. We present encapsulated behavior of power management features in the form of classes and their interaction in the form of sequence diagrams. A proper graphical representation of these complex power management processes can give the user a capability to manage the complexity, tight performance and power constraints in the system. These features can then be used by the different application and device vendors to evolve test cases for verifying the compatibility among the various devices and the applications being used with this OS in their prototype.

Section 2 provides typical software architecture of power manager in an RTOS. Section 3 discusses the power states in the RTOS. Section 4 describes the proposed model for power management in RTOS. Section 5 concludes this research.

2. SOFTWARE ARCHITECTURE of POWER MANAGER in RTOS

Figure 3 shows the layered software architecture of an RTOS. The top layer consists of user defined applications and OS supported applications. OS supported application include user interfaces and client-services. Embedded system designers and developers design their applications and link them with some dynamic link library (Core DLL library). Multimedia, networking, and other applications, along with device drivers are supported at the OS layer. Below this layer there is OS kernel that handles the major OS functionality such as scheduling, task management, memory management, I/O management among other. Under this layer, there is a original equipment manufacture (OEM) layer.

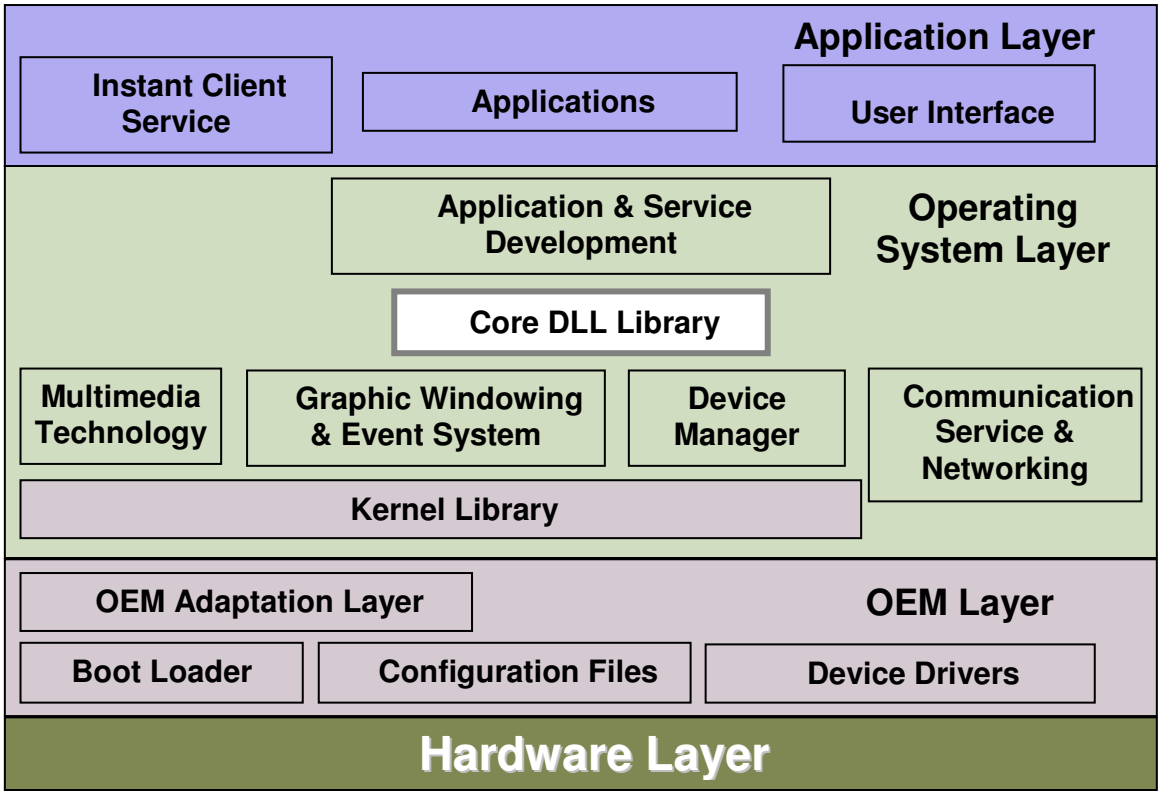


Figure 3: Layered Software Architecture of OS

OEM layer usually supports all the OEM hardware and contain device drivers for various hardware devices originally supported by OS. This layer also contains boot loader and configuration manager which loads OS every time we start this system. This part of the OS is write-protected and application/device/embedded system developers are not allowed to alter the content of this layer.

We have abstracted power management functionality of an OS in Figure 4. An integral part of OS-level power management is the policy manager. Policy Manager is responsible for the efficient management of the various applications and devices running on the portable system. This is one typical scenario of OS and is very similar to WindowsCE.NET. Other RTOS may or may not follow the similar architecture. Figure 4 shows the interaction of devices, applications and battery with the power manager (PM). The PM acts as a mediator between devices, applications, and the processor. Information from the various interfaces about their power status is then collected and managed by PM. On the basis of the collected information, the entire system is put into the lowest possible power state for a particular application.

The PM also co-ordinates the different devices, system and processor states to decrease the overall system power consumed. While maintaining critical resources in the system and monitoring processor utilization to ensure its operation at the lowest possible state, it also provides a means for the drivers to inter-communicate about their power states. The Software architecture of the PM is responsible for providing the services to the device drivers of the system, to get notified, and respond upon power changes.

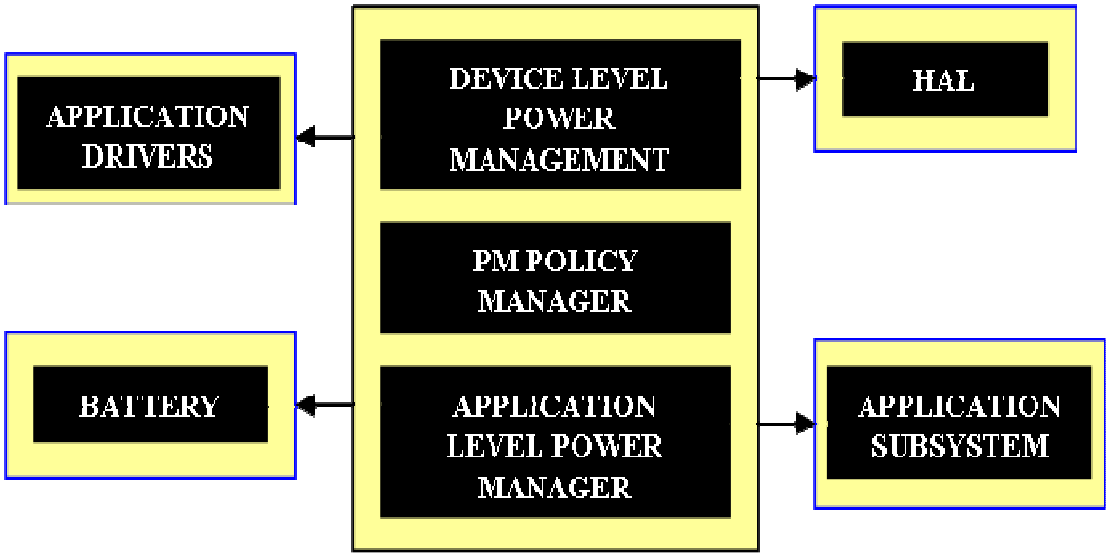


Figure 4: Interaction of power manager with devices and applications

Power management with power manageable hardware comprises one or more layers of software. Hardware specific power management software and operating system policy manager, which are in between the hardware independent software interface, are also defined. This creates a layered cooperative environment through software interfaces, and allows applications, operating systems, device drivers and the PM to work together, thereby reducing the power consumed. The higher-level application software is therefore able to use PM without any inkling of the hardware interface as the details of the hardware are masked by the PM. This leads to increased usage due to extended system battery life.

3. POWER STATES

In any system, under certain circumstances some hardware components are always in an idle state. This applies to the devices, to the processor and to the applications. Under these conditions, when no task is

being executed, the particular device or application can be put into a lower power state. In an RTOS, there are different power states associated with applications, devices and the processor. Thus, it is possible that a processor may be in sleeping state, a device in soft off and an application in full-on state. Figure 5 shows the state transitions for a complete system. Power managed devices receive power state change notifications in the form of I/O control codes (IOCTLs). This diagram separates device power states from the system power states.

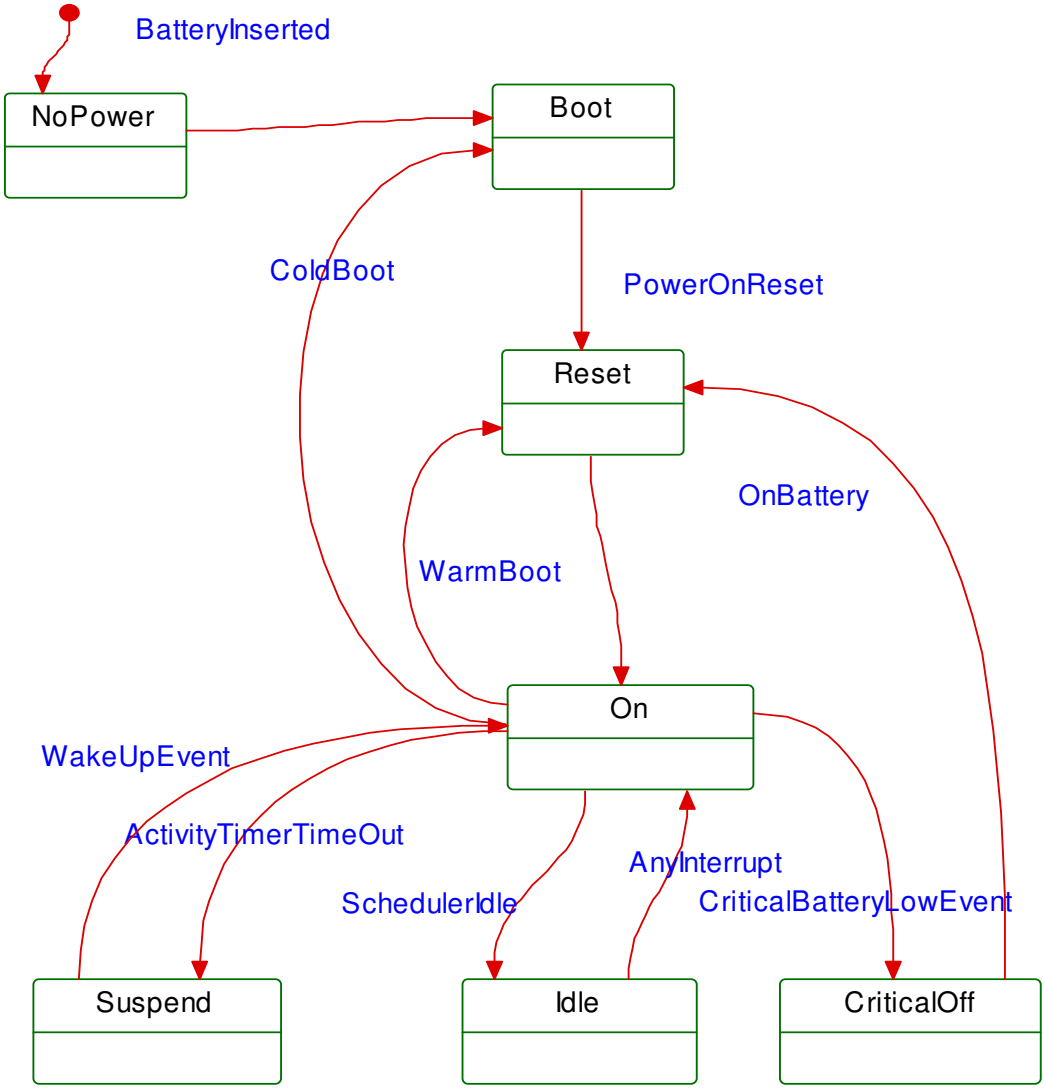


Figure 5: System Power States

There are seven predefined power states in a typical system. 1) The system is said to be in the “NoPower” state (S0), when the system has no power. 2) Upon the insertion of the battery the system moves to “Boot” state (S1). 3) The “On” state (S2) used for the normal operation, is a state in which the system dispatches user mode (application) threads for execution. Dynamic Frequency Management (DFM) and Dynamic Voltage Management (DVM) optimizations are done in this state. 4) The “Idle” state (S3) is a low-wake-latency sleep state. In this state, system context is maintained by the hardware while there is no loss of system context in the CPU or peripheral devices. Upon detection of some events such as user activity, the system moves back to the On state. 5) The “Suspend” state (S4), is a low-wake-latency sleep state where all system contexts are lost except system memory. 6) The “Critical Off” State

(S5) is a non-volatile sleep state. In this state system context is saved and restored when needed. Here the operating system saves the necessary information in the non-volatile memory and tags the corresponding context markers. 7) Finally, the "Reset" state (S6)is also referred to as soft restart state. In this state the system contexts are properly saved before being lost. All the data and other user information are also stored in memory before the system restarts itself. These system states are shown in Figure 5. This state diagram can also be considered as the state transition diagram of the active class [37] Policy Manager discussed in the next section.

A system may be in any of the above-explained seven states depending upon the task, which is being executed on the device. However, transition among states can consume some time depending on the thread executed by the PM.

The PM makes state transition decisions according to same power management policy, which is discussed in the next section. The power states of devices and processor are based on the same structure as that of the system. However, different devices may be in different power states, while the system and processor are in another state. For instance, a device can be in the Off state while the system is in the On state.

4. A MODEL for POWER MANAGEMENT

We have developed a UML model for power management features. The model uses three Manager (controller) classes that coordinate the basic power related functions.

4.1. Policy Manager

The Policy Manager constantly monitors battery state. It also orchestrates all system-level and device-level state changes. These notifications are passed to applications polymorphically using the interface specified in the ApplicationDriverPowerManager (APM) and to the devices using the interface specified in DeviceDriverPowerManager (DDPM). On detecting a low battery state, the Policy Manager decides to force the system into the Idle state. It sends a notification to the APM and DDPM, which in turn notifies all application and drivers registered with it respectively.

Figure 6 depicts the class diagram of power management in an RTOS. It can be seen from the figure that Policy Manager and Battery are singleton class whereas the DDPM and APM are abstract class associated with the Policy Manager. The relationships shown are used to notify drivers and applications about the various system level state changes. For example, when a new application is plugged in, the drivers are notified through the policy manager notification interface regarding the different system level state changes that may occur. The "BootInitilizer" class is for loading important application and drivers while the system is booted.

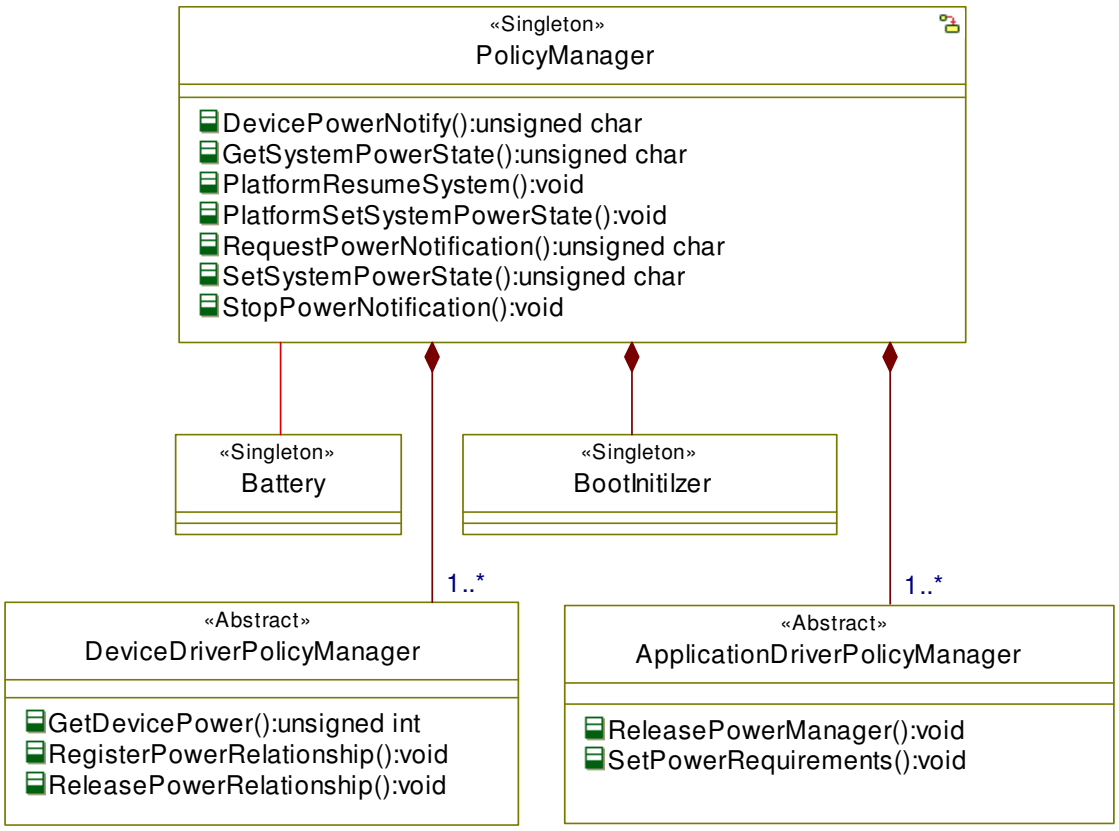


Figure 6: UML Class Diagram of Power Management

4.2. Device Driver Power Manager

Figure 7 depicts several possible concrete subclasses of the DeviceDriverPolicyManager, each associated with a specific device driver such as camera, keyboard, display, headset among others. These devices (referred as subclasses in UML) are presented here just as examples.

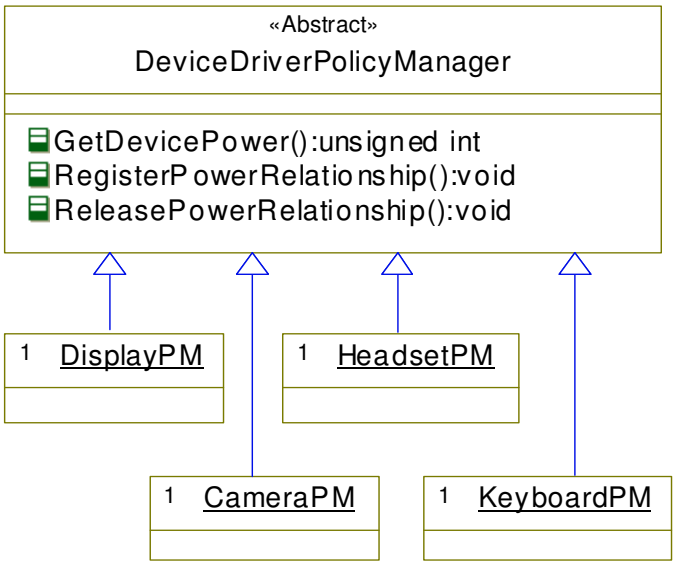


Figure 7: UML Class Diagram of DDPM

In a specific embedded system, some of these subclasses may not be present, but these subclasses may instead have other subclasses contained in them. Various applications, services and device drivers are notified upon the (dis)appearance of device interfaces by the DDPM Interface notification. This feature of RTOS can be regarded as similar to the plug and play of Windows OS. Using the device policy manager's interface, the PM can receive and set specific capabilities of the device driver. However, to be compatible with this power management framework, the device driver must support all the power management states. The DDPM's interaction with the Policy Manager is shown in the sequence diagram of Figure 8.

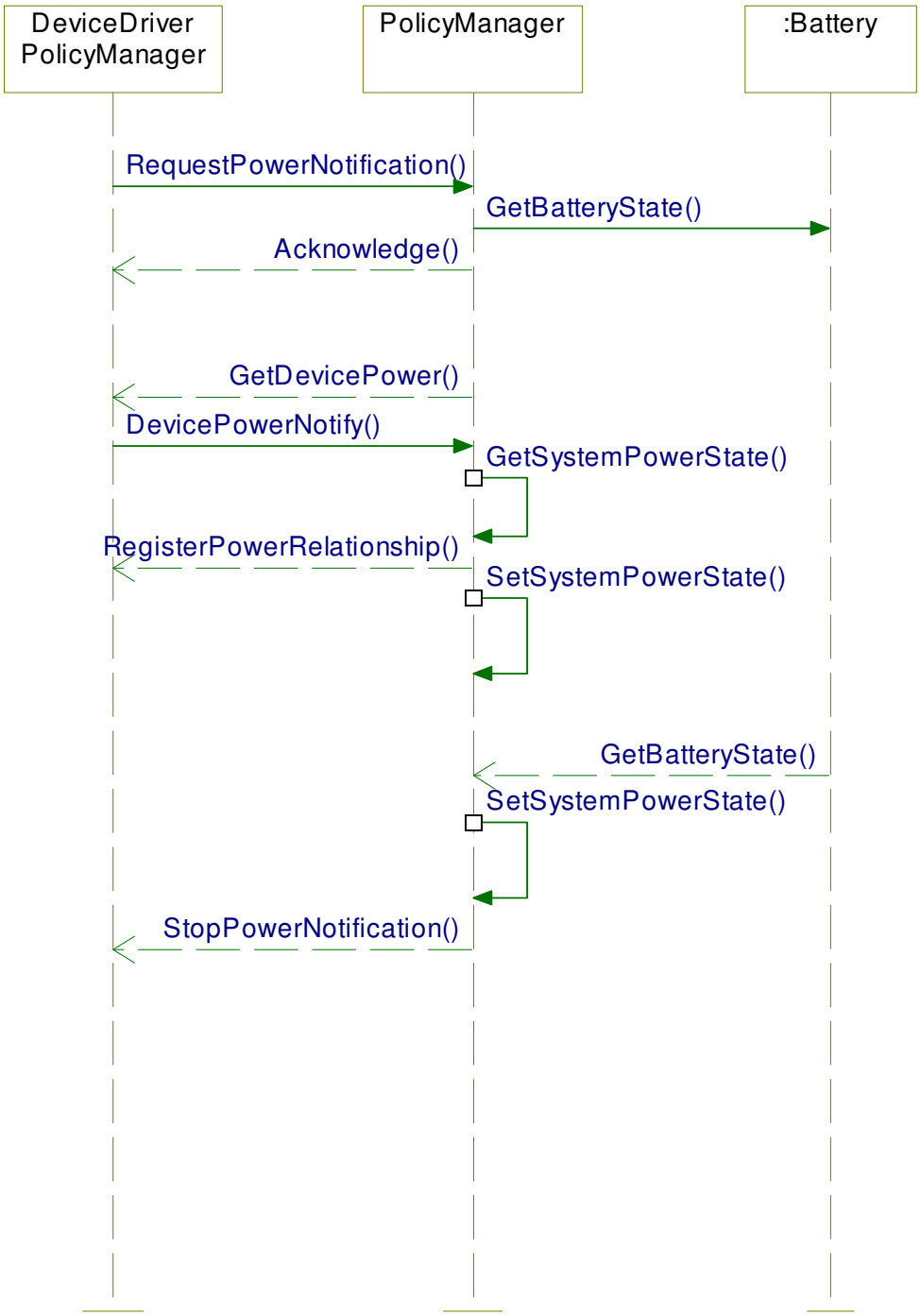


Figure 8: Sequence Diagram of Device Driver Interaction with the Power Manager

Initially, all the device drivers (DeviceDriverPowerManager) register themselves with the Policy Manager through RequestPowerNotification() and receive an acknowledgement. The PM reads a list of device classes from the registry and uses RequestPowerNotifications() to determine when devices of that class are loaded. In order to for a device to get activated in the system, the device finds out its current power state by GetDevicePower() and then notifies the policy manager to change its state. DevicePowerNotifiy() informs the device about the change in its power state. Once the power state of the device has been changed the system changes its previous power state to a new power state in order to accommodate the change in the power state of the device. The Policy Manager constantly monitors battery status. Upon detection of an idle state, the Policy Manager sends a query, to transfer into idle state. Upon the acceptance of the query by the device, the Policy Manager changes the system state. Concurrently, if the Policy Manager detects a low battery state, it notifies the DeviceDriverPolicyManager, to change the device state of all devices to idle. As the device state of the device changes to idle state, an acknowledgement is sent to the Policy Manager, which puts the system state to idle.

4.3. Application Driver Power Manager

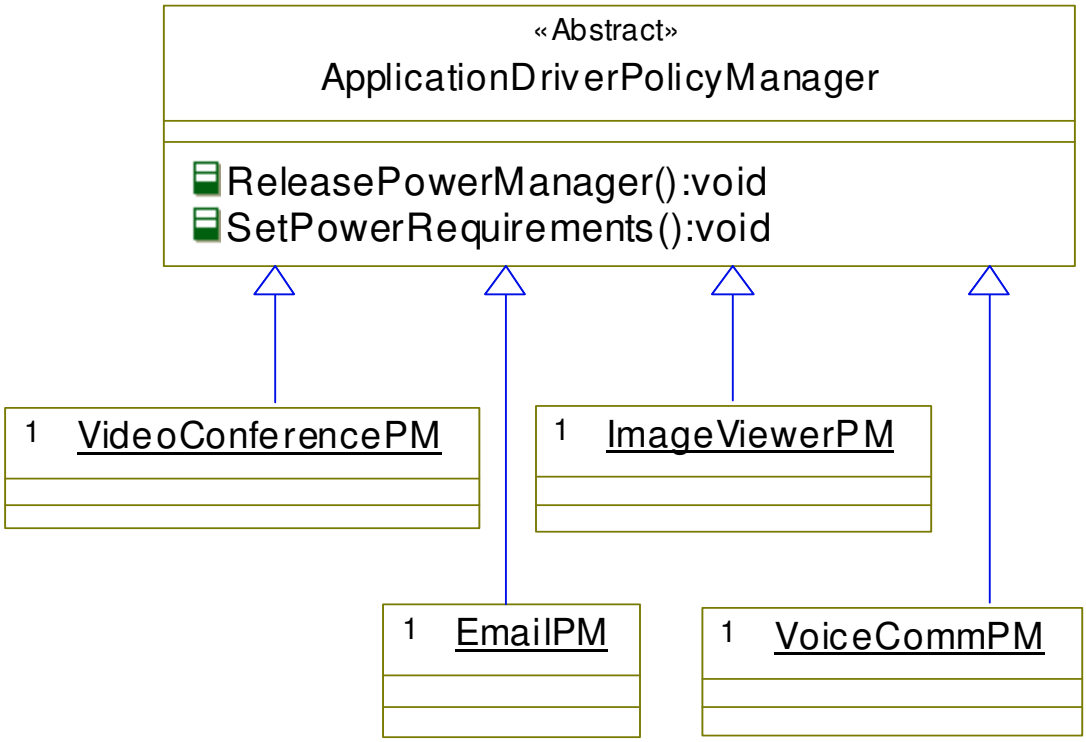


Figure 9: Class Diagram APM

Class diagram for the APM is shown in Figure 9. This Figure depicts some applications such as video conferencing, voice communication, email, audio, and multimedia among others, associated with APM class. Various system level power changes are notified to the applications using the interface specified by APM which in turn is specified by PM. Specific guidelines are set for devices, which applications can request to set system power level by means of DDPM. On specific devices or systems as a whole, a reduced power consuming state can be set if the applications request the power manager to change the device's power state. High-performance "power-smart" applications can also use battery status data provided by PM to provide the best experience for the user by lowering down performance (e.g., lowering frame-rates) in order to preserve battery.

Figure 10 shows the interaction of the APM with the PM and battery as a sequence diagram. Initially the applications register themselves (RequestPowerNotification()) with the PM and receive an acknowledgement. APM notifies the Policy Manager that an application has a specific device power requirement and sets is accordingly (SetPowerRequirement()). The application also requests the power notification for the specific device drivers it needs in order to execute. The system responds to its request by changing the power state of those device drivers (SetDevicePower()). Once the application power requirements are fulfilled the policy manager updates the system power state (GetSystemPowerState(), SetSystemPowerStstate()). The policy manager constantly monitors the battery status. Upon the detection of an idle state event, the APM gets the system power state (GetSystemPowerState()) and device power state (GetDevicePowerState()). Power states of drivers are then forced into sleep state by the Policy Manager. Concurrently, if the Policy Manager detects a low battery state, it notifies the ApplicationDriverPolicyManager, to change the application state of all applications to standby followed by the change in the system state to standby state.

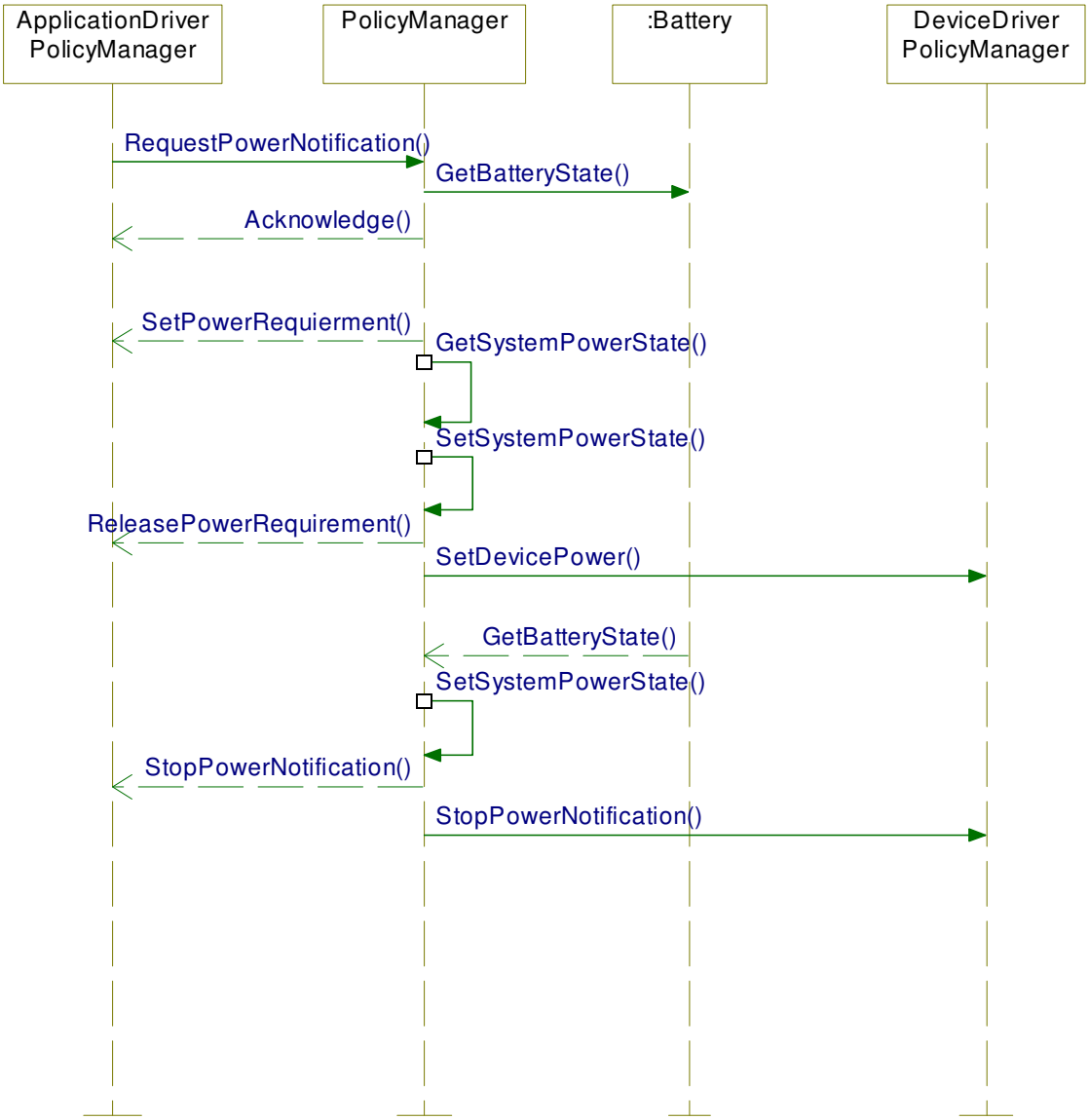


Figure 10: Sequence Diagram of Application Driver Power Manager Interaction with Power Manager

5. CONCLUSION

We have provided an abstract model for implementing power management in an embedded RTOS. System level power management can be designed using an OO approaches We recommend the interfaces for interactions between PM and the devices power manager, as well as PM and application power manager. We present relationships and interactions of these managers with each using UML class diagrams, sequence diagrams, and state charts. This abstract object-oriented representation of power management clarifies the operation of the power manager in conjunction with applications, devices and the processors for the developers of applications and devices. If the operating system and the device drivers are designed using this framework, the task of managing power within applications can be significantly simplified resulting into longer battery life at run time. On the other hand, well defined interfaces also simplify the task of power management in device drivers so that a more granular power management strategy can be used, thus lowering overall power needs of the system. The object-oriented power management interface also simplifies testing for power management scenarios, thus making it possible to test for cases that were not tested earlier. It is our hope that the proposed framework will foster development of embedded systems that are more power efficient, easy to maintain, and faster to develop.

6. REFERENCES

[1] A. Agarwal, S. Rajput, A. S. Pandya, "*Power management System for Embedded RTOS: An Object Oriented Approach*", IEEE Canadian Conference on Electrical and Computer Engineering, May 2006, pp. 2305-2309

[2] Y. H. Lu, L. Benini, G. D. Micheli, "*Power aware operating systems for interactive systems*", IEEE Transactions on Very Large Scale Integration Systems, Volume 10, Issue 2, April 2002, pp. 119-134

[3] S. Vaddagiri, A.K. Santhanam, V. Sukthankar, and M. Iyer, "*Power management in linux-based systems*," Linux Journal, March 2004, http://www.linuxjournal.com/article.php?sid=6699.

[4] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, H. Franke, "*Reducing disk power consumption in servers with drpm*" IEEE Computer, Volume 36, Issue 12, December 2003, pp. 59-66

[5] "*Adaptive Power Management for Mobile Hard Drives*", IBM, April 1999, http://www.almaden.ibm.com/almaden/pbwhitepaper.pdf

[6] H.S. Choi, D. Y. Huh, "*Power Electronics Specialists Conference 2005*", 36th IEEE Power Electronics Specialists Conference, 2005, pp. 2817-2822

[7] L. Benini, R. Hodgson., P. Siegel, "*System-level power estimation and optimization*", International Symposium on Low Power Electronics and Design, IEEE Conference, pp. 173-178, August 1998.

[8] L.S. Brakmo, D.A. Wallach, M.A. Viredaz, "*uSleep: A technique for reducing energy consumption in handheld device*", 2nd IEEE International Conference Proceeding on Mobile Systems Applications, and Services, June 2004.

[9] Y.-H. Lu, G. D. Micheli, "*Comparing system-level power management policies*", IEEE Design & Test of Computers special issue on Dynamic Power Management of Electronic Systems, pages 10–19, March/April 2001.

[10] Sinha A., Chandrakasan A. P., "*Energy efficient real-time scheduling [Microprocessors]*", *IEEE International Conference of Computer Aided Design*, pp. 458-463, November 2001.

[11] Y.-H. Lu, E.-Y. Chung, T. ˇSimuni´c, L. Benini, G. D. Micheli, "*Quantitative comparison of power management algorithms*", IEEE Conference on Design Automation and Test in Europe, March 2000, pp. 20-26.

[12] Q. Qiu, Q. Wu, M. Pedram, "*Dynamic power management in mobile multimedia system with guaranteed quality-of-service*", IEEE Design Automation Conference, Vol. 49, pp 834-849. June 2001.

[13] Udani, S., Smith, J., "*Power management in mobile computing*", Department of Computer Information Sciences, Technical Report, University of Pennsylvania, February 1998.

[14] Yung-Hsiang Lu and Giovanni De Micheli, "*Adaptive hard disk power management on personal computers*", Great Lakes Symposium on VLSI, 1999, pp. 50–53.

[15] "Embedded Processor Performance Parameter Data-Sheet"www.intel.com

[16] Burd T. D., Brodersen R. W., "*Energy efficient CMOS microprocessor design*" In Proceedings of the 28th Annual Hawaii International Conference on System Sciences. Volume 1: Architecture, IEEE Computer Society Press, 1995, pp. 288-297.
[17] Zimmermann, R., and Fichtner, W., "*Low power logic styles: CMOS versus pass-transistor logic*", IEEE Journal of Solid State Circuits, Vol. 32, pp. 1079-1089.
[18] A. Weissel, F. Bellosa, "*Process cruise control: event-driven clock scaling for dynamic power management*", IEEE Proceedings of the International Conference on Compilers, Architecture, and Synthesis for Embedded Systems, October 2002.
[19] D. Grunwald, P. Levis, C. Morrey III, M. Neufeld, K. Farkas, "*Policies for dynamic clock scheduling*", Symposium on Operating Systems Design and Implementation, October 2000, pp 78-86
[20] K. Lahiri, A. Raghunathan, "*Power analysis of system-level on-chip communication architectures*", IEEE International Conference on Hardware/Software Codesign and System Synthesis (CODES + ISSS), 2004, pp. 236-241.
[21] Y. J. Lu, A. C. W. Wai, L. L. Wei Fan, B. K. Lok, P. Hyunjeong, K. Joungho, "*Hybrid analytical modeling method for split power bus in multilayer package*", IEEE Transactions on Electromagnetic Compatibility, Volume 48, Issue 1, February 2006, pp. 82-94.
[22] C. Le, P. Nathaniel, L. H. Yung, "*Joint power management of memory and disk under performance constraints*", IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems, Volume 25, issue 12, December 2006, pp. 2697-2711.
[23] M. Fukashi, H. Isamu, G. Takayuki, N. Hideyuki, I. Takashi, S. Hiroki, D. Katsumi, A. Kazutami, "*A configurable enhanced TTRAM macro for system-level power management unified memory*", IEEE Journal of Solid-State Circuits, Volume 42, Issue 4, April 2007, pp. 852-861.
[24] Q. Zhu, Y. Zhou, "*Power aware storage cache management*", IEEE Transactions on Computers, Volume 54, issue 5, May 2005, pp. 587-602.
[25] A. Sagahyroon, M. Karunaratne, "*Impact of cache optimization techniques on energy management*", IEEE Canadian Conference on Electrical and Computer Engineering, Volume 4, May 2004, pp. 1831-1833
[26] L. Benini, S. K. Shuklam, R. K. Gupta, "*Tutorial: architectural system level and protocol level techniques for power optimization for networked embedded systems*", 18th International Conference on VLSI Design, January 2005, pp. 18
[27] C. Grelu, N. Baboux, .A. Bianchi, C. Plossu, "*Low switching losses devices architectures for power management applications integrated in a low cost 0.13/spl mu/m CMOS technology*", 35th IEEE Proceedings of Solid-State Device research Conference, September 2005, pp 477-480
[28] J. Byoun, P. L. Chapman, "*Central power management unit as portable power management architecture based on true digital control*", IEEE Workshop on Computers in Power Electronics, August 2004, pp. 69-73
[29] L. Benini, A. Bogliolo, G. D. Micheli, "*A survey of design techniques for system-level dynamic power management*", IEEE Transactions on very large Scale Integration Systems, Volume 8, Issue 3, June 2000, pp. 299-316
[30] T. Simunic, S. P. Boyd, P. Glynn, "*Managing power consumption in network on chips*", IEEE Transactions on Very large Scale Systems, Volume 12, Issue 1, January 2004, pp. 96-107
[31] D. Monticelli, "*System approaches to power management*", 17th Annual IEEE Conference and Exposition on Applied Power Electronics, Volume 1, March 2002, pp. 3-7
[32] Y. H. Lu, T. Simuni´c, G. D. Micheli, "*Software controlled power management*", IEEE International Workshop on Hardware/Software Codesign, 1999, pp. 157–161.
[33] R. M. Passos, C. J. N. Coelho, A. A. F. Loureiro, R. A. F Minim, "*Dynamic power management in wireless sensor networks: an application-driven approach*", 2nd Annual IEEE Conference on Wireless on-demand Network Systems and Services, January 2005, pp. 109-118
[34] "Roadmap Architects – The Technology Working Groups" http://public.itrs.net
[35] S. Kumar, A. Jantsch, J-P. Soininen, M. Forsell, M. Millberg, J. Oberg, K. Tiensyrja, and A. Hemani, "*A network on chip architecture and design methodology*", IEEE Computer Society Annual symposium on VLSI, April 2002, 117-124
[36] Martin Timmerman, "RTOS Evaluations", 2000, http://www.dedicated-systems.com
[37] Jenson Douglas E., "*Real-time design pattern robust scalable architecture for real time systems*", Boston, Addition-Wesley, 2002.

Hybrid Genetic Algorithm for Multicriteria Scheduling with Sequence Dependent Set up Time

Ashwani Dhingra *ashwani_dhingra1979@rediff.com*
Department of Mechanical Engineering
University Institute of Engineering & Technology
Maharshi Dayanand University
Rohtak, 124001, Haryana INDIA

Pankaj Chandna *pchandna08@gmail.com*
Department of Mechanical Engineering
National Institute of Technology
Kurukshetra, 136119, Haryana INDIA

Abstract

In this work, multicriteria decision making objective for flow shop scheduling with sequence dependent set up time and due dates have been developed. Multicriteria decision making objective includes total tardiness, total earliness and makespan simultaneously which is very effective decision making for scheduling jobs in modern manufacturing environment. As problem of flow shop scheduling is NP hard and to solve this in a reasonable time, four Special heuristics (SH) based Hybrid Genetic Algorithm (HGA) have also been developed for proposed multicriteria objective function. A computational analysis upto 200 jobs and 20 machines problems has been conducted to evaluate the performance of four HGA's. The analysis showed the superiority of SH1 based HGA for small size and SH3 based HGA for large size problem for multicriteria flow shop scheduling with sequence dependent set up time and due dates.

Keywords: Flow shop scheduling, Genetic algorithm, Sequence dependent set up time, Total tardiness, Total earliness, makespan

1. INTRODUCTION

Scheduling in manufacturing systems is typically associated with allocating a set of jobs on a set of machines in order to achieve some objectives. It is a decision making process that concerns the allocation of limited resources to a set of tasks for optimizing one or more objectives. Manufacturing system is classified as job shop and flow shop and in the present work; we have dealt with flow shop. In a job shop, set of jobs is to be scheduled on a set of machines and there is no restriction of similar route on the jobs to follow, however in flow shop environment all the jobs have to follow the similar route. Classical flow shop scheduling problems are mainly concerned with completion time related objectives (e.g. flow time and makespan) and aims to reduce production time and enhance productivity and facility utilization. In modern manufacturing and operations management, on time delivery is a significant factor as for the survival in the competitive markets and effective scheduling becomes very important in order to meet customer requirements as rapidly as possible while maximizing the productivity and facility utilization. Therefore, there is a need of scheduling system which includes multicriteria decisions like makespan, total flow time, earliness, tardiness etc.

Also, flow shop scheduling problems including sequence dependent set up time (SDST) have been considered the most renowned problems in the area of scheduling. Ali Allahverdi et al [1] investigated the survey of flow shop scheduling problems with set up times and concluded that considering flow shop scheduling problems is to obtain remarkable savings when setup times are included in scheduling. Sequence dependent setup times are usually found in the condition where the multipurpose facility is available on multipurpose machine. Flow shop scheduling with sequence dependent set up time can be found in many industrial systems like textile industry, stamping plants, chemical, printing, pharmaceutical and automobile industry etc. Several researchers have considered the problem of flow shop scheduling with single criterion and very few have dealt with multicriterion decision making including sequence dependent set up.

The scheduling literature also revealed that the research on flow shop scheduling is mainly focused on bicriteria without considering sequence dependent set up time .Very few considered flow shop scheduling including sequence dependent set up time with multicriteria. Rajendran [2] has implemented a heuristic for flow shop scheduling with multiple objectives of optimizing makespan, total flow time and idle time for machines. The heuristic preference relation had been proposed which was used as the basis to restrict the search for possible improvement in the multiple objectives. Ravindran et al. [3] proposed three heuristic similar to NEH for flow shop scheduling with multiple objectives of makespan and total flow time together and concluded that proposed three heuristic yields good results than the Rajendran heuristic CR[2].Gupta et al.[4] considered flow shop scheduling problem for minimising total flow time subject to the condition that the makespan of the schedule is minimum. Sayin and Karabati [5] minimized makespan and sum of completion times simultaneously in two machines flow shop scheduling environment. Branch and bound procedure was developed that iteratively solved single objective scheduling problems until the set of efficient solutions was completely enumerated. Danneberg et al. [6] addressed the permutation flow shop scheduling problem with setup times and considered makespan as well as the weighted sum of the completion times of the jobs as objective function. For solving such a problem, they also proposed and compared various constructive and iterative algorithms. Toktas et al. [7] considered the two machine flow shop scheduling by minimizing makespan and maximum earliness simultaneously. They developed a branch & bound and a heuristic procedure that generates all efficient solutions with respect to two criteria. Ponnambalam et al. [8] proposed a TSP GA multiobjective algorithm for flow shop scheduling, where a weighted sum of multiple objectives (i.e. minimizing makespan, mean flow time and machine idle time) was used. The proposed algorithm showed superiority which when applied to benchmark problems available in the OR-Library. Loukil et al. [9] proposed multiobjective simulated annealing algorithm to tackle the multiobjective production scheduling problems. They considered seven possible objective functions (the mean weighted completion time, the mean weighted tardiness, the mean weighted earliness, the maximum completion time (makespan), the maximum tardiness, the maximum earliness, the number of tardy jobs). They claimed that the proposed multiobjective simulated annealing algorithm was able to solve any subset of seven possible objective functions. Fred Choobineh et al. [10] proposed tabu search heuristic with makespan, weighted tardiness and number of tardy jobs simultaneously including sequence dependent setup for n jobs on a single machine. They illustrated that as the problem size increases, the results provided by the proposed heuristic was optimal or nearer to optimal solutions in a reasonable time for multiobjective fitness function considered. Rahimi Vahed and Mirghorbani [11] developed multi objective particle swarm optimization for flow shop scheduling problem to minimize the weighted mean completion time and weighted mean tardiness simultaneously. They concluded that for large sized problem, the developed algorithm was more effective from genetic algorithm. Noorul Haq and Radha Ramanan[12]used Artificial Neural Network (ANN) for minimizing bicriteria of makespan and total flow time in flow shop scheduling environment and concluded that performance of ANN approach is better than constructive or improvement heuristics. Lockett and Muhlemann [13] proposed branch and bound algorithm for scheduling jobs with sequence dependent setup times on a single processor to minimize the total number of tool changes. Proposed algorithm was suitable for only small problems which was the major limitation of that algorithm. Gowrishankar et al. [14] considered *m*-machine flow shop

scheduling with minimizing variance of completion times of jobs and also sum of squares of deviations of the job completion times from a common due date. Blazewicz et al. [15] proposed different solution procedures for flow shop scheduling for two machine problem with a common due date and weighted latework criterion. Eren [16] considered a bicriteria *m*-machine flowshop scheduling with sequence dependent setup times with objective of minimizing the weighted sum of total completion time and makespan. He developed the special heuristics for fitness function considered and proved that the special heuristic for all number of jobs and machines values was more effective than the others. Erenay et al. [17] solved bicriteria scheduling problem with minimizing the number of tardy jobs and average flowtime on a single machine. They proposed four new heuristics for scheduling problem and concluded that the proposed beam search heuristics find efficient schedules and performed better than the existing heuristics available in the literature. Naderi et al. [18] minimized makespan and maximum tardiness in SDST flow shop scheduling with local search based hybridized the simulated annealing to promote the quality of final solution.

Therefore, in modern manufacturing system, production cost must be reduced in order to survive in this dynamic environment which can be done by effective utilisation of all the resources and completion of production in shorter time to increase the productivity also simultaneously considering due and early dates of the job. As minimisation of makespan with not meeting the due date is of no use for an industry since there is loss of market competitiveness, loss of customer, tardiness and earliness penalty etc.

Hence, for the today's need, we have considered the flowshop scheduling problem with sequence dependent set up time with tricriteria of weighted sum of total tardiness, total earliness and makespan which is very effective decision making in order to achieve maximum utilization of resources in respect of increasing the productivity and meeting the due dates so as the customer good will and satisfaction. We also proposed hybrid genetic algorithm in which initial seed sequence is obtained from heuristic similar to NEH [19]. As classical NEH considered processing times for makespan minimization and proposed heuristic also works on multicriteria objective function (i.e. weighted sum of total tardiness, total earliness and makespan).

2. STATEMENT OF PROBLEM

We have considered the flow shop scheduling problem with sequence dependent setup times and due dates associated to jobs in which the objective is to minimize the multicriteria decision making for manufacturing system including total tardiness, total earliness and makespan simultaneously. Various assumptions, parameters and multicriteria objective function considered has been illustrated below:

Assumptions

- Machines never break down and are available throughout the scheduling period.
- All the jobs and machines are available at time Zero.
- All processing time on the machine are known, deterministic and finite.
- Set up times for operations are sequence dependent and are not included in processing times
- Pre-emption is not allowed.
- Each machine is continuously available for assignment, without significant division of the scale into shifts or days and without any breakdown or maintenance.
- The first machine is assumed to be ready whichever and whatever job is to be processed on it first.
- Machines may be idle.
- Splitting of job or job cancellation is not allowed.
- It is associated with each job on each machine i.e. the time required to bring a given machine to a state, which allows the next job to commence and are immobilized to the machines.

Parameters

i	Index for Machines	$i=1,2,3.....m$	C_j	Completion time of job 'j'
j	Index for Jobs	$j=1,2,3.....n$	d_j	Due date of job'j'
α	Weight for Total tardiness	$\alpha \geq 0$	T_j	Tardiness of job 'j'
β	Weight for Total earliness	$\beta \geq 0$	E_j	Earliness of job'j'
γ	Weight for makespan	$\gamma \geq 0$ $\alpha + \beta + \gamma = 1$		

Multicriteria objective function

Multicriteria decision making objective function proposed in this work has based on realistic environment for manufacturing system (i.e. minimizing weighted sum of total tardiness, total earliness and makespan).The significance of all the three objective function (Individual or combined) are explained below:-

S.No	Objective Function	Significance
(i)	Total Tardiness (T_j)	Total tardiness (T_j) is a due date related performance measure and it is considered as summation of tardiness of individual jobs. If maximum jobs are to be completed in time but few jobs left which is overdue as of improper scheduling than minimizing total tardiness reflects that situation so that all the jobs will be completed in time. Not meeting the due dates may cause loss of customer, market competitiveness and termed as tardiness penalty.
(ii)	Total Earliness (E_j)	Total earliness (E_j) is also a due date related performance measure but reflects early delivery of jobs and it is considered as summation of earliness of individual jobs. If the jobs are produced before the due dates, than it also creates problem of inventory for an organization or it may cause penalty to the industry in terms of inventory cost and termed as earliness penalty.
(iii)	Makespan (C_{max})	Makespan is also a performance criterion which is defined as completion time of last job to be manufactured. In scheduling it is very important as to utilize maximum resources and increase productivity. Minimization of makespan achieves the goal of an industry.
(iv)	Multicriteria (weighted sum of total tardiness, total earliness and makespan)	As minimization of all the three performance criteria are important for an industry in the dynamic environment of markets, upward stress of competition, earliness and tardiness penalty and overall for Just in Time (JIT) Manufacturing and increasing productivity. So, for achieving this, there is a requirement of scheduling system which considered multicritreia decision making. To congregate this, we have developed this multicriteria objective function, which may be very effective decision making tool for scheduling jobs in the dynamic environment.

Therefore, the formulation of multicriteria objective function is stated below:
Total weighted tardiness which reflects the due dates of jobs to be scheduled as considered for minimization of late deliverance of jobs and has defined as:

$$\sum_{j=1}^{n} T_j$$

Where $T_j = C_j - d_j$ if $C_j - d_j \geq 0$

$= 0$ otherwise

Total earliness which reflects the early delivery of jobs to be scheduled as early delivery of jobs seems to be harmful and defined as:

$$\sum_{j=1}^{n} E_j$$

Where $E_j = d_j - C_j$ if $d_j - C_j \geq 0$

$= 0$ otherwise

Another commonly performance measure is completion time of last job i.e makespan (C_{max}) which has been used for maximum utilization of resources to increase productivity.
Therefore in the present work, for the requirement of Just in Time (JIT) manufacturing in terms of earliness and tardiness and also for increasing productivity, we have proposed the multi criteria decision making objective function including all the above three performance measure i.e. weighted sum of total tardiness, total earliness and makespan simultaneously for flow shop scheduling with sequence dependent set up times and has been framed as:

$$Min\left[\alpha\sum_{j=1}^{n} T_j + \beta\sum_{j=1}^{n} E_j + \gamma C_{\max}\right]$$

3. HYBRID GENETIC ALGORITHM (HGA)

Genetic algorithm is the optimization technique which can be applied to various problems, including those that are NP-hard. The genetic algorithm does not ensured an optimal solution; however it usually provides good approximations in a reasonable amount of time as compared to exact algorithms. It uses probabilistic selection as a basis for evolving a population of problem solutions. An initial population is created and subsequent generations are created according to a pre-specified breeding and mutation methods inspired by nature.

A genetic algorithm must be initialized with a starting population. Generating of initial population may be varied: feasible only, randomized, using some heuristics etc. Simple genetic algorithm generates initial population randomly and limitation of that algorithm is that if the initial solution is better than solution provided by the algorithm may be of good quality and if it is inferior than final results may not be better in a reasonable time. As flow shop scheduling belongs to NP hard and there is large search space to be searched in flow shop scheduling for finding optimal solutions and hence it is probable that random generation of initial solutions provides relatively weak results. For this, initial feasible solution is obtained by some heuristics for judgement of optimality in a very reasonable time. Generation of initial sequence with some heuristics and than that sequence is used as the initial sequence along with population as the procedure of simple genetic algorithm and called as Hybrid Genetic Algorithm (HGA).

Outline of Hybrid Genetic Algorithm(HGA)

The Hybrid Genetic Algorithm (HGA) acts as globally search technique which is similar to simple genetic algorithm with only deviation of generation of initial solution. In HGA, initial feasible solution is generated with the help of some heuristics and than this initial sequence has been used along with the population according to population size for the executing the procedure of simple genetic algorithm. The proposed HGA is described as:-

Step 1: Initialization and evaluation

a) The algorithm begins with generation of initial sequence with special heuristics (SH) called as one of the chromosome of population as described in section 3.2.
b) Generation of (Ps-1) sequences randomly as per population size (Ps).
c) Combining of initial sequence obtained by special heuristics with randomly generated sequence to form number of sequences equal to population size (Ps).

Step2: Reproduction

The algorithm then creates a set of new populations. At each generation, the algorithm uses the individuals in the current generation to generate the next population. To generate the new population, the algorithm performs the following steps:

a) Scores each member of the current population by computing fitness (i.e. weighted sum of total tardiness, total earliness and makespan simultaneously).
b) Selects parents based on the fitness function (i.e. multicriteria decision making).
c) Some of the individuals in the current population that have best fitness are chosen as *elite* and these elite individuals are utilized in the next population.
d) Production of offspring from the parents by *crossover* from the pair of parents or by making random changes to a single parent (*mutation)*.
e) Replaces the current population with the children to form the next generation.

Step3: Stopping limit

The algorithm stops when time limit reaches to $n \times m \times 0.25$ seconds.

Proposed Special Heuristic(SH)

The Special Heuristic (SH), the procedure which is similar to NEH [19] has been developed to solve the multicriteria flow shop scheduling with due dates and sequence dependent set up times for instances upto 200 jobs and 20 machines developed by Taillord [19]. Procedure of SH is described as below:

Step 1. Generation of initial sequence.

Step2. Set k = *2*. Pick the first two jobs from the initial sequence and schedule them in order to minimize the weighted sum of total tardiness, total earliness and makespan. As if there are only two jobs. Set the better one as the existing solution.

Step 3. Increment k by 1. Generate *k* candidate sequences by introducing the first job in the residual job list into each slot of the existing solution. Amongst these Candidates, select the better one with the least partial minimization of the weighted sum of total tardiness, total earliness and makespan simultaneously. Update the selected partial solution as the new existing solution.

*Step 4***.** If *k = n*, a feasible schedule has been found and stop. Otherwise, go to step 3.

Special heuristics SH1, SH2, SH3 and SH4 are obtained by using the EDD, LDD, EPDD and LPDD sequences respectively, in step 1 of proposed special heuristics. They have described below:

a) *Earliest Due Date (EDD):-* Schedule the jobs initially as per ascending order of due dates of jobs. (Kim, 1993).
b) *Latest Due Date (LDD):-* Arrange the jobs initially as per descending order of the due dates of jobs.

c) *Earliest processing time with due dates (EPDD):-* Schedule the jobs according to ascending order of $\left[\sum_{i=1}^{m} P_{ij} + d_j\right]$.

d) *Latest processing time with due dates (LPDD*):- Arrange the jobs according to descending order of $\left[\sum_{i=1}^{m} P_{ij} + d_j\right]$.

Parameters Settings

- *Population Size (Ps):* Population size refers to the search space i.e. algorithm has to search the specified number of sequences and larger the sequence, more the time is needed to execute the process of genetic algorithm. As in flow shop scheduling number of possible sequences is equal to n!. Therefore, if the population size is equal to n! than application of genetic algorithm has no use. So, larger the initial population that is created, the more likely the best solution from it will be closer to optimal but at the cost of increased execution time. So, in the present work it is set to 50 irrespective the size of problem to solve in a reasonable time.
- *Crossover function:* Crossover is the breeding of two parents to produce a single child. That child has features from both parents and thus may be better or worse than either parent as per fitness function. Analogous to natural selection, the more fit the parent, the more likely the generation have. Different types of crossover have used in literature and after having experimental comparison , we have found that the order crossover(OX) provides the best results for the multicriteria problem considered among the partially matched crossover(PMX), Order crossover (OX), Cycle crossover (CX) and single point crossover (SPX). So, in the present work, we have applied the order crossover (OX).
- *Mutation function:* For each sequence in the parent population a random number is picked and by giving this sequence a percent chance of being mutated. If this sequence is picked for mutation then a copy of the sequence is made and operation sequence procedure reversed. Only operations from different jobs will be reversed so that the mutation will always produces a feasible schedule. From the experiment, it is found that reciprocal exchange (RX) proves to be good with combination of order crossover (OX) and hence been used.
- *Elite Count:* The best sequences found should be considered in subsequent generations. At a lowest, the only best solution from the parent generation needs to be imitating to the next generation thus ensuring the best score of the next generation is at least as better as the previous generation. Here elite is expressed as number of sequences. In this work, we have fixed the elite count as two means that we clone the top two sequences that have least fitness function for the next generation.
- *Crossover fraction:* It is the fraction for which crossover has to perform on the parents as per population size in each generation. This is fixed to 0.8 i.e crossover should be done on 80% of total population size.
- *Mutation Fraction:* It is also used as fraction and specified for which process of mutation has to perform on the parents as per population size in each generation. This is fixed to 0.15 i.e. Mutation should be done on 15% of total population size.
- *Stopping condition:* Stopping condition is used to terminate the algorithm for certain numbers of generation. In this work, for fair comparison among different SH based Hybrid Genetic algorithm, we have used time limit base stopping criteria. So, the algorithm stops when maximum time limit reaches $n \times m \times 0.25$ seconds.

4. RESULTS AND DISCUSSIONS

In this study, flow shop scheduling with sequence dependent set up time and due dates of jobs have been considered in flow shop environment. The multicriteria objective function including weighted sum of total tardiness, total earliness and makespan has been proposed which is very

effective decision making model for achieving industry as well as customer goals through scheduling the jobs. As the problems belong to NP hard, so we have also developed hybrid genetic algorithm, in which initial feasible sequence is obtained from special heuristic. In the present work, we have developed four special heuristics (SH) and hybrid with genetic algorithm and termed special heuristic based hybrid genetic algorithms HGA (SH1), HGA (SH2), HGA (SH3) and HGA (SH4) as stated in section 3.2. Problems upto 200 jobs and 20 machines developed by Taillord [20] in flow shop environment with sequence dependent set up time and due dates have been solved from the proposed hybrid genetic algorithm. Computational and experimental study for the entire four hybrid genetic algorithm has also been made for comparison. All experimental tests are conducted on a personal computer with P IV, core 2duo processor and 1 GB Ram. As proposed HGA generate initially only one sequence by SH and remaining as per population size (Ps) randomly, so for each size of problem, we have run HGA five times for taking final average to compensate randomness. Also for reasonable comparison, stopping limit of HGA has been fixed to time limit based criteria which is $n \times m \times 0.25$ seconds. Comparison of all the four hybrid genetic algorithm has been done by calculating Error which is computed as:-

$$\text{Error (\%)} = \frac{Average_{sol} - Best_{sol}}{Best_{Sol}} X100$$

Where $Average_{sol}$ is the average solution obtained by a given algorithm and $Best_{sol}$ is the best solution obtained among all the methods or the best known solution. Lesser the error, better the results obtained. $Best_{sol}$ can be found from the results obtained by running HGA five times for a particular problem and Average solution is the final average solution produced by the algorithm for all the five runs. The Error for all the four HGA have been compared for five, ten and twenty machines problems with four sets of weights (0.33, 0.33, 0.33), (0.25,0.25,0.5) , (0.5, 0.25, 0.25) and (0.25, 0.25, 0.5) for multicriteria decision making objective function and have been shown in Figure 1, Figure 2 and Figure 3.

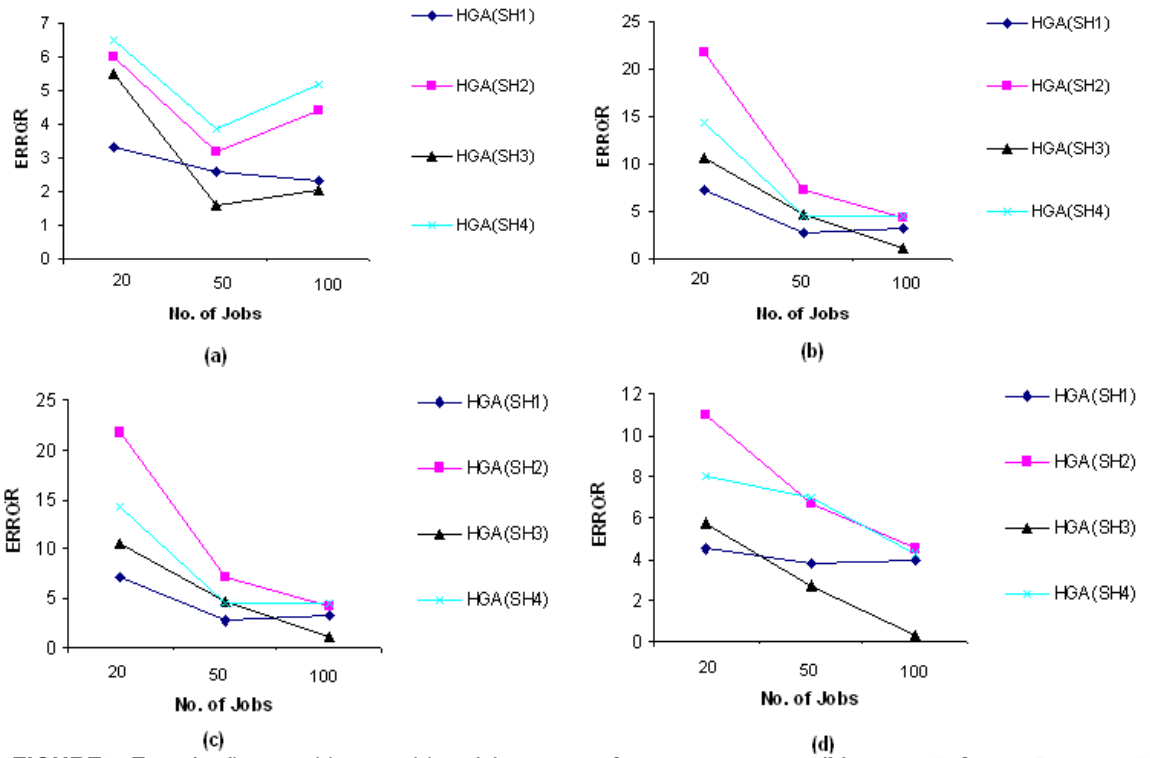


FIGURE1: Error for five machines problem **(a)** α = 0.33, β = 0.33 & γ = 0.33 **(b)** α = 0.25, β = 0.25 & γ = 0.5 **(c)** α = 0.5, β = 0.25 & γ = 0.25 **(d)** α = 0.25, β = 0.25 & γ = 0.5.

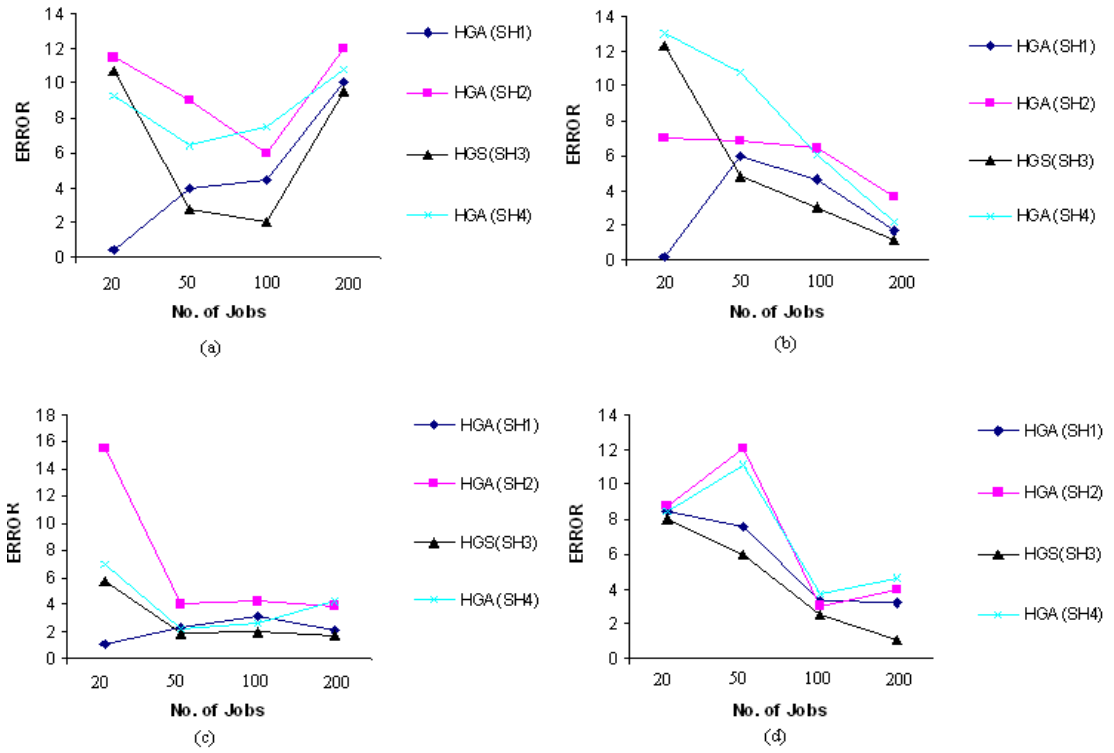


FIGURE 2: Error for ten machines problem **(a)** α = 0.33, β = 0.33 & γ = 0.33 **(b)** α = 0.25, β = 0.25 & γ = 0.5 **(c)** α = 0.5, β = 0.25 & γ = 0.25 (d) α = 0.25, β = 0.25 & γ = 0.5

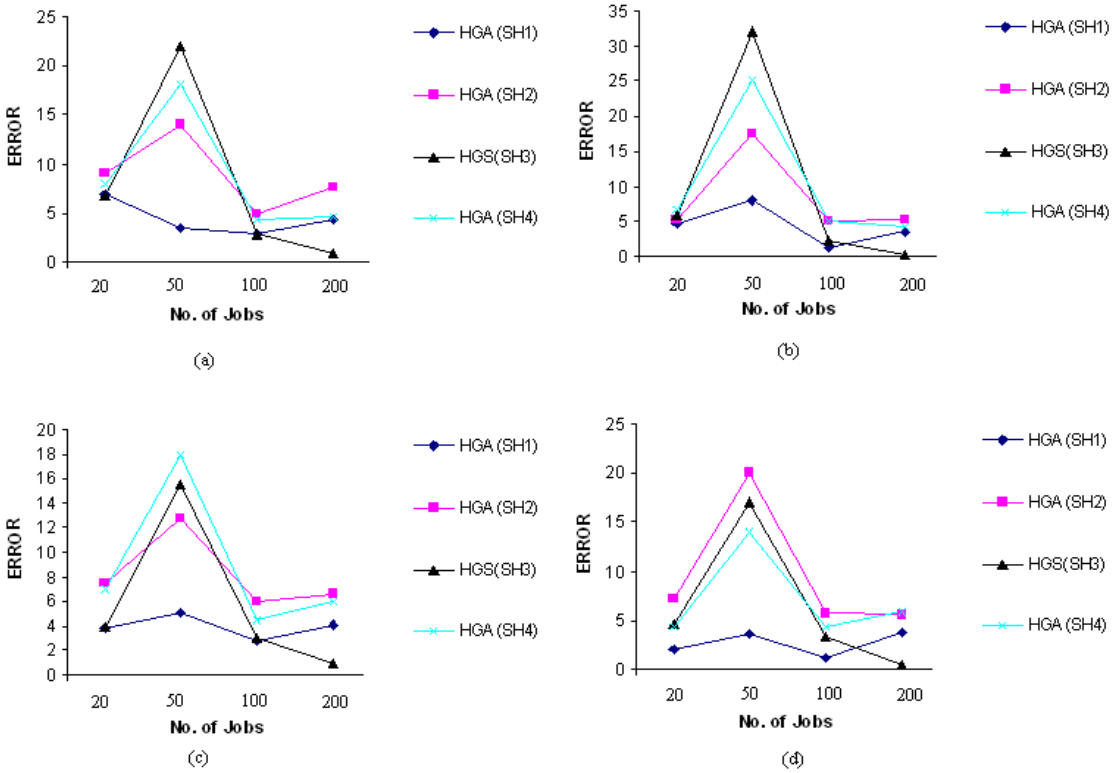


FIGURE 3: Error for twenty machines problem **(a)** α = 0.33, β = 0.33 & γ = 0.33 **(b)** α = 0.25, β = 0.25 & γ = 0.5 **(c)** α = 0.5, β = 0.25 & γ = 0.25 **(d)** α = 0.25, β = 0.25 & γ = 0.5

From the analysis for 5, 10 and 20 machines problems as shown in Figure 1, Figure 2 and Figure 3, performance of proposed SH1 based HGA upto 20 jobs and SH3 based HGA as the jobs size increases, showed superiority over others for all the four sets of weight values ((0.33, 0.33, 0.33), (0.25,0.25,0.5) , (0.5, 0.25, 0.25) and (0.25, 0.25, 0.5)) for multicriteria decision making in flow shop scheduling under sequence dependent set up time i.e. weighted sum of total tardiness, total earliness and makespan.

5. CONCLUSIONS

In the present work, we have framed multicriteria decision making for flow shop scheduling with weighted sum of total tardiness, total earliness and makespan under sequence dependent set up time and also proposed four special heuristic based hybrid genetic algorithms. Computational analysis has also been done for comparing the performance of proposed four HGA's i.e. HGA (SH1), HGA (SH2), HGA (SH3) and HGA (SH4). The HGA's have been tested upto 200 jobs and 20 machines problems in flow shop scheduling as derived by Taillord [20] for all the four weight values (α, β and γ) for fitness function (i.e. (0.33, 0.33, 0.33), (0.25, 0.25, 0.5), (0.5, 0.25, 0.25) and (0.25, 0.25, 0.5)). From the analysis it has been concluded that the proposed HGA(SH1) for smaller and HGA(SH3) for larger job size problems showed superiority over other for 5,10 and 20 machines problem for multicriteria decision making flow shop scheduling (i.e. weighted sum of total tardiness, total earliness and makespan) under sequence dependent set up time and due dates.

6. REFERENCES

1. Allahverdi A., Ng C.T., Cheng T.C.E. and Kovalyov M.Y. *"A survey of scheduling problems with setup times or costs".* European Journal of Operation Research, 187(3):985–1032,2008
2. Rajendran C. *"Theory and methodology heuristics for scheduling in flow shop with multiple objectives"* .European Journal of Operation Research , 82 (3):540 –555,1995
3. Ravindran D., Noorul Haq A., Selvakuar S.J. and Sivaraman R. *" Flow shop scheduling with multiple objective of minimizing makespan and total flow time".* International Journal of Advanced Manufacturing Technology, 25:1007–1012,2005
4. Gupta J.N.D., Venkata G., Neppalli R. and Werner F. *"Minimizing total flow time in a two-machine flowshop problem with minimum makespan".* International Journal of Production Economics, 69:323–338,2001
5. Sayin S. and Karabati S. *"A bicriteria approach to the two-machine flow shop scheduling problem".* European Journal of Operation Research, 113:435–449,1999
6. Danneberg D., Tautenhahn T., Werner F. *"A comparison of heuristic algorithms for flow shop scheduling problems with setup times and limited batch size".* Mathematical and Computer Modelling, 29 (9):101–126,1999
7. Toktas B., Azizoglu M. and Koksalan S.K. *"Two-machine flow shop scheduling with two criteria: Maximum earliness and makespan".* European Journal of Operation Research, 157(2): 286–295,2004
8. Ponnambalam S. G., Jagannathan H., Kataria M. and Gadicherla A. *"A TSP-GA multi-objective algorithm for flow shop scheduling".* International Journal of Advanced Manufacturing Technology, 23: 909–915,2004
9. Loukil T., Teghem J. and Tuyttens D. *"Solving multi-objective production scheduling problems using metaheuristics".* European Journal of Operation Research , 161(1) :42–61,2005
10. Fred Choobineh F., Mohebbi E. and Khoo, H. *"A multi-objective tabu search for a single-machine scheduling problem with sequence dependent setup times".* European Journal of Operational Research, 175, 318–337,2006
11. Rahimi Vahed R and Mirghorbani S. M. *"A multi-objective particle swarm for a flow shop scheduling problem".* Combinatorial Optimization, 13 (1):79–102,2007

12. Noorul Haq and Radha Ramanan T. *"A bicriterian flow shops scheduling using artificial neural network".* International Journal of Advanced Manufacturing Technology, 30: 1132–1138, 2006
13. Lockett A. G. and Muhlemann A. P. *"Technical notes: a scheduling problem involving sequence dependent changeover times".*Operation Research, 20: 895-902,1972
14. Gowrisankar K, Chandrasekharan Rajendran and Srinivasan G. *"Flow shop scheduling algorithm for minimizing the completion time variance and the sum of squares of completion time deviations from a common due date".* European Journal of Operation Research, 132:643–665,2001
15. Blazewicz J, Pesch E, Sterna M and Werner F. *" A comparison of solution procedures for two-machine flow shop scheduling with late work criterion".* Computers and Industrial Engineering, 49:611–624,2005
16. Eren, T. *"A bicriteria m-machine flow shop scheduling with sequence-dependent setup times".* Applied Mathematical Modelling,2009 (In press), doi: 10.1016/j.apm.2009.04.005
17. Erenay F.S., et al. *"New solution methods for single machine bicriteria Scheduling problem: Minimization of average flow time and number of tardy jobs".* European Journal of Operational Research,2009 (In press), doi:10.1016/j.ejor.2009.02.014
18. Naderi B., Zandieh M. and Roshanaei V. *"Scheduling hybrid flowshops with sequence dependent setup times to minimize makespan and maximum tardiness".* International Journal of Advanced Manufacturing Technology, 41:1186–1198,2009
19. Nawaz M., Enscore E. and Ham I. *"Heuristic Algorithm for the M-Machine N-Job Flow Shop Sequencing Problem"*. Omega, 11: 91–95,1983
20. Taillard E. *"Benchmarks of basic scheduling problems"*. European Journal of Operation Research, 64:278–285, 1993

Hybrid Genetic Algorithm for Multicriteria Scheduling with Sequence Dependent Set up Time

Ashwani Dhingra *ashwani_dhingra1979@rediff.com*
Department of Mechanical Engineering
University Institute of Engineering & Technology
Maharshi Dayanand University
Rohtak, 124001, Haryana INDIA

Pankaj Chandna *pchandna08@gmail.com*
Department of Mechanical Engineering
National Institute of Technology
Kurukshetra, 136119, Haryana INDIA

Abstract

In this work, multicriteria decision making objective for flow shop scheduling with sequence dependent set up time and due dates have been developed. Multicriteria decision making objective includes total tardiness, total earliness and makespan simultaneously which is very effective decision making for scheduling jobs in modern manufacturing environment. As problem of flow shop scheduling is NP hard and to solve this in a reasonable time, four Special heuristics (SH) based Hybrid Genetic Algorithm (HGA) have also been developed for proposed multicriteria objective function. A computational analysis upto 200 jobs and 20 machines problems has been conducted to evaluate the performance of four HGA's. The analysis showed the superiority of SH1 based HGA for small size and SH3 based HGA for large size problem for multicriteria flow shop scheduling with sequence dependent set up time and due dates.

Keywords: Flow shop scheduling, Genetic algorithm, Sequence dependent set up time, Total tardiness, Total earliness, makespan

1. INTRODUCTION

Scheduling in manufacturing systems is typically associated with allocating a set of jobs on a set of machines in order to achieve some objectives. It is a decision making process that concerns the allocation of limited resources to a set of tasks for optimizing one or more objectives. Manufacturing system is classified as job shop and flow shop and in the present work; we have dealt with flow shop. In a job shop, set of jobs is to be scheduled on a set of machines and there is no restriction of similar route on the jobs to follow, however in flow shop environment all the jobs have to follow the similar route. Classical flow shop scheduling problems are mainly concerned with completion time related objectives (e.g. flow time and makespan) and aims to reduce production time and enhance productivity and facility utilization. In modern manufacturing and operations management, on time delivery is a significant factor as for the survival in the competitive markets and effective scheduling becomes very important in order to meet customer requirements as rapidly as possible while maximizing the productivity and facility utilization. Therefore, there is a need of scheduling system which includes multicriteria decisions like makespan, total flow time, earliness, tardiness etc.

Also, flow shop scheduling problems including sequence dependent set up time (SDST) have been considered the most renowned problems in the area of scheduling. Ali Allahverdi et al [1] investigated the survey of flow shop scheduling problems with set up times and concluded that considering flow shop scheduling problems is to obtain remarkable savings when setup times are included in scheduling. Sequence dependent setup times are usually found in the condition where the multipurpose facility is available on multipurpose machine. Flow shop scheduling with sequence dependent set up time can be found in many industrial systems like textile industry, stamping plants, chemical, printing, pharmaceutical and automobile industry etc. Several researchers have considered the problem of flow shop scheduling with single criterion and very few have dealt with multicriterion decision making including sequence dependent set up.

The scheduling literature also revealed that the research on flow shop scheduling is mainly focused on bicriteria without considering sequence dependent set up time .Very few considered flow shop scheduling including sequence dependent set up time with multicriteria. Rajendran [2] has implemented a heuristic for flow shop scheduling with multiple objectives of optimizing makespan, total flow time and idle time for machines. The heuristic preference relation had been proposed which was used as the basis to restrict the search for possible improvement in the multiple objectives. Ravindran et al. [3] proposed three heuristic similar to NEH for flow shop scheduling with multiple objectives of makespan and total flow time together and concluded that proposed three heuristic yields good results than the Rajendran heuristic CR[2].Gupta et al.[4] considered flow shop scheduling problem for minimising total flow time subject to the condition that the makespan of the schedule is minimum. Sayin and Karabati [5] minimized makespan and sum of completion times simultaneously in two machines flow shop scheduling environment. Branch and bound procedure was developed that iteratively solved single objective scheduling problems until the set of efficient solutions was completely enumerated. Danneberg et al. [6] addressed the permutation flow shop scheduling problem with setup times and considered makespan as well as the weighted sum of the completion times of the jobs as objective function. For solving such a problem, they also proposed and compared various constructive and iterative algorithms. Toktas et al. [7] considered the two machine flow shop scheduling by minimizing makespan and maximum earliness simultaneously. They developed a branch & bound and a heuristic procedure that generates all efficient solutions with respect to two criteria. Ponnambalam et al. [8] proposed a TSP GA multiobjective algorithm for flow shop scheduling, where a weighted sum of multiple objectives (i.e. minimizing makespan, mean flow time and machine idle time) was used. The proposed algorithm showed superiority which when applied to benchmark problems available in the OR-Library. Loukil et al. [9] proposed multiobjective simulated annealing algorithm to tackle the multiobjective production scheduling problems. They considered seven possible objective functions (the mean weighted completion time, the mean weighted tardiness, the mean weighted earliness, the maximum completion time (makespan), the maximum tardiness, the maximum earliness, the number of tardy jobs). They claimed that the proposed multiobjective simulated annealing algorithm was able to solve any subset of seven possible objective functions. Fred Choobineh et al. [10] proposed tabu search heuristic with makespan, weighted tardiness and number of tardy jobs simultaneously including sequence dependent setup for n jobs on a single machine. They illustrated that as the problem size increases, the results provided by the proposed heuristic was optimal or nearer to optimal solutions in a reasonable time for multiobjective fitness function considered. Rahimi Vahed and Mirghorbani [11] developed multi objective particle swarm optimization for flow shop scheduling problem to minimize the weighted mean completion time and weighted mean tardiness simultaneously. They concluded that for large sized problem, the developed algorithm was more effective from genetic algorithm. Noorul Haq and Radha Ramanan[12]used Artificial Neural Network (ANN) for minimizing bicriteria of makespan and total flow time in flow shop scheduling environment and concluded that performance of ANN approach is better than constructive or improvement heuristics. Lockett and Muhlemann [13] proposed branch and bound algorithm for scheduling jobs with sequence dependent setup times on a single processor to minimize the total number of tool changes. Proposed algorithm was suitable for only small problems which was the major limitation of that algorithm. Gowrishankar et al. [14] considered *m*-machine flow shop

scheduling with minimizing variance of completion times of jobs and also sum of squares of deviations of the job completion times from a common due date. Blazewicz et al. [15] proposed different solution procedures for flow shop scheduling for two machine problem with a common due date and weighted latework criterion. Eren [16] considered a bicriteria *m*-machine flowshop scheduling with sequence dependent setup times with objective of minimizing the weighted sum of total completion time and makespan. He developed the special heuristics for fitness function considered and proved that the special heuristic for all number of jobs and machines values was more effective than the others. Erenay et al. [17] solved bicriteria scheduling problem with minimizing the number of tardy jobs and average flowtime on a single machine. They proposed four new heuristics for scheduling problem and concluded that the proposed beam search heuristics find efficient schedules and performed better than the existing heuristics available in the literature. Naderi et al. [18] minimized makespan and maximum tardiness in SDST flow shop scheduling with local search based hybridized the simulated annealing to promote the quality of final solution.

Therefore, in modern manufacturing system, production cost must be reduced in order to survive in this dynamic environment which can be done by effective utilisation of all the resources and completion of production in shorter time to increase the productivity also simultaneously considering due and early dates of the job. As minimisation of makespan with not meeting the due date is of no use for an industry since there is loss of market competitiveness, loss of customer, tardiness and earliness penalty etc.

Hence, for the today's need, we have considered the flowshop scheduling problem with sequence dependent set up time with tricriteria of weighted sum of total tardiness, total earliness and makespan which is very effective decision making in order to achieve maximum utilization of resources in respect of increasing the productivity and meeting the due dates so as the customer good will and satisfaction. We also proposed hybrid genetic algorithm in which initial seed sequence is obtained from heuristic similar to NEH [19]. As classical NEH considered processing times for makespan minimization and proposed heuristic also works on multicriteria objective function (i.e. weighted sum of total tardiness, total earliness and makespan).

2. STATEMENT OF PROBLEM

We have considered the flow shop scheduling problem with sequence dependent setup times and due dates associated to jobs in which the objective is to minimize the multicriteria decision making for manufacturing system including total tardiness, total earliness and makespan simultaneously. Various assumptions, parameters and multicriteria objective function considered has been illustrated below:

Assumptions

- Machines never break down and are available throughout the scheduling period.
- All the jobs and machines are available at time Zero.
- All processing time on the machine are known, deterministic and finite.
- Set up times for operations are sequence dependent and are not included in processing times
- Pre-emption is not allowed.
- Each machine is continuously available for assignment, without significant division of the scale into shifts or days and without any breakdown or maintenance.
- The first machine is assumed to be ready whichever and whatever job is to be processed on it first.
- Machines may be idle.
- Splitting of job or job cancellation is not allowed.
- It is associated with each job on each machine i.e. the time required to bring a given machine to a state, which allows the next job to commence and are immobilized to the machines.

Parameters

i	Index for Machines	$i=1,2,3.....m$	C_j	Completion time of job 'j'
j	Index for Jobs	$j=1,2,3.....n$	d_j	Due date of job'j'
α	Weight for Total tardiness	$\alpha \geq 0$	T_j	Tardiness of job 'j'
β	Weight for Total earliness	$\beta \geq 0$	E_j	Earliness of job'j'
γ	Weight for makespan	$\gamma \geq 0$ $\alpha + \beta + \gamma = 1$		

Multicriteria objective function

Multicriteria decision making objective function proposed in this work has based on realistic environment for manufacturing system (i.e. minimizing weighted sum of total tardiness, total earliness and makespan).The significance of all the three objective function (Individual or combined) are explained below:-

S.No	Objective Function	Significance
(i)	Total Tardiness (T_j)	Total tardiness (T_j) is a due date related performance measure and it is considered as summation of tardiness of individual jobs. If maximum jobs are to be completed in time but few jobs left which is overdue as of improper scheduling than minimizing total tardiness reflects that situation so that all the jobs will be completed in time. Not meeting the due dates may cause loss of customer, market competitiveness and termed as tardiness penalty.
(ii)	Total Earliness (E_j)	Total earliness (E_j) is also a due date related performance measure but reflects early delivery of jobs and it is considered as summation of earliness of individual jobs. If the jobs are produced before the due dates, than it also creates problem of inventory for an organization or it may cause penalty to the industry in terms of inventory cost and termed as earliness penalty.
(iii)	Makespan (C_{max})	Makespan is also a performance criterion which is defined as completion time of last job to be manufactured. In scheduling it is very important as to utilize maximum resources and increase productivity. Minimization of makespan achieves the goal of an industry.
(iv)	Multicriteria (weighted sum of total tardiness, total earliness and makespan)	As minimization of all the three performance criteria are important for an industry in the dynamic environment of markets, upward stress of competition, earliness and tardiness penalty and overall for Just in Time (JIT) Manufacturing and increasing productivity. So, for achieving this, there is a requirement of scheduling system which considered multicritreia decision making. To congregate this, we have developed this multicriteria objective function, which may be very effective decision making tool for scheduling jobs in the dynamic environment.

Therefore, the formulation of multicriteria objective function is stated below:
Total weighted tardiness which reflects the due dates of jobs to be scheduled as considered for minimization of late deliverance of jobs and has defined as:

$$\sum_{j=1}^{n} T_j$$

Where $T_j = C_j - d_j$ if $C_j - d_j \geq 0$

$= 0$ otherwise

Total earliness which reflects the early delivery of jobs to be scheduled as early delivery of jobs seems to be harmful and defined as:

$$\sum_{j=1}^{n} E_j$$

Where $E_j = d_j - C_j$ if $d_j - C_j \geq 0$

$= 0$ otherwise

Another commonly performance measure is completion time of last job i.e makespan (C_{max}) which has been used for maximum utilization of resources to increase productivity.
Therefore in the present work, for the requirement of Just in Time (JIT) manufacturing in terms of earliness and tardiness and also for increasing productivity, we have proposed the multi criteria decision making objective function including all the above three performance measure i.e. weighted sum of total tardiness, total earliness and makespan simultaneously for flow shop scheduling with sequence dependent set up times and has been framed as:

$$Min\left[\alpha\sum_{j=1}^{n} T_j + \beta\sum_{j=1}^{n} E_j + \gamma C_{\max}\right]$$

3. HYBRID GENETIC ALGORITHM (HGA)

Genetic algorithm is the optimization technique which can be applied to various problems, including those that are NP-hard. The genetic algorithm does not ensured an optimal solution; however it usually provides good approximations in a reasonable amount of time as compared to exact algorithms. It uses probabilistic selection as a basis for evolving a population of problem solutions. An initial population is created and subsequent generations are created according to a pre-specified breeding and mutation methods inspired by nature.

A genetic algorithm must be initialized with a starting population. Generating of initial population may be varied: feasible only, randomized, using some heuristics etc. Simple genetic algorithm generates initial population randomly and limitation of that algorithm is that if the initial solution is better than solution provided by the algorithm may be of good quality and if it is inferior than final results may not be better in a reasonable time. As flow shop scheduling belongs to NP hard and there is large search space to be searched in flow shop scheduling for finding optimal solutions and hence it is probable that random generation of initial solutions provides relatively weak results. For this, initial feasible solution is obtained by some heuristics for judgement of optimality in a very reasonable time. Generation of initial sequence with some heuristics and than that sequence is used as the initial sequence along with population as the procedure of simple genetic algorithm and called as Hybrid Genetic Algorithm (HGA).

Outline of Hybrid Genetic Algorithm(HGA)

The Hybrid Genetic Algorithm (HGA) acts as globally search technique which is similar to simple genetic algorithm with only deviation of generation of initial solution. In HGA, initial feasible solution is generated with the help of some heuristics and than this initial sequence has been used along with the population according to population size for the executing the procedure of simple genetic algorithm. The proposed HGA is described as:-

Step 1: Initialization and evaluation

a) The algorithm begins with generation of initial sequence with special heuristics (SH) called as one of the chromosome of population as described in section 3.2.
b) Generation of (Ps-1) sequences randomly as per population size (Ps).
c) Combining of initial sequence obtained by special heuristics with randomly generated sequence to form number of sequences equal to population size (Ps).

Step2: Reproduction

The algorithm then creates a set of new populations. At each generation, the algorithm uses the individuals in the current generation to generate the next population. To generate the new population, the algorithm performs the following steps:

a) Scores each member of the current population by computing fitness (i.e. weighted sum of total tardiness, total earliness and makespan simultaneously).
b) Selects parents based on the fitness function (i.e. multicriteria decision making).
c) Some of the individuals in the current population that have best fitness are chosen as *elite* and these elite individuals are utilized in the next population.
d) Production of offspring from the parents by *crossover* from the pair of parents or by making random changes to a single parent (*mutation)*.
e) Replaces the current population with the children to form the next generation.

Step3: Stopping limit

The algorithm stops when time limit reaches to $n \times m \times 0.25$ seconds.

Proposed Special Heuristic(SH)

The Special Heuristic (SH), the procedure which is similar to NEH [19] has been developed to solve the multicriteria flow shop scheduling with due dates and sequence dependent set up times for instances upto 200 jobs and 20 machines developed by Taillord [19]. Procedure of SH is described as below:

Step 1. Generation of initial sequence.

Step2. Set k = *2*. Pick the first two jobs from the initial sequence and schedule them in order to minimize the weighted sum of total tardiness, total earliness and makespan. As if there are only two jobs. Set the better one as the existing solution.

Step 3. Increment k by 1. Generate *k* candidate sequences by introducing the first job in the residual job list into each slot of the existing solution. Amongst these Candidates, select the better one with the least partial minimization of the weighted sum of total tardiness, total earliness and makespan simultaneously. Update the selected partial solution as the new existing solution.

*Step 4***.** If *k = n*, a feasible schedule has been found and stop. Otherwise, go to step 3.

Special heuristics SH1, SH2, SH3 and SH4 are obtained by using the EDD, LDD, EPDD and LPDD sequences respectively, in step 1 of proposed special heuristics. They have described below:

a) *Earliest Due Date (EDD):-* Schedule the jobs initially as per ascending order of due dates of jobs. (Kim, 1993).
b) *Latest Due Date (LDD):-* Arrange the jobs initially as per descending order of the due dates of jobs.

c) *Earliest processing time with due dates (EPDD):-* Schedule the jobs according to ascending order of $\left[\sum_{i=1}^{m} P_{ij} + d_j\right]$.

d) *Latest processing time with due dates (LPDD*):- Arrange the jobs according to descending order of $\left[\sum_{i=1}^{m} P_{ij} + d_j\right]$.

Parameters Settings

- *Population Size (Ps):* Population size refers to the search space i.e. algorithm has to search the specified number of sequences and larger the sequence, more the time is needed to execute the process of genetic algorithm. As in flow shop scheduling number of possible sequences is equal to n!. Therefore, if the population size is equal to n! than application of genetic algorithm has no use. So, larger the initial population that is created, the more likely the best solution from it will be closer to optimal but at the cost of increased execution time. So, in the present work it is set to 50 irrespective the size of problem to solve in a reasonable time.
- *Crossover function:* Crossover is the breeding of two parents to produce a single child. That child has features from both parents and thus may be better or worse than either parent as per fitness function. Analogous to natural selection, the more fit the parent, the more likely the generation have. Different types of crossover have used in literature and after having experimental comparison , we have found that the order crossover(OX) provides the best results for the multicriteria problem considered among the partially matched crossover(PMX), Order crossover (OX), Cycle crossover (CX) and single point crossover (SPX). So, in the present work, we have applied the order crossover (OX).
- *Mutation function:* For each sequence in the parent population a random number is picked and by giving this sequence a percent chance of being mutated. If this sequence is picked for mutation then a copy of the sequence is made and operation sequence procedure reversed. Only operations from different jobs will be reversed so that the mutation will always produces a feasible schedule. From the experiment, it is found that reciprocal exchange (RX) proves to be good with combination of order crossover (OX) and hence been used.
- *Elite Count:* The best sequences found should be considered in subsequent generations. At a lowest, the only best solution from the parent generation needs to be imitating to the next generation thus ensuring the best score of the next generation is at least as better as the previous generation. Here elite is expressed as number of sequences. In this work, we have fixed the elite count as two means that we clone the top two sequences that have least fitness function for the next generation.
- *Crossover fraction:* It is the fraction for which crossover has to perform on the parents as per population size in each generation. This is fixed to 0.8 i.e crossover should be done on 80% of total population size.
- *Mutation Fraction:* It is also used as fraction and specified for which process of mutation has to perform on the parents as per population size in each generation. This is fixed to 0.15 i.e. Mutation should be done on 15% of total population size.
- *Stopping condition:* Stopping condition is used to terminate the algorithm for certain numbers of generation. In this work, for fair comparison among different SH based Hybrid Genetic algorithm, we have used time limit base stopping criteria. So, the algorithm stops when maximum time limit reaches $n \times m \times 0.25$ seconds.

4. RESULTS AND DISCUSSIONS

In this study, flow shop scheduling with sequence dependent set up time and due dates of jobs have been considered in flow shop environment. The multicriteria objective function including weighted sum of total tardiness, total earliness and makespan has been proposed which is very

effective decision making model for achieving industry as well as customer goals through scheduling the jobs. As the problems belong to NP hard, so we have also developed hybrid genetic algorithm, in which initial feasible sequence is obtained from special heuristic. In the present work, we have developed four special heuristics (SH) and hybrid with genetic algorithm and termed special heuristic based hybrid genetic algorithms HGA (SH1), HGA (SH2), HGA (SH3) and HGA (SH4) as stated in section 3.2. Problems upto 200 jobs and 20 machines developed by Taillord [20] in flow shop environment with sequence dependent set up time and due dates have been solved from the proposed hybrid genetic algorithm. Computational and experimental study for the entire four hybrid genetic algorithm has also been made for comparison. All experimental tests are conducted on a personal computer with P IV, core 2duo processor and 1 GB Ram. As proposed HGA generate initially only one sequence by SH and remaining as per population size (Ps) randomly, so for each size of problem, we have run HGA five times for taking final average to compensate randomness. Also for reasonable comparison, stopping limit of HGA has been fixed to time limit based criteria which is $n \times m \times 0.25$ seconds. Comparison of all the four hybrid genetic algorithm has been done by calculating Error which is computed as:-

$$\text{Error (\%)} = \frac{Average_{sol} - Best_{sol}}{Best_{Sol}} X100$$

Where $Average_{sol}$ is the average solution obtained by a given algorithm and $Best_{sol}$ is the best solution obtained among all the methods or the best known solution. Lesser the error, better the results obtained. $Best_{sol}$ can be found from the results obtained by running HGA five times for a particular problem and Average solution is the final average solution produced by the algorithm for all the five runs. The Error for all the four HGA have been compared for five, ten and twenty machines problems with four sets of weights (0.33, 0.33, 0.33), (0.25,0.25,0.5) , (0.5, 0.25, 0.25) and (0.25, 0.25, 0.5) for multicriteria decision making objective function and have been shown in Figure 1, Figure 2 and Figure 3.

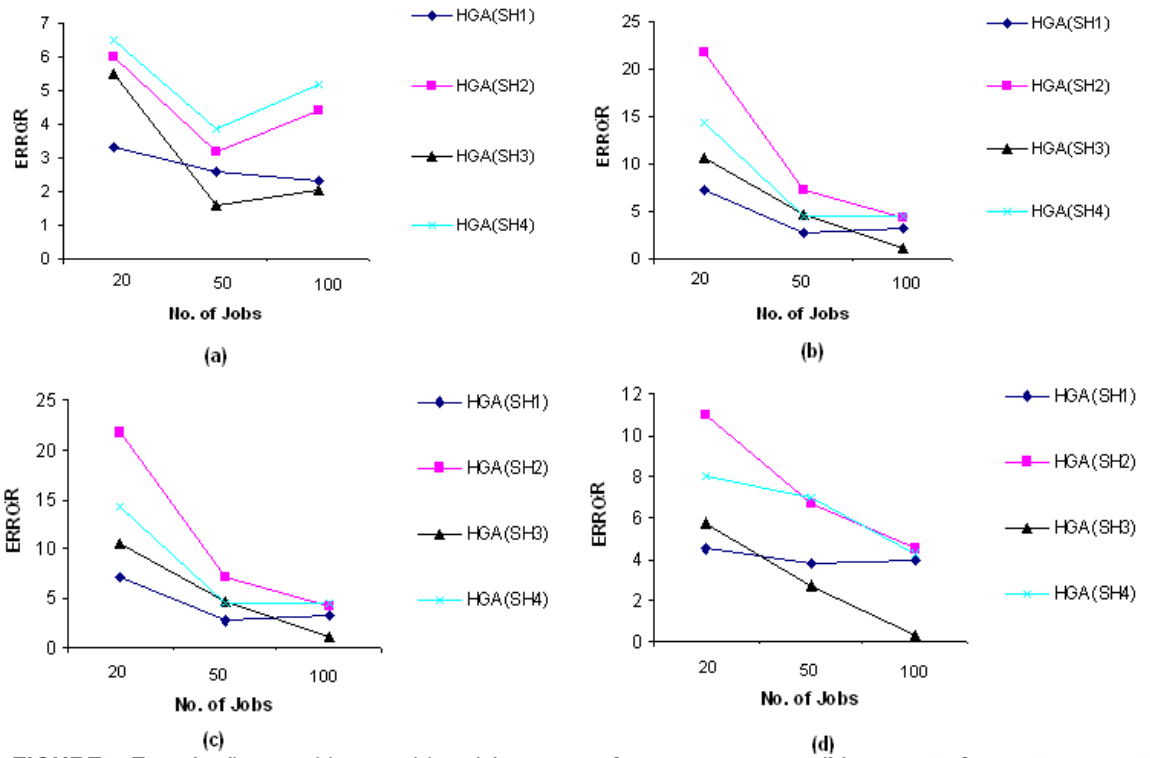


FIGURE1: Error for five machines problem **(a)** α = 0.33, β = 0.33 & γ = 0.33 **(b)** α = 0.25, β = 0.25 & γ = 0.5 **(c)** α = 0.5, β = 0.25 & γ = 0.25 **(d)** α = 0.25, β = 0.25 & γ = 0.5.

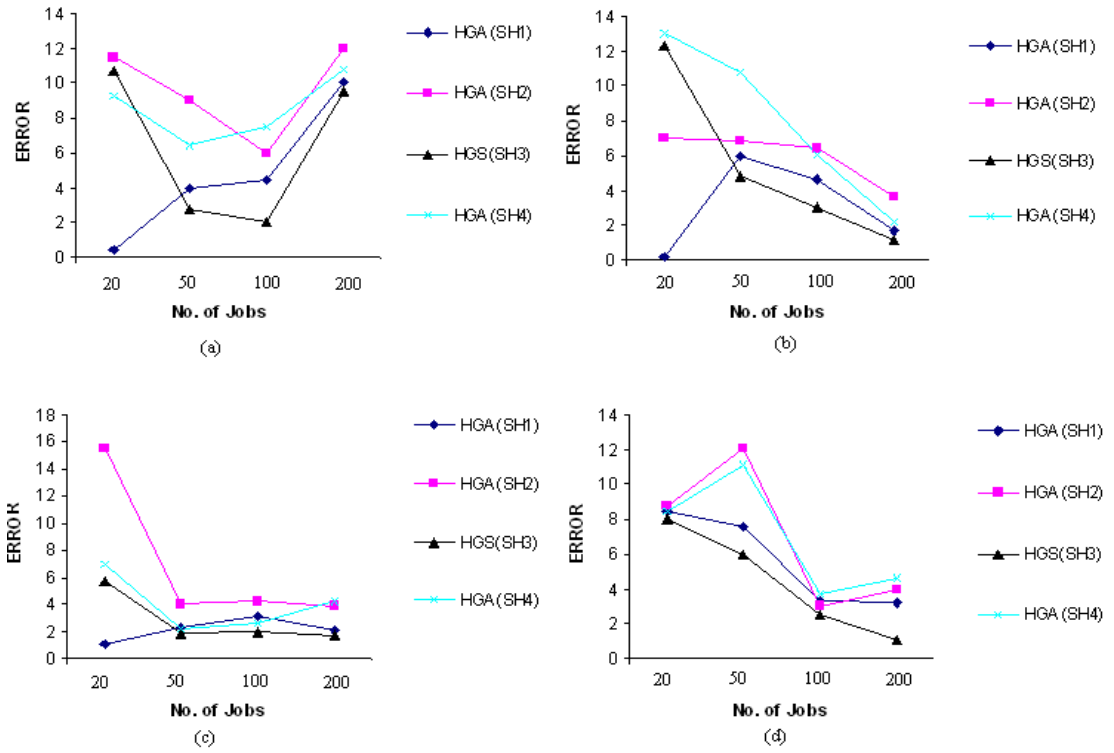


FIGURE 2: Error for ten machines problem **(a)** α = 0.33, β = 0.33 & γ = 0.33 **(b)** α = 0.25, β = 0.25 & γ = 0.5 **(c)** α = 0.5, β = 0.25 & γ = 0.25 (d) α = 0.25, β = 0.25 & γ = 0.5

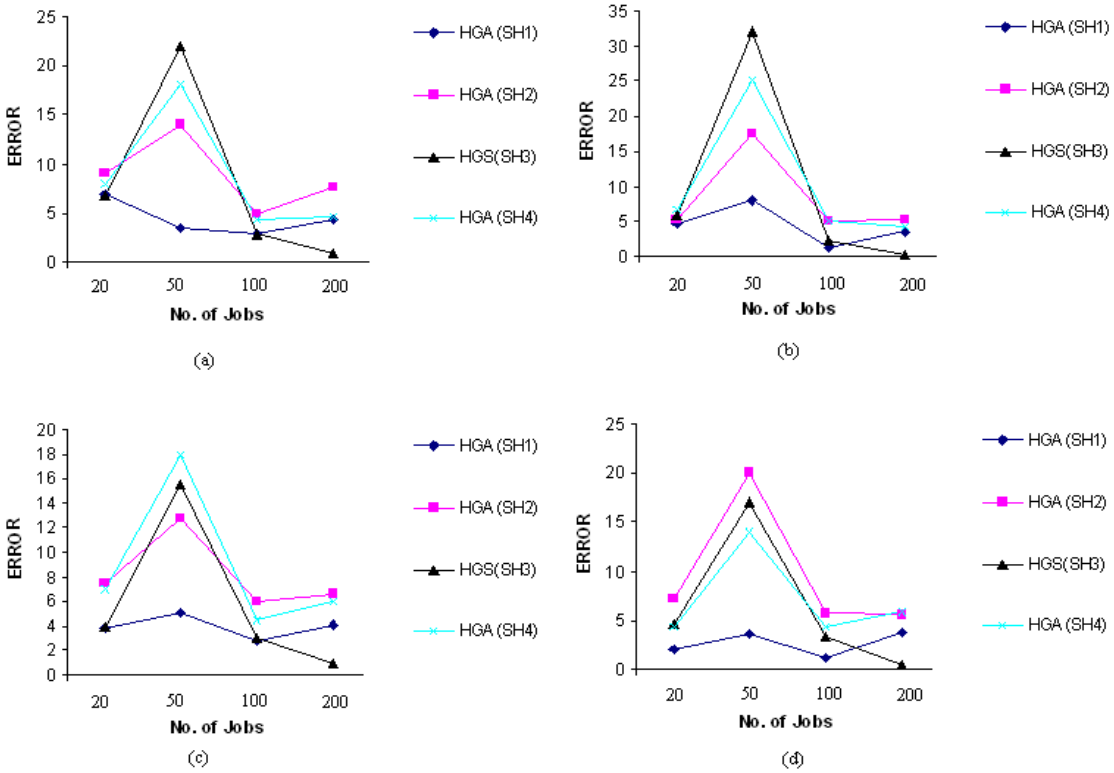


FIGURE 3: Error for twenty machines problem **(a)** α = 0.33, β = 0.33 & γ = 0.33 **(b)** α = 0.25, β = 0.25 & γ = 0.5 **(c)** α = 0.5, β = 0.25 & γ = 0.25 **(d)** α = 0.25, β = 0.25 & γ = 0.5

From the analysis for 5, 10 and 20 machines problems as shown in Figure 1, Figure 2 and Figure 3, performance of proposed SH1 based HGA upto 20 jobs and SH3 based HGA as the jobs size increases, showed superiority over others for all the four sets of weight values ((0.33, 0.33, 0.33), (0.25,0.25,0.5) , (0.5, 0.25, 0.25) and (0.25, 0.25, 0.5)) for multicriteria decision making in flow shop scheduling under sequence dependent set up time i.e. weighted sum of total tardiness, total earliness and makespan.

5. CONCLUSIONS

In the present work, we have framed multicriteria decision making for flow shop scheduling with weighted sum of total tardiness, total earliness and makespan under sequence dependent set up time and also proposed four special heuristic based hybrid genetic algorithms. Computational analysis has also been done for comparing the performance of proposed four HGA's i.e. HGA (SH1), HGA (SH2), HGA (SH3) and HGA (SH4). The HGA's have been tested upto 200 jobs and 20 machines problems in flow shop scheduling as derived by Taillord [20] for all the four weight values (α, β and γ) for fitness function (i.e. (0.33, 0.33, 0.33), (0.25, 0.25, 0.5), (0.5, 0.25, 0.25) and (0.25, 0.25, 0.5)). From the analysis it has been concluded that the proposed HGA(SH1) for smaller and HGA(SH3) for larger job size problems showed superiority over other for 5,10 and 20 machines problem for multicriteria decision making flow shop scheduling (i.e. weighted sum of total tardiness, total earliness and makespan) under sequence dependent set up time and due dates.

6. REFERENCES

1. Allahverdi A., Ng C.T., Cheng T.C.E. and Kovalyov M.Y. *"A survey of scheduling problems with setup times or costs".* European Journal of Operation Research, 187(3):985–1032,2008
2. Rajendran C. *"Theory and methodology heuristics for scheduling in flow shop with multiple objectives"* .European Journal of Operation Research , 82 (3):540 –555,1995
3. Ravindran D., Noorul Haq A., Selvakuar S.J. and Sivaraman R. *" Flow shop scheduling with multiple objective of minimizing makespan and total flow time".* International Journal of Advanced Manufacturing Technology, 25:1007–1012,2005
4. Gupta J.N.D., Venkata G., Neppalli R. and Werner F. *"Minimizing total flow time in a two-machine flowshop problem with minimum makespan".* International Journal of Production Economics, 69:323–338,2001
5. Sayin S. and Karabati S. *"A bicriteria approach to the two-machine flow shop scheduling problem".* European Journal of Operation Research, 113:435–449,1999
6. Danneberg D., Tautenhahn T., Werner F. *"A comparison of heuristic algorithms for flow shop scheduling problems with setup times and limited batch size".* Mathematical and Computer Modelling, 29 (9):101–126,1999
7. Toktas B., Azizoglu M. and Koksalan S.K. *"Two-machine flow shop scheduling with two criteria: Maximum earliness and makespan".* European Journal of Operation Research, 157(2): 286–295,2004
8. Ponnambalam S. G., Jagannathan H., Kataria M. and Gadicherla A. *"A TSP-GA multi-objective algorithm for flow shop scheduling".* International Journal of Advanced Manufacturing Technology, 23: 909–915,2004
9. Loukil T., Teghem J. and Tuyttens D. *"Solving multi-objective production scheduling problems using metaheuristics".* European Journal of Operation Research , 161(1) :42–61,2005
10. Fred Choobineh F., Mohebbi E. and Khoo, H. *"A multi-objective tabu search for a single-machine scheduling problem with sequence dependent setup times".* European Journal of Operational Research, 175, 318–337,2006
11. Rahimi Vahed R and Mirghorbani S. M. *"A multi-objective particle swarm for a flow shop scheduling problem".* Combinatorial Optimization, 13 (1):79–102,2007

12. Noorul Haq and Radha Ramanan T. *"A bicriterian flow shops scheduling using artificial neural network".* International Journal of Advanced Manufacturing Technology, 30: 1132–1138, 2006
13. Lockett A. G. and Muhlemann A. P. *"Technical notes: a scheduling problem involving sequence dependent changeover times".*Operation Research, 20: 895-902,1972
14. Gowrisankar K, Chandrasekharan Rajendran and Srinivasan G. *"Flow shop scheduling algorithm for minimizing the completion time variance and the sum of squares of completion time deviations from a common due date".* European Journal of Operation Research, 132:643–665,2001
15. Blazewicz J, Pesch E, Sterna M and Werner F. *" A comparison of solution procedures for two-machine flow shop scheduling with late work criterion".* Computers and Industrial Engineering, 49:611–624,2005
16. Eren, T. *"A bicriteria m-machine flow shop scheduling with sequence-dependent setup times".* Applied Mathematical Modelling,2009 (In press), doi: 10.1016/j.apm.2009.04.005
17. Erenay F.S., et al. *"New solution methods for single machine bicriteria Scheduling problem: Minimization of average flow time and number of tardy jobs".* European Journal of Operational Research,2009 (In press), doi:10.1016/j.ejor.2009.02.014
18. Naderi B., Zandieh M. and Roshanaei V. *"Scheduling hybrid flowshops with sequence dependent setup times to minimize makespan and maximum tardiness".* International Journal of Advanced Manufacturing Technology, 41:1186–1198,2009
19. Nawaz M., Enscore E. and Ham I. *"Heuristic Algorithm for the M-Machine N-Job Flow Shop Sequencing Problem"*. Omega, 11: 91–95,1983
20. Taillard E. *"Benchmarks of basic scheduling problems"*. European Journal of Operation Research, 64:278–285, 1993

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