Stochastic Identical Parallel Machine Scheduling to Minimize the Number of Risky Jobs

Conference Paper · May 2012


Gokhan Egilmez
University of New Haven


Gürsel A. Süer
Ohio University

Proceedings of the 2012 Industrial and Systems Engineering Research Conference
G. Lim and J.W. Herrmann, eds.

Stochastic Identical Parallel Machine Scheduling to Minimize the Number of Risky Jobs

Gökhan Eğilmez©, Can Çelikbilek and Gürsel A. Süer
Industrial and Systems Engineering Department
Russ College of Engineering and Technology
Ohio University
Athens, OH, 45701, USA
ge298307@ohio.edu©, cc340609@ohio.edu and suer@ohio.edu
Tel: (740) 593-1542, Fax (740) 593-0778

Abstract
Parallel machine scheduling is one of the widely studied problems in machine scheduling literature. In this paper, identical parallel machine scheduling problem is considered in the presence of probabilistic processing times and deterministic due dates. In classical scheduling problems, a job is called tardy if it is completed after its due date, otherwise it's called early. However, in probabilistic concept, a job has a probability of tardiness from 0 to 1. To capture this situation, jobs having probability of tardiness between 0 and 1 are called "risky". If the probability is equal to 0 or 1, job is called early or tardy, respectively. The objective is to minimize the number of risky jobs by keeping the number of tardy jobs minimum. A stochastic non-linear mathematical model is developed. Experimentation is performed with various problem sets jobs. As a result, the proposed model decreased the number of risky jobs and provided a safer schedule than the deterministic model.
Key words: stochastic scheduling, parallel machine, risky job, probability of tardiness

1. Introduction
Manufacturing scheduling is a critical decision making process in production environments. In manufacturing scheduling, mainly two types of systems are considered. *Deterministic scheduling* is a type of scheduling problem that provides certain information about the parameters such as processing times, due dates, release dates, etc. On the other hand, *stochastic scheduling* deals with problems when at least one of the parameters is not known with certainty [1]. Scheduling under stochastic conditions is highly complex and more difficult than deterministic problems. However stochastic scheduling reflects the reality more precisely.

Scheduling continues to be a popular research field in the industrial engineering world. This is evident from the amount of articles that have been published in the literature. A significant of research in the scheduling field has been addressed to the problem of minimizing penalties related with tardiness. Mainly single machine scheduling problems with NP hard have been studied with tardiness cost [2]. For example, Karp [3] studied the minimization of weighted number of tardy jobs and Du and Leung [4] studied the minimization of total tardiness in deterministic scheduling.

Many researchers have studied the stochastic single machine scheduling problem with the objective of minimizing the expected number of tardy jobs [5]. Soroush and Fredenhall [5] studied single machine stochastic scheduling in order to analyze the impact of varying processing times for earliness-tardiness costs. Moreover, Balut [6] studied single machine scheduling problem with normally distributed processing times and different due dates in order to minimize the number of tardy jobs. Similarly, Pinedo [7] discussed a study which incorporates independent exponential processing times and Jang [8] studied a similar problem in which processing times follow normal distribution. Moreover, Lin and Lee [9] considered stochastic single machine scheduling problem with known distributions of random processing times and due dates. The objective of the study was to formulate three different models to minimize lateness. Seo, Klein and Jang [10] studied stochastic single machine scheduling problem with normal processing time of the jobs and deterministic due dates and developed a non-linear integer programming model in order to minimize the expected number of tardy jobs. Sarin, Erel and Steiner [11] studied stochastic single machine problem which jobs have stochastic normal processing times and common due dates. The objective of this study was to minimize the expected incompletion cost.

Parallel machine stochastic scheduling has been also addressed in various works in stochastic machine scheduling literature. Pinedo and Emmons [12] studied stochastic scheduling of jobs that have random processing times, random due dates and weights that are assigned to parallel machines in order to minimize the expected total weight of tardy jobs. In that study, optimal deterministic schedules and optimal preemptive and non-preemptive dynamic schedules were proposed.

Moreover, Weiss [13] developed various new formulae for stochastic processing times of the jobs on parallel machines for expected flow time and weighted flow time. Along with these formulae Smith's Rule is used as a meta-heuristics approach for this problem. Weber [14] discussed about the minimization of makespan or flow time by considering the priority of a job. In this study, the processing times follow exponential distribution. Forst [15] also studied bicriteria stochastic scheduling in a single machine in order to minimize the sum of expected weighted tardiness and the expected total weighted flowtime.

This paper discusses identical parallel machine stochastic scheduling problem with jobs having a deterministic due date and stochastic processing times in order to minimize the number of risky jobs by keeping the number of tardy jobs minimum. In classical scheduling problems, a job is said to be tardy when the completion time C_i of a job is greater than its due date D_i. The tardiness is measured by C_i-D_i. If a job is not tardy, the tardiness value is adjusted to 0. Therefore, the tardiness of a job is shown as max (C_i-D_i, 0). Jobs with a probability of tardiness value between 0 and 1 are called "risky". The objective of this study is to develop a stochastic non-linear mathematical model that minimizes the number of risky jobs in a parallel machine scheduling environment.

The literature review that has been done so far shows that, the problem of this study will be unique in the sense of developing a stochastic non-linear mathematical model that achieves minimum number of risky jobs in a parallel machine scheduling environment. The section 2 briefly discusses the problem definition and statement, section 3 discusses the methodology and the developed mathematical models. Then, section 5 mentions the experimentation results and finally in section 6, concluding remarks about this study and future work will be stated.

2. Problem Statement

Minimizing the expected number of tardy jobs is one of the most studied problems in stochastic scheduling literature. According to Pinedo [7], minimizing the expected number of tardy jobs on a single machine is only traceable with exponentially distributed processing times. The scheduling problems become more complex when other processing time distributions are used. In this study, there are n jobs with deterministic due dates to be scheduled on m parallel machines and each job *j* has the processing time which is assumed to be normally distributed random variable with mean (μ_j) and standard deviation (σ_j) along with deterministic due date (D_j).

Assume that job *i* is scheduled on machine *m*. The expected completion time of job i scheduled on k^{th} position on machine m is $C_{[k,m]}$ as shown in Equation 1. The variance of k^{th} job on machine m is shown in Equation 2.

$$C_{[k,m]} = \sum_{j=1}^{k} \mu_{[j]} \tag{1}$$

$$\sigma^2{}_{[k,m]} = \sum_{j=1}^{k} \sigma^2{}_{[j]} \tag{2}$$

By using the expected completion time and variance information, the probability of tardiness for job *i* scheduled on k^{th} position on machine m is $PrT_{[k,m]}$ is shown in Equation 3. If the probability of tardiness is equal to 0 (a probability value which is less than 0.001 is assumed as 0), the corresponding job is called "early". If the probability value is equal to 1, it is called "tardy". If the probability value is between 0 and 1, the job is classified as "risky" since it has a positive risk of being tardy.

$$PrT_{[k,m]} = p\left(Z_{[k,m]} \leq \frac{(C_{[k,m]} - D_j)}{\sqrt{\sigma^2{}_{[k,m]}}}\right) \tag{3}$$

Minimizing the number of tardy jobs is considered in majority of stochastic scheduling works. However, minimizing the number of risky jobs along with tardy jobs is more crucial since the probability of tardiness is greater than 0 for a risky job. In this paper, the problem studied is sequencing n jobs on m parallel machines such that the number of risky jobs is minimized.

3. Methodology

A stochastic non-linear mathematical model is developed to minimize the number of risky jobs in a parallel machine environment. Chance-constraint programming is used as transformation method.
The general representation of chance-chance constraint programming is as follows [16]:

Objective function:

$$\text{Maximize } z = \sum_{j=1}^{n} c_j * x_j \tag{4}$$

Subject to

$$P\left\{\sum_{j=1}^{n} a_{ij}x_j \leq b_i\right\} \geq (1 - \alpha_i), i = 1, 2 \dots m;\ x_j \geq 0, \text{for all j} \tag{5}$$

The objective function of general representation of problem is given in Equation 4. And the transformation of stochastic constraint is given in Equation 5. There are three cases the transformation method requires different treatment in terms of independency between a_{ij} and b_i.

Case 1) a_{ij} is normally distributed and b_i is constant
Case 2) a_{ij} is constant and b_i is normally distributed
Case 3) a_{ij} and b_i are normally distributed and independent.

Case 1 fits to the approach used in this study since processing times are normally distributed and due dates are constant. To illustrate case 1, the following example is given.

Example:
An example problem is derived to illustrate the transformation performed with stochastic programming. Consider following chance constraint problem

$$\text{Maximize } z = 6 * x_1 + 2 * x_2 + x_3 \tag{6}$$

subject to

$$P\{a_{11} * x_1 + a_{12} * x_2 + a_{13} * x_3 \leq 11\} \geq 0.95 \tag{7}$$

with all $x_j \geq 0$.

Suppose that a_{1j}'s are independent normally distributed random variables with the following means and variances:

$E\{a_{11}\} = 2$, $E\{a_{12}\} = 4$, $E\{a_{13}\} = 6$, $Var\{a_{11}\} = 16$, $Var\{a_{12}\} = 9$, $Var\{a_{13}\} = 25$

From standard normal distribution tables, $K_{\alpha=0.05} = 1.645$. The stochastic constraint can be transformed into following equivalent deterministic constraint:

$$2 * x_1 + 4 * x_2 + 6 * x_3 + 1.645 * \sqrt{16 * x_1^2 + 9 * x_2^2 + 25 * x_3^2} \leq 11 \tag{8}$$

The stochastic problem is transformed into deterministic equivalent and the proposed mixed-integer non-linear mathematical model is solved with Lingo. The notation is as follows:

Notation:

Indices:

j	*Job index*
k	*Position index*
m	*Machine index*

Parameters:

n	*Total number of jobs*
M	*Number of machines*
μP_j	*Mean process time of job j*
σPj^2	*Variance of process time of job j*
D_j	*Due date of job j*

Decision Variables:

X_{jmk}	*1 if job j is assigned to the k^{th} position on machine m, 0 otherwise.*
C_{mk}	*Completion time of the job in k^{th} position in machine m*
T_{mk}	*Tardiness value of the job in k^{th} position in machine m*
nT_{mk}	*1 if job in k^{th} position in machine m is tardy, 0 otherwise*
PrT_{mk}	*Probability of tardiness value of the job in k^{th} position in machine m*

The objective is to minimize the number of tardy jobs and the total probability of tardiness as shown in Equation 9. The reason both the number of tardy jobs *(nT)* and total probability of tardiness are kept in the objective function is that if *nT* is omitted, the number of tardy jobs might increase while reducing total probability of tardiness. Equation 10 guarantees that each job is assigned to a position on a machine. Equation (11) guarantees that each position in each cell can be assigned at most one job. Jobs are enforced to be assigned consecutively by Equation (12). Equations (13.a) and (13.b) determine the completion time of the job in that position. Equation (14) calculates the tardiness value of a job. Equation (15) calculates the probability of tardiness (*PrT*) for a job. Finally, Equation (16) counts the tardy jobs.

Objective Function:

$$\min Z = \sum_{m=1}^{M} \sum_{k=1}^{n} (nT_{mk} + PrT_{mk}) \tag{9}$$

Subject to:

$$\sum_{m=1}^{M} \sum_{k=1}^{n} X_{jmk} = 1 \quad j = 1 \dots n \tag{10}$$

$$\sum_{j=1}^{n} X_{jmk} \leq 1 \quad \text{for } m = 1 \dots M \;\; k = 1 \dots n \tag{11}$$

$$\sum_{j=1}^{n} X_{ijmk} \geq \sum_{j=1}^{n} X_{jm(k+1)} \qquad \text{for } m = 1 \dots M, k = 1 \dots n-1 \tag{12}$$

$$C_{m1} = \sum_{j=1}^{n} X_{jm1} * \mu P_j \quad m = 1 \ldots M \tag{13.a}$$

$$C_{mk} - C_{m(k-1)} \geq \sum_{j=1}^{n} X_{jmk} * \mu P_j \qquad for\ m = 1 \ldots M, k = 2 \ldots n \tag{13.b}$$

$$C_{mk} - \sum_{j=1}^{n} X_{jmk} * D_j \leq T_{mk} \qquad for\ m = 1 \ldots M, \quad k = 1 \ldots n \tag{14}$$

$$PrT_{mk} = p\left(Z_{mk} \leq \frac{\left(C_{mk} - \sum_{j=1}^{k} X_{jmk} * D_j \right)}{\sqrt{\sum_{j=1}^{k} X_{jmk} * {\sigma_{p_{ij}}}^2}} \right) \qquad for\ m = 1 \ldots M, \quad k = 1 \ldots n \tag{15}$$

$$T_{mk} \leq R * nT_{mk} \qquad for\ m = 1 \ldots M, \quad k = 1 \ldots n \tag{16}$$

Definition of Variables:
$X_{jmk} \in \{0,1\}, W_{mk} \in \{0,1\}, Y_{mk} \in \{0,1\},\ nT_{mk} \in \{0,1\}, C_{mk} \geq 0,\ T_{mk} \geq 0, 0 \leq PrT_{mk} \leq 1$

4. Experimentation and Results

Experimentation is performed with datasets with 12 jobs and 18 jobs. Lingo 13.0 64mb version is used for the experimentation. Jobs are assumed as independent of each other. Each job has probabilistic processing time and deterministic due date. The processing time of each job is normally distributed. The standard deviation for each job is taken as 10 % of mean processing time. Deterministic model is used as benchmark to illustrate the impact of considering processing time variance in scheduling. Deterministic model can be obtained as excluding the stochastic constraint (Constraint 10) from the proposed model. Mean processing times are used in the experimentation with deterministic model. Both of the results from deterministic and proposed stochastic models are compared. Since there is no probabilistic concept in deterministic model, the tardiness probability for each job is calculated manually based on the obtained optimal schedule from deterministic approach. After the optimal schedule is obtained and probability of tardiness for each job is calculated, jobs are classified as three groups, namely: “early”, “risky” and “tardy”. Jobs are classified “early” if the probability of tardiness is 0 (the probability value that is less than 0.001 is assumed as 0). If the probability of tardiness is between 0 and 1, the corresponding job is classified as “risky” job. If the probability of tardiness is 1, the job is called “tardy”.

4.1. Results of Experiment with Problem 1

Mean and variance of processing times and due dates are shown in Table 1. In this experiment, there are 12 independent jobs to be processed on two parallel machine shop floor configuration. The results are given in detail in this experimentation for the ease of understanding the concept.

Table 1: Data – 12 Jobs on 2 Parallel Machines

Job	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Mean Processing Time	198	216	104	135	141	234	214	126	115	214	126	115
Variance of Processing Time	392	467	108	182	199	548	458	159	132	458	159	132
Due Date	230	312	461	655	434	914	459	493	536	571	817	952

The optimal schedule obtained from deterministic model is shown in Table 2. According to Table 2, there are 6 jobs classified as early, 4 jobs classified as risky and 2 jobs classified as tardy. On the other hand, according to the optimal schedule obtained from stochastic model (given in Table 3), there are 2 risky, 2 tardy and 8 early jobs obtained. The comparison is shown in Table 4 and illustrated in Figure 1.

Table 2: The Optimal Schedule Obtained from Deterministic Model

Machine - 1	**2**	**7**	**6**	**12**	**1**	**10**
Processing Time	216	214	234	115	198	214
Due Date	312	459	914	952	230	571
Completion Time	216	430	664	779	977	1191
Type of Job	Early	**Risky**	Early	Early	**Tardy**	**Tardy**
Probability of Tardiness	0.00	**0.17**	0.00	0.00	**1.00**	**1.00**
Machine - 2	**3**	**5**	**9**	**8**	**4**	**11**
Processing Time	104	141	115	126	135	126
Due Date	461	434	536	493	655	817
Completion Time	104	245	360	486	621	747
Type of Job	Early	Early	Early	**Risky**	**Risky**	**Risky**
Probability of Tardiness	0.00	0.00	0.00	**0.39**	**0.11**	**0.01**

Table 3: Optimal Schedule Obtained from Stochastic Model

Machine - 1	**2**	**10**	**4**	**6**	**12**		
Processing Time	216	214	135	234	115		
Due Date	312	571	655	914	952		
Completion Time	216	430	565	799	914		
Type of Job	Early	Early	Early	Early	**Risky**		
Probability of Tardiness	0.00	0.00	0.00	0.00	**0.18**		
Machine - 2	**3**	**5**	**8**	**9**	**11**	**7**	**1**
Processing Time	104	141	126	115	126	214	198
Due Date	461	434	493	536	817	459	230
Completion Time	104	245	371	486	612	826	1024
Type of Job	Early	Early	Early	**Risky**	Early	**Tardy**	**Tardy**
Probability of Tardiness	0.00	0.00	0.00	**0.02**	0.00	**1.00**	**1.00**

Table 4: Comparison of deterministic and stochastic approach

Comparison Subject	**Deterministic Approach**	**Stochastic Approach**
The number of early jobs (*nE*)	6	8
The number of risky jobs (*nR*)	4	2
The number of tardy jobs (*nT*)	2	2
Total Probability of Tardiness *(Total PrT)*	2.68	2.20

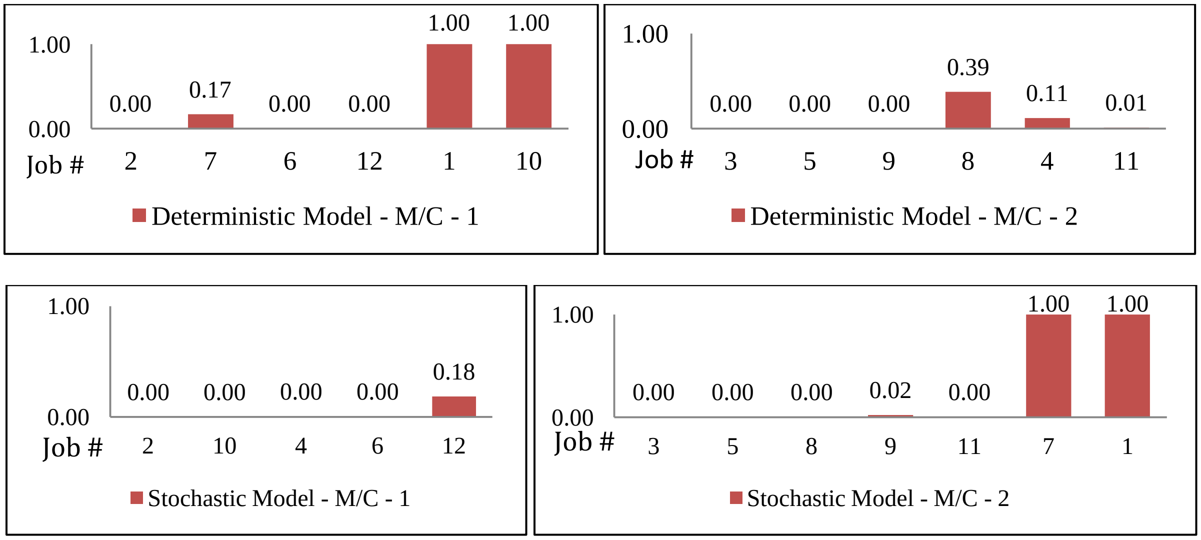

Figure 1: Comparison in terms of probability of tardiness

According to the results, the proposed stochastic approach decreased the number of risky jobs and kept the number of tardy jobs the same thus increased the number of early jobs. The total probability of tardiness decreased from 2.68 to 2.20.

4.2. Results of Experiment with Problem 2

In this experiment, the number of jobs increased to 18. The optimal schedules obtained from both deterministic and stochastic approaches are shown in Figure 2.

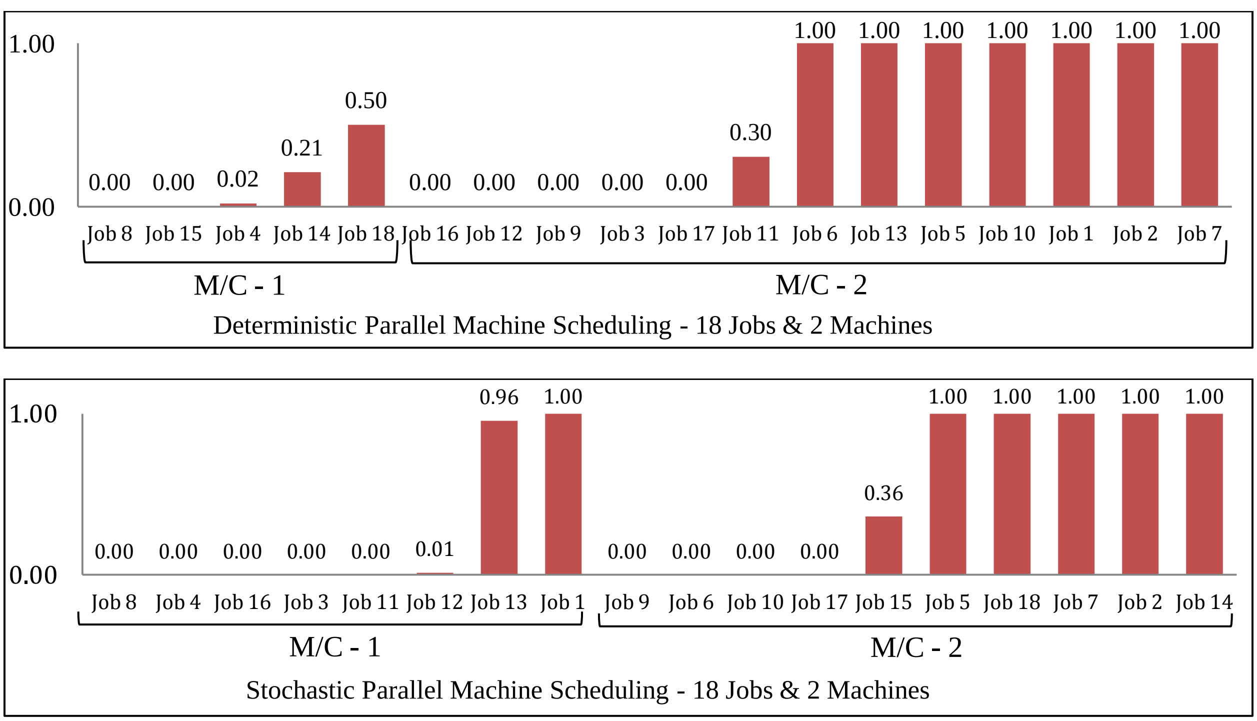

Figure 2: Comparison in terms of probability of tardiness

According to the results of deterministic approach, there are 7 tardy, 4 risky and 7 early jobs obtained. The probability of tardiness values of risky jobs vary from 0.02 to 0.5. On the other hand, the proposed stochastic approach provided 6 tardy, 3 risky and 9 early jobs. The comparison is given in Table 5. According to Table 5, the risky jobs decreased from 4 to 3 and tardy jobs decreased from 7 to 6. The number of early jobs increased from 7 to 9. Total probability of tardiness is decreased from 8.03 to 7.33.

Table 5: Comparison of deterministic and stochastic approach

Performance Measure	Deterministic Approach	Proposed Stochastic Approach
nE	7	9
nR	4	3
nT	7	6
Total PrT	8.03	7.33

5. Conclusion

In this paper, a stochastic parallel machine scheduling problem is addressed. Jobs are assumed to have deterministic due dates and normally distributed processing times. The objective is to minimize the number of risky jobs by keeping the number of tardy jobs minimum. A new classification of jobs is introduced. Jobs are grouped into three, namely: tardy, risky and early depending on the probability of tardiness. If the probability of tardiness is equal to 1, corresponding job is called “tardy”. On the other hand, a job is called “early” if the probability of tardiness is equal to 0. Jobs with probability of tardiness between 0 and 1 are called “risky” due to carrying a positive risk of tardiness. To assume that a job is early, a threshold value of 0.001 is used and if the probability of tardiness is less than the threshold, it is assumed as 0 thus job is called “early”.

A stochastic non-linear mathematical model is developed and experimented with Lingo optimization software on a dual-core 1.73 Ghz and 2 Gb ram desktop computer. Two datasets are used in the experimentation: 12 jobs and 18 jobs on a two parallel machine shop. The proposed model is compared with deterministic model. In both of the configurations, proposed approach decreased the number of risky jobs. The number of tardy jobs is kept the same in 12-job dataset and decreased by 1 in 18-job dataset. Therefore, the number of early jobs increased and the total probability of tardiness reduced significantly in both experiments. Due to the computational complexity, larger datasets could not be experimented. The main contribution of this approach is that scheduler is able to view the individual probability of tardiness for each job in an optimal sequence obtained via the proposed approach. Based on the probability of tardiness values, risky jobs can be identified and due dates can be modified to reduce the probability of tardiness to satisfactory levels. Due date modification heuristics, other performance measures such as total tardiness, the maximum tardiness etc., job splitting and metaheuristics are other directions that current research can be extended. These directions are left as future work.

References

1. Sarin, Subnash C., Balaji Nagarajan, and Lingrui Liao. 2010, “Stochastic Scheduling Expectation-Variance Analysis of a Schedule”. New York: Cambridge University Press.
2. W. Jang and C. M. Klein, 2002,“Minimizing the expected number of tardy jobs when processing times are normally distributed,” *Operations Research Letters*, vol. 30, no. 2, pp. 100-106.
3. R.M. Karp, 1972, “Reducibility among combinatorial problems”, in: R.E. Miller, J.W. Thatcher (Eds.), *Complexity of Computer Computations*, Plenum Press, New York, pp. 85–103.
4. J. Du, J.Y. Leung, 1990, “Minimizing total tardiness on one processor is NP-hard”, Math. Oper. Res. 3, 483–495.
5. H.M. Soroush, L.D. Fredendall, 1994,“The stochastic single machine scheduling problem with earliness and tardiness costs”, *European Journal of Operational Research* 77, 287–302.
6. S. J. Balut, 1973, “Scheduling to minimize the number of late jobs when set-up and processing times are uncertain”, *Management Science*, vol. 19, no. 11, pp. 1283-1288.
7. M. Pinedo, 1983, “Stochastic scheduling with release dates and due dates”, *Operations Research*, vol. 31, no. 3, pp. 559–572.
8. W. Jang, 2002, “Dynamic scheduling of stochastic jobs on a single machine”, *European Journal of Operational Research* 138, pp. 518–530.

9. Lin, C., & Lee, C. 1995, "Single-machine stochastic scheduling with dual criteria". *IIE Transactions* 27, 244–249.
10. D. Seo, C. Klein, and W. Jang, 2005, "Single machine stochastic scheduling to minimize the expected number of tardy jobs using mathematical programming models", *Computers & Industrial Engineering*, vol. 48, no. 2, pp. 153-161.
11. Subhash C. Sarin, Erdal Erel, George Steiner, 1991, "Sequencing jobs on a single machine with a common due date and stochastic processing times", *European Journal of Operational Research*, Volume 51, Issue 2, Pages 188-198.
12. M.Pinedo, H.Emmons, 1990, "Theory and Methodology Scheduling stochastic jobs with due dates on parallel machines", *European Journal of Operational Research*, vol. 47, pp. 49-55.
13. G. Weiss, 1990, "Approximation results in parallel machines stochastic scheduling," *Annals of Operations Research*, vol. 26, no. 1, pp. 195–242.
14. R. Weber, 1982, "Scheduling Jobs with Stochastic Processing Requirements on Parallel Machines to Minimize Makespan or Flowtime," *Journal of Applied Probability*, vol. 19, no. 1, pp. 167-182.
15. F. G. Forst, 1995, "Bicriterion stochastic scheduling on one or more machines," *European journal of operational*
16. Taha, H., A. (1996). Operations Research: An Introduction (p. 916). New Delhi: Prentice Hall; 6 Har/Dis edition.