Iranian Journal of Operations Research
Vol. 12, No. 1, 2021, pp. 109-126

Implementation of clustering on a multi objective single machine batch scheduling problem (A case study in Iran)

H. Haddad [1*]

Batch scheduling is among the important problems in industrial engineering and has been widely attendant in practical applications. Clustering is the set of observation assignment into some subsets so that the observations in the same cluster are similar in some sense and the similarity of generated clusters is very low. Clustering is considered as one of the approaches in unsupervised learning and a common technique for statistical data analysis which has been applied in many fields, including machine learning, data mining and etc. This paper studies a case study in Iran Puya company (as a home appliance maker company in Iran). In the production line of refrigerator of the current company, a cutting machine is identified as a bottleneck that can process several iron plates simultaneously. In this regard a good scheduling on this cutting machine improves the effectiveness of production line in terms of cost and time. The objective is to minimize the total tardiness and maximizing the job values when the deteriorated jobs are delivered to each customer in various size batches. Based on these assumptions a mathematical model is proposed and two hybrid algorithms based on simulation annealing and clustering methods are offered for solving it and the results are compared with the global optimum values generated by Lingo 10 software. Based on the effective factors of the problem, a number of sensitivity analyses are also implemented including number of jobs and rate of deterioration. Accordingly, the running time grows exponentially when the number of jobs increases. However the rate of deterioration could not affect the running time. Computational study demonstrates that using clustering methods leads an specified improvements in total costs of company between 15 to 41 percent.

Keywords: *Batch scheduling; single machine; deterioration; job values; clustering*

Manuscript was received on 11/28/2021, revised on 01/12/2022 and accepted for publication on 01/25/2022.

1. Introduction

Batch processing is considered as one of the unsupervised approaches so as to improve the production planning and organization in which jobs could be sent to consumer nodes individually or be held in production unit to be sent by more jobs. In this regard a real case study in Iran is reviewed in Iran Puya home appliance company which manufactures wide ranges of appliances involving oven, refrigerator and etc. According to production planning reviews, a bottleneck is identified on the place of a cutting machine that processes n jobs as a single machine. The objective is to minimize the total tardiness and delivery costs and maximizing the job values, where all of jobs are delivered to some customers within some batches. Each job may be transmitted to its customer right after its completion or be held to be sent with more jobs. Batching could accelerate the job process and decrease the delivery costs to customers. In order to accomplish the batching level as a systematic approach, some clustering methods involving Hierarchical clustering and K-Means algorithm are employed and the present jobs are batched based on the similarity of processing times and due dates. The main advantage of the proposed action is that the required changes of machine for processing different jobs will be minimized and the jobs with common due dates will be sent to consuming nodes with lower values of lateness. The remaining parts of this paper

[1] Department of Management, Torbat -e Heydariyeh Branch, Islamic Azad University, Torbat-e Heydariyeh, Iran (Email: anooshomidi@gmail.com)

are as follows: in section 2 the literature of problem is reviewed, in section 3 the proposed model is presented and the variables and parameters are introduced. In section 4 the solution approach is offered and in section 5 the computational studies are presented.

2. Literature Review

Single machine scheduling problem have a rich literature in the field of production scheduling and several researches have been done by considering several objectives and limitations. Tang et al. (2017) studied a two-agent scheduling problem with deteriorating jobs on a single machine when jobs are processed on machine within some batches. They considered two objective functions including minimizing makespan for an agent and number of tardy jobs for another agent. They considered some bounded and unbounded cases for the problem and proposed some optimal algorithms so as so solve them. The proposed method was able to achieve the global optimal solutions in short run times. Zhou et al. (2020) presented a single machine batch scheduling when the arrival times for jobs were dynamic and the time of use policy for electricity was considered for proposing a model with efficient energy consumption. Based on their research the mismatch between demand and supply is a challenging issue. In this regard a demand response approach was used for aligning the demand and supply of electricity. They presented some heuristic approaches for finding near optimal solutions in reasonable run times. Rostami (2021) proposed a new problem in this field involving capacitated batch delivery and pickup systems. In this regard a comprehensive mathematical model was presented for production scheduling, delivering the products to multiple customers and picking up the end of life products in order to recycle and reused in manufacturing plants. A three-staged branch and bound method was employed so as to solve the instances optimally. Zhang et al. (2021) studied a generalized single machine batch scheduling problem with just in time consideration so as to minimize the earliness and tardiness costs simultaneously. They investigated some intrinsic properties of optimal solution and proposed a span limited tree search for achieving near optimal solutions. Zheng et al. (2021) studied the problem of single machine batch scheduling with dual setup times. They stated that setup time depends not only on batch size also on technological characteristics and also considered two types of setup costs including quantitative and technological inspired from autoclave modeling production. They presented a MILP model with objective of makespan minimization. Then a two stage approximation algorithm was established for solving the problem. Another related research in this field is Pei et al. (2017) in which the sequence setup time and learning effect were considered.
According to literature of scheduling, the problems those combine scheduling and delivery costs are complex and achieving global optimal solutions is impossible using traditional optimization methods (Mazdeh et al. (2011 a)).

Based on how the batches are processed on the machine, the single machine batch scheduling problem would considered as serial or parallel batching. In serial batching, the processing time of each batch is equal to the completion time of the last job assigned to it. Ng et al. (2002 b) considered a single machine serial batch scheduling to minimize maximum lateness with precedence constraints between the batches and proposed an o(n^2) polynomial time algorithm. They (2003 a) also considered the same problem with objective of total completion time, release date and precedence constraints and presented an o(n^5) polynomial algorithm.

Yuan et al. (2007) studied a serial batch scheduling when each batch contained exactly k jobs and the objective was to minimize the total weighted completion time. They proved that the problem is NP-hard and presented a polynomial algorithm with some simplified rules. Another research is related to Ishii et al. (2010) where the size of each batch was considered as a fuzzy number and three objectives including makespan, maximum flow time and the satisfaction level were considered. A Lagrange relaxation was proposed for the problem that generated non-dominant solutions. Cheng et al. (2001 b) presented a linear programming model for the problem by considering the resource dependent setup and processing times so as to minimize the total weighted resource consumption and achieved the optimum solution in polynomial time. Melouk et al. (2004) proposed a simulation annealing approach for minimizing the makespan with limited machine capacity and compared their results with CPLEX solver. Based on their comparison, SA overcame the results of CPLEX in all instances. Cheng and Kovalyov (2000) accomplished an extensive research in the field of regular cost functions in single machine batch scheduling and considered several objectives including number of late jobs, maximum lateness, total tardiness, total weighted completion time and total weighted tardiness and classified the complexity of them. Furthermore, a dynamic programming approach was proposed for solving the objectives optimally.

Ng et al. (2004 c) investigated the problem with two external resources which were used to compress the setup and job processing times linearly. The objective function was to minimize the total job completion time and the problem was solved in polynomial time.

On the other hand, the processing time of a parallel batch is calculated by the maximum processing time of jobs that are apparent to that batch. Nong et al. (2008 a) studied a single machine parallel batching problem to minimize makespan with consideration of family setup costs and release date. A polynomial approximation scheme was developed by some simplification rules and an algorithm with worst case ratio of 2.5 was proposed. An extensive review in the field of parallel batching literature were studied by Potts and Kovalyov (2000).

In recent years, researchers have considered the delivery cost as a secondary objective called "combined optimization batch delivery problem". Mahdavi et al. (2011 a) studied the problem of minimizing the number of weighted tardy jobs with delivery cost and used a simulation annealing approach to solve the problem. Tian et al. (2007) studied the problem of minimizing the sum of total weighted flow times with delivery cost. They proved that the problem is NP-hard and presented an optimal algorithm. Mahdavi et al. (2011 c) presented a branch and bound approach to solve the problem of minimizing the weighted sum of flow times with delivery cost. Another research was related to Mahdavi et al. (2007 b) which presented a branch and bound to minimize flow time and delivery cost simultaneously.

Deterioration means that the processing time of jobs increases over time. This assumption is widely applied in scheduling problems. Browne and Yechiali (1990) were first researchers which presented the concept of deterioration in which the processing time of a job was a function of its starting time. Mosheiov (1994), who was the first to apply simple linear deteriorating for single machine scheduling problem, considered this problem to minimize the makespan, total completion time, total weighted completion time, total weighted waiting time, and total tardiness, number of tardy jobs, maximum lateness and maximum tardiness.

Wu et al. (2009) presented a branch-and-bound scheme and two heuristic algorithms in a single-machine to minimize the makespan. Wang et al. (2009) solved the single machine scheduling problems with learning effect and deteriorating jobs to minimize the total weighted completion time and the maximum lateness. Wang and Wang (2010) considered a single machine scheduling problem with deteriorating jobs in which the processing time of a job was defined as a simple linear function of its starting time to minimize the total weighted earliness penalty subject to no tardy jobs. For more researches in the field of using deterioration in scheduling problems see Husang et al. (2010), Cheng and Ji (2010) and AL-Anzi et al. (2007). For more studies in terms of application of scheduling in real industry problems see Niksirat (2018) and Musavi et al. (2017).

Table 1- the review of current studies

Authors	Objective function	Solution approach
Tang et al. (2017)	Makespan and number of tardy jobs	Optimal algorithms
Zhou et al. (2020)	Makespan and electricity consumption	Particle Swarm Optimization
Rostami (2021)	Maximum lateness	Branch and Bound
Zhang et al. (2021)	Makespan and delivery costs	Tree search and genetic algorithm

The current researches in the field of single machine batch scheduling have been addressed in above table. In this regard the main distinction of our research in comparison to mentioned articles could be stated as following:

- In all the mentioned papers, the batching procedure is implemented as a random procedure so as to optimize the objective functions. However, in this paper this issue is done as a unsupervised systematic procedure using clustering and data science approaches.
- The proposed objective function in this paper is unique and there is not any related research which consider tardiness, job values in makespan and delivery costs simultaneously.

In this paper, we consider a batch scheduling problem, where the objective is to minimize the total tardiness and delivery costs and maximizing the job values simultaneously which based on our knowledge, is not considered in the literature. Furthermore, the real data sets are used for solving the problem based on the information of Iran Puya company. Afterwards, two hybrid algorithms based on simulation annealing and clustering algorithms are offered to find the near optimal solutions. In order to check the verification of SA, some data sets of problem are solved optimally, using Lingo software, and the results are compared with each other. Some sensitivity analyses are also done based on important factors of the problem.

The main highlights of the current research could be mentioned as follows:

- Reviewing and modelling a real case study in Iran (Iran Puya company) so as to improving the effectiveness of production line
- Implementation of clustering to forming the batches (Based on our knowledge this item is not considered in literature so far.)
- Presenting a mathematical model based on deterioration and job values in makespan
- Improvement the production and transmission costs between 15 till 41 percent.

3. Problem formulation

Iran Puya is one of existing home appliance makers in Iran that manufactures several type of appliances like oven, refrigerator and etc. In this research the production line of refrigerator have been studied and in this regard a bottleneck was identified in the place of a cutting machine. This machine processes n jobs with no preemption and all of the jobs are available for processing at time zero. The completed jobs can be delivered to the customer immediately after completion or be awaited next jobs to be delivered as a batch. According to the number of jobs, N batches are considered and the jobs are assigned to the batches. Clearly, any batch that does not have any jobs will be omitted and then the sequence of batches is determined so that the proposed objective function is minimized.

The decision variables and parameters of the problem are listed below:

N	Number of jobs ready to be scheduled
p_i	The normal processing time of job i
d_i	The due date of i-th job
$p_{bat}(k)$	The processing time of batch k
$D_{bat}(k)$	The due date of batch k
n_k	Number of jobs assigned to batch k
t_j	The tardiness of job where scheduled on j-th position
c_i	The completion time of job where scheduled on j-th position
α	The rate of deterioration
B	The decreasing rate of job value
s_k	The shipping cost of batch k
$y_{ik} = 0 \; or \; 1$	It is 1 if job i is assigned to batch k and otherwise it is 0
$x_{kj} = 0 \; or \; 1$	It is 1 if batch k is located on jth position in the sequence and Otherwise it is 0
$a_k = 0 \; or \; 1$	It is 1 if there is at least one job assigned to batch k and Otherwise it is 0

And the proposed model is presented as follows:

$$\min z = \sum_k s_k a_k + \sum_k t_k - \sum_k e^{-\beta c_k} \tag{1}$$

St:

$$c_j = c_{j-1} + \sum_k x_{kj}(p_{bat}(k).\, j^{\alpha} \qquad j=1,2,\ldots,N \tag{2}$$

$$c_0 = 0 \tag{3}$$

$$p_{bat}(k) = \max_i(y_{ik}.\, p_i) \qquad K=1,2,\ldots,N \tag{4}$$

$$n_k = \sum_i y_{ik} \qquad K=1,2,\ldots,N \tag{5}$$

$$t_j = max\left\{0, c_j - \sum_k D_{bat}(k)\, x_{kj}\right\} \qquad j=1,2,\ldots,N \tag{6}$$

$$D_{bat}(k) = \max_i(y_{ik}.\, d_i) \qquad n_k \neq 0,\; K=1,2,\ldots,N \tag{7}$$

$$\sum_j x_{kj} = 1 \qquad k=1,2,\dots,N \qquad (8)$$

$$\sum_k x_{kj} = 1 \qquad j=1,2,\dots,N \qquad (9)$$

$$\sum_k y_{ik} = 1 \qquad i=1,2,\dots,N \qquad (10)$$

$$x_{kj} = 0 \; or \; 1 \qquad (11)$$

$$y_{ik} = 0 \; or \; 1 \qquad (12)$$

Equation (1) introduces the objective function where the first term corresponds to the delivery cost, the second term corresponds to the sum of tardiness and the third term is related to the sum of job values. Constraint (2) states that the completion time of each batch is equal to the completion time of prior batches plus the sum of processing times of jobs assigned to it. Constraint (3) mentions that the machine and all the jobs are available at time zero. Constraint (4) declares that the completion time of each batch is equal to the sum of processing times of jobs assigned to it. Constraint (5) determines how many jobs are located on each batch. Constraint (6) clarifies that the tardiness of each batch is equal to the gap between the time that the processing of that batch has been completed and its due date. Clearly, the value of tardiness must be positive. Equation (7) shows how the due date of each batch is calculated. Constraint (8) states that each batch is processed only once at each sequence and constraint (9) determines each position can be assigned to just one batch. Constraint (10) mentions each job can be assigned to exactly one batch to be processed and constraint (11), (12) shows that x and y are binary variables.

The following items could be summarized as the novelty of proposed mathematical model:

As above mentioned, the literature of single machine batch scheduling is rich and several researches have been studied the problem with variant objectives and constraints. However, the current research presents a real industry case study for production planning in Iran Puya company so as to achieve the managerial interests and objectives.

The current model minimizes two objectives including tardiness and delivery costs and maximizes job values in makespan simultaneously which is unique in the literature.

In all batch scheduling problems, batching procedure is implemented as a random algorithm to minimize the objective function. But in the current study, this procedure is done as a systematic manner using clustering approach.

4. Solution approach

As mentioned, the considered problem is NP-hard; hence it is not possible to use ordinary optimization methods for solving the problem. In this case, two hybrid algorithms, based on simulation annealing meta- heuristic (SA) and some clustering algorithms so as to reach near optimal solutions in reasonable run times, are presented.

4.1. Simulation annealing

Simulated annealing (SA) has inspired from cooling procedure of solids and it has been applied widely to solve many optimization problems. It has been indicated that SA is considered as a Marcov chain (Van Laarhoven 1988), so its sensitivity to the initial solution is much less than the other meta-heuristics and has a great ability to avoid getting in local optimum solutions.

In this paper, the algorithm starts from an initial temperature termed as T0, in which two random strings are generated as Figure 1.

2	2	1	3	4

1	3	4	2

Fig. 1- the coding scheme of proposed SA

The first string corresponds to assigning the jobs to batches and the other is related to find a suitable scheduling of generated batches.

Based on this fact, it is considered that there are five available jobs for processing on the machine. The first string shows that the first and the second jobs are assigned to batch 2, the third job is located on batch 1 and forth and fifth jobs are assigned to batches 3 and 4 respectively. Since number 5 is not mentioned in this string, so the fifth batch is remained empty and will be removed.

In second string, a random schedule of all the batches that contain at least one job is generated.

This procedure continues to the point that the equilibrium condition occurs. In this problem, the equilibrium condition occurs when the gaps between proposed objectives functions in consecutive iterations in a certain temperature are as less as possible. This condition can be demonstrated by:

$$\Delta Z = \sum_{i=k-j}^{k} Z_{[i+1]} - Z_{[i]} <= \varepsilon \quad (13)$$

Where Zi+1 is the value of objective function in i+1-th iteration of algorithm, j is the number of objectives that are considered to calculate the total gap and is a very small number.

By reaching the equilibrium in the temperature, the temperature decreases and the procedure starts from the lower temperature and continues until reaching the next equilibrium.

Neighborhood search is also implemented by swapping two randomly selected positions in the second string (batch string) in order to meet other nodes of solution space.

4.2. Clustering

Clustering is the process for assigning a set of items into subsets (called clusters) such that observations in the same cluster are similar in some sense and the similarity of generated clusters is little. Clustering is an approach of unsupervised learning and a common technique for statistical data analysis that has been applied in many fields. Clustering contains several algorithms and methods. In this paper we only consider the Hierarchical clustering and K- means clustering in order to assign the jobs to a batch based on some similarities.

A. Hierarchical Clustering (HC)

Hierarchical Clustering is classified to agglomerative and divisive schemes. In agglomerative clustering, the assignment starts with N clusters where N is equal to the number of items in order that each batch contains only one item. In each step, two items which have more similarities are merged with each other and a unique batch is generated. This procedure continues until the number of clusters is equal to a predefined number. On the other hand, divisive scheme starts the assignment procedure by adding all the items in a batch. In each step, an item that has the least similarity with other items is removed from the batch and assigned to another batch. This procedure continues until that the number of clusters becomes equal to the desired number. The identification of similarities between the items can be done in different ways, which is what distinguishes single linkage from complete linkage and average linkage clustering.

B. K-means clustering (KC)

The basic concept of K-means clustering (MacQueen, 1967) is very similar to the Hierarchical clustering. However, the K-means method considers the assignments as an optimization problem. Based on this regard, first a number is defined as the desired number of clusters. Then each batch is occupied with a random item. These items are correspondences to the center of their clusters and are termed as K-centroids. The next step is related to assigning the remaining items in the clusters according to the maximum similarity between the remaining items and the centeroids. By adding a new item in each cluster, the centeriod values are recalculated and this procedure continues until that the centroid values no longer change. Although it could be proved that the procedure will always terminate, the k-means algorithm does not necessarily find the most optimal configuration, corresponding to the global objective function which is minimum. The algorithm is also significantly sensitive to the initially randomly selected cluster centers. The k-means algorithm can be run multiple times to reduce this effect. The presented clustering algorithms are efficient for assigning the items into batches, but have some disadvantages. One of the most important facts of these disadvantages is that the desired batch number could not be calculated by the algorithm and must be given to it as an input data, While in proposed SA, the value of batches is calculated by the algorithm optimally. Since SA has an important role in finding the solution, therefore, its parameters is calibrated using a Taguchi approach and its results are offered in the following.

4.3. The hybrid algorithms description

In this section, the performances of proposed hybrids are described. In both of the hybrids, the procedure begins with running the proposed SA to determine the suitable value for number of batches. Afterwards, one of the clustering algorithms is run in order to assign the jobs into batches, based on similarity rules that are considered as processing times and due dates of them. Then the generated batches are scheduled by the second string presented in SA section.

A. SA-HC hybrid

First, the SA algorithm is run and the numbers of batches are determined. Then jobs are assigned to batches using Hierarchical Clustering approach. The linkage method is considered as the average between items and the distance measure is defined as squared Euclidean. As mentioned, the similarity of items is calculated based on the processing times and due dates of them. It means that the jobs that have close values of processing time and due date are joined to a unit batch. Afterwards, the generated batches are scheduled. The performance of proposed hybrid is presented in figure 2 in more detail:

```
1. Start with the SA and determine the number of batches. ( N_final)
2. Consider N virtual batches. (N is equal to the number of jobs)
3. Assign each job to a batch randomly so that each batch occupies only one batch.
4. while N < >  N_final DO
5.        Merge two batches that has the closest similarities
6.        N=N-1
7. loop
```

```
8. Enter the generated batches in the SA by a random schedule. (state k)
9. T=T_0
10. While T<T_F DO
11.     Do while the equilibrium condition has not occurred
12.         Generate a neighborhood ( state j)
13.         Z=f(j)-f(k)
14.         If z <0 then k=j
15.         Else if random (0,1) < exp(-Δz/T)
16.         K=j
17.     loop
18.   T= T.(cooling rate)
19.   Loop
20. End
```

Fig. 2. The pseudo-code of SA-HC

B. SA-KC hybrid

Similar to the SA-HC hybrid, the hybrid begins with the starting of SA algorithm and the number of batches is determined. Then using K-means clustering jobs are assigned to batches. The similarity of items is checked based on the processing times and due dates of them. finally the generated batches are scheduled. The performance of proposed hybrid is presented in figure 3 in more details:

```
1. Start with the SA and determine the number of batches. ( N_final)
2. Assign N jobs to each batch randomly as the center of them. (center(i))
3. while the center^new(j) –center (j) < >0 do
4.      assign the remaining jobs to the generated batches based on the maximum similarity
        Recalculate the center of each batch (center^new(j))
5. loop
6. Enter the generated batches in the SA by a random schedule. (state k)
7. T=T_0
8. While T<T_F DO
9.      Do while the equilibrium condition has not occurred-
10.         Generate an neighborhood ( state j)
11.         Z=f(j)-f(k)
12.         If z <0 then k=j
13.         Else if random (0,1) < exp(-Δz/T)
14.         K=j
15.     loop
16.   T= T.(cooling rate)
17.   Loop
18. End
```

Fig. 3. The pseudo-code of SA-KC

3.5. Calibrating the SA parameters

In order to calibrate the proposed SA, a Taguchi approach is used which attempts to identify controllable factors (control factors) that minimize the effect of the noise factors.
The effective factors and their levels are also described in Table 2.

Table 2-The Taguchi experiment inputs

Factor	symbol	levels	Type	Degree of freedom
Initial temperature	A	3	A(1)=500 A(2)=1500 A(3)=3000	2
Number of objectives that are considered to calculate the total gap in each temperature	B	3	B(1)=3 B(2)=4 B(3)=5	2

The associated degree of freedom for these two factors is equal to 8. According to Taguchi standard table of orthogonal array, an L_9 pattern fulfils all the minimum necessary requirements should be used.
Three important measures are considered to evaluate the factors as robust measure, the average responses for each combination of control factors and the variability in the responses due to the noise (standard deviation). The results are depicted in figures 4, 5 and 6 respectively.

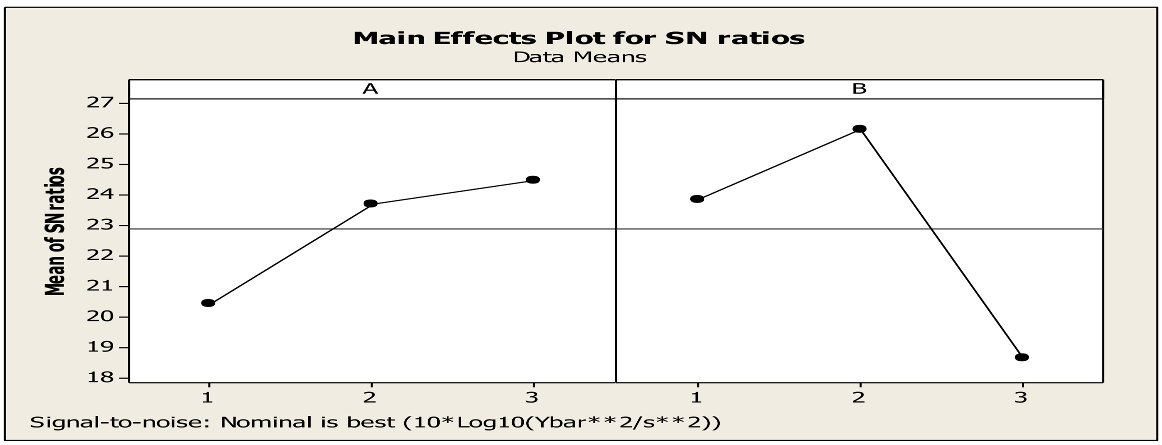

Fig 4- the results for responses based on S/N ratio

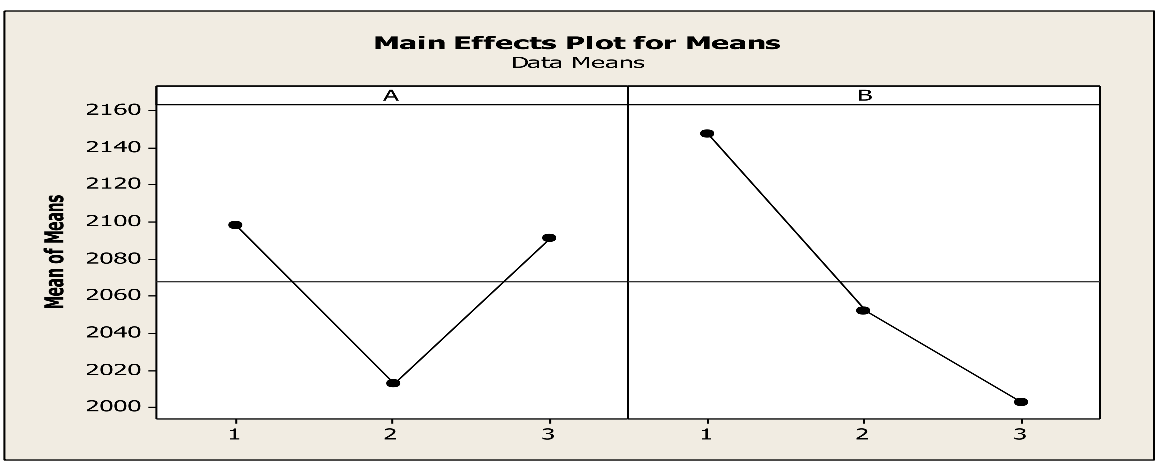

Fig 5- the results for responses based on means

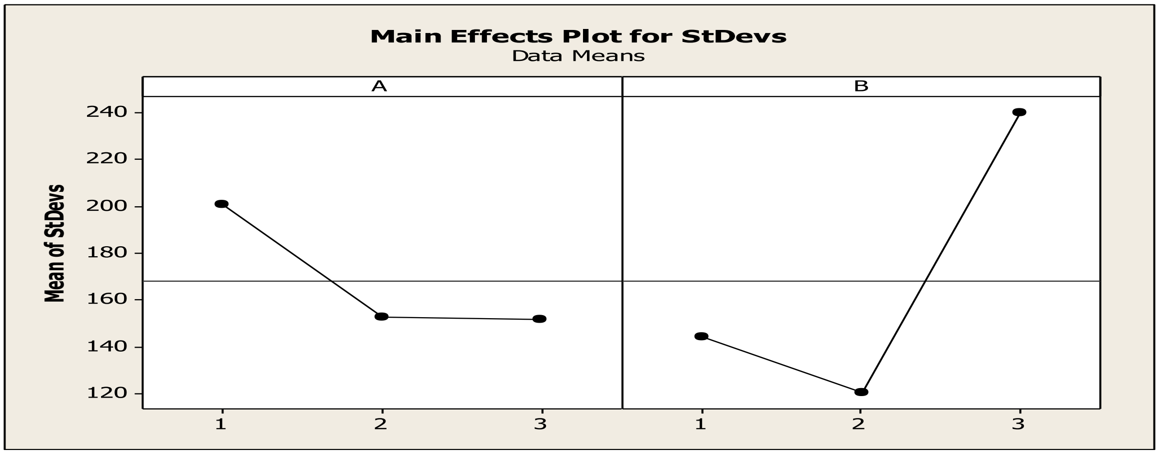

Fig 6- the results for responses based on standard deviations

A measure of robustness is used to identify control factors that reduce variability in a product or process by minimizing the effects of uncontrollable factors. Fig. 4 indicates the robustness of each combination of factors. Clearly, it is desired to select a pair of factors that generate the maximum robustness. Therefore, based on this figure, A(3) and B(2) are selected.
Fig. 5 shows the average responses for each combination of control factors. Since the objective function is minimization, the minimum value for this measure is desired, so A (2) and B (3) are selected.
Finally, Fig. 6 shows the variability in the responses due to the noise that is desired to be minimum, hence A(3) and B(2) are selected. Based on the mentioned measures, the most efficient combination of the proposed factors are A(3) and B(2) which better satisfies the response values.

5. Computational experiment

All the instances for this problem were coded by Matlab software and were run on personal computer with CORE I 7 processor and 4 GB of RAM. The required data were employed from real information achieved from Iran Puya company.
The instances are also solved using Lingo 10 software to determine the efficiency and capability of proposed algorithms to reach the global optimum.
To make the sensitivity analysis, some important factors including problem dimension (number of considered batches) and rate of deterioration are considered.

5.1. Sensitivity analysis based on the problem dimension

In order to implement the analysis first, some instances with small dimensions are considered and the results are compared to equivalent Lingo results. Lingo is incapable of solving the problem for greater than 6 jobs because of immense complexity of the problem. Rate of deterioration and rate of job values are also considered as 0.2 and 0.002 respectively in this section. Table 3 depicts the comparison of the results of SA, hybrids and Lingo based on the problem dimension.

Table 3- sensitivity analysis based on small problem dimensions

Number of tasks	Number of batches	Simple SA			SA-KC			SA-HC			Lingo	
		VOF	Run time	%GAP	VOF	Run time	%GAP	VOF	Run time	%GAP	VOF	Run time

			(sec)			(sec)						(sec)
4	1	232.34	3	0	232.34	-	0	232.34	-	0	232.34	90
5	2	371.56	12	2.37	371.56	1	2.37	371.56	1	2.37	362.93	348
6	2	420.97	17	16.20	362.26	3	0	365.20	3	0.08	362.26	1643
7	4	877.96	63	-	801.78	25	-	761.76	25	-	-	-
8	5	1284.6	57	-	1284.63	19	-	1233.75	16	-	-	-
9	4	1533.74	58	-	1438.94	9	-	1438.94	9	-	-	-
10	4	2096.07	52	-	1485.77	10	-	1825.22	10	-	-	-
15	8	5930.45	60	-	4952.93	13	-	5010.34	16	-	-	-

Where the second column indicates the number of batches and the third to fifth columns show the performance of single SA including VOF (Value of Objective Function), run time in seconds and the percentage of gap between them and global optimum.
For each instance, the SA was run 5 times and the best obtained solution was considered as the best VOF. The other columns also present the results of proposed hybrids and global optimums by Lingo.
The gap of results between SA and global optimum is calculated by:

$$\%gap = \frac{VOF(SA) - VOF(global)}{VOF(global)} * 100 \tag{14}$$

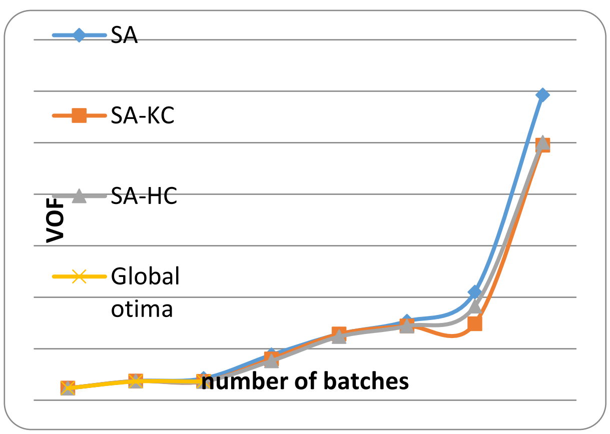

Fig. 7- the VOF sensitivity analysis based on small problem dimensions

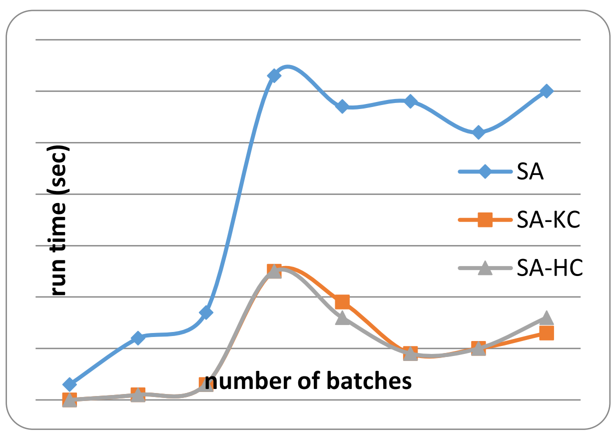

Fig. 8- the run time sensitivity analysis based on small problem dimensions

Figures 7 and 8 depict the comparison of the results and run times respectively. According to the results of table 2 and depicted figures, by using the clustering methods not only the value of objective function improves, but also the required run time for solving the problem decreases. For small dimensions of problem, both hybrids perform very efficient and present the solutions with low values of error in a reasonable time. The comparison of SA and proposed hybrids for solving the problem in larger scales are presented in table 4.

Table 4- sensitivity analysis based on medium and large problem dimensions

Number of tasks	Number of batches	Simple SA		SA-KC		SA-HC	
		VOF	Run time (sec)	VOF	Run time (sec)	VOF	Run time
25	16	18966	66	14075	16	11864	22
30	14	20542	78	16706	16	14723	19
40	23	50101	106	41026	23	43033	24
50	30	73794	138	60552	25	66015	30

Based on the results from table 4, a better solution is achievable in lower run times when a hybrid algorithm is employed. Two proposed hybrids act drastically in comparison to the simple SA and are able to find solutions for medium and large scales.

5.2. Sensitivity analysis based on the rate of deterioration

The effect of deterioration is evaluated by dimension of 10 jobs and value of 0.002 for the rate of job values. Table 4 illustrates the results of SA and Lingo for several rates of deterioration.

Table 5- sensitivity analysis based on rate of deterioration for N=10

Rate of deterioration	Number of batches	Simple SA		**SA-KC**		**SA-HC**	
		VOF	Run time (sec)	VOF	Run time (sec)	VOF	Run time
0.1	5	1818	33	1220	10	1208	10
0.2	5	1849	28	1554	10	1517	10
0.3	5	2322	33	1991	10	1818	10
0.4	5	2968	31	2495	10	2437	9
0.5	5	3153	32	3036	10	2979	10
0.6	5	3928	30	3698	8	3363	8
0.7	5	4489	31	4473	10	4283	8
0.8	5	5797	30	5180	10	5071	7
0.9	5	6699	31	5180	8	5748	8

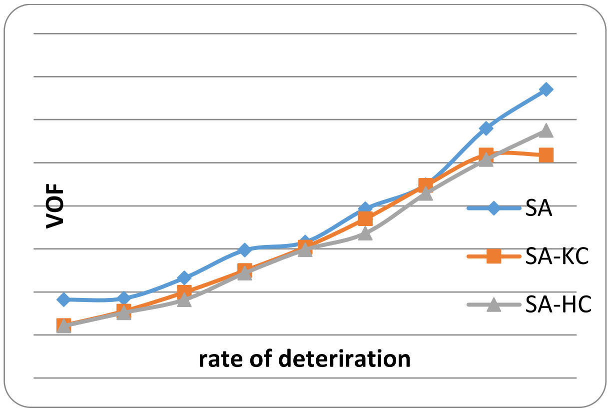

Fig.9- the VOF sensitivity analysis based on rate of deterioration for N=10

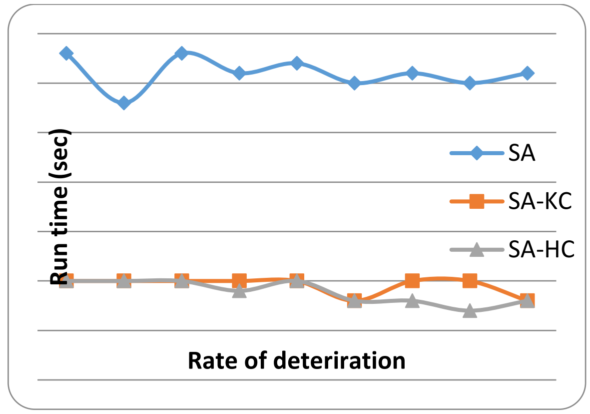

Fig.10- the run time sensitivity analysis based on rate of deterioration for N=10

According to depicted plots, the rate of deterioration is highly effective on the performance of SA and hybrids and clearly changes their final solutions. But it could not affect the required running time for solving the problem.

5.3. The comparison of the two proposed hybrid

In all of the previous sections, the results of both proposed hybrids were much better than single SA, especially for large scales problems. In this section, it is desired to compare the results of hybrid algorithms in a fair way. For this purpose, their performance is compared in equal time intervals. Each algorithm is run 5 times and –the- best, average and the worst solutions are recorded. Table 5 summarizes this information.

Table 6- the comparison of hybrids performances

Run time (sec)	Alg.	Best	Avg.	Worst	SD	SD/Avg.
2.5	SA-HC	5515	6246	6690	445.61	0.07
	SA-KC	5447	6240	6657	484.65	0.07
5	SA-HC	4959	5237	5737	294.17	0.05
	SA-KC	5626	5651	5933	186.31	0.03
7.5	SA-HC	4959	5030	5189	104.06	0.02
	SA-KC	5177	5467	5660	188.60	0.03
10	SA-HC	4959	4974.2	5035	33.98	0.0068
	SA-KC	5177	5178	5180	1.64	0.0003

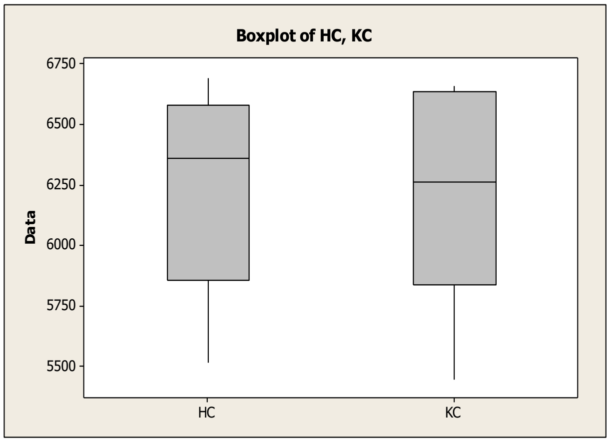

Fig.11-The performances of hybrids in 2.5 seconds

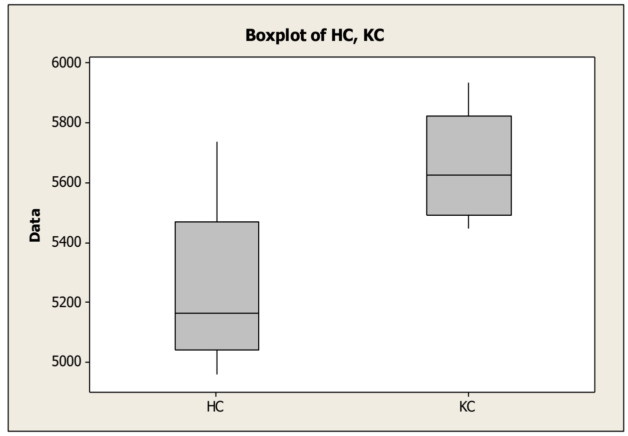

Fig12.The performances of hybrids in 5 seconds

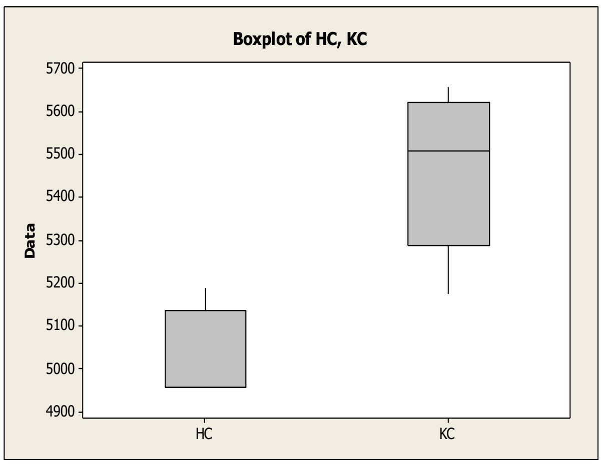

Fig13.The performances of hybrids in 7.5 seconds

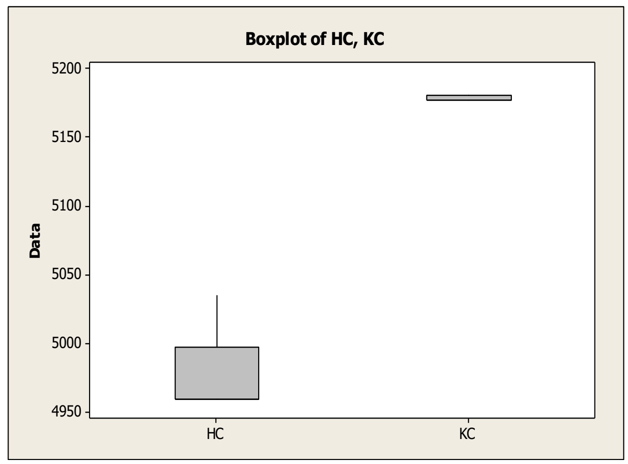

Fig14. The performances of hybrids in 10 seconds

Based on the results of table 5 and depicted plots, the performance of HC is better than KC, except the case of 2.5 second run time. On the other hand, the standard deviations of both algorithms decrease as time passes. So, it could be conclude that the HC algorithm is able to act more rational in terms of getting convergent in final steps of algorithm. One reason for better performance of HC than KC is that the number of clusters have to be determined in KC as an input before starting the algorithm. But the number of clusters in HC in determined within the main procedure of algorithm and it is related to the total error of assignment.

5.4. The comparison of propose methods to the real industry values

The ultimate objective of the current study is to improve the effectiveness of proposed case study. So, in this section some samples of problem is solved and the objective function is compared to the real values.

Table 7- the comparison of output values to the real objectives

Instance number	Number of jobs	HC	KC	Real operational cost	% improvement
		Value of objective	Value of objective		

		function	function		
1	15	7262	7488	8856	22
2	30	11773	12521	13658	16
3	50	9175	9232	11568	26
4	100	7460	7532	9965	36
5	120	10999	10999	15456	41

As depicted in Table 7 using two approaches could improve the operation process in range of 22% to 41%. So it is ideal to employ them within production planning procedure.

6. Conclusion and future research

This paper studied the problem of single machine batch scheduling with objective of minimizing the total tardiness and delivery costs and maximizing the job values simultaneously and a mathematical model was presented for this purpose. The jobs were also considered deteriorated where their processing times were dependent on the sequence in which the process was accomplished. In order to solve the proposed model, two hybrid algorithms were proposed based on the clustering approach and their performances were compared to global optimum for small dimensions of problem generated with Lingo software. Based on these comparisons, all the presented methods were able to find near optimal solutions in most reasonable run times. A Taguchi approach was used to calibrate the initial parameters of solution approaches for tuning of their performance and Afterwards, a number of sensitivity analyses were implemented based on effective factors of the problem, including problem dimension and rate of deterioration. According to the proposed computational experiments, implementation of clustering methods could improve the quality of achieved solutions dramatically in terms of objective function and running times. Based on the computational results, the complexity of problem and required running time increases exponentially and Lingo solver is incapable to solve the problem for instances with greater of 6 jobs. On the other hand the different values for rate of deterioration could not affect the running times. For a better comparison between two proposed methods, the achieved solution for each algorithm was presented within a constant time frame including 2.5, 5, 7.5 and 10 seconds that accordingly HC algorithm performed better than KC. Due to the immense complexity of single machine batch scheduling, the current study faced with some limitations including the finding of exact solutions and increasing the running times with the growth of job numbers.

For future researches, the solution method can be considered as some exact algorithms like branch and bound, branch and cut and Lagrangian relaxation in order to reach the global optimum even for medium and larger scales of jobs. Furthermore some tools similar to Artificial Neural Networks could be employed to improve the effectiveness of any hybrid algorithms to search the solution space.

References

Al-Anzi FS, Allahverdi A, Kovalyov MY., 2007. Batching deteriorating items with applications in computer communication and reverse logistics. European Journal of Operational Research 182, 1002–1011.

Browne S, Yechiali U., 1990. Scheduling deteriorating jobs on a single processor, Operations Research 38, 495–498.

Cheng TCE, Ji M., 2010. Batch scheduling of simple linear deteriorating jobs on a single machine to minimize makespan. European Journal of Operational Research 202, 90–98.

Cheng TCE, Janiak A, Kovalyovv M., 2001. Single machine batch scheduling with resource dependent setup and processing times. European Journal of Operational Research 135 (1)- 177-183.

Cheng TCE, Kovalyov MY., 2000. Single machine batch scheduling with sequential job processing. IIE Transactions 33 (5), 413-420, DOI: 10.1023/A:1011005314354.

Husang H, Wang JB, Wang XR., 2010. A generalization for single-machine scheduling with deteriorating jobs to minimize earliness penalties, Int J Adv Manuf Technol 47, 1225–1230.

Ishii H, Xuesong Li, Masuda T., 2010. Single machine batch scheduling problem with fuzzy batch size. 40th International Conference on Computers and Industrial Engineering. Awaji, 1-5

Johnson S. (1967).,Hierarchical Clustering Schemes. *Psychometrika*, 2, 241-254.

Lu, LF. Yuan JJ., 2007. The single machine batching problem with identical family setup times to minimize maximum lateness is strongly NP-hard. European Journal of Operational Research 177, 1302–1309

MacQueen J., (1967). Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, University of California Press, 281-297.

Mahdavi Mazdeh M, Hamidinia A, Karamouzian A. 2011. A mathematical model for weighted tardy jobs scheduling problem with a batched delivery system. International Journal of Industrial Engineering Computations 2.

Mahdavi Mazdeh M, Sarhadi M , Hindi KS., 2007. A branch-and-bound algorithm for single-machine scheduling with batch delivery minimizing flow times and delivery costs. European Journal of Operational Research 183, 74–86

Mahdavi Mazdeh M, Shashaani S , Ashouri A , Hindi KS., 2011, Single-machine batch scheduling minimizing weighted flow times and delivery costs. Applied Mathematical Modelling 35, 563–570.

Melouk S, Damodaran P, Chang P., 2004. Minimizing makespan for single machine batch processing with non-identical job sizes using simulated annealing. International Journal of Production Economics 87 (2), 141-147.

Mosheiov G., 1994. Scheduling jobs under simple linear deterioration, Computers and Operations Research 21, 653–659.

Musavi MM Tavakkoli-Moghaddam R. Rayat F. A Bi-Objective Green Truck outing and Scheduling Problem in a Cross Dock with the Learning Effect. Iranian Journal of Operations Research Vol. 8, No. 1, 2017, pp. 02-14

Niksirat M. Efficient Algorithm Based on Network Flows for Two-Depot Bus Scheduling Problem. Iranian Journal of Operations Research. Vol. 9, No. 1, 2018, pp. 49-61 DOI: 10.29252/iors.9.1.49

Ng CT, Cheng TCE, Yuan JJ, Liu ZH., 2003. On the single machine serial batching scheduling problem to minimize total completion time with precedence constraints, release dates and identical processing times. Operations Research Letters 31, 323 – 326.

Ng CT, Cheng TCE, Yuan JJ., 2002. A note on the single machine serial batching scheduling problem to minimize maximum lateness with precedence constraints. Operations Research Letters 30, 66 – 68.

Ng CT, Cheng TCE, Kovalyov M., 2004. Single machine batch scheduling with jointly compressible setup and processing times. European Journal of Operational Research 153 (1), 211-219.

Nong Q, Ng CT, Cheng TCE., 2008. The bounded single-machine parallel-batching scheduling problem with family jobs and release dates to minimize makespan. Operations Research Letters 36, 61 – 66.

Potts CN, Kovalyov MY., 2000. Scheduling with batching: a review. European Journal of Operational Research 120, 228–249.

Pei J , Liu X, Liao B , Pardalos P , Kong M. Single-machine Scheduling with Learning Effect and Resource-dependent Processing Times in the Serial-batching Production, *Applied Mathematical modelling* (2017), doi: 10.1016/j.apm.2017.07.028

Rostami M. Minimizing maximum tardiness subject to collect the EOL products in a single machine scheduling problem with capacitated batch delivery and pickup systems. (2021) Computers & Industrial Engineering 161 (2021) 107634

Tang L, Zhao X, Liu J, Joseph Y.-T. Leung, Competitive two-agent scheduling with deteriorating jobs on single parallel-batching machine, *European Journal of Operational Research* (2017),doi:10.1016/j.ejor.2017.05.019

Tian Ji, Fu R, Yuan J., 2007. On-line scheduling with delivery time on a single batch machine. Theoretical Computer Science 374, 49–57.

Van Laarhoven P.J.M, Aarts E.H.,1988. Simulated Annealing: Theory and Applications, Kluwer Academic Publishers, Dordrecht, 1988.

Wang D, Wang JB., 2010. Single-machine scheduling with simple linear deterioration to minimize earliness penalties, Int J Adv Manuf Technol 46, 285–290, DOI 10.1007/s00170-009-2086-8.

Wang JB, Huang X, Wang XY, Yin N, Wang L., 2009. Learning effect and deteriorating jobs in the single machine scheduling problems, Applied Mathematical Modeling 33, 3848–3853.

Wu CC, Shiau YR, Lee, LH, Lee WC., 2009. Scheduling deteriorating jobs to minimize the makespan on a single machine, Int J Adv Manuf Technol 44, 1230–1236, DOI 10.1007/s00170-008-1924-4

Yuan JJ, Liu Z.H, Ng C.T, Cheng T.C.E., 2004. The unbounded single machine parallel batch scheduling problem with family jobs and release dates to minimize makespan. Theoretical Computer Science 320, 199–212.

Zhang H, Wu F, Yang Z. Hybrid approach for a single-batch-processing machine scheduling problem with a just-in-time objective and consideration of non-identical due dates of jobs. Computers and Operations Research 128 (2021) 105194

Zhou S, Jin M, Du N, Energy-efficient scheduling of a single batch processing machine with dynamic job arrival times, Energy (2020), doi: https://doi.org/10.1016/j.energy.2020.118420

Zheng S, Xie N, Wu Q. Single batch machine scheduling with dual setup times for autoclave molding manufacturing. Computers & Operations Research 133 (2021) 105381