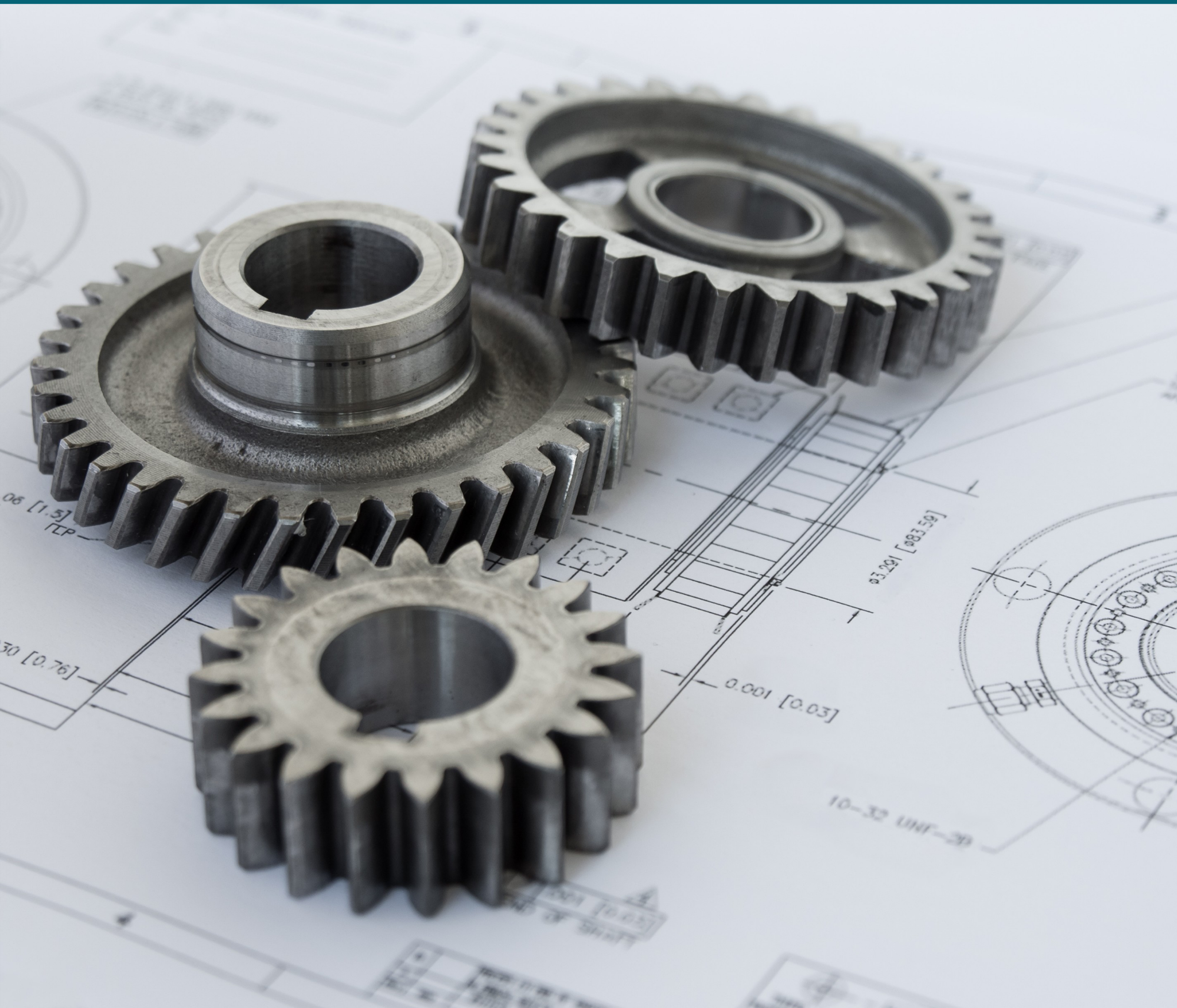
REA
Research Expansion Alliance

RJARIE

JOURNAL OF APPLIED RESEARCH ON INDUSTRIAL ENGINEERING



Print ISSN: 2538-5110
Online ISSN: 2676-6167
www. journal-aprie.com

Journal of Applied Research on Industrial Engineering
www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 1–154.

TABLE OF CONTENTS

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 1–10.

Paper Type: Research Paper

A Compromise Solution for the Neutrosophic Multiobjective Linear Programming Problem and its Application in Transportation Problem

Elham Hosseinzadeh[1,*], **Javad Tayyebi**[2]

[1] Department of Mathematics, Kosar University of Bojnord, Bojnord, Iran; e.hosseinzadeh@kub.ac.ir.

[2] Department of Industrial Engineering, Birjand University of Technology, Birjand, Iran; javadtayyebi@birjandut.ac.ir.

Citation:

Hosseinzadeh, E., & Tayyebi, J. (2023). A compromise solution for the neutrosophic multiobjective linear programming problem and its application in transportation problem. *Journal of applied research on industrial engineering, 10*(1), 1-10.

Received: 12/02/2022 Reviewed: 15/03/2022 Revised: 09/05/2022 Accepted: 01/07/2022

Abstract

Neutrosophic set theory plays an important role in dealing with the impreciseness and inconsistency in data encountered in solving real-life problems. The current paper focuses on the Neutrosophic Fuzzy Multi-Objective Linear Programming Problem (NFMOLPP), where the coefficients of the objective functions, constraints and right-hand side parameters are single-valued trapezoidal Neutrosophic Numbers (NNs). From the viewpoint of complexity of the problem, a ranking function of NNs is proposed to convert the problem into equivalent MOLPPs with crisp parameters. Then suitable membership functions for each objective are formulated using their lowest and highest value. With the aim of linear programming techniques, a compromise optimal solution of NFMOLPP is obtained. The main advantage of the proposed approach is that it obtains a compromise solution by optimizing truth-membership, indeterminacy-membership, and falsity-membership functions, simultaneously. Finally, a transportation problem is introduced as an application to illustrate the utility and practicality of the approach.

Keywords: Multiobjective programming problem, Neutrosophic set, Single valued trapezoidal, Neutrosophic number, Indeterminacy membership functions.

1 | Introduction

Optimizing more than one commensurable and/or conflicting objective function under a set of well-defined constraints is termed as Multi-Objective Programming Problems (MOPPs). Most often, many real-world applications, such as transportation, supplier selection, inventory control, supply chain planning, etc., take the form of MOLPs. While dealing with multiple objectives, it is not always possible to obtain a single solution that optimizes each objective function, efficiently. However, a compromise solution is possible that satisfies each objective, simultaneously. Therefore, the concept of compromise solutions is an important aspect and leads in search of the global optimality criterion. In the past few decades, a tremendous amount of research has been presented in the context of multiobjective optimization techniques.

Corresponding Author: e.hosseinzadeh@kub.ac.ir

https://doi.org/10.22105/jarie.2022.328580.1451

A major disadvantage of fuzzy sets is its inability to efficiently represent imprecise and inconsistent information as it considers only the truth membership function [29]. Intuitionistic fuzzy set is a modification of fuzzy sets, which considered both the truth and falsity membership functions [4]. But it still had some drawbacks in depicting human-like decision making. In the last decade, a large number of studies on fuzzy and Intuitionistic fuzzy multi-objective optimization techniques have been presented. Among these studies, we can mention the works of Mahajan and Gupta [18], Borovička [6], Ahmadini and Ahmad [3], Yu et al. [26], and Rizk-allah et al. [20].

In 1998, a new type of sets called the neutrosophic set was introduced by Smarandache [22] to deal with decision making problems which involved incomplete, inconsistent and indeterminate information. Here indeterminacy is considered as an independent factor, which has a major contribution in decision making. Neutrosophic set helps in human-like decision making by considering truth, falsity and indeterminacy membership functions.

Ye et al. [24] presented some new operations of NNs to make them suitable for engineering applications. They proposed a neutrosophic function involving NNs. Then, they used it to solve neutrosophic linear programming problems [24], [25]. Ye et al. [23] analysed joint roughness coefficient taking the help of NN functions. NN generalized weighted power averaging operator formulated by Liu and Liu [17] and it is applied to multi-attribute group decision making in NN environment. Maiti et al. [19] proposed a goal programming strategy to solve multi-level Multi-Objective Linear Programming Problem (MOLPP) with NNs.

Recently, Deli and Şubaş [10] suggested a novel ranking method for single-valued NN and they applied it to multi-criteria decision-making problems. Ahmad et al. [1] discussed the energy-food-water nexus security management through neutrosophic modeling and optimizing approaches. Also, they presented a study on supplier selection problem with Type-2 fuzzy parameters and solved it using an interactive neutrosophic optimization algorithm [2]. Wang et al. [28] proposed a novel method to solve multiobjective linear programming problems with triangular NN. Kumar Das et al. [8] presented a novel lexicographical-based method for linear programming problems with trapezoidal NN. Their method uses a lexicographical order.

As a special instance of linear programming problems, many authors focused on solving transportation problems in fuzzy environment, such as the multi-objective case [11], [14], the case with fractional objectives [13], [16], the inverse version [12], the problem with heptagonal and pentagonal fuzzy numbers [9], [15], and the problem with fuzzy variables [7].

This paper attempts to formulate and to solve the multiobjective linear programming problem with Single-Valued Trapezoidal Neutrosophic (SVTN) parameters. This problem has mixed constraints, in which the coefficients of objectives, the coefficients of constraints and right-hand sides of constraints are SVTN numbers. A new method to find a compromise optimal solution of NFMOLP problem is proposed. In the proposed method, the accuracy function is used to transfer the NFMOLPP into equivalent crisp MOLPP. Finally, we apply the approach for a transportation problem to show its utility and performance.

The rest of this paper is organized as follows: In Section 2 basic concepts and algebra operations of NN are reviewed. Section 3 deals with modelling multiobjective linear programming problems with neutrosophic fuzzy parameters. In Section 4, a solution method for obtaining a compromise solution of NFMOLPP is introduced. In Section 5, the entire solution procedure is summarized in the form of an algorithm. In Section 6 a transportation problem as an application is presented. Finally, some concluding remarks are reported in Section 7.

Hosseinzadeh and Tayyebi| J. Appl. Res. Ind. Eng. 10(1) (2023) 1-10

2 | Preliminaries

In this section, some basic concepts and definitions on neutrosophic sets and single-valued trapezoidal NN are reviewed from the literature.

Definition 1 ([13]). Let X be a nonempty set. A neutrosophic set (NS) $\tilde{A}^N$ is defined as

$$\widetilde{A}^N = \{\langle x\colon T_A\ x), I_A\ x), F_A(x)\rangle, x\ \in X\}, T_A\ x), I_A\ x), F_A\ x) \in]0^-, 1^+[\ \}.$$

Where $T_A\ x), I_A\ x)$ and $F_A(x)$ are called truth-membership function, indeterminacy-membership function and falsity-membership function, respectively, and there is no restriction on the summation of them, so $0^- \leq T_A\ x) +\ I_A\ x) + F_A\ x) \leq 3^+$, and $]0^-, 1^+[$ is non-standard unit interval.

Since it is difficult to apply NSs to practical problems, Wang et al. [27] introduced the concept of a Single Valued Neutrosophic Set (SVNS), which is an instance of a NS and can be used in real scientific and engineering applications.

Definition 2. Let $T_a, I_a, F_a\ \in [0,1]$, then a Single-Valued Trapezoidal Neutrosophic Number (SVTNN) $a^N = \langle\ a_1, a_2, a_3, a_4);\ T_a, I_a, F_a\rangle$ is a special NS on the real numbers $\mathbb{R}$, whose truth, indeterminacy and falsity membership functions are given as follows:

$$\mu_a\ x) = \begin{cases} T_a\left(\dfrac{x - a_1}{a_2 - a_1}\right), & a_1 \leq x\ \leq a_2, \\ T_a, & a_2 \leq x\ \leq a_3, \\ T_a\left(\dfrac{x - a_4}{a_4 - a_3}\right), & a_3 < x\ \leq a_4, \\ 0, & \text{oherwise.} \end{cases}$$

$$\lambda_a\ x) = \begin{cases} \dfrac{a_2 - x + I_a(x - a_1)}{a_2 - a_1}, & a_1 \leq x\ \leq a_2, \\ I_a, & a_2 \leq x\ \leq a_3, \\ \dfrac{x - a_3 + I_a(a_4 - x)}{a_4 - a_3}, & a_3 < x\ \leq a_4, \\ 1, & \text{oherwise.} \end{cases}$$

$$\nu_a\ x) = \begin{cases} \dfrac{a_2 - x + F_a(x - a_1)}{a_2 - a_1}, & a_1 \leq x\ \leq a_2, \\ F_a, & a_2 \leq x\ \leq a_3, \\ \dfrac{x - a_3 + F_a(a_4 - x)}{a_4 - a_3}, & a_3 < x\ \leq a_4, \\ 1, & \text{oherwise,} \end{cases}$$

where, T_a, I_a and F_a are the maximum truth, minimum indeterminacy, and minimum falsity membership degrees, respectively.

Definition 3. Let $\tilde{a}^N = \langle(a_1,\ a_2,\ a_3,\ a_4);\ T_{\tilde{a}}, I_{\tilde{a}}, F_{\tilde{a}}\rangle$ and $\tilde{b}^N = \langle(b_1, b_2, b_3, b_4);\ T_{\tilde{b}}, I_{\tilde{b}}, F_{\tilde{b}}\rangle$ be two arbitrary SVTNNs and $\gamma \neq 0$ be any real number, then

I. $a^N +\ b^N = \langle\ a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4);\ T_a\ \wedge T_b, I_a\ \vee I_b, F_a\ \vee F_b\rangle$.

II. $a^N -\ b^N = \langle\ a_1 - b_4, a_2 - b_3, a_3 - b_2, a_4 - b_1);\ T_a\ \wedge T_b, I_a\ \vee I_b, F_a\ \vee F_b\rangle$.

III. $\gamma a^N = \begin{cases} \langle\ \gamma a_1, \gamma\ a_2, \gamma a_3, \gamma a_4);\ T_a, I_a, F_a\rangle, & \gamma\ > 0, \\ \langle\ \gamma a_4, \gamma a_3, \gamma a_2, \gamma a_1);\ T_a, I_a, F_a\rangle, & \gamma <\ 0. \end{cases}$

Definition 4. Let $a^N = \langle\ a_1,\ a_2,\ a_3,\ a_4);\ T_a, I_a, F_a\rangle$ be a SVTNN. Then, the score function $S(a^N)$ and accuracy function $A(a^N)$ of a SVTNN are respectively defined as follows:

I. $S\left(a^N\right) = \frac{1}{16}\, a_1 + \, a_2 + \, a_3 + \, a_4)\left(T_a \, + \, 1 - I_a\right) + \, \, 1 - F_a)\big).$

II. $A\left(a^N\right) = \frac{1}{16}\, a_1 + \, a_2 + \, a_3 + \, a_4)\left(T_a \, + \, 1 - I_a\right) - \, \, 1 - F_a)\big).$

Definition 5. Suppose $a^N = \langle\, a_1,\, a_2,\, a_3,\, a_4);\, T_a, I_a, F_a\rangle$ and $b^N = \langle\, b_1, b_2, b_3, b_4);\, T_b, I_b, F_b\rangle$ be any two SVTNNs. Then, we define a ranking method as follows:

I. If $S\left(a^N\right) > S\left(b^N\right)$ then $a^N > b^N$.

II. If $S\left(a^N\right) = S\left(b^N\right)$ and if $A\left(a^N\right) > A\left(b^N\right)$ then $a^N > b^N$, $A\left(a^N\right) < A\left(b^N\right)$ then $a^N < b^N$, $A\left(a^N\right) = A\left(b^N\right)$ then $a^N = b^N$.

Theorem 1. Let $g:\ \mathbb{S} \to \mathbb{R}, \mathbb{S} \subset \mathbb{R}^n$ be a real valued function. If g is a convex function, then $\{x:\ g(x) \leq c, \forall\, c \in \mathbb{R}\}$ is a convex set and if g is a concave function, then $\{x:\ g(x) \geq c, \forall\, c \in \mathbb{R}\}$ is a convex set.

3 | Problem Formulation

The general form of a MOLPP with k objectives can be described as follows:

$$\begin{aligned}
&\min \quad Z\ x) = [Z_1\ x), Z_2\ x), \dots, Z_r\ x)],\\
&\text{s.t.} \quad \sum_{j=1}^{n} a_{ij}x_j \geq b_i,\ \ i = 1,2,\dots,m_1,\\
&\sum_{j=1}^{n} a_{ij}x_j \leq b_i,\ \ i = \ m_1+1,\ m_1+2,\dots,m_2,\\
&\sum_{j=1}^{n} a_{ij}x_j = b_i,\ \ i = \ m_2+1,\ m_2+2,\dots,m,\\
&x_j \geq 0,\ \ j = 1,2,\dots,n,
\end{aligned} \tag{1}$$

where $Z_k\ x) = \sum_{k=1}^{r} c_{kj}\, x_j, k = 1,2,\cdots,r$ is the k-th objective function.

Definition 6. Let S be the set of all feasible solutions for *Eq. (1)*. A point x^* is said to be an efficient or Pareto optimal solution of *Eq. (1)* if there does not exist any $x \in\ S$ such that, $Z_k(x^*) \geq Z_k(x)$ for every k, and $Z_k(x^*) > Z_k(x)$ for at least one k.

If all the parameters of *Problem (1)* are uncertain, and they can be represented by SVTNNs, then *Problem (1)* becomes a Neutrosophic Fuzzy Multi-Objective Linear Programming Problem (NFMOLPP) as follows:

$$\begin{aligned}
&\min \quad Z^N\ x) = \left[Z_1^N\ x), Z_2^N\ x), \dots, Z_r^N\ x)\right],\\
&\text{s.t.} \quad \sum_{j=1}^{n} a_{ij}^N x_j \geq b_i^N,\ \ \ i = 1,2,\dots,m_1,\\
&\sum_{j=1}^{n} a_{ij}^N x_j \leq b_i^N,\ \ \ i = \ m_1+1,\ m_1+2,\dots,m_2,\\
&\sum_{j=1}^{n} a_{ij}^N x_j = b_i^N,\ \ \ i = \ m_2+1,\ m_2+2,\dots,m,\\
&x_j \geq 0,\ \ \ \ j = 1,2,\dots,n,
\end{aligned} \tag{2}$$

where $\tilde{Z}_k^N = \sum_{j=1}^{n} (\tilde{c}_{kj})^N\, x_j, k = 1, 2, \dots, r.$

Using accuracy function which is linear, *Problem (2)* is converted into the following crisp MOLPP:

$$\begin{aligned}
&\min \quad Z'(x) = [Z_1'(x), Z_2'(x), \dots, Z_r'(x)],\\
&\text{s.t.} \quad \sum_{j=1}^{n} a_{ij}' x_j \geq b_i', \quad i = 1,2,\dots,m_1,\\
&\sum_{j=1}^{n} a_{ij}' x_j \leq b_i', \quad i = m_1+1,\ m_1+2,\dots,m_2,\\
&\sum_{j=1}^{n} a_{ij}' x_j = b_i', \quad i = m_2+1,\ m_2+2,\dots,m,\\
&x_j \geq 0, \quad j = 1,2,\dots,n.
\end{aligned} \tag{3}$$

Where $Z_k'(x) = A\left(Z_k^N\, x\right) = \sum_{j=1}^{n} A\left((c_{kj})^N\right) x_j,\ \forall k = 1,2,\dots,r;\ b_i' = A(b_i^N)$ and $a_{ij}' = A(a_{ij}^N)$ for all $i = 1,\dots,m, j = 1,\dots,n$.

Theorem 2 ([21]). An efficient solution for crisp *MOPP (3)* is an efficient solution for *NFMOLPP (2)*.

Thus, solving the NFMOLPP *Model (2)* is equivalent to solving the crisp MOLPP *Model (3)*.

4 | Solution Method

In this section, we restrict our attention to NFMOLPP and present an approach to solve it.

By using the definition of the fuzzy decision proposed by Bellman and Zadeh [5], we can characterize the fuzzy decision set D as follows:

$$D = Z \cap C,$$

where Z and C are fuzzy goals and fuzzy constraints, respectively.

In a similar manner, we also introduce the neutrosophic decision set D^N, which consider neutrosophic goals and constraints as follows:

$$D^N = \left(\bigcap_{k=1}^{r} Z_k\right) \cap \left(\bigcap_{i=1}^{m} C_i\right).$$

Where

$$\begin{aligned}
\mu_{D^N}(x) &= \min\left\{\mu_{Z_1}(x), \dots, \mu_{Z_r}, \mu_{C_1}, \dots, \mu_{C_m}, \forall\ x \in X\right\},\\
\lambda_{D^N}(x) &= \max\left\{\lambda_{Z_1}(x), \dots, \lambda_{Z_r}, \lambda_{C_1}, \dots, \lambda_{C_m}, \forall\ x \in X\right\},\\
\nu_{D^N}(x) &= \max\left\{\nu_{Z_1}(x), \dots, \nu_{Z_r}, \nu_{C_1}, \dots, \nu_{C_m}, \forall\ x \in X\right\},
\end{aligned}$$

are the truth, indeterminacy and the falsity membership functions of neutrosophic decision set D^N, respectively.

In the sequal, we solve the multi-objective programming problem by considering one objective function at a time and ignoring the others. Then, we find minimum and maximum values of each objective function. Let L_k be the minimum value and U_k be the maximum value of Z_k, i.e.,

$$U_k = \max\ [Z_k(x)] \text{ and } L_k = \min\ [Z_k(x)] \quad \forall k = 1,2,\dots,r. \tag{4}$$

The bounds for the k-th objective function under the neutrosophic environment can be obtained as follows:

$$U_k^{\mu} = U_k, L_k^{\mu} = L_k, \quad \text{for truth membership,} \tag{5}$$

$$U_k^{\lambda} = U_k^{\mu} + s_k,\ L_k^{\lambda} = L_k^{\mu}, \quad \text{for indeterminacy membership,} \tag{6}$$

$$U_k^{\nu} = U_k^{\mu},\quad L_k^{\nu} = L_k^{\mu} + t_k, \quad \text{for falsity membership,} \tag{7}$$

Hosseinzadeh and Tayyebi| J. Appl. Res. Ind. Eng. 10(1) (2023) 1-10

where $s_k, t_k \in (0,1)$ are predetermined real numbers prescribed by decision-makers. For each objective function, consider truth membership function $\mu_k(Z_k\ x))$, indeterminacy membership function $\lambda_k(Z_k\ x))$ and falsity membership function $\nu_k(Z_k\ x))$ as the following functions:

$$\mu_k\big(Z_k\ x)\big) = \begin{cases} 1, & Z_k\ x) \le L_k^{\mu}, \\ \frac{U_k^{\mu} - Z_k\ x)}{U_k^{\mu} - L_k^{\mu}}, & L_k^{\mu} \le Z_k\ x) \le U_k^{\mu}, \\ 0, & Z_k\ x) \ge U_k^{\mu}. \end{cases} \tag{8}$$

$$\lambda_k\big(Z_k\ x)\big) = \begin{cases} 0, & Z_k\ x) \le L_k^{\nu}, \\ \frac{Z_k\ x) - L_k^{\lambda}}{U_k^{\lambda} - L_k^{\lambda}}, & L_k^{\lambda} \le Z_k\ x) \le U_k^{\lambda}, \\ 1, & Z_k\ x) \ge U_k^{\nu}. \end{cases} \tag{9}$$

$$\nu_k\big(Z_k\ x)\big) = \begin{cases} 0, & Z_k\ x) \le L_k^{\nu}, \\ \frac{Z_k\ x) - L_k^{\nu}}{U_k^{\nu} - L_k^{\nu}}, & L_k^{\nu} \le Z_k\ x) \le U_k^{\nu}, \\ 1, & Z_k\ x) \ge U_k^{\nu}. \end{cases} \tag{10}$$

Since, decision maker wants to maximize the range of acceptance and to minimize the range of rejection, we are looking for a solution with the maximum degree of membership and the minimum degree of nonmembership.

In this regard, according to the concept of fuzzy decision set [5], an optimal compromise solution can be selected as the design for which it maximizes the minimum truth degree (acceptance) and minimize the maximum indeterminacy (rejection up to some extent) and a falsity (rejection) degree by taking all objectives, simultaneously. Therefore, according to the fuzzy decision of Belman and Zadeh [16], we have to solve the following multiobjective programming problem:

$$\begin{aligned} &\text{Maximize} \quad (\min\{\mu_1\big(Z_1\ x)\big), \dots, \mu_r\big(Z_r\ x)\big)\}), \\ &\text{Minimize} \quad (\max\{\lambda_1\big(Z_1\ x)\big), \dots, \lambda_r\big(Z_r\ x)\big)\}), \\ &\text{Minimize} \quad (\max\{\nu_1\big(Z_1\ x)\big), \dots, \nu_r\big(Z_r\ x)\big)\}), \\ &\quad \text{s.t.} \qquad \text{all the constraints of } \ 3). \end{aligned} \tag{11}$$

Suppose that $\alpha = \min_{k=1,\dots,r} \mu_k\big(Z_k\ x)\big)$, $\beta = \max_{k=1,\dots,r} \lambda_k\big(Z_k\ x)\big)$ and $\gamma = \max_{k=1,\dots,r} \nu_k(Z_k\ x))$.

Therefore, *Problem (11)* can be rewritten in the form of

$$\begin{aligned} &\text{Maximize} \quad \alpha, \\ &\text{Minimize} \quad \beta, \\ &\text{Minimize} \quad \gamma, \\ &\text{s.t.} \ \ \mu_k\big(Z_k\ x)\big) \ge \alpha, \qquad k = 1, \dots, r, \\ &\lambda_k\big(Z_k\ x)\big) \le \beta, \qquad k = 1, \dots, r, \\ &\nu_k\big(Z_k\ x)\big) \le \gamma, \qquad k = 1, \dots, r, \\ &\alpha \ge \beta, \quad \alpha \ge \gamma, \quad \alpha + \beta + \gamma \le 3, \\ &\alpha, \beta, \gamma \in \ 0,1). \end{aligned} \tag{12}$$

All the constraints of *Eq. (3)*.

Using the weighted sum method and by setting the *Relations (8)*, *(10)* and *(9)*, the *Problem (12)* can be formed into the following equivalent problem:

$$\begin{aligned} &\text{Max} \ \ w_1\alpha - w_2\beta - w_3\gamma, \\ &\text{s.t.} \ \ Z_k\ x) + \big(U_k^{\mu} - L_k^{\mu}\big)\alpha \le U_k^{\mu}, \qquad k = 1, \dots, r, \\ &Z_k\ x) - \big(U_k^{\lambda} - L_k^{\lambda}\big)\beta \le L_k^{\lambda}, \qquad k = 1, \dots, r, \\ &Z_k\ x) - \ U_k^{\nu} - L_k^{\nu})\gamma \le L_k^{\nu}, \qquad k = 1, \dots, r, \\ &\alpha \ge \beta, \quad \alpha \ge \gamma, \quad \alpha + \beta + \gamma \le 3, \qquad \alpha, \beta, \gamma \in \ 0,1). \end{aligned} \tag{13}$$

All the constraints of *Eq. (3)*.

Theorem 3. If $(\hat{x},\hat{\alpha},\beta,\hat{\gamma})$ is a unique optimal solution of *Problem (13)*, then $(\hat{x},\hat{\alpha},\beta,\hat{\gamma})$ is an efficient solution for *Problem (3)*.

Proof: On contrary, suppose that $(\hat{x},\hat{\alpha},\beta,\hat{\gamma})$ is not an efficient solution for *Problem (3)*. Then, there exists a feasible solution $x^* \neq \hat{x}$ to *Problem (3)*, such that $Z_k\ x^*) \leq\ Z_k(\hat{x})$ for all $k = 1, \dots, r$ and $Z_k\ x^*) <\ Z_k(\hat{x})$ for at least one k. Therefore, $\frac{Z_k\ x^*)-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}} \leq \frac{Z_k\ \hat{x})-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}}$ for all $k = 1, \dots, r$ and $\frac{Z_k\ x^*)-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}} < \frac{Z_k\ \hat{x})-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}}$ for at least one k. Thus, $\max_k\left(\frac{Z_k\ x^*)-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}}\right) \leq (<) \max_k(\frac{Z_k\ \hat{x})-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}})$. Let $\gamma^* = \max_k(\frac{Z_k\ x^*)-L_k^{\nu}}{U_k^{\nu}-L_k^{\nu}})$ then $\gamma^* \leq\ <)\hat{\gamma}$. Similiarly, consider that Let $\beta^* = \max_k(\frac{Z_k\ x^*)-L_k^{\lambda}}{U_k^{\lambda}-L_k^{\lambda}})$ then $\beta^*\ \leq\ <)\beta$.

In the same manner, we have $\frac{U_k^{\mu}-Z_k\ x^*)}{U_k^{\mu}-L_k^{\mu}} \geq \frac{U_k^{\mu}-Z_k\ \hat{x})}{U_k^{\mu}-L_k^{\mu}}$ for all $k = 1, \dots, r$ and $\frac{U_k^{\mu}-Z_k\ x^*)}{U_k^{\mu}-L_k^{\mu}} > \frac{U_k^{\mu}-Z_k\ \hat{x})}{U_k^{\mu}-L_k^{\mu}}$ for at least one k.

Hence, $\min_k\left(\frac{U_k^{\mu}-Z_k\ x^*)}{U_k^{\mu}-L_k^{\mu}}\right) \geq\ >) \min_k(\frac{U_k^{\mu}-Z_k\ \hat{x})}{U_k^{\mu}-L_k^{\mu}})$. Let $\alpha^* = \min_k(\frac{U_k^{\mu}-Z_k\ x^*)}{U_k^{\mu}-L_k^{\mu}})$ this gives $(\hat{\alpha}-\beta-\hat{\gamma}) < (\alpha^*-\beta^*-\gamma^*)$ which means that the solution is not unique optimal. This contradicts the fact that $(\hat{x},\hat{\alpha},\beta,\hat{\gamma})$ is the unique optimal solution of *Eq. (13)*. Hence, it is an efficient solution of *Eq. (3)*.

5 | Compromise Solution Algorithm for NFMOLPP

In this section, we summarize the compromise solution procedure developed in Section 4 as the following algorithm.

Step 1. Formulate the NFMOLPP as given in *Problem (1)*.

Step 2. Transform the NFMOLPP into crisp MOLPP as given in *Problem (3)* by the accuracy function.

Step 3. Find an optimal solution of each single objective LPP and determine the upper and lower bounds by using *Eq. (4)*.

Step 4. Using U_k and L_k, obtain the upper and lower bounds for truth, indeterminacy and falsity membership function as given in *Eqs. (5)-(7)*.

Step 5. Use linear membership functions as given in *Eqs. (9)-(10)* and transform the optimization *Problem (12)* to crisp programming model as in *Eq. (13)*.

Step 6. Solve crisp programming *Problem (13)* using suitable techniques or software packages.

6 | Numerical Example

To illustrate the application of the proposed approach for a real-life transportation problem, the following numerical example is considered. Since, the parameters of transportation problem vary due to various uncertain situation like weather condition, traffic condition, petroleum price, the crisp value of parameters cannot deal the situation properly. To address this situation, we express parameters by SVTNNs.

Example 1. Consider a transportation problem in which we have two objectives with 2 sources and 3 destinations. The cost of transportation per vehicle is denoted by C_1^N appeared in the first objective and amount of carbon dioxide (CO_2) emission per vehicle C_2^N is appeared in the second objective. The neutrosophic fuzzy parameters related to this example are summarized in *Table 1*. The supply of two origins and the demand of three destinations are all SVTN numbers given as follows:

Table 1. Neutrospphic fuzzy parameters for $c_{ij}^{1)N}$ if and $c_{ij}^{2)N}$.

$c_{11}^{(1)N} = \langle (20,30,40,50);0.8,0.3,0.6\rangle$	$c_{21}^{(1)N} = \langle (45,55,65,75);0.8,0.5,0.3\rangle$
$c_{12}^{(1)N} = \langle (50,60,70,80);0.6,0.4,0.3\rangle$	$c_{22}^{(1)N} = \langle (55,65,90,105);0.7,0.4,0.5\rangle$
$c_{13}^{(1)N} = \langle (80,90,110,120);0.7,0.2,0.5\rangle$	$c_{23}^{(1)N} = \langle (30,40,60,70);0.9,0.5,0.3\rangle$
$c_{11}^{(2)N} = \langle (8,12,14,18);0.7,0.4,0.3\rangle$	$c_{21}^{(2)N} = \langle (25,35,40,50);0.9,0.2,0.4\rangle$
$c_{12}^{(2)N} = \langle (30,40,45,55);0.8,0.2,0.6\rangle$	$c_{22}^{(2)N} = \langle (11,16,20,25);0.6,0.3,0.5\rangle$
$c_{13}^{(2)N} = \langle (18,24,30,36);0.6,0.2,0.5\rangle$	$c_{23}^{(2)N} = \langle (18,26,32,40);0.7,0.3,0.4\rangle$

$$a_1^N = \langle\ 60,80,100,120);0.8,0.3,0.4\rangle,\quad a_2^N = \langle\ 45,65,85,105);0.7,0.3,0.5\rangle,$$
$$b_1^N = \langle\ 35,55,75,95);0.6,0.2,0.5\rangle,\quad b_2^N = \langle\ 20,30,40,50);0.9,0.4,0.6\rangle,$$
$$b_3^N = \langle\ 50,60,70,80);0.6,0.2,0.7\rangle.$$

Now, the mathematical formulation of the problem can be stated as follows:

$$\begin{aligned}
&\min \quad Z_1^N = \sum_{i=1}^{2}\sum_{j=1}^{3} c_{ij}^{1)N} x_{ij},\\
&\min \quad Z_2^N = \sum_{i=1}^{2}\sum_{j=1}^{3} c_{ij}^{2)N} x_{ij},\\
&\text{s.t} \quad \sum_{j=1}^{3} x_{ij} \le a_i, \qquad i = 1{,}2,\\
&\sum_{i=1}^{2} x_{ij} \ge b_j, \qquad j = 1{,}2{,}3,\\
&x_{ij} \ge 0 \quad \forall\ i = 1,2\ \&\ j = 1,2,3.
\end{aligned} \tag{14}$$

Using the notion of accuracy *Function (4)*, the crisp version of *Problem (14)* can be stated as follows:

$$\begin{aligned}
&\min \quad Z_1' = 27.125\,x_{11} + 40.625x_{12} + 75x_{13} + 45x_{21} + 55.125x_{22} + 37.5x_{23},\\
&\min \quad Z_2' = 9.1\,x_{11} + 34\,x_{12} + 19.575\,x_{13} + 31.875\,x_{21} + 12.66\,x_{22} + 21.75\,x_{23},\\
&\text{s.t} \quad x_{11} + x_{12} + x_{13} \le 65.25,\\
&x_{21} + x_{22} + x_{23} \le 54.375,\\
&x_{11} + x_{21} \ge 47.125,\\
&x_{12} + x_{22} \ge 27.125,\\
&x_{13} + x_{23} \ge 50.375,\\
&x_{ij} \ge 0 \quad \forall\ i = 1,2\ \&\ j = 1,2,3.
\end{aligned} \tag{15}$$

The above problem is solved by taking only one objective function and neglecting the others. The solution sets are obtained as follows:

$$\begin{aligned}
&z_1 = 4212.281,\\
&x_{11} = 47.125,\ x_{12} = 18.125,\ x_{13} = 0,\ x_{21} = 0,\ x_{22} = 9,\ x_{23} = 45.375.\\
&z_2 = 1718.097,\\
&x_{11} = 47.125,\ x_{12} = 0,\ x_{13} = 18.125,\ x_{21} = 0,\ x_{22} = 27.125,\ x_{23} = 27.25.
\end{aligned}$$

For each objective, the best and worst values are given as:

$$U_1 = 5154.781,\quad L_1 = 4212.281,\ \text{and}\ U_2 = 2145.394,\quad L_2 = 1718.097.$$

After constructing *Problem (13)* using linear membership functions defined in *Relations (8)-(10)* and considering $w_1 = w_2 = w_3 = \frac{1}{3}$, we solved it by Lingo software, the following solution is obtained:

$$\begin{aligned}
&\alpha = 0.541,\quad \beta = 0.495,\quad \gamma = 0.499,\\
&x_{11} = 47.125,\quad x_{12} = 7.172,\quad x_{13} = 10.843,\\
&x_{21} = 0,\quad x_{22} = 14.843,\quad x_{23} = 39.532.
\end{aligned}$$

7 | Conclusion

In this paper an effective modeling and optimization framework for the NFMOLPP is presented, where the coefficients of the objective functions, constraints and right-hand side parameters are single-valued trapezoidal NN. In the proposed method, a ranking function of NNs is used to convert the NFMOLPP into an equivalent crisp MOLPP. Then, using the best and worst values of objectives, an appropriate membership function for each objective function is defined to avoid decision deadlock situation in hierarchical structure. In this regard, according to the concept of fuzzy decision set, an optimal compromise solution is selected as the design which it maximizes the degree of acceptance and minimizes the degree of rejection upto some extent and rejection degree by taking all objectives simultaneously. The proposed approach can be used to solve real-world problems arising in industries and business organizations with imprecise and contradictory information. Finally, a transportation problem has been discussed to show the applicability of proposed approach.

Conflicts of Interest

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

[1] Ahmad, F., Ahmad, S., Zaindin, M., & Adhami, A. Y. (2021). A robust neutrosophic modeling and optimization approach for integrated energy-food-water security nexus management under uncertainty. *Water*, *13*(2), 121.

[2] Ahmad, S., Ahmad, F., & Sharaf, M. (2021). Supplier selection problem with type-2 fuzzy parameters: a neutrosophic optimization approach. *International journal of fuzzy systems*, *23*, 755-775.

[3] Ahmadini, A. A. H., & Ahmad, F. (2021). Solving intuitionistic fuzzy multiobjective linear programming problem under neutrosophic environment. *AIMS mathematics*, *6*(5), 4556-4580.

[4] Atanassov, K. T. (1986). Intuitionistic fuzzy sets. Fuzzy sets and systems, *20*(1), 87-96.

[5] Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. *Management science*, *17*(4), B-141. https://doi.org/10.1287/mnsc.17.4.B141

[6] Borovička, A. (2020). New complex fuzzy multiple objective programming procedure for a portfolio making under uncertainty. *Applied soft computing*, *96*, 106607. https://doi.org/10.1016/j.asoc.2020.106607

[7] Das, S. K. (2022). An approach to optimize the cost of transportation problem based on triangular fuzzy programming problem. *Complex & intelligent systems*, *8*(1), 687-699. https://doi.org/10.1007/s40747-021-00535-2

[8] Kumar Das, S., Edalatpanah, S. A., & Dash, J. K. (2021). A novel lexicographical-based method for trapezoidal neutrosophic linear programming problem. *Neutrosophic sets and systems*, *46*(1), 12. https://digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1944&context=nss_journal

[9] Das, S. K. (2020). Application of transportation problem under pentagonal neutrosophic environment. *Journal of fuzzy extension& applications*, *1*(1), 27- 41. https://doi.org/10.22105/jfea.2020.246633.1001

[10] Deli, I., & Şubaş, Y. (2017). A ranking method of single valued neutrosophic numbers and its applications to multi-attribute decision making problems. *International journal of machine learning and cybernetics*, *8*(4), 1309-1322. https://doi.org/10.1007/s13042-016-0505-3

[11] Khalifa, A. E. W. (2019). A signed distance method for solving multi-objective transportation problems in fuzzy environment. *International journal of research in industrial engineering*, *8*(3), 274-282.

[12] Khalifa, H. A. E. W., Kumar, P., & Mirjalili, S. (2021). A KKM approach for inverse capacitated transportation problem in neutrosophic environment. *Sādhanā*, *46*(3), 1-8.

[13] Khalifa, H. A. E. W., Kumar, P., & Alharbi, M. G. (2021). On characterizing solution for multi-objective fractional two-stage solid transportation problem under fuzzy environment. *Journal of intelligent systems*, *30*(1), 620-635.

[14] Khalifa, H., Elhenawy, M., Masoud, M., Bhuiyan, H., & Sabar, N. R. (2021). On multi-objective multi-item solid transportation problem in fuzzy environment. *International journal of applied and computational mathematics*, *7*(1), 24. https://doi.org/10.1007/s40819-021-00961-3

[15] Khalifa, H. A. E. W. (2020). Goal programming approach for solving heptagonal fuzzy transportation problem under budgetry constraint. *Operations research and decisions*, *30*(1), 85-96.

[16] Khalifa, H. A. E. W. (2019). Fuzzy compromise approach for solving interval-valued fractional multi-objective multi-product solid transportation problems. *Journal of system management*, *5*(2), 1-20.

[17] Liu, P., & Liu, X. (2018). The neutrosophic number generalized weighted power averaging operator and its application in multiple attribute group decision making. *International journal of machine learning and cybernetics*, *9*, 347-358.

[18] Mahajan, S., & Gupta, S. K. (2021). On fully intuitionistic fuzzy multiobjective transportation problems using different membership functions. *Annals of operations research*, *296*, 211-241.

[19] Maiti, I., Mandal, T., & Pramanik, S. (2020). Neutrosophic goal programming strategy for multi-level multi-objective linear programming problem. *Journal of ambient intelligence and humanized computing*, *11*(8), 3175-3186.

[20] Rizk-Allah, R. M., Abo-Sinna, M. A., & Hassanien, A. E. (2021). Intuitionistic fuzzy sets and dynamic programming for multi-objective non-linear programming problems. *International journal of fuzzy systems*, *23*, 334-352. https://doi.org/10.1007/s40815-020-00973-z

[21] Singh, S. K., & Yadav, S. P. (2015). Modeling and optimization of multi objective non-linear programming problem in intuitionistic fuzzy environment. *Applied mathematical modelling*, *39*(16), 4617-4629. https://www.sciencedirect.com/science/article/pii/S0307904X15002681

[22] Smarandache, F. (1999). A unifying field in logics: neutrosophic logic. In *Philosophy* (pp. 1-141). American Research Press.

[23] Ye, J. (2017). Bidirectional projection method for multiple attribute group decision making with neutrosophic numbers. *Neural computing and applications*, *28*, 1021-1029. https://doi.org/10.1007/s00521-015-2123-5

[24] Ye, J. (2018). Neutrosophic number linear programming method and its application under neutrosophic number environments. *Soft computing*, *22*, 4639-4646. https://doi.org/10.1007/s00500-017-2646-z

[25] Ye, J. (2014). Prioritized aggregation operators of trapezoidal intuitionistic fuzzy sets and their application to multicriteria decision-making. *Neural computing and applications*, *25*, 1447-1454. https://doi.org/10.1007/s00521-014-1635-8

[26] Yu, G. F., Li, D. F., Liang, D. C., & Li, G. X. (2021). An intuitionistic fuzzy multi-objective goal programming approach to portfolio selection. *International journal of information technology & decision making*, *20*(05), 1477-1497. https://doi.org/10.1142/S0219622021500395

[27] Wang, H., Smarandache, F., Zhang, Y., & Sunderraman, R. (2010). *Single valued neutrosophic sets*. Infinite Study.

[28] Wang, Q., Huang, Y., Kong, S., Ma, X., Liu, Y., Das, S. K., & Edalatpanah, S. A. (2021). A novel method for solving multiobjective linear programming problems with triangular neutrosophic numbers. *Journal of mathematics*, *2021*, 1-8. https://doi.org/10.1155/2021/6631762

[29] Zadeh, L. A. (1965) Fuzzy Sets. *Information control*, *8*(3), 338-353. http://dx.doi.org/10.1016/S0019-9958(65)90241-X

Hosseinzadeh and Tayyebi| J. Appl. Res. Ind. Eng. 10(1) (2023) 1-10

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 11–24.

Paper Type: Research Paper

Analyzing Behavioral Patterns of Bus Passengers Using Data Mining Methods (Case Study: Rapid Transportation Systems)

Amir Daneshvar[1], Fariba Salahi[2,*], Maryam Ebrahimi[1], Bijan Nahavandi[3]

[1] Department of Information Technology Management, Faculty of Management, Electronic Branch, Islamic Azad University, Tehran, Iran; daneshvar.amir@gmail.com; mar.ebrahimi@gmail.com.

[2] Department of Industrial Management, Faculty of Management, Electronic Branch, Islamic Azad University, Tehran, Iran; salahi_en@yahoo.com.

[3] Department of Industrial Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran; bnahavandi@gmail.com.

Citation:

Daneshvar, A., Salahi, F., Ebrahimi, M., & Nahavandi, B. (2023). Analyzing behavioral patterns of bus passengers using data mining methods (case study: rapid transportation systems). *Journal of applied research on industrial engineering, 10*(1), 11-24.

Received: 26/08/2021 Reviewed: 01/10/2021 Revised: 09/11/2021 Accepted: 10/01/2022

Abstract

The aim of analyzing passengers' behavioral patterns is providing support for transportation management. In other words, to improve services like scheduling, evacuation policies, and marketing, it is essential to understand spatial and temporal patterns of passengers' trips. Smart Card Automated Fare Collection System (SCAFCS) makes it possible to utilize data mining tools for the purpose of passengers' behavioral pattern analysis. The specific goal of this research is to obtain functional information for passenger's behavioral pattern analysis in city express bus which is called BRT, and classification of passengers to improve performance of bus fast transportation system. Additionally, it is attempted to predict usage and traffic status in a line through predicting passenger's behavior in a bus line. This paper applies smart card data to provide combinational algorithms for clustering and analysis based on data mining. To this end, we have used a combination of data mining methods and Particle Swarm Optimisation (PSO) algorithm and leveraged multivariate time series prediction to estimate behavioral patterns. Results show that price and compression ratio features are the most influencing features in the separability of transportation smart card data. According to the Pareto front, four features: card identification number, bus identification number, bus line number, and charge times influence clustering criteria.

Keywords: Smart card, E-ticket, Artificial intelligence, Particle swarm optimisation, Behavioral pattern.

1 | Introduction

Urban society's health depends on macro policy makings in urban planning and especially transportation planning; thus, paying attention to public transportation is essential in daily life and is considered necessary for improving the quality of life in urban areas. Paying attention to this issue would reduce most people's problems significantly [1].

Transportation is one of the main milestones of urban development, used to transport people and goods between different spaces and locations. By extension and expansion of metropolises and increasing demand for fast and cost-efficient intercity transportation, the need for efficient urban systems that can transport a large volume of passengers resulted in the emergence of metro and urban express buses, which are called BRT. Currently, this system is considered one of the most

Corresponding Author: salahi_en@yahoo.com

https://doi.org/10.22105/jarie.2021.302560.1373

critical servers. According to this point, we should take care that if ignoring passengers' short-term and long-term behavioral patterns and not serving them efficiently, we would face a reduction in using these services. Subsequently, there is the possibility of using personal cars, urban traffic, and environmental problems [2]. To establish a comprehensive and appropriate transportation system, designers should predict transportation demand by altering system characteristics and how people use it. Trip demand models, significantly predicting trip characteristics and application of transportation facilities under different scenarios, are proposed [3]. To realistic predictions in trip demand modeling, we need to use behavioral features. Nowadays, these needs have become more serious, since long-term guidelines have been replaced by short-term policies like work scheduled and remote communication and trip spatial-temporal patterns of passengers.

The proper understanding of separate passengers' behaviors helps transportation officials evaluate their current services and adjust their strategies to improve efficiency. For example, suppose that a company provides tickets applicable for one month; this would let passengers ensure consistency of their fares for one month. Evaluating their behavior would help estimate traffic and income charges. In most transportation devices, issues like not paying the cost or stealing exist. We should locate these inappropriate behaviors by passenger's behavioral analysis. One of the main problems in this area is the large number of outliers that increase errors in the clustering section. Also, in practice, it would result that the algorithm would not converge by applying a large amount of data. Another problem is the existence of multiple start points, which result in a difference in obtained results from the same data.

To overcome these problems, we should simultaneously improve clustering performance and the correct data set management. Thus, we should apply appropriate preprocessing, including correlation reduction and elimination of outliers, and simultaneously, we should use proper models and powerful clustering methods. Hence in this paper, to achieve this goal utilizing smart cards data, we have proposed efficient combined algorithms for clustering and analysis based on data mining. Generally, we could help to improve transportation system performance and efficient management in this method, utilizing operations like data cleaning, data labeling and clustering, classification, multiobjective optimisation, and intuitive analysis.

The current study has utilized data mining and artificial intelligence to determine how we could predict passengers' behavioral patterns using urban express buses operating smart card (e-ticket) data?

2 | Literature Review

A review of internal and external research shows that planning on efficient transportation systems is essential for countries. As proposed in [4], a linear programming model minimize the set of line set up costs, maintenance, repair costs, fuel consumption, and the time of bus entrance to the station. The proposed model is simulated in GAMS software, and sensitivity analysis has been performed on that. According to the proposed model's high complexity and computational time, a combinational meta-heuristic method based on Particle Swarm Optimisation (PSO) and Ant Colony Optimization (ACO) was developed to solve it and simulated in MATLAB software. Results show that fuel cost and spatial distance between stations have the most influence on the welfare of passengers. Also, in terms of model complexity and solution time, the number of stations has more effect than the number of lines and buses.

The study of [5] is about studying the nature and dimensions of express buses' full passenger line using analyzing strengths and weaknesses, opportunities and threats, and evaluating the feasibility of concurrent usage of bus lines. The performance of Vahed Company bus drivers has been evaluated in [6]. The results have shown a meaningful correlation between danger and actual accidents of drivers with error, violations, age, and license type of drivers.

Daneshvar et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 11-24

Research on the topic of dynamic modeling of urban transportation systems has been performed in [7]. Its purpose was to provide a dynamic model for urban transportation systems regarding a stable urban transportation system. Using this model, they have understood the complex and dynamic nature of the urban transportation systems and evaluated the influence of policies. In this paper, they have utilized the systems dynamics approach to propose a comprehensive and integrated model for the Tehran urban transportation system regarding the concept of the stable urban transportation system and evaluated the system's dynamic behavior. Furthermore, the simulated model provides policies for an increase in urban transportation systems' efficiency and improve parameters of urban traffic. This model has been analyzed using Vensim software and Tehran city data. Results showed that combined components of proposed policies greatly influence enhancing the efficiency of the transportation system and statutes of Tehran city traffic parameters.

In recent years, many types of research have been performed on using smart card data in public transportation. Obtained data of smart cards have been used in many applications like analyzing demand amount and planning and others [8]. It has been used in [9], [10], a smart card Automatic Fare Collection (AFC) system to define various consumers and measure their trip habits; analyses were based on patterns in days, weeks, and seasons. Also, in many works, they have used the clustering method as a data mining method in e-ticket information classification. Among these, the scaled k-means++ enhanced method has been used for data with large volumes and a high changing rate [11]. The stages of data mining in this paper are mention below:

I. Data filtering.
II. Data clustering (using k-means and HAC algorithms).
III. Describing groups.

In this work, we have used features of time, date, and type of card and correspond line direction (25452) of a smart card; after clustering, we obtained four groups of clusters. Results have shown that peak usage of these individuals is in business center closure, and these individuals usually use transportation systems to round trips between work and home [10]. In [9], daily data of 4.47 million cards are used for various analyses. In these works, only three-parameter of the card number, fare type, and date and time. They have used k-means clustering to classify data. According to the data vision and performed distance calculations, the number of clusters has been set to three. Results show that the first cluster has a transaction peak on a typical day, and we could say that registered trips in this cluster have extended during fay. The second cluster corresponds to morning trips; since post-morning transactions were very low, many transactions have been reported between 9 to noon. In the third cluster, most trips have registered in post morning and evening. Also, the transaction percentage of the three clusters shows that on working days (Monday to Friday), individuals in the first cluster had the most transactions. On weekend days, individuals in the second and third clusters had the most transportation system usage.

In [11], smart card data and GPS are used to provide a method for estimating the original destination of the OD matrix of Santiago city public transportation. Considering get-off time and location information, they have used two datasets of one week and in different periods to estimate more than 80% of transactions. In [12], a method is proposed based on Markoff save and Bayesian decision-making tree to infer stop stations of bus passengers. This research has to utilize smart card data and data received from GPS installed in buses. In [13], a method is proposed for data extraction to monitor the real-time performance of transportation systems based on smart card data and GPS data.

In [14], a method is proposed for determining significant places based on users' mobile phones (active android). Detection of inappropriate behaviors somehow is an outlier's detection problem. Many methods have been proposed to determining these outliers, which we could mention various techniques based on classification, distance-based on clustering, and others among them. In the following, we have classified performed methods in three main classes, and we will generally evaluate works in these three groups.

2.1 | Strategy Level Studies

In this section, performed works are proportional to long-term planning. Many researchers suggest that we should utilize smart card information's in long-term planning. For example, the study of [10] has mentioned that analysis of smart card data results in a better understanding of users' behaviors. In *Table 1*, a review of performed studies at the strategic level has been presented. The table shows that most of these works are on users' properties and classification without having users' previous private information. In all performed works using smart card data, knowledge of the date, time, and card number are available. We could calculate adequate statistics using them. Of course, since usually information of user person is not available, we will use traditional methods like surveys to compensate for these defects. As it is mentioned earlier, we could use smart cards to measure users' loyalty.

Table 1. Studies that used smart card data for strategic transportation planning.

Reference	Data	Type of Analysis or Data Application	Advantages of Obtained Results
Agard et al. [9]	Data of stations, time, location, and type of card	Definition of regular users and measuring their trip habits. analysis using daily, weekly, or seasonally variations	A better understanding of user behaviors
Alfred Chu et al. [15]	Estimated get-off points and type of card	Temporal-spatial image of the network. Driver Assistance Benefits Implication (DABI)	Improve network lines' geometry and stations schedule. Adopting transportation networks to users demands
Park et al. [16]	Historical data	Estimating of future behavioral trends. Creating future demand matrix	Long-term service adoption Network expansion and predicting requirements of network expansion
Trépanier and Morency [17], Trépanier et al. [18], [19], Trépanier and Vassiviere [20]	Trips transactions using the card. Dates of begin and end of trips	Computing intervals in which users have used cards	Modeling user's loyalty using smart card

This paper will propose an efficient combinational algorithm in clustering and analysis based on smart card data. A novel aspect of this work is utilizing a combination of data mining and artificial intelligence methods to obtain a maximum of useful information and evaluate results obtained from using multivariate time series prediction to estimate behavioral patterns, which has not been done previously. Most of the presented works have used single and straightforward algorithms to perform data mining (evaluation of outliers, clustering, and optimizing). Also, we attempt to predict the status of usage and traffic on a bus line, using passenger behavior prediction in that line.

2.2 | Comparison of Studies

A classification model in order to detect unusual events in the daily life behavior patterns of users of mobile phones using pattern recognition is developed in [21]. The machine learning model was used in [22] to analyze online learning behavior of students. In [23], artificial intelligence approaches useful for behavioral modeling are reviewed among which machine-learning methods sound to be more effective. In [24], behavior analysis with the use of machine learning was introduced.

Table 2. Comparison of studies.

Reference	Pattern Recognition	Machine Learning	Simulation
Ahn and Han [21]	✓		
Yan and Au [22]		✓	
Osoba and Davis [23]		✓	
Ceja [24]		✓	
The present study	✓	✓	✓

Recently, some studies have been published ablout forecasting through using Artificial Neural Networks (ANNs). For example, Aliahmadi et al. [25] compared the use of traditional methods and neural networks in time series forecasting. The study of [26] used trend analysis on the time series data and regression analysis to find the cause-and-effect relation among variables. In [27], ANN approaches are applied in change point analysis.

3 | Research Methodology

First, in this section, we have provided a flowchart of the proposed method. As mentioned, to manage and analyze the behavior of passengers in a transportation system, we could utilize smart card data to have a helpful efficiency. A way to achieve this goal is through data mining techniques. Generally, in these methods, we help improve transportation systems' performance and efficient management, using operations like data cleaning, labeling and clustering of data, classification, multi-object optimisation, and intuitive analyses. Finally, we could consider the data mining model the same as the model shown in *Fig. 1*. In this paper, we use several stages, which have been mentioned below:

3.1 | Data Preprocessing

In this stage, we eliminate the noises of data using a median filter. We will convert properties like time, date, and card numbers to integer numbers in data preprocessing to facilitate clustering. Additionally, we have normalized data to reach meaningful results in neural networks.

3.2 | Dimension Reduction and Feature Selection

In the first layer, we have performed a feature selection procedure. We have to utilize multi-object optimisation methods to perform unsupervised feature selection. The proposed method for this layer is to use a multi-object PSO algorithm as a multi-object optimisation method. Proposed object functions are the number of features η_f and statistical validation criteria of space. We could show that the value of these criteria would be reduced with a reduction in various data. The main stages of the Multi Objective Particle Swarm Optimisation (MOPSO) algorithm have mentioned below:

1. Generation of primary population and initialization of speed and location of each particle (in initialization, we will set particle speed vector equal to zero, and we will select location vector randomly).
2. Calculation of cost functions for particle.
3. Finding non-dominant members of the population and storing them in the archive.
4. Generation of hypercubes in object space and positioning them so that these hypercubes represent coordination system and the cost function of particle represent coordination of each particle.
5. Each of the particle randomly chooses a leader from the archive and move toward it.

If the termination condition is not meet, go to Stage 3 and otherwise terminate the algorithm.

Regarding feature selection, we have encoded each feature's existing with "1" and the lack of that feature with "0". Hence, in the application of the PSO algorithm, feature selection of particle' components location vector is in terms of 0 and 1. In Coello et al. [28], a jump based on uniform distribution has been proposed to improve the algorithm's performance. In this study, we suggest including jump and intersection operators in the BMOPSO algorithm. The procedure is that we apply the jump operator on 20% of particle with the most cost, and (after excluding the previous 20%) we apply the intersection operator on 60% of particle with the most price. In the intersection operator, we randomly choose two-point of two-particle vector (parents), and then we exchange values of two particle between these intersection points.

In the proposed method, a one-bit jump means converting 0 to 1 and, inversely, a jump probability of pm. We randomly choose a number between 0 and 1, and then if this number is lower than pm, we change the corresponding bit, and otherwise, we will not change the bit. We apply jump on bits forming particle independently. Using intersection and jump operators has resulted in good improvement in the proposed algorithm of multi-object optimisation.

In the next stage, we pass selected features along with the number of clusters to the clustering algorithm, and then we calculate space statistics using these separated data. In this section, we have used the k-means algorithm as a clustering algorithm. In multi-object PSO algorithm or BMOPSO, each particle contains $\eta_f + k_c$ bit in which η_f is the number of features and k_c is the number of clusters, and have encoded in 4-bit form $k \in \{2, \dots, 17\}$.

3.3 | Clustering of Smart Card Data

Matrix X is original data with a dimension of $N \times d$, in which N is the number of observations and d is the number of features. The output of the first layer is matrix X with dimensions of $N \times R$, in which R≤d is selected features. The second layer is responsible for clustering data in the K cluster. The proposed method in this layer is a combination of Rough K-means clustering algorithm and PSO optimisation algorithm. Hence, first, we choose cluster centers in the PSO algorithm to calculate cluster centers according to defined objective function; then, we will apply these centers to the Rough k-means clustering algorithm as the initial average. In applying the Rough K-means algorithm, we use the ratio of distances instead of the distance difference, as suggested in the following equation. Using distance ratio had produced a better result for outliers clustering. Additionally, we have used a single-object particle swarm algorithm to determine w1 parameters and ε threshold values.

$$\frac{\mathrm{d}\ x_n, m_i)}{\mathrm{d}\left(x_n, m_j\right)} \leq \xi, \qquad (1)$$
$$\mathrm{d}\ x_n, m_i) - \mathrm{d}\left(x_n, m_j\right) \leq \xi.$$

3.4 | Predicting Status of Some Lines and Stations

We use the delayed equation and function of time series. We can utilize linear models like ARIMAX and nonlinear models like ANNs to find this function. In this paper, we have used GDMH neural network to predict two-variable time series. Generally, to expect single-variable time series, we could say that representation of time series is a function of delayed samples, and our purpose is to find this function. In other words, with this assumption that:

$$t) = f\left(x\ t - d_1), x\ t - d_2), \dots, x\ t - d_n)\right), \qquad (2)$$
$$\mathrm{Delay} = \{d_1, d_2, \dots, d_n\}.$$

The main goal is obtaining function f, which describes time series in the best possible way. In this paper, we have used GMDH neural network to predict two-variable time series. In this type of neural network, suppose x is input and y is the model's output; the relation between input and output is mentioned below:

$$y = a_0 + \sum_{i=1}^{m} a_i\ f_i. \qquad (3)$$

In this stage, we test the trained network and then evaluate obtained results to predict lines status.

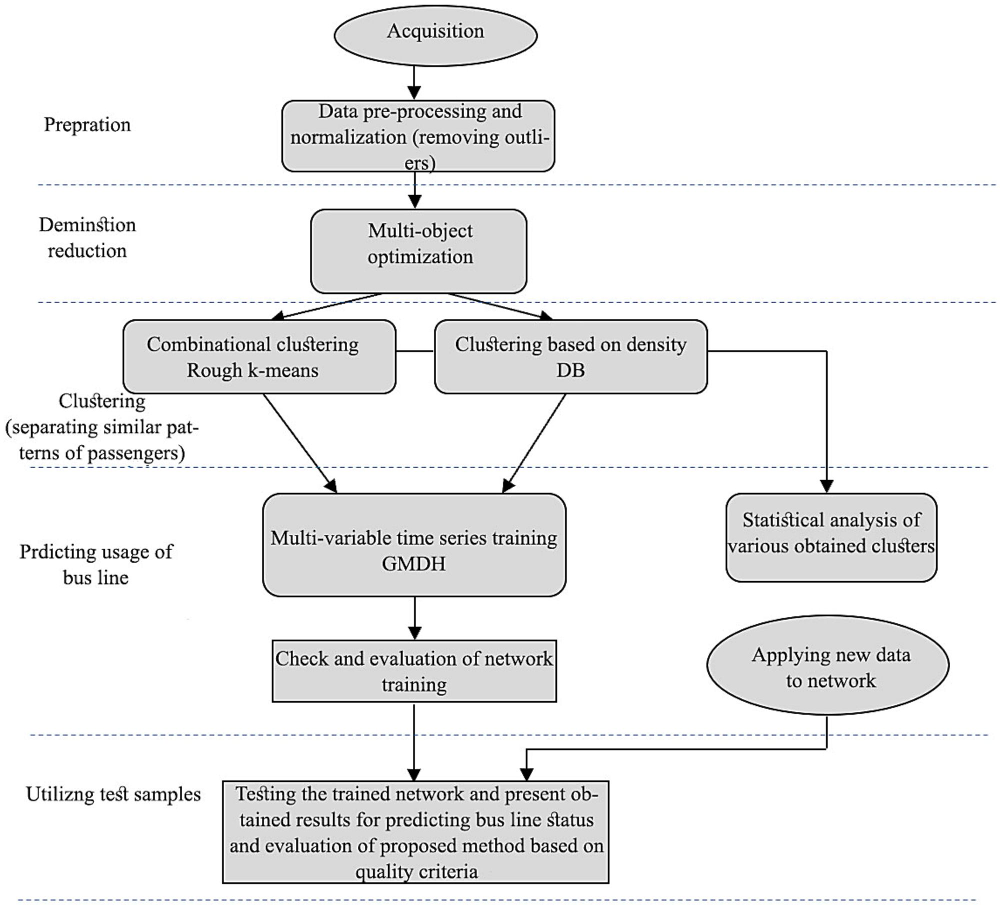


Fig. 1. A general framework for evaluation of behavioral patterns of the passenger.

This study has been performed utilizing smart card data and collected information from the automatic transportation system of Isfahan Card. The obtained data is related to lines (1) to (3) of express bus and lines (9), (21), (86), (91), and (4) of the urban bus, which all are in the Isfahan metropolis. The total number of stations correspond to these eight lines is 151 stations.

The given range for datasets is one month (October 2017) and includes (6, 571, 123) samples that have been acquired from 880,000 smart cards. Each of traffics has an identification number for passengers in a time interval of traffic and name and specifications of traffic station and name and specification of destination and information of the vehicle. Additionally, in case of traffic with the bus (internal or external), the direction of the trip is mentioned. AFC requires that passengers have extended their smart cards. To perform batch analysis of passengers, first, we should reconstruct the trip chain according to performed traffics with a smart card. If we use raw traffics to describe trip behavior, passengers involved in multi-stage trips will be considered more active than single sage passengers, which bias the classification process.

Trip chains are reconstructed by a two-stage method. The first stage is based on the following assumptions:

I. Assumption of the closest station: the passenger will place in a station nearest to his living location for specific traffic, and the beginning of the trip is from there.
II. Daily symmetry assumption: the last station is on symmetry to begin station or is close to it. In each traffic, we will predict the distance to all available stations in a line, and the nearest station is considered as the destination of the current trip. If the selected destination and following locations have a reasonable distance for walking, a location close to the current location is considered as the next station; otherwise, destination recognition will not be performed.

4 | Simulation Results

This study has been performed utilizing smart card data and collected information from the automatic transportation system of the Isfahan Card service. As it is mentioned, the user data are related to 8 bus lines with 985,448 samples (after excluding 11% of trips with low transactions) from 151 bus stations in Isfahan city. Additionally, these data have seven feature which is mentioned below:

I. Card Identification number (Card-ID).
II. Bus Identification number (Bus-ID).
III. Transaction Type (Trn-Type).
IV. Charge numbers (Ch-num).
V. Charge amount (Ch-amount).
VI. Cord number (card-num).
VII. Bus Line number (Bus-Line).

4.1 | Applying Feature Selection Algorithm

In the first step of applying data to the feature selection algorithm, the main features and the number of clusters between 2 to 15 are involved in the form of 4-bit gray code. Thus $n_f = 7$ and $k_c = 4$, and each particle will be in an 11-dimensions particle swar algorithm. Use parameters in the BMOPSO optimisation algorithm are mentioned in *Table 3*.

Table 3. Parameters in the BMOPSO optimisation algorithm.

γ	Mu	C1	C2	Population	W	α	β	Iteration
2	0.1	1	2	100	0.45	0.1	2	10

Fig. 2 shows the estimated Pareto front and other members after five epochs of the feature section algorithm with primary data. Dominated members of the population who form the Pareto front are highlighted with red color. Aside from overwhelmed members, several selected features are mentioned.

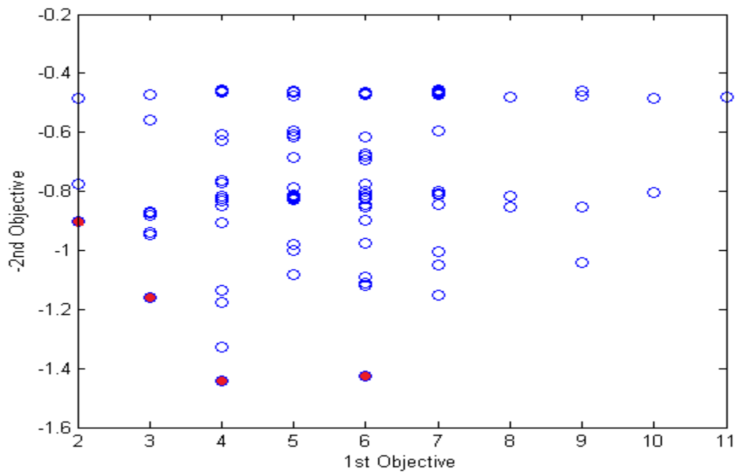


Fig. 2. Estimated Pareto front along with other members after five epochs of the feature selection algorithm.

As evident from the figure, minimizing two objective functions of several features and free space statistics validation criteria shows that price and density ratio features are the most significant features in the superstability of transportation smart card data. Finally, according to simulated results in ranking feature selection, we could consider the features as below:

I. Card-ID.
II. Bus-ID.
III. Bus-Line.

Daneshvar et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 11-24

IV. Ch-num.
V. Card-num.
VI. Trn-Type.
VII. Ch-amount.

Also obtained results for feature numbers of 4 to 7 with cluster number 1/100 show an optimized number of clusters is 9. According to the obtained Pareto front, we conclude that only 4 of the most influencing features in clustering criteria and 9 clusters will not have significant error. In the clustering layer, we have used only these four features.

4.2 | Data Clustering

Used parameters for the PSO algorithm are mentioned in *Table 4*. Also, paraments of Rough k-means algorithm are obtained from imperialist competitive optimisation algorithm with cost function of Davis Bolden clustering criteria. *Fig. 3* used parameters in the PSO algorithm.

Table 4. Parameters in the PSO algorithm.

Iteration (MAX)	Population	# of Empires	A	B	ζ	Mu
1000	100	15	1	2	0.1	0.1

Fig. 3 shows the training curve for the proposed algorithm and PSO-Rough k-means algorithm. The proposed method outperforms the PSO method, at least in cost function minimization.

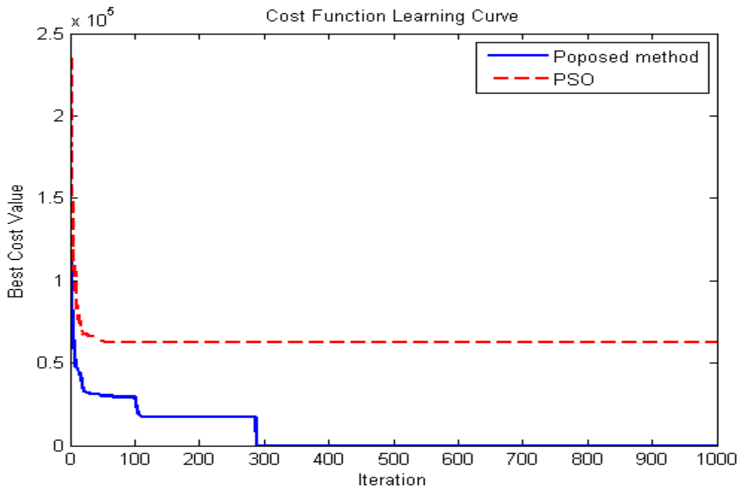


Fig. 3. Training curve of the proposed method and PSO-Rough k-means.

First, obtained clusters are studied according to the performed transactions' number in each station during days of the week (except Friday, excluded data in this analysis.)

Evaluating the activity of these stations shows that there are two general categories in clusters (*Fig. 4*). The first category includes clusters in which the number of daily transactions has had some peaks and are usually used during all hours of the day. Six clusters from 9 clusters have these properties, including 75.5% of stations (114 stations). The second category comprises clusters in which the amount of transactions in corresponding statins in a half-day is different from another half-day. These states include clusters with 24.5% (37 stations). And we could say that users use these stations to their workplace in the center of the city. *Fig. 4* shows a state of stations with usage peak in hours 7 to 8 in the morning.

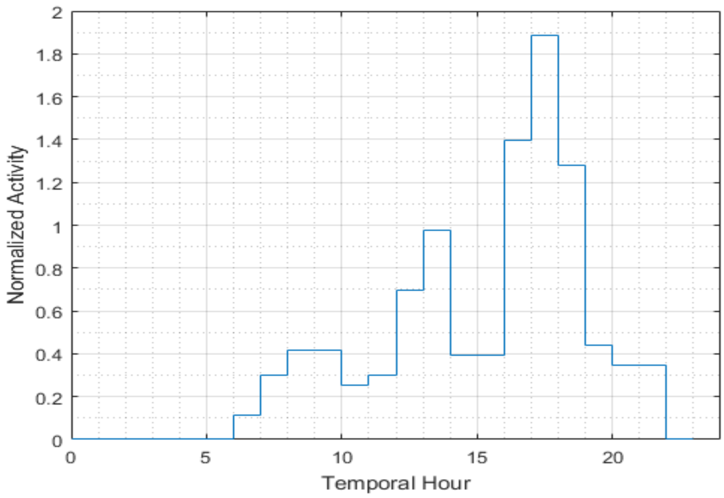


Fig. 4. Plots of two clusters that had usage peak in the morning.

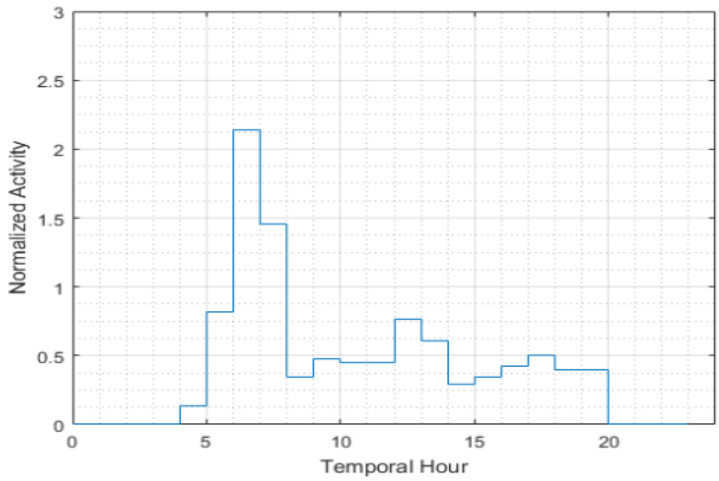


Fig. 5. Cluster plot which shows usage peak in post morning.

Fig. 5 shows the reverse state of two previous clusters which usage peak is in the evening in hours 17 to 18. A number of these stations are industrial areas, and many of them are in the city center. By studying this cluster, we notice that these stations correspond to individuals who have used these statins way back from industrial or official centers toward their homes. No significant usage from these stations is recorded for the rest of the day.

Fig. 6 shows the distribution of all clusters on days of the week. The aim is to evaluate passengers' behaviors during days of the week. As it is evident from this figure, in business days (Saturday to Wednesday), most of the usage was related to individuals if fifth, seventh, eighth, and ninth clusters, and in weekends (Thursday and Friday), most uses were related to first and third and sixth clusters. According to the plot, we could say that fifth, seventh, eighth, and ninth clutters related to users who most of them are a student it is the transportation system for going (or returning) from the workplace. On the other hand, first, third, and sixth clusters include users who have used public transportation to spend their weekends in most visited areas or found to the suburbs of Isfahan (by going to available terminals in the city) on their weekends. The second and fourth clusters have almost similar distribution on both business days and weekends. These clustered include industrial areas users who also go to their workplace as shifts in the last days of weeks.

Daneshvar et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 11-24

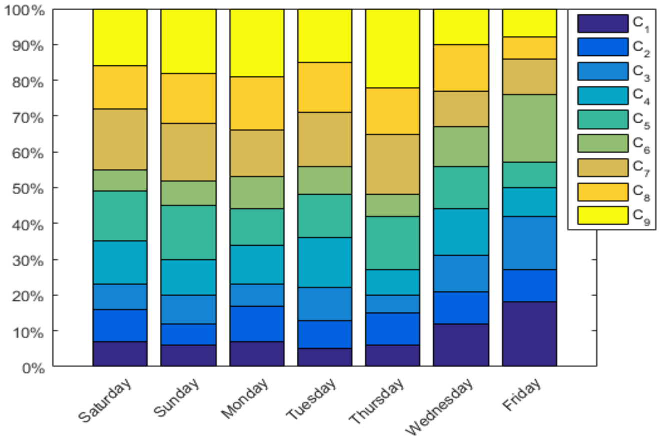


Fig. 6. Usage percent plot of each cluster during days of the week.

4.3 | Station Prediction

In this section, we have discussed the prediction of two different bus stations in terms of clustering. The number of prepared temporal samples for a week in each station is 2016. Delay vectors have been considered as below, and prediction of one day earlier is desired. Corresponding delays have been considered for a half-hour, one hour, one day, and two days.

Delays = [6 12 288 576].

Time series related to specified station from line 91 is shown in *Fig. 7*. Moreover, *Fig. 8* has shown the actual time series and estimated time series for training and test data station correspond to line 91, respectively. We should note that 80% of data is used for training, and 20% is used for testing.

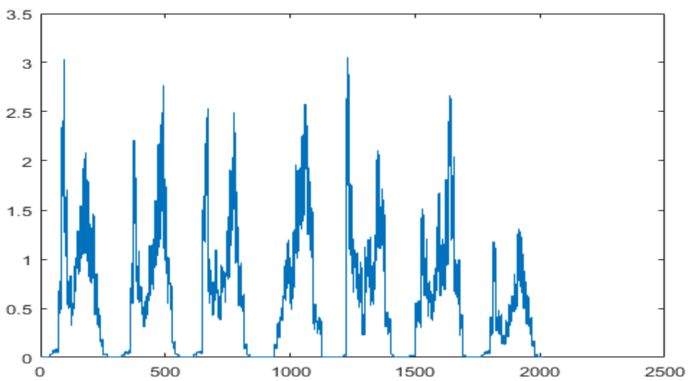


Fig. 7. Time series related to specified stations from line 91.

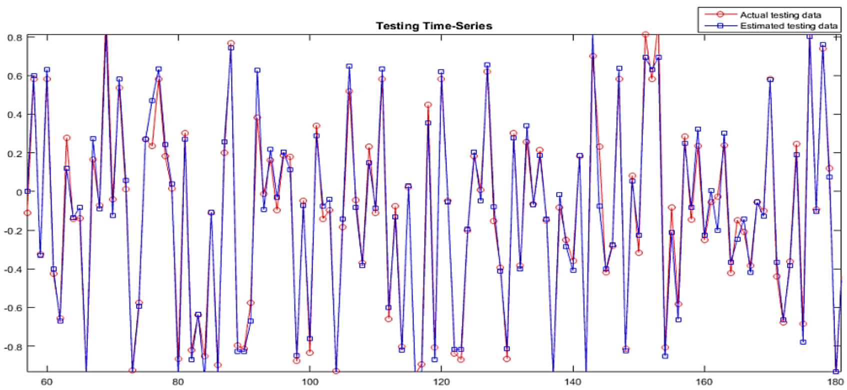


Fig. 8. Actual and estimate time series for test data in station related to line 91.

Daneshvar et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 11-24

Fig. 9 has shown regression for training data and test data, respectively. As it is evident from these figures, prediction performance in this state is significantly more robust.

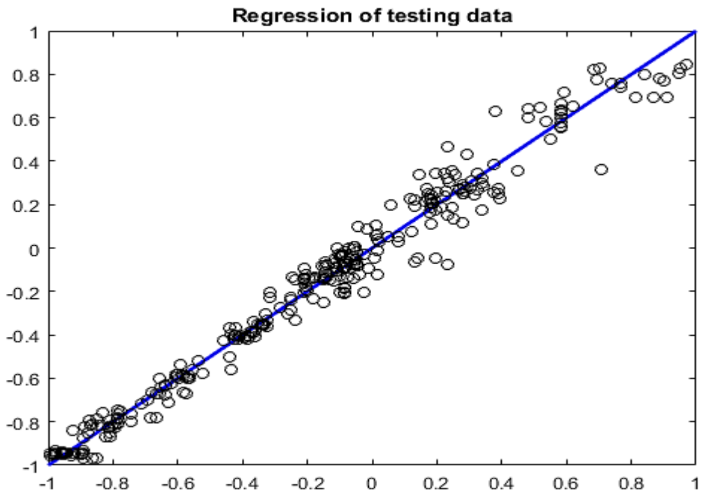


Fig. 9. Regression for test data.

5 | Conclusion

In this paper, we analyzed obtained data from the smart card of express transportation systems using data mining and artificial intelligence techniques. Three main axes considered in this analysis include:

I. Clustering stations and evaluating their activity statues.
II. Clustering from users (passengers) pint of view and comment usage behavior of passengers during the week.
III. Predicting status of two different stations using two-variable time series.

Obtained results for data related to one month included in 9 clusters. Stations have been placed in three general statuses: stations with high transactions in the morning, stations with the high transaction in the post morning, and stations with transactions during the day. This category has devoted most stations to itself. Results have shown that price and density ratio features are the most influencing features inseparability of transportation smart card data. According to obtained Pareto front, it has shown that the use of four influencing features includes card identification number, bus identification number, bus line number, and charge number in clustering criteria, will not generate a significant error and in the clustering layer, we have only used from these features. In the passenger's behavior section, the percentage of each cluster during days of the week is presented. The first category was clustered a whole number of transactions during the day had some peaks and generally, these stations have been used during all hours of a day. This category includes 6 clusters which form 75.5% of stations (114 stations). The second category includes clusters whose transactions in the different stations in half of the day were different from another half of day, which cost 24% station (37 stations). In the last step, the status of two distinct stations in form of time series is considered and the prediction of one day ahead has been predicted using a neural network. Behavior extraction of old stations is an important step for the planning of managers for different stations in terms of usage amount, which could be utilized in infrastructure plans. Also, we could utilize analysis of passenger's behavior and identification and separation of stations in terms of time which passenger traffic volume is high, to planning transportation systems. Finally, appropriate moving of buses and appropriate scheduling in stations which have passenger traffic load, lead to improve statue and reduce passenger volume.

For future research, we could use feature combination methods like principal component analysis PCA or independent component analysis in feature selection. The data analysis section could use required preprocessing processes like eliminating outliers and filling lost features before applying them to the main algorithm. Also, we could increase the number of objective functions.

References

[1] Soltani, A. (2012). Public transportation in response to a common challenge in developing countries. *Quarterly report of Fars province building engineering organization*, *16*(55), 1-71. https://www.fceo.ir/NewsShow.Aspx?NewsId=Boxn5Q5U4vUN3ZVhMBrxgA

[2] Poorahmad, A., Omranzadeh, B. (2012). Evaluating and presenting strategies for developing the BRT transportation system in the metropolis of Tehran using the SWOT model. *Quarterly journal of urban research and planning*, *3*(11), 17-36. (**In Persian**). https://www.sid.ir/paper/220118/fa

[3] Guo, J. Y., & Bhat, C. R. (2001). *Representation and analysis plan and data needs analysis for the activity-travel system* (No. 0-4080-1). University of Texas at Austin. Center for Transportation Research. https://rosap.ntl.bts.gov/view/dot/4865

[4] Jabbari, E., & Saeedi, S. (2018). Solving the public transportation problem using linear programming and hybrid PS-ACO meta-heuristic method. *Journal of transportation engineering, 11*(2), 353-371. (**In Persian**). https://www.magiran.com/paper/2072392?lang=en

[5] Momeni, E., & Darvishzadeh, M. M. (2017). Investigating the nature and dimensions of passenger lines and the feasibility of using high-speed bus lines simultaneously with it in Tehran. *Rahvar research studies*, *6*(23), 43-72. (**In Persian**). https://ecc.isc.ac/showJournal/25525/101776/1316268

[6] Tavakoli Kashani, A., Mansoori Kargar, H. (2015). The effect of bus Drivers behavior of Tehran Branch Company on urban accidents. *Rahvar research studies, 5*(16), 39-62. (**In Persian**). https://www.noormags.ir

[7] Ahmadvand, A. M., Mohammadiani, Z., and Khodadadi, H. (2015). Dynamic modeling of sustainable urban transportation system in order to improve and improve traffic. *Management research in Iran*, *19*(2), 52-31. (**In Persian**). http://ensani.ir/fa/article/download/345444

[8] Pelletier, M. P., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: a literature review. *Transportation research part c: emerging technologies*, *19*(4), 557-568.

[9] Agard, B., Partovi Nia, V., & Trépanier, M. (2013). Assessing public transport travel behaviour from smart card data with advanced data mining techniques. *World conference on transport research* (Vol. 13, pp. 15-18). http://datawisdom.ca/paper/2013-Transport.pdf

[10] Agard, B., Morency, C., & Trépanier, M. (2006). Mining public transport user behaviour from smart card data. *IFAC Proceedings volumes*, *39*(3), 399-404.

[11] Dzikrullah, F., Setiawan, N. A., & Sulistyo, S. (2016). Implementation of scalable k-means++ clustering for passengers temporal pattern analysis in public transportation system (BRT Trans Jogja case study). *6th international annual engineering seminar (InAES)* (pp. 78-83). IEEE. https://doi.org/10.1109/INAES.2016.7821911

[12] Ma, X. L., Wang, Y. H., Chen, F., & Liu, J. F. (2012). Transit smart card data mining for passenger origin information extraction. *Journal of Zhejiang university science C*, *13*(10), 750-760. https://doi.org/10.1631/jzus.C12a0049

[13] Ma, X., & Wang, Y. (2014). Development of a data-driven platform for transit performance measures using smart card and GPS data. *Journal of transportation engineering*, *140*(12), 04014063. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000714

[14] Isaacman, S., Becker, R., Cáceres, R., Kobourov, S., Rowland, J., & Varshavsky, A. (2010). A tale of two cities. *Proceedings of the eleventh workshop on mobile computing systems & applications* (pp. 19-24). https://doi.org/10.1145/1734583.1734589

[15] Alfred Chu, K. K., Chapleau, R., & Trepanier, M. (2009). Driver-assisted bus interview: Passive transit travel survey with smart card automatic fare collection system and applications. *Transportation research record*, *2105*(1), 1-10.

[16] Park, J. Y., Kim, D. J., & Lim, Y. (2008). Use of smart card data to define public transit use in Seoul, South Korea. *Journal of the transportation research record*, *2063*(1), 3-9. https://doi.org/10.3141/2063-01

[17] Trépanier, M., & Morency, C. (2010). Assessing transit loyalty with smart card data. *12th world conference on transport research*. Lisbon. Portugal. https://www.researchgate.net

[18] Trépanier, M., Morency, C., & Agard, B. (2009). Calculation of transit performance measures using smartcard data. *Journal of public transportation*, *12*(1), 79-96.

[19] Trépanier, M., Tranchant, N., & Chapleau, R. (2007). Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of intelligent transportation systems*, *11*(1), 1-14.

[20] Trépanier, M., & Vassiviere, F. (2008). Democratized smartcard data for transit operator. *15th world congress on intelligent transport systems and its America's 2008 annual meetingits americaerticoits japantranscore*. New York NY, United States.
[21] Ahn, J., & Han, R. (2013). Personalized behavior pattern recognition and unusual event detection for mobile users. *Mobile information systems*, *9*(2), 99-122. DOI: 10.3233/MIS-130155
[22] Yan, N., & Au, O. T. S. (2019). Online learning behavior analysis based on machine learning. *Asian association of open universities journal*, 14(2), 97-106
[23] Osoba, O., & Davis, P. K. (2019). An artificial intelligence/machine learning perspective on social simulation: New data and new challenges. *Social-behavioral modeling for complex systems*, 443-476. https://doi.org/10.1002/9781119485001.ch19
[24] Ceja, E. G. (2021). *Behavior analysis with machine learning using R*. Chapman and Hall/CRC.
[25] Aliahmadi, A., Jafari-Eskandari, M., Mozafari, A., & Nozari, H. (2016). Comparing linear regression and artificial neural networks to forecast total productivity growth in Iran. *International journal of information, business and management*, *8*(1), 93-113.
[26] Udoka, C. G. (2020). The impact of passenger's traffic on exchange rate and economic growth in Nigerian aviation industry. *International journal of research in industrial engineering*, *9*(4), 364-378.
[27] Habibi, R. (2021). Poisson regression model with change points. *Journal of applied research on industrial engineering*, *8*(Special Issue), 1-8. https://doi.org/10.22105/jarie.2021.229051.1149
[28] Coello, C. A. C., Lamont, G. B., & Van Veldhuizen, D. A. (2007). *Evolutionary algorithms for solving multi-objective problems* (Vol. 5, pp. 79-104). New York: Springer. https://link.springer.com/book/10.1007/978-0-387-36797-2

Daneshvar et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 11-24

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

JARIE

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 25-44.

Paper Type: Research Paper

Solving a Multi-Objective Mathematical Model for Aggregate Production Planning in a Closed-Loop Supply Chain under Uncertain Conditions

Ommolbanin Yousefi[1,*], Saeed Rezaeei Moghadam[1], Neda Hajheidari[1]

[1] Department of Industrial Engineering, Faculty of Malek-e-Ashtar University of Technology, Isfahan, Iran; yousefi_1302@yahoo.com; kasra_kk2002@yahoo.com; Nedahajheidari@gmail.com.

Citation:

Yousefi, O., Rezaeei Moghodam, S., & Hajheidari, N. (2023). Solving a multi-objective mathematical model for aggregate production planning in a closed-loop supply chain under uncertain conditions. *Journal of applied research on industrial engineering, 10*(1), 25-44.

Received: 02/03/2021 Reviewed: 10/04/2021 Revised: 20/06/2021 Accepted: 03/08/2021

Abstract

One of the most important decisions taken in a supply chain is the issue of Aggregate Production Planning (APP) where a program-within a medium time-range- is determined for optimum manufacturing of all products using shared equipment and resources. This research presents a multi-objective model that helps the decision makers to make such decisions. The proposed model comprises four main objectives, the first one of which considers minimizing costs (including costs of manufacturing product, supplying, maintenance, inventory stock shortage, and expenditures related to man power). The second objective is defined as maximizing customers' satisfaction. Minimizing suppliers' satisfaction makes up the third objective and maximizing the quality of the manufactured products constitutes the fourth objective. In this model, the demand parameter is investigated under uncertain conditions; hence, other parameters influenced by this parameter are also presented under uncertain conditions occurring within three differing scenarios. This model is solved through LP-metric and the LINGO v14.0.1.55 software. At first the model is solved by means of numerical example; then it is solved by the actual data that are related to a military industry. Finally, process, variables like inventory level, overtime work hours etc, are valued with the help of closed-loop supply chain of the proposed model.

Keywords: Customers and suppliers' satisfaction, Aggregate production planning, Closed-loop supply chain, Multi-objective mathematical planning model.

1 | Introduction

In the industrial world of today, managers follow two strategic objectives: 1) optimizing operational efficiency of their organization, 2) extending and developing relations with other organizations. One of the relationships which bring about integration and unity among organizations is the concept of supply chain management [1]. Hence, the way the supply chain is designed has become a strategic decision in the process of supply chain management. The design plays a crucial role in the proper performance of supply chain. The problem that makes the design of the supply chain procedure all the more significant is the way waste products are managed efficiently. Put it differently, it is required that in the design of the supply chain, particular attention should be paid to the returned products, a problem which demands the creation of reverse supply chain.

Corresponding Author: yousefi_1302@yahoo.com

https://doi.org/10.22105/jarie.2021.275841.1265

The problem that makes the design of the supply chain procedure all the more significant is the way waste products are managed efficiently. Put it differently, it is required that in the design of the supply chain, particular attention should be paid to the returned products, a problem which demands the creation of reverse supply chain [2].

A suitable supply chain is a competitive advantage for companies and plants and help them to survive in competitive market [3]. An investigation made of supply chain models indicates that a huge portion of research is devoted to studying progressive supply chain. However, since 2005, when reverse supply chain was introduced, we have witnessed a plethora of research studies in the field [4]. While the forward supply chain concerns the flow from raw materials to end products and from manufacturer to consumer, the reverse supply chain concerns the reverse flow from consumer to manufacturer. In addition, some of supply chain concern both of reverse and forward flow. The name of this supply chain is closed-loop [5]. Recycling and reconstructing the products which are spending final stage of their life cycle are importantne in a reverse supply chain. In this regard, after gathering and inspecting the returned products, they are partitioned into recyclable and non-recyclable products. The recyclable products are carried to the recycle centers, where - based on their observable qualities - they undergo re-manufacturing (repair) process. Sometimes separation operation is executed on the returned products and reusable parts are utilized in manufacturing operations. The non-recyclable products are transported to extermination centers where they undergo safe elimination procedure [6]. Uncertainty is one of the key factors in the reverse supply chain that must be controlled, thus, the company could optimize the supply chain function [7]. What is obtained through reverse supply chain has great leverage on producers' programming. In fact, through implementing the above supply chain, there would be economic gains in production costs, usage of new facilities, and optimum exploitation of available resources, all playing effective roles in producer's decision making in the design of Aggregate Production Planning (APP). APP is a process which determines the optimal level of production and stock inventory to meet the demands for product on a long-term basis while simultaneously considers the capacity limitation of the means and resources [6]. One of the main decisions for effectiveness and responsiveness of manufacturing and supply chain systems is APP. In addition, it can help to determine the best way to utilize resources to meet forecasted demand [8].

In this research, investigation is carried out on designing and solving a mathematical model for APP in a closed-loop supply chain of a specific military industry. Military products are usually made up of chemical, mechanical, and electronic components. Inspection of the products in the supply chain of the mentioned industry is of a demolition type, that is, in case where the quality of the products is not confirmed by the employer (IRGC military force, IRI-Iranian-Army, law enforcement units, and foreign purchasers), they are end masse retuned to the supplier. The returned products are either demolished in the reconstruction units or delivered to the producer after reconstruction. Also, in case of the non-usage of the products by the customer after technical warranty expiration (10-15 years), they are dispatched to the Repair and Maintenance (RM) unit so that after undergoing correctional jobs, they are re-dispatched to the customer or producer.

The objective of the present research is to demonstrate how decision making on a specified product's manufacturing and supplying process can help producers in the field. To this end, the producer can manufacture the required products on his/her own plant. Accordingly, he/she should make his decision in the light of the capacity of available means and resources, production expenditures, and the quality of the produced commodity, what amount of products to produce at regular working hours and what amount to produce at overtime working hours. In his/her APP, he/she might also decide on outsourcing the production of a portion of his/her required products to outside suppliers. Such planning becomes of utmost importance as he should make his decision based on such requisite indices and criteria as expenditure, quality level, and prioritization- what amount of each product to delegate from each customer. Along this line, in the proposed model, a win-win relation with the suppliers is deemed essential. Hence, in the model offered, the optimization of the customers' satisfaction is taken into account so that - by considering customers' prioritization - the shortage rate of the unmet demands on the part of the supplier is kept at minimum. It should be noted that most supply chain models consider minimizing costs, while the supply chain of the proposed model encompasses maximizing customers and suppliers' satisfaction, too. Further, the proposed supply chain, unlike other supply chains, comprises two centers-depot and RM.

Thus, in the design of the extended applied model proposed in this research study, such cases as determining the contribution of suppliers, reconstruction centers, RM measures, production at regular hours, and overtime manufacture of each of the products as well as the amount of dispatched products to each of the customers are among decisions taken in the proposed model. Moreover, such objectives as minimizing producer's costs including production expenditures, cost of retaining and inventory deficit, costs related to supplying products through outsourcing, maximizing the quality of the manufactured products at regular time, overtime, and production by suppliers or procuring products from repair, maintenance and reconstruction centers, where each one has a distinct quality are among parameters considered in the advanced model. Also, special attention is paid to the assessment of suppliers and customers so that optimum satisfaction of these two groups is provided. Therefore, in view of the particular attention paid by the authorities of the concerned military industry to the issue of APP, the current research (regarded as a proper response to the need of that industry) was conducted in the format of a multi-objective mathematical model built on implications of APP in closed-loop supply chain while paying due attention and regard to each of the ingredients of the closed-loop chain. Furthermore, primary exchanges of views with the authorities in the above-named industry and conducting a survey of available research on theoretical foundations have given new dimensions and extensions to the subject under study. To proceed with the research, we first provide a review of the research background. Then, the proposed mathematical model is introduced. Next, the solution to the model is explained. Finally, the model is solved given the extracted data from the concerned industry.

2 | Literature Review

Hafezalkotob et al. [9] developed the cooperative APP. This planning help to decrease the production costs and workforce and inventory costs. These costs constitute a large fraction of the operating costs of many manufacturing plants. In addition, this research quantifies the cost saving and synergy of different coalitions of production plants. The research accomplished by Masud and Hwang [10], on the issue of APP resulted in a model with multiple objectives where the concept of APP with resource limitation is raised and investigated via genetic meta-heuristic algorithm. In their model, such objectives as maximizing profits, bringing costs to a minimum, minimizing the amount of stock inventory, minimizing goods shortage, maximizing the usage of existing means, and minimum amount of overtime work are among factor taken into consideration. Also, references are made to such parameters as man power working hours for manufacturing each unit of product, time to use machine for manufacturing each unit of product, cost of manufacturing each unit of product, overtime cost of manufacturing each product unit, machine capacity at regular manufacturing hour for each unit of product and so on. Ghahremani-Nahr et al. [3] proposed a mathematical model of a multi-product multi-period multi echelon closed-loop supply chain network design under uncertainty. In this paper, the quantities of products and raw material transported between the supply chain entities in each period by considering different transportation mode, the number and locations of the potential facilities, the shortage of products in each period, and the inventory of products in warehouses and plants with considering discount and uncertainty parameters are determined. In addition, the robust possibility optimization approach was used to control the uncertainty parameter.

Cheraghalikhani et al. [11] conducted a literature review on APP. They accomplished a comprehensive classification of APP from two perspectives. In the first perspective, they considered the level of uncertainty existing in the APP model in addition to the number of objective functions that a model contains, whereas, in the second perspective, besides primary issues in APP models, further issues are considered e.g., multiple product item, labor characteristics, degree of DM satisfaction from solution, product characteristics, setup, multiple manufacturing plant, time value of money, financial concepts, supply chain concepts as well as multiple product market.

Hatefi et al. [12] developed a novel mathematical model. In their model, network design decisions integrate in both forward and reverse flows drawing upon reliability concepts. Reliability concepts confront with resource disruptions. Owing to the importance of the role of hybrid distribution-collection resources, in both forward and reverse flows, the authors assumed that they might be randomly disrupted. Random resource disruptions give rise to risks that might be related to the epistemic uncertainties in the model parameters. The proposed model preserves an integrated forward-reverse logistics network against them. To deal with random resource disruptions, two effectively reliable strategies are considered: 1) locating reliable and unreliable hybrid resources to deal with disruption strikes, 2) Unreliable hybrid resources might lose a percentage of their capacities because they are permitted to be partially disrupted. In the end, several numerical experiments have been proposed along with sensitivity analysis. These experimentations clarify the importance, applicability, and effectiveness of the developed model.

Ghorbani et al. [13] proposed a fuzzy goal programing-based approach. Through their approach, they solve a multi-objective mathematical model of reverse SC design while considering three objective functions. Objective functions minimize the recycling cost of the product and the rate of the waste made through the recycle process. In the end, a numerical example is conducted to illustrate the effectiveness of the model. Rivaz et al. [14] suggested a new model based on fuzzy goal programming. They focused on MOTPs (a special type of multi objective programming problems). Given that, there does not usually exist an optimal solution that would simultaneously satisfy all objectives in multi objective problems, the best way in this situation is seeking suitable compromise solutions for such problems. In addition, they tried to vary the weights in the new model and obtain the different solutions.

Khalifa [15] surveyed Multi-criteria De Novo Linear Programming (F-MDNLP) problems. In this research, the fuzzy goal programming approach has considered as a suitable approach to obtaining α-optimal compromise solution and to achieve satisfactory results for the DM. According to this topic, author tried to use of fuzzy goal programming approach. In this approach the decision maker's role only was the evaluation of the $\alpha -$ efficient solutions to limit the influences of his/ her incomplete knowledge about the problem domain.

Baykasoglu [16] made an investigation of APP with multiple objectives within tabu search meta-heuristic algorithm. In this study, APP is defined as programming for middle-term capacity of a 2 to 18-month planning span. However, in the light of the industry type and the organization products, the timing can change and encompass longer spans. In this model, such decision variables as product inventory in each period, returned products, number of work force at each period, and the profit level are presented and discussed.

Leung et al. [17] developed a decision back-up system for solving multi-objective mixed- integer model of APP - through adopting an ideal planning method. The model addresses overall manufacturing products, manufacturing units, and manufacturing periods, minimizing work force in the factory in the periods under study, bringing inventory shortage to a minimum, thus minimizing the returned product levels etc.

Gholamian et al. [18] produced one research on APP of multi-products, multi-objectives, and a total of seven units in a supply chain under uncertain conditions adopting a phase approach of multi-objective optimization. In their model, phase parameters include cost of each regular and overtime working hours, cost related to suppliers as against each unit of raw material, cost of transportation from presenter, cost of raw material procured by supplier, cost of hiring, firing and training of the personnel, cost of holding product inventory, cost of transporting goods to customers, cost of holding raw material, penalty cost of deficit in product dispatched to customer, sale price of each unit of product to customers as well as the number of requests made by customer. Also definite parameters include maximum product procured from supplier, machine-hour expended for manufacturing each product unit, maximum machine capacity, warehouse space for each product unit, warehouse space for each unit of primary material, maximum available space in warehouse, available regular and overtime work hour, number of available workforce, required delay time for carrying raw material, and permissible shortage, In this model, decision variables consist of number of products produced at regular and overtime work hours, number of products by suppliers, number of personnel, number of primary materials, quality level of personnel, number of end products sent to customers, inventory of final product as well as deficit in product inventory.

Mirzapour et al. [19] made a study of multi-objective robust optimization model of multi-product APP in a supply chain under uncertain conditions. In this research, the supply chain includes numerous suppliers, producers, and customers and a discussion on multi-period, multi-product APP under uncertain conditions is presented. This model proposes a multi-objective non-linear programing scheme for the first time for a new mixed-integer within robust optimization approach while simultaneously considering conflicting objectives in a supply chain under uncertain conditions. The first objective includes minimizing production cost, hiring, firing and training costs, cost related to primary material, cost of holding product inventory as well as transportation and shortage costs. The second objective concentrates on minimizing total maximum shortage rate among customers' place of residence throughout the designated period paying special attention to customers' satisfaction. Also, taken into consideration in this study is the work level, laborers' productivity, over time, contractual work, storage capacity, and time parameters. Eventually, the proposed model is solved as an integer-programming model.

Rahimi et al. [20] proposed a robust optimization model for multi-objective multi-period supply chain planning under uncertainty considering quantity discounts. In this model the current profit and company's expected profit maximize respectively, by making a balance between the total costs of the supply chain and the distributor company's revenues of selling products and by, introducing brands and taking the risk of loss on it.

Zanjani et al. [21] developed a multi-objective Robust Mixed-Integer Linear Programming (RMILP) model. This model is related to Hybrid Flow Shop (HFS) scheduling. They focused on this topic because it has good adaptability with most real-world applications including innumerable cases of uncertainty of parameters that would influence jobs processing when the schedule is executed. The developed model is able to assign a set of jobs to available machines in order to obtain the best trade-off between two objectives including total tardiness and make span under uncertain parameters.

Mirzapour Al-e-hashem et al. [22] conducted a study on an efficient algorithm for solving robust multi-objective APP under uncertainty circumstances. In their research study, they presented a multi-objective model for solving the problem of an APP extended over a few periods of multi-products for a middle-term perspective under uncertain conditions. In this model, the first objective is defined as minimizing expected overall value and cost of inventory quantity, cost of overtime and contractual work, returned orders, machinery and warehouse capacities. The second objective function is expressed as minimizing shortages among all customers' regions. Finally, the third objective function considers maximizing laborers' productivity, weighted mean of productivity level in all factories throughout the designated period. At this stage, the model is solved through a genetic algorithm where the obtained results demonstrate the model's efficiency

Mirzapour Al-e-hashem et al. [23] seek to develop an APP model in a green supply chain for several time periods and multi-products in a green supply chain over a middle term perspective assuming demand uncertainty. The proposed model highlights such features as transportation costs, relations between delay time to delivery and transportation costa, and the discount rate for encouraging manufacturer for higher number of orders. This model for the first time employs a non-linear mixed-integer programming.

Mulvey [24] has introduced a robust optimization framework which includes two robust types. This method consists of robust solution (a solution almost optimal in all scenarios) and robust model (a model having almost plausible answer in all scenarios). In this method, optimization is generally defined as a penalty objective function, which situation is considered both for the robust model and the robust solution.

The objective function is also weighted by parameters of uncertainty and in the objective function and by the afore-mentioned restrictions. The robust optimization method presented by Mulvey is, in fact, a model developed from randomized planning. The method comes about as a result of replacing the expected classic (traditional) minimization of cost function with a penalty objective function having clear references to changeable costs.

In what follows, the robust optimization method is briefly explained [25]. Consider the following linear programing model which includes randomized parameters:

$$\mathrm{Min}\, c^T x + d^T y. \tag{1}$$

$$\text{Subject to:}$$
$$Ax = b. \tag{2}$$
$$Bx + Cy = e. \tag{3}$$
$$x, y \geq 0. \tag{4}$$

In the above model, the following assumptions are observed:

x: Decision variables to be determined under model's uncertain conditions.

B, C, e: Randomized matrix of technology coefficient and right-hand values.

N: Set of scenario n ε N {1, 2,…} relative to under randomized model's parameters of uncertainty character, with any scenario, is a subset of n ε N. For sets of scenarios *Pn* = (*n* ∑ *Pn=1*) and probability of all scenarios (dn, Bn, Cn, en), values 1 to n are scenario N members.

B, C, e explained above, under uncertain conditions are in Bn, Cn, en forms for any scenario n ε N. Also, y defined as control variable whose variables are under concerned scenarios.

Therefore, yn for any n causes n implausibility. If the model is plausible, then σn holds true. After all, parameters of uncertainty nature indicate non-plausibility of the model under any scenario n.

If the model is plausible, then σn equals zero; otherwise σn is of positive value, based on *Eq. (7)*.

Robust optimization model is formulated as follows:

$$\mathrm{Min}\,\sigma(x, y_1, y_2, ...y_n) + w \times p(\sigma_1, \sigma_2, ...\sigma_n). \tag{5}$$

$$\text{Subject to:}$$
$$Ax = b. \tag{6}$$
$$B_n x + C_n y_n + \sigma_n = e_n \text{ for all } n \in N. \tag{7}$$
$$x \geq 0,\ y_n \geq 0,\ \sigma_n \geq 0 \text{ for all } n \in N. \tag{8}$$

The first part of the model indicates robust solution, the important decision of the decider is not "DISLIKE", what is intended is to reduce costs and risk level. While the second part indicates robust model intending to prevent non-plausibility of the model.

Ψn = f(x , y) expresses profit or cost function for scenario n.

High variance for Ψn = f (x , yn) denotes that the solution involves high risk decision. In other words, a minor change in the value of parameters can bring about a major change in the function value [22].

$$\sum_{n\in N} P_n \Psi_n + \lambda \sum_{n\in N} P_n (\Psi_n + \sum_{nl\in N} P_{nl} \Psi_{nl})^2 = \sigma. \tag{9}$$

Table 1 also summarizes these researches. According to this table, all of other mentioned researches parallel with this research. Present study is an APP that is done in a closed-loop supply chain. In addition, this study is considered all of objectives of supply chain and uncertain conditions and, was done at a high-tech industry.

Table 1. A summary of done studies.

Authors	Years	Objectives		Process		Type of Demand		Type of Supply Chain	
		One	Multi	Supply Chain	Producer	Uncertain	Certain	Reverse	Forward
Ghahremani-Nahr et al. [3]	2020	✓		✓		✓		✓	✓
Jang and Chung [8]	2020	✓		✓		✓			✓
Rivaz et al. [14]	2020		✓	✓			✓		✓
Rahimi et al. [20]	2018		✓	✓		✓			✓
Gholamian et al. [18]	2015		✓		✓		✓		
Mirzapour et al. [22]	2012		✓		✓	✓			
Mirzapour et al. [19]	2011		✓		✓	✓			✓
Pan and Nagi [25]	2010		✓	✓		✓			✓
Baykasoglu [16]	2001		✓		✓		✓		
Mulvey [24]	1995	✓		✓		✓			✓
Masud and Hwang [10]			✓		✓		✓		

3 | The Proposed Mathematical Model

This model is a supply chain that has three levels- producer, consumers and a center for reconstruction (depot), RM. In this chain, a producer starts out by sending several merchandises to customers. The manufacturer produces a part of customers' demands at regular and overtime work hours, whilst the remaining demands are delegated to outside suppliers. Eventually the goods delivered to the customers, in case they are defective, are returned by customers to the depot center, where, after undergoing correctional procedures, are sent back to the producer, such that in later cycles, they are re-sent to the customers.

Additionally, when the expiry data of the products' warranty approaches, they are shipped to the depot center by the customers, and if possible, after undergoing necessary repairs and corrections, are re-sent to the customers; otherwise, the products are de-assembled and returned to the producer. So, the supply chain of the proposed model is reverse and forward, simultaneously and it is a closed-loop supply chain. A graphic representation of this chain is provided in *Fig. 1*.

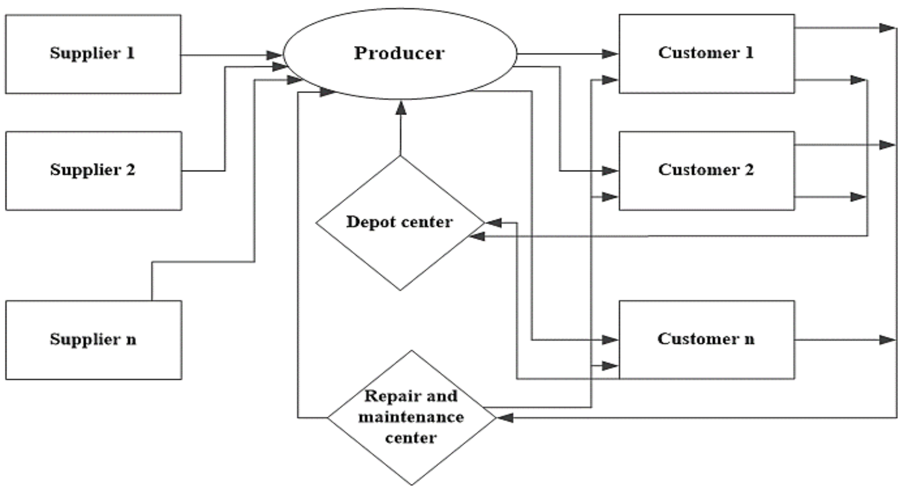


Fig. 1. Network with eight vertices.

3.1 | Assumptions of the Proposed Model

I. The products are produced and sold in closed loop three- level supply chain. The chain consists of several suppliers, a producer, and several customers as well as centers for depot and RM as well.
II. In case of non-usage of the products, after a few years, they are transferred to the RM center by the customer.
III. In the RM center, after rendering repair and correctional work on the products, they are delivered to the customer or de-assembled and shipped to the producer.
IV. The returned products to the depot center (by the customer) are either demolished or de-assembled and dispatched to the producer.
V. What is dealt with in the supply chain of the proposed model is production and sale of a product composed of several components by itself.
VI. Not all customers are equally important and some relative to others enjoy higher importance.
VII. The cost of RM, the capacity and the quality of production at regular and overtime work hours are different.
VIII. The suppliers, as regard the price and the delivery time of the product, are different.
IX. It is intended that a win-win relation between the producer and the supplier is established.
X. The products are produced and sold in a closed loop three-level supply chain in which several suppliers, a producer and several customers are involved. The chain also includes a center for depot and one for RM.
XI. The products not used for several years by the customer are returned to the RM center.
XII. In the RM center, the products are repaired, after which they are returned to the customer or converted into spare parts and shipped to the producer.
XIII. In the depot center, the products returned by the customer, are exterminated or de-assembled and dispatched to the producer in the form of spare parts.
XIV. In the supply chain of our model, several products are manufactured and presented for sale.
XV. The capacity, the cost, and the quality of production at regular and overtime work hours of supplying merchandise by the suppliers, the RM center and the reconstruction center are different.
XVI. The forecast demand includes uncertainty.
XVII. The manufacturing cost of one unit of the product at regular work hours includes uncertainty.
XVIII. The manufacturing cost of one unit of the product at overtime work hours includes uncertainty.
XIX. The cost of supplying one unit of the product from the suppliers includes uncertainty.
XX. The cost of one laborer at regular work hours includes uncertainty.
XXI. The cost of laborer at overtime work hours includes uncertainty.
XXII. The hiring cost of one instance of work force includes uncertainty.
XXIII. The firing cost of one instance of work force includes uncertainty.
XXIV. The holding cost of one unit of merchandise in the depot center warehouse includes uncertainty.
XXV. The holding cost of one unit of merchandise RM center includes uncertainty.
XXVI. The holding cost of one unit of merchandise in the warehouse of the producer's center includes uncertainty.
XXVII. The sale price of the product includes uncertainty.
XXVIII. The capacity of holding the merchandise in the producer's center includes uncertainty.
XXIX. The deficit rate of the product includes uncertainty.

3.2 | Indices

i (i = 1,2,…..I): denotes i^{th} product.

k (k = 1,2,…..K): denotes k^{th} customer.

t (t = 1,2,…..T): denotes t^{th} period.

j (j = 1,2,….J): denotes j^{th} supplier.

$n \varepsilon N$: denotes n^{th} scenario.

Yousefi et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 25-44

3.3 | The Model Parameters

d_{iktn}: Forecast demand of i^{th} product at t^{th} period for k^{th} customer under n^{th} scenario.

α_{iktn}: Percentage of returned i^{th} product by k^{th} customer to depot center at t^{th} period under n^{th} scenario.

β_{iktn}: Percentage of returned i^{th} product by k^{th} customer to RM center at t^{th} period under n^{th} scenario.

CAPP: Capacity for holding merchandise at producer center.

CAPD: Capacity for holding merchandise at depot center.

CAPM: Capacity for holding merchandise at RM center.

CPR_{in}: Cost of producing one unit of i^{th} product at regular work hours under n^{th} scenario.

CPO_{in}: Cost of producing one unit of i^{th} product at overtime work hours under n^{th} scenario.

CD_{in}: Cost of producing one unit of i^{th} product from depot center under n^{th} scenario.

CM_{in}: Cost of producing one unit of i^{th} product from RM center under n^{th} scenario.

CSC_{ijn}: Cost of supplying one unit of i^{th} product from j^{th} supplier under n^{th} scenario.

CLR_{tn}: Cost of one laborer at t^{th} period at regular work hours under n^{th} scenario.

CLO_{tn}: Cost of one laborer at t^{th} period at overtime work hour under n^{th} scenario.

HC_{tn}: Cost of hiring manpower at t^{th} period under n^{th} scenario.

FC_{tn}: Cost of firing manpower at t^{th} period under n^{th} scenario.

HIP_{itn}: Cost of holding one unit of i^{th} product at t^{th} period in producer's warehouse under n^{th} scenario.

HID_{itn}: Cost of holding one unit of i^{th} product in depot center at t^{th} period under n^{th} scenario.

HIM_{itn}: Cost of holding one unit of i^{th} product at t^{th} period in RM center warehouse under n^{th} scenario.

π_{iktn}: Cost of shortage of one unit of i^{th} product for k^{th} customer at t^{th} period under n^{th} scenario.

QR_{it}: Production Quality Coefficient (QC) of i^{th} product at t^{th} period at regular work hours.

QO_{it}: Production quality coefficient of i^{th} product at t^{th} period at overtime work hours.

QSC_{ijt}: Production quality coefficient of i^{th} product by j^{th} supplier at t^{th} period.

QD_{it}: Production quality coefficient of i^{th} product at t^{th} period by the depot center.

QM_{it}: Production quality coefficient of i^{th} product at t^{th} period by the RM center.

WC_k: Worth coefficient of k^{th} customer.

WSC_j: Worth coefficient of j^{th} supplier.

Yousefi et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 25-44

MW_t: Maximum work force available at t^{th} period.

MOT_t: Maximum overtime work hour available at t^{th} period.

TW: Maximum work hour needed.

TP_i: Total person-hour rate needed for producing i^{th} product (at regular and overtime work hours).

γ_t: Percentage of permissible change in human work force at t^{th} period.

MSC_{ijt}: Maximum permissible supply of i^{th} product from j^{th} supplier at t^{th} period.

Piktn: Sale price of i^{th} product to k^{th} customer at t^{th} period under n^{th} scenario.

$CPRD_{in}$: Cost of producing one unit of i^{th} product at regular work hours in the depot center under n^{th} scenario.

$CPOD_{in}$: Cost of producing one unit of i^{th} product overtime work hours in the depot center under n^{th} scenario.

$CPRM_{in}$: Cost of producing one unit of i^{th} product at regular work hours in the RM center under n^{th} scenario.

$CPOM_{in}$: Cost of producing one unit of i^{th} product overtime work hours in the RM center under n^{th} scenario.

P_n: Probability of any scenario.

A_n: Designed scenario for parameters with uncertainty in the first objective function.

θ_n : Variable used for costs variation linearization.

λ_n: Weight on solution's variance.

3.4 | Decision Variables

B_{ikt}: Deficit amount (back order) of i^{th} product at t^{th} period for the k^{th} customer.

X_{it}: Amount of producing i^{th} family products at regular work hour production at t^{th} period.

Y_{it}: Amount of producing i^{th} family products at overtime work hour production at t^{th} period.

ZD_{it}: Amount of supplying i^{th} family products by the depot center at t^{th} period.

ZM_{it}: Amount of supplying i^{th} family products by the RM center at t^{th} period.

F_{ikt}: Amount of i^{th} family shipped product for the k^{th} customer at t^{th} period.

SC_{ijt}: Amount of i^{th} family products that are procured by j^{th} supplier at t^{th} period.

OT_t: Overtime work hours needed at t^{th} period.

IP_{it}: Inventory level of i^{th} family product at the end of t^{th} period at the producer's site.

WL_t: Number of work laborers needed at t^{th} period.

HL_t: Number of hired laborers at t^{th} period.

FL_t: Number of fired laborers at t^{th} period.

ZC_{ikt}: Number of i^{th} family product shipped for k^{th} customer at t^{th} period from the RM center.

IM_{it}: Inventory level of i^{th} family product at the end of t^{th} period in the RM center.

ID_{it}: Inventory level of i^{th} family product at the end of t^{th} period in the depot center.

XD_{it}: Amount of producing i^{th} family product at regular work hours at t^{th} period in the depot center.

YD_{it}: Amount of producing i^{th} family product at overtime work hours at t^{th} period in the depot center.

XM_{it}: Amount of producing i^{th} family product at regular work hours at t^{th} period in the RM center.

YM_{it}: Amount of producing i^{th} family product at overtime work hours at t^{th} period in the RM center.

3.5 | Mathematical Model

$$\begin{aligned}\text{Min} z_1 = E + \lambda_1[(P_{n1} \times A_{n1}) - E + 2\theta_{n1}] + \lambda_2[(P_{n2} \times A_{n2}) - E + 2\theta_{n2}] + \\ \lambda_3[(P_{n3} \times A_{n3}) - E + 2\theta_{n3}].\end{aligned} \tag{10}$$

$$\begin{aligned}\text{Max} z_2 = \sum_t \sum_i (X_{it} \times QR_{it} + Y_{it} \times QO_{it}) + \sum_t \sum_i \sum_j (SC_{ijt} \times QSC_{ijt}) + \\ \sum_t \sum_i (ZD_{it} \times QD_{it}) + \sum_t (ZM_{it} \times QM_{it}).\end{aligned} \tag{11}$$

$$\text{Min} z_3 = \sum_{t_k} \max(WC_k \times \sum_i B_{ikt}). \tag{12}$$

$$\text{Max} z_4 = \sum_{t_j} \min(WCS_j \times \sum_i SC_{ijt}). \tag{13}$$

Subject to:

$$IP_{i(t-1)} + X_{it} + Y_{it} + \sum_j SC_{ijt} + ZD_{it} + ZM_{it} + \sum_k B_{ik(t-1)} = \sum_k B_{ik(t-1)} + \sum_k F_{ikt} + IP_{it} \quad \text{for all } i, t, \tag{14}$$

$$ID_{it} = ID_{i(t-1)} + \sum_k \alpha_{iktn} \times F_{ikt} - ZM_{it} - ZD_{it} \quad \text{for all } i, t, n, \tag{15}$$

$$IM_{it} = IM_{i(t-1)} + \sum_k \beta_{iktn} \times F_{ikt} - ZM_{it} - ZC_{ikt} \quad \text{for all } i, t, n, \tag{16}$$

$$\sum_i IP_{it} \le CAPP \quad \text{for all } t, \tag{17}$$

$$\sum_i ID_{it} \le CAPD \quad \text{for all } t, \tag{18}$$

$$\sum_i IM_{it} \le CAPM \quad \text{for all } t, \tag{19}$$

$$WL_t \le MW_t \quad \text{for all } t, \tag{20}$$

$$WL_t = WL_{(t-1)} + HL_t + FL_t \quad \text{for all } t, \tag{21}$$

$$HL_t \times FL_t = 0, \tag{22}$$

$$(IP_{it} + IM_{it}) \times \sum_k B_{ikt} = 0 \quad \text{for all } i, t, \tag{23}$$

$$OT_t \le MOT_t \quad \text{for all } t, \tag{24}$$

$$\sum_i TP_i \times X_{it} \le TW \quad \text{for all } t, \tag{25}$$

$$\sum_i TP_i \times Y_{it} \leq OT \quad \text{for all } t, \tag{26}$$

$$FL_t + HL_t \leq \gamma_{(t-1)} \times WL_{(t-1)} \quad \text{for all } t, \tag{27}$$

$$SC_{ijt} \leq MSC_{ijt} \quad \text{for all } i, j, t, \tag{28}$$

$$SC_{ijt} \leq SC_{(i-1)jt} \quad \text{for all } i, j, t, \tag{29}$$

$$B_{ikt} = B_{ik(t-1)} + d_{iktn} - F_{ikt} - ZC_{ikt} \quad \text{for all } i, k, t, n, \tag{30}$$

$$ZD_{it} \leq CAPD \quad \text{for all } i, t, \tag{31}$$

$$ZM_{it} \leq CAPM \quad \text{for all } i, t, \tag{32}$$

$$ZC_{ikt} \leq CAPM \quad \text{for all } i, k, t, \tag{33}$$

$$F_{ikt} + ZC_{ikt} \leq d_{iktn} \quad \text{for all } i, k, t, n, \tag{34}$$

$$A_{n1} - [(P_{n1} \times A_{n1}) + (P_{n2} \times A_{n2}) + (P_{n3} \times A_{n3})] + \theta_{n1} \geq 0, \tag{35}$$

$$A_{n2} - [(P_{n1} \times A_{n1}) + (P_{n2} \times A_{n2}) + (P_{n3} \times A_{n3})] + \theta_{n2} \geq 0, \tag{36}$$

$$A_{n3} - [(P_{n1} \times A_{n1}) + (P_{n2} \times A_{n2}) + (P_{n3} \times A_{n3})] + \theta_{n3} \geq 0, \tag{37}$$

$$Y_{it}, ZD_{it}, ZM_{it}, F_{ikt}, SC_{ijt}, OT_t, IP_{it}, WL_t, HL_t, FL_t, ZC_{ikt}, IM_{it}, XD_{it}, YD_{it}, XM_{it}, YM_{it}, ID_{it} \geq 0, \tag{38}$$

$$i = 1,2,3, \quad k = 1,2,3, \quad j = 1,2,3, \quad t = 1,2,3, \quad n = 1,2,3.$$

Eq. (10) shows the first objective function of the problem defined for minimizing the costs. The costs relate to the following cases: producing a unit of product at regular and overtime work hours, supplying a unit of product by the suppliers, the RM and the depot centers, an individual laborer at regular work hour, an individual laborer at overtime work hour, hiring and firing human work force, holding a unit of product in the producer's place, RM, and depot warehouses, shortage in unit of product for customer, and the cost of forecast demand. The function is written in the robust form according to the Mulvey's method. *Eq. (11)* represents the model's second objective function defined for maximizing the QC. QC embraces the following instances: sum of production QC at regular work hours, production QC at overtime work hours, QC of received product from suppliers, QC of received product from depot center, and QC of received product from RM center. *Eq. (12)* shows the third objective function of the problem which includes minimizing maximum shortage among customer and customers' importance coefficient. *Eq. (13)* is a display of the model's fourth objective function whose purpose is maximizing minimum rate of supplying product from suppliers. *Eq. (14)* expresses the producer's inventory balance. *Eq. (15)* denotes the inventory balance at depot center. *Eq. (16)* denotes the inventory balance at RM center. The capacity for holding the product at the producer's center is indicated by *Eq. (17)*. *Eq. (18)* shows the capacity for holding the product at depot center. *Eq. (19)* represents the capacity for holding the product at RM center. *Eq. (20)* represents the limitations of maximum work force available. *Eq. (21)* is an indication of balance in the producer's human work force. In *Eq. (22)* demonstrates the hiring or firing of personnel at each period. *Eq. (23)* shows the inventory or shortage of each product at each period. *Eq. (24)* represents limitations in overtime work ceiling. *Eq. (25)* displays the time for manufacturing the product is less at each period of available regular time. *Eq. (26)* indicates that the time for manufacturing the product is less at overtime work hours. The percentage of permissible changes in human work force at each period is shown in *Eq. (27)*. Maximum purchase of the producer from supplier's product is indicated in *Eq. (28)*. *Eq. (29)* indicates Maximum purchase of the product from suppliers at each period. *Eq. (30)* shows the balance in the shortage of the producer's product in relation to the shortage of the previous period, the rate of the product dispatched from the producer to the customer and the RM at each period. *Eq. (31)* and *Eq. (32)* points to maximum product supplied from RM and depot centers. *Eq. (33)* and *Eq. (34)* demonstrates maximum product shipped from RM center and depot center to the customer at each period. *Eqs. (35)-(37)* show the robust limitations of the model. Lastly, *Eq. (38)* represents non-negativity of the decision variables. The linearized form of *Limitations*

Yousefi et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 25-44

(20) and *(21)* are indicated in *Eq. (39)* and *Eq. (40)*, respectively. The variables XHLt and XFLt are binary. If the new human work forces' hire or fire occurs, the values of XHLt and XFLt will be equal to 1, respectively:

$$HL_t \le MW_t \times XHL_t. \tag{39}$$

$$XHL_t + XFL_t = 1. \tag{40}$$

The linearized form of *Limitations (17)* and *(23)* are indicated in *Eqs. (41)-(43)*.

$$\sum_k B_{ikt} \le [\sum_{t=1} \sum_k D_{ikt}] \times \theta_{it} \quad \text{for all } i, t. \tag{41}$$

$$IP_{it} \le CAPP \times (1 - \theta_{it}) \quad \text{for all } t. \tag{42}$$

$$\theta = \begin{cases} 1, & \sum_k B_{ikt} > 0, \\ 0, & IP_{it} > 0. \end{cases} \tag{43}$$

Eqs. (44)-(46) show the designed scenarios in the Mulvey method.

$$\begin{aligned} A_{n1} = {} & \sum_t \sum_i (CPR_{in1} \times X_{it} + CPO_{in1} \times Y_{it} + CPRD_{in1} \times XD_{it} + \\ & CPOD_{in1} \times YD_{it} + CPRM_{in1} \times XM_{it} + CPOM_{in1} \times YM_{it}) + \\ & \sum_t \sum_j \sum_i (CSC_{ijn1} \times SC_{ijt}) + \sum_t \sum_i (CD_{in1} \times ZD_{it} + CM_{in1} \times ZM_{it}) + \\ & \sum_t (CLR_{tn1} \times WL_t + CLO_{tn1} \times OT_t) + \sum_t (HL_t \times HC_{tn1} + FL_t \times FC_{tn1}) + \\ & \sum_t \sum_i (IP_{it} \times HIP_{itn1} + ID_{it} \times HID_{itn1} + IM_{it} \times HIM_{itn1}) + \sum_t \sum_i \sum_k (B_{ikt} \times \pi_{iktn1}) - \\ & \sum_t \sum_i \sum_k (F_{ikt} \times P_{iktn1}). \end{aligned} \tag{44}$$

$$\begin{aligned} A_{n2} = {} & \sum_t \sum_i (CPR_{in2} \times X_{it} + CPO_{in2} \times Y_{it} + CPRD_{in2} \times XD_{it} + \\ & CPOD_{in2} \times YD_{it} + CPRM_{in2} \times XM_{it} + CPOM_{in2} \times YM_{it}) + \\ & \sum_t \sum_j \sum_i (CSC_{ijn2} \times SC_{ijt}) + \sum_t \sum_i (CD_{in2} \times ZD_{it} + CM_{in2} \times ZM_{it}) + \\ & \sum_t (CLR_{tn2} \times WL_t + CLO_{tn2} \times OT_t) + \sum_t (HL_t \times HC_{tn2} + FL_t \times FC_{tn2}) + \\ & \sum_t \sum_i (IP_{it} \times HIP_{itn2} + ID_{it} \times HID_{itn2} + IM_{it} \times HIM_{itn2}) + \sum_t \sum_i \sum_k (B_{ikt} \times \pi_{iktn2}) - \\ & \sum_t \sum_i \sum_k (F_{ikt} \times P_{iktn2}). \end{aligned} \tag{45}$$

$$\begin{aligned} A_{n3} = {} & \sum_t \sum_i (CPR_{in3} \times X_{it} + CPO_{in3} \times Y_{it} + CPRD_{in3} \times XD_{it} + \\ & CPOD_{in3} \times YD_{it} + CPRM_{in3} \times XM_{it} + CPOM_{in3} \times YM_{it}) + \\ & \sum_t \sum_j \sum_i (CSC_{ijn3} \times SC_{ijt}) + \sum_t \sum_i (CD_{in3} \times ZD_{it} + CM_{in3} \times ZM_{it}) + \\ & \sum_t (CLR_{tn3} \times WL_t + CLO_{tn3} \times OT_t) + \sum_t (HL_t \times HC_{tn3} + FL_t \times FC_{tn3}) + \\ & \sum_t \sum_i (IP_{it} \times HIP_{itn3} + ID_{it} \times HID_{itn3} + IM_{it} \times HIM_{itn3}) + \sum_t \sum_i \sum_k (B_{ikt} \times \pi_{iktn3}) - \\ & \sum_t \sum_i \sum_k (F_{ikt} \times P_{iktn3}). \end{aligned} \tag{46}$$

Eqs. (47)-(49) present the mathematical expectations of designed scenarios in the suggested model and *Eq. (50)* presents the mathematical expectations of total scenarios.

$$E_1 = P_{n1} \times A_{n1}. \tag{47}$$

$$E_2 = P_{n2} \times A_{n2}. \tag{48}$$

$$E_3 = P_{n3} \times A_{n3}. \tag{49}$$

$$E = E_1 + E_2 + E_3. \tag{50}$$

4 | Solution for Proposed Model

Considering the multi-objectivity of the proposed model in the present research, attempt is made to find a Pareto optimal solution to the model. That is to say, a Pareto answer represents a decisive and effective solution from among available responses. One of the most common methods for solving multi-objective problems in arriving at Pareto optimal responses is invoking the LP metric method. In this method whose relevant calculations are readily observable in *Eq. (51)*, the model - in a single objective form - is first solved for each of the objective functions; then the set of obtained answers considering the type of objective function in terms of minimization or maximization are placed in *Eq. (51)* as illustrated below:

$$LP:\{w_1(\frac{z_1-z_1^*}{z_1^{nadir}-z_1^*})^p+w_2(\frac{z_2^*-z_2}{z_2^*-z_2^{nadir}})^p+w_3(\frac{z_3-z_3^*}{z_3^{nadir}-z_3^*})^p+w_1(\frac{z_4^*-z_4}{z_4^*-z_4^{nadir}})^p\}^{\frac{1}{p}}. \quad (51)$$

The assumptions related to the above relation are summarized below:

- *1≤ P ≤∞, whose value determines the degree of emphasis towards existing deviations, as the bigger the latter value, the more the emphasis placed on the biggest deviation.*
- *Wi: Weight considered for i^th^ objective function (i=1, 2, 3, 4).*
- *Zi: i^th^ objective function of the problem (i=1, 2, 3, 4).*
- *fi: Optimal answer obtained through solving the model as against i^th^ objective function.*
- *Zinadir: Anti-ideal answer as against i^th^ objective function.*

5 | The Proposed Model Results

The model proposed was solved in LP metric method making use of software v14.0.1.55 on a Windows-7 system with the specifications RAM 300 HZ 2, 20, GB. The model was solved given real data from the industry in question.

5.1 | Solving the Model in the Case under Study

As mentioned above, the problem was solved using real data taken from the industry under study. In the supply chain of the latter industry, five different products are manufactured.

At the first level and the last level of the chain, four suppliers are placed and the product is dispatched to four customers. In this regard, some research studies consider a three-month period. The problem parameters - in view of the obtained information from the above-mentioned industry - are presented in *Table 2* through *Table 15*. *Table 16* provides some of the obtained Pareto optimal answers. In order to acquire the answers related to each row in this table, the following steps are taken:

I. Optimize each of the objective functions separately -taking into account the model's constraints - once as maximizing and second as minimizing in the LINGO software.
II. Write LP metric relation connected with *Eq. (51)* utilizing the results from previous stage.
III. Optimize the obtained function from the previous stage, taking constraints into consideration by means of the LINGO software.
IV. Extract optimal values of decision variables using the obtained solution from previous stage.
V. Compute the value of each of objective functions as against optimal decision variables obtained from previous stage.

The values acquired from the last stage are variable Pareto optimal answers presented in columns 2 through 4 of *Table 16*. As can be seen, these values are related to the objective functions. Values connected with decision variables of each row, are those same values obtained from the fourth stage in the above-mentioned stages. In practice, after selecting one of Pareto optimal responses - by the decision maker/s in the industry in question - the values related to the decision variables can easily be provided.

Table 2. Sales price, forecast demand, and cost of shortage of one unit of product in each scenario.

k	i	π_{iktn}/ t=1	π_{iktn}/ t=2	π_{iktn} / t=3	d_{iktn}/ t=1	d_{iktn}/ t=2	d_{iktn}/ t=3	p_{iktn}/ t=1	p_{iktn} / t=2	p_{iktn}/ t=3
1	1	3	2	1	370	290	100	610	700	460
		2	5	7	339	280	230	500	555	420
		3	6	6	275	240	150	450	500	400
	2	3	2	1	370	290	100	610	700	460
		2	5	7	339	280	230	500	555	420
		3	6	6	275	240	150	450	500	400
	3	3	2	1	370	290	100	610	700	460
		2	5	7	339	280	230	500	555	420
		3	6	6	275	240	150	450	500	400
2	1	3	2	1	395	300	130	620	710	470
		2	5	7	349	290	240	510	570	423
		3	6	6	295	250	190	465	512	411
	2	3	2	1	395	300	130	620	710	470
		2	5	7	349	290	240	510	570	423
		3	6	6	295	250	190	465	512	411
	3	3	2	1	395	300	130	620	710	470
		2	5	7	349	290	240	510	570	423
		3	6	6	295	250	190	465	512	411
3	1	3	2	7	445	350	180	670	760	520
		2	5	6	399	345	290	570	620	480
		3	6	6	335	304	243	510	580	460
	2	3	2	7	445	350	180	670	760	520
		2	5	6	399	345	290	570	620	480
		3	6	6	335	304	243	510	580	460
	3	3	2	7	445	350	180	670	760	520
		2	5	6	399	345	290	570	620	480
		3	6	6	335	304	243	510	580	460

Table 3. Percentage of returned product by customer to depot and rapiar and maintenance center in each scenario.

Scenario	i	k	α_{iktn}/ t= 1	α_{iktn}/ t= 2	α_{iktn}/ t= 3	β_{iktn}/ t= 1	β_{iktn}/ t= 2	β_{iktn}/ t= 3
1	1	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
	2	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
	3	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
2	1	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
	2	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
	3	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
3	1	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
	2	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07
	3	1	0.07	0.06	0.07	0.07	0.06	0.07
		2	0.01	0.05	0.19	0.03	0.07	0.19
		3	0.08	0.01	0.07	0.08	0.02	0.07

Table 4. Cost of producing at regular and overtime hours from depot, and RM centers.

Scenario	i	CPO_{in}	CPR_{in}	CD_{in}	CM_{in}
1	1	90	90	80	70
	2	110	90	80	90
	3	120	100	90	90
2	1	100	100	90	80
	2	120	100	90	100
	3	130	150	100	100
3	1	150	150	140	130
	2	160	150	140	150
	3	170	200	150	150

Table 5. Capacity for holding merchandise at producer, depot, and RM centers and maximum work hour needed.

CAPP	CAPD	CAPM	TW
15000	10000	10000	60

Table 6. Total person-hour rate needed for goods (at regular and overtime work hours).

i	Ti
1	200
2	200
3	200

Table 7. Cost of holding goods at producer's warehouse, depot, and RM centers in each scenario.

Scenario	i	HIP_{itn} /t= 1	HIP_{itn} /t= 2	HIP_{itn} /t= 3	HID_{itn} /t= 1	HID_{itn} /t= 2	HID_{itn} /t= 3	HIM_{itn}/t= 1	HIM_{itn}/t= 2	HIM_{itn}/t= 3
1	1	85	90	90	25	28	26	20	23	22
	2	90	92	93	27	29	27	22	23	25
	3	92	98	100	29	31	31	25	24	24
2	1	85	90	90	25	28	26	20	23	22
	2	90	92	93	27	29	27	22	23	25
	3	92	98	100	29	31	31	25	24	24
3	1	85	90	90	25	28	26	20	23	22
	2	90	92	93	27	29	27	22	23	25
	3	92	98	100	29	31	31	25	24	24

Table 8. Production quality coefficient at regular and overtime work hours in depot, and RM centers.

i	Qd_{it}/ t= 1	Qd_{it}/ t= 2	Qd_{it}/ t= 3	Qm_{it}/ t= 1	Qm_{it}/ t= 2	Qm_{it}/ t= 3	Qr_{it}/ t= 1	Qr_{it}/ t= 2	Qr_{it}/ t= 3	QO_{it}/ t= 1	QO_{it}/ t= 2	QO_{it}/ t= 3
1	0.97	0.92	0.96	0.92	0.96	0.95	0.98	0.97	0.98	0.97	0.97	0.97
2	0.97	0.98	0.95	0.96	0.98	0.95	0.97	0.97	0.98	0.97	0.97	0.97
3	0.98	0.93	0.96	0.9	0.97	0.98	0.97	0.98	0.97	0.97	0.98	0.97

Table 9. Maximum allowable supply of goods from supplier.

j	i	MSC_{ijt}/ t= 1	MSC_{ijt}/ t= 2	MSC_{ijt}/ t= 3
1	1	136	110	147
	2	167	83	64
	3	170	90	70
2	1	123	53	83
	2	141	41	65
	3	133	70	85
3	1	95	70	110
	2	125	100	96
	3	90	90	90

Table 10. Worth coefficient of suppliers and customers.

WSC_j/ j= 1	WSC_j/ j= 2	WSC_j/ j= 3	WC_k/ k= 1	WC_k/ k= 2	WC_k/ k= 3
0.8	0.6	0.5	0.9	0.6	0.8

Table 11. Production cost at regular and overtime work hours in depot and RM centers in each scenario.

Scenario	i	$CPOM_{in}$	$CPRM_{in}$	$CPOD_{in}$	$CPRD_{in}$
1	1	100	90	105	80
	2	100	90	105	80
	3	100	90	105	80
2	1	110	100	120	90
	2	110	100	120	90
	3	160	170	120	90
3	1	120	200	170	140
	2	120	200	170	140

Table 12. Cost of supplying a unit of product from supplier in each scenario.

Scenario	i	CSC_{ijn}/ j= 1	CSC_{ijn}/ j= 2	CSC_{ijn}/ j= 3
1	1	460	700	610
	2	420	555	500
	3	400	500	450
2	1	470	710	620
	2	423	570	510
	3	411	512	465
3	1	520	760	670
	2	480	620	570
	3	460	580	510

Table 13. Cost of manpower at regular and overtime work hours, and hiring and firing cost of one instance of human work force in each scenario.

Scenario	t	CLR_{tn}	CLO_{tn}	HC_{tn}	FC_{tn}
1	1	150	190	50	70
	2	120	195	50	80
	3	135	190	50	90
2	1	170	249	60	70
	2	190	250	60	80
	3	210	280	60	90
3	1	210	290	80	80
	2	232	270	90	90
	3	240	295	95	100

Table 14. Maximum work force available and overtime work hour and percentage of allowable change in human work force in each period.

t	MW_t	MOT_t	γ_t
1	50	55	0.2
2	50	56	0.2
3	50	57	0.2

Table 15. Production quality coefficient of product by supplier.

j	i	QSC_{ijt}/t= 1	QSC_{ijt}/t= 2	QSC_{ijt}/t= 3
1	1	0.78	0.8	0.62
	2	0.86	0.65	0.92
	3	0.68	0.58	0.88
2	1	0.95	0.82	0.51
	2	0.82	0.94	0.53
	3	0.72	0.62	0.55
3	1	0.93	0.76	0.66
	2	0.64	0.89	0.96
	3	0.52	0.6	0.63

Table 16. Answer proceeding from solving model as against objective functions.

	Z_1	Z_2	Z_3	Z_4
p=1	-30602741.4508065	9924770.23363924	7330.80000000000	240.452199760426
p=2	-37240491.6050168	15806267.7158352	15721.5000000000	257.995403149302
p=3	-32559383.6529256	12188671.2093891	15675.3000000000	254.248064660243
p=4	-29473058.8500060	11323108.6526589	7382.40000000000	243.818014726437
Mean	-32468918.89	12310704.45	11527.5	249.1284206
Variance	8808471415452.9	4725470550249.4	17397006. 4	52.1

Table 17. The obtained value for some of decision variables.

Variable	Value	Explanation
B211	1792	Deficit amount (back order) of second product at first period for the first customer.
X21	2601	Amount of producing second family products at regular work hour production at first period.
Y21	6865	Amount of producing second family products at overtime work hour production at first period.
ZD21	222	Amount of supplying second family products by the depot center at first period.
ZM21	8753	Amount of supplying second family products by the RM center at first period
F211	351	Amount of second family shipped product for the first customer at first period.
SC211	100	Amount of second family products that are procured by first supplier at first period.
OT1	42	Overtime work hours needed at first period.
IP21	23	Inventory level of second family product at the end of first period at the producer's site
WL1	10	Number of work laborers needed at first period.
HL1	4	Number of hired laborers at first period.
FL1	2	Number of fired laborers at first period.
ZC211	1055	Number of second family product shipped for first customer at first period from the RM center.
IM21	100	Inventory level of second family product at the end of first period in the RM center
ID21	101	Inventory level of second family product at the end of first period in the depot center.
XD21	9793	Amount of producing second family product at regular work hours at first period in the depot center
YD21	7186	Amount of producing second family product at overtime work hours at first period in the depot center.
XM21	7186	Amount of producing second family product at regular work hours at first period in the RM center.
YM21	9619	Amount of producing second family product at overtime work hours at first period in the RM center.

Pareto-optimal solutions are determined in the *Table 16*. As shown in this table, the solutions are non-dominate. For example, for P=2 and P=4, the best value of cost (first objective) is obtained for P=2 whereas the best value of suppliers' satisfaction (third objective) is obtained for P=4. In addition, the best value of customers' satisfaction (second objective) and the quality of the manufactured products (fourth objective) are for P=2. So, although P=2 optimizes the three objective functions, but the third function objectives is optimized for P=4. In addition, the mean and variance of objective functions are determined in *Table 16*. According to this table the first objective function has the max variance. So, this function, among 4 functions, has the max of dispersion. It should be noted that we can find more solutions for more values of P.

Table 17 shows the value of some of decision variables. Given that there are many variables or indices, we cannot propose all of them and this table is just an example and a part of outputs of model.

6 | Conclusion

This research study presents a multi-objective mathematical model for APP in a closed-loop supply chain under uncertain conditions. Worthy of note in the design of the model – formulated as a nonlinear planning scheme- is the particular attention it pays to creating a depot center, and a center for RM while simultaneously taking into account satisfying customers and suppliers as well as giving particular attention to the quality of the manufactured products and various costs and expenditures. The demand and the parameters related to the demand include uncertainty and the objectives of the model consist of minimizing costs, maximizing the product quality provided by suppliers and the product produced by the producer at regular and overtime work hours together with minimizing the sum total weight of maximum shortage among customers accompanied by minimizing the overall total weight of minimum rate of supplying goods from suppliers with a view to establishing a win-win relation.

The proposed model is first solved by a numerical example, then it is solved by the actual data taken from a closed-loop supply chain related to a specified military industry; and finally, all of variables are valued with the help of the closed-loop supply chain of the proposed model. As is evident from *Table 16*, the Pareto answers related to the problem are established. Also, in order to validate the finding, the above model is investigated with the help of the data taken from a numerical example at greater dimensions. The acquired results contribute greatly to the supply chain in attaining higher profits, better decision-making criteria and an increased level of rendering services to customers. The model can be applied in APP for various industries. Future research works might also add other parameters to the model and uncertain conditions for parameters of uncertain nature can be applied in view of the prevailing conditions of each industry. Furthermore, in proportion to the model becoming more complicated, meta-heuristic algorithms can be invoked to solve the model. It should be noted that proposed objectives and limitations are not limited to these cases and may there are some of new objectives and limitations. In addition, there are many methods for validation the model like LP- metric, epsilon-constraint and etc. So, choice of suitable method for validation was a limitation in this research. Furthermore, there are many methods to deal with uncertain conditions like fuzzy programming, chance programming, and sensitivity analyses, robust and etc. So, choice of suitable method was a limitation, too. Moreover, if there are many indices, like period or product, the model will become complicated and we con not solve it with and have to salve it with meta-heuristic algorithms.

References

[1] Chen, C. L., & Lee, W. C. (2004). Multi-objective optimization of multi-echelon supply chain networks with uncertain product demands and prices. *Computers & chemical engineering*, *28*(6-7), 1131-1144.

[2] Pokharel, S., & Mutha, A. (2009). Perspectives in reverse logistics: a review. *Resources, conservation and recycling*, *53*(4), 175-182.

[3] Ghahremani-Nahr, J., Nozari, H., & Najafi, S. E. (2020). Design a green closed loop supply chain network by considering discount under uncertainty. *Journal of applied research on industrial engineering*, *7*(3), 238-266.

[4] Teimoori, E., Shafiian Bejestani, J., & Kalb Khani, K. (2002). Simultaneous selecting and developing of suppliers in fuzzy state. *The first national conference of logistics and supply chains,* Tehran, Iran. (**In Persian**). https://civilica.com/doc/8764/

[5] Hansen, Z. N. L., Larsen, S. B., Nielsen, A. P., Groth, A., Gregersen, N. G., & Ghosh, A. (2018). Combining or separating forward and reverse logistics. *The international journal of logistics management*, *29*(1), 216-236. https://doi.org/10.1108/IJLM-12-2016-0299

[6] Zhang, R., Zhang, L., Xiao, Y., & Kaku, I. (2012). The activity-based aggregate production planning with capacity expansion in manufacturing systems. *Computers & industrial engineering*, *62*(2), 491-503.

[7] Paduloh, P., Djatna, T., Sukardi, S., & Muslich, M. (2020). Uncertainty models in reverse supply chain: a review. *International journal supply chain management*, *9*(2), 139-149.

[8] Jang, J., & Do Chung, B. (2020). Aggregate production planning considering implementation error: a robust optimization approach using bi-level particle swarm optimization. *Computers & industrial engineering*, *142*, 106367. https://doi.org/10.1016/j.cie.2020.106367

[9] Hafezalkotob, A., Chaharbaghi, S., & Lakeh, T. M. (2019). Cooperative aggregate production planning: a game theory approach. *Journal of industrial engineering international*, *15*, 19-37. https://doi.org/10.1007/s40092-019-0303-0

[10] Masud, A. S., & Hwang, C. L. (1980). An aggregate production planning model and application of three multiple objective decision methods. *International journal of production research*, *18*(6), 741-752. https://doi.org/10.1080/00207548008919703

[11] Cheraghalikhani, A., Khoshalhan, F., & Mokhtari, H. (2019). Aggregate production planning: a literature review and future research directions. *International journal of industrial engineering computations*, *10*(2), 309-330.

[12] Hatefi, S. M., Jolai, F., Torabi, S. A., & Tavakkoli-Moghaddam, R. (2015). A credibility-constrained programming for reliable forward–reverse logistics network design under uncertainty and facility disruptions. *International journal of computer integrated manufacturing*, *28*(6), 664-678. https://doi.org/10.1080/0951192X.2014.900863

[13] Ghorbani, M., Arabzad, S. M., & Tavakkoli-Moghaddam, R. (2014). A multi-objective fuzzy goal programming model for reverse supply chain design. *International journal of operational research*, *19*(2), 141-153.

[14] Rivaz, S., Nasseri, S. H., & Ziaseraji, M. (2020). A fuzzy goal programming approach to multiobjective transportation problems. *Fuzzy information and engineering*, *12*(2), 139-149. https://doi.org/10.1080/16168658.2020.1794498

[15] Khalifa, H. A. (2018). On solving fully fuzzy multi-criteria De Novo programming via fuzzy goal programming approach. *Journal of applied research on industrial engineering*, *5*(3), 239-252. DOI: 10.22105/jarie.2018.148642.1054

[16] Baykasoglu, A. D. İ. L. (2001). MOAPPS 1.0: aggregate production planning using the multiple-objective tabu search. *International journal of production research*, *39*(16), 3685-3702. https://doi.org/10.1080/00207540110061607

[17] Leung, S. C., Wu, Y., & Lai, K. K. (2003). Multi-site aggregate production planning with multiple objectives: a goal programming approach. *Production planning & control*, *14*(5), 425-436. https://doi.org/10.1080/0953728031000154264

[18] Gholamian, N., Mahdavi, I., & Tavakkoli-Moghaddam, R. (2016). Multi-objective multi-product multi-site aggregate production planning in a supply chain under uncertainty: fuzzy multi-objective optimisation. *International journal of computer integrated manufacturing*, *29*(2), 149-165. https://doi.org/10.1080/0951192X.2014.1002811

[19] Mirzapour Al-E-Hashem, S. M. J., Malekly, H., & Aryanezhad, M. B. (2011). A multi-objective robust optimization model for multi-product multi-site aggregate production planning in a supply chain under uncertainty. *International journal of production economics*, *134*(1), 28-42.

[20] Rahimi, E., Paydar, M. M., Mahdavi, I., Jouzdani, J., & Arabsheybani, A. (2018). A robust optimization model for multi-objective multi-period supply chain planning under uncertainty considering quantity discounts. *Journal of industrial and production engineering*, *35*(4), 214-228. https://doi.org/10.1080/21681015.2018.1441195

[21] Zanjani, B., Amiri, M., Hanafizadeh, P., & Salahi, M. (2021). Robust multi-objective hybrid flow shop scheduling. *Journal of applied research on industrial engineering*, *8*(1), 40-55.

[22] Mirzapour Al-e-hashem, S. M. J. M., Aryanezhad, M. B., & Sadjadi, S. J. (2012). An efficient algorithm to solve a multi-objective robust aggregate production planning in an uncertain environment. *The international journal of advanced manufacturing technology*, *58*, 765-782. https://doi.org/10.1007/s00170-011-3396-1

[23] Mirzapour Al-e-hashem, S. M. J., Baboli, A., & Sazvar, Z. (2013). A stochastic aggregate production planning model in a green supply chain: considering flexible lead times, nonlinear purchase and shortage cost functions. *European journal of operational research*, *230*(1), 26-41.

[24] Mulvey, J. M., Vanderbei, R. J., & Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations research*, *43*(2), 264-281.

[25] Pan, F., & Nagi, R. (2010). Robust supply chain design under uncertain demand in agile manufacturing. *Computers & operations research*, *37*(4), 668-683.

Yousefi et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 25-44

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 45-55.

Paper Type: Research Paper

Lean at Home: Applying 5 Whys and Lean PFMEA to Home Projects

Robert S. Keyser[1,*], Parisa Pooyan[1]

[1] Department of Systems and Industrial Engineering, Kennesaw State University, Marietta, GA, USA; rkeyser@kennesaw.edu; ppooyan@kennesaw.edu.

Citation:

Keyser, R. S., & Pooyan, P. (2023). Lean at home: applying 5 whys and lean PFMEA to home projects. *Journal of applied research on industrial engineering, 10*(1), 45-55.

Received: 20/12/2021 Reviewed: 23/01/2022 Revised: 09/03/2022 Accepted: 20/04/2022

Abstract

Root Cause Analysis (RCA) techniques are often applied to problems in the workplace; however, they may also prove very useful to home projects. This research explores the application of two RCA techniques in home projects: 1) 5 Whys to determine the root cause of a home air conditioning unit that runs continuously but does not cool, 2) an innovative Lean PFMEA to repair a John Deere riding mower that starts, then stops. Employing the 5 Whys technique led to the discovery of incorrect color-coded wiring from the original air conditioning unit to the thermostat. Lean PFMEA enabled a correct diagnosis and resolution of the mower start/stop issue via a Kaizen event, grass clippings in the fuel line, which was remedied by cleaning the fuel tank and replacing the fuel lines, fuel filter, and carburetor. These techniques provide Lean methodological approaches to problem-solving, which often leads to reduced homeowner aggravation, repair time, and repair expense.

Keywords: 5 Whys, Lean PFMEA, Home projects, Lean, Root cause analysis techniques.

1 | Introduction

John Krafcik is credited with making popular the term 'Lean', named after the world-famous Toyota production system, which focuses on waste elimination, shorter lead times, frequent changeovers, smaller lot sizes, and reduced inventories [1]. Lean was initially implemented in the automotive industry with great success, and the adoption of Lean production systems soon spread across many manufacturing industries worldwide. Today, the Lean philosophy is used in many service industries as well, such as healthcare [2]-[4], education [5], [6], and restaurants [7], yielding similar successes.

Lean can also be practiced in home applications and the literature reveals sparse research in this area. For example, Keyser [8] demonstrates how Root Cause Analysis (RCA) techniques, such as 5 Whys, Cause-and-Effect diagram, and Fault Tree analysis can be used to diagnose problems in home projects, Martinez [9] explores the possibility of applying Lean thinking in the Personal and Household Services (PHS) sector in the Czech Republic, and Brown et al. [10] studied the application of lean principles in the construction of Quadrant homes, resulting in the construction of a completed house in exactly 54 days.

Corresponding Author: rkeyser@kennesaw.edu

https://doi.org/10.22105/jarie.2022.320848.1410

This research discusses examples of how the 5 Whys and Lean PFMEA techniques can be used to diagnose problems encountered with home projects. More specifically, these techniques are used to identify possible causes leading to a failure. Once all possible causes were identified, steps were then undertaken by the homeowner to remedy home project issues.

2 | Literature Review

'Lean' philosophy was first introduced by John Krafcik [11], and became popular as a world-class manufacturing and management theory by Womack et al. [1] in their best-selling book, "the machine that changed the world". Lean progressively offers an interconnected socio-technical framework with a primary focus on making changes based on the needs and purpose of the enterprise [12], [13]. As such, the Lean skills can provide an ideal platform to provide sustainable production lines with fewer physical effort, less resources, less time, more quantity [1].

The new paradigm follows the basic philosophy of Lean thinking as the conversion of waste into customer-defined value. This is achieved by identifying value from customer standpoint, classifying the value stream while eliminating all steps with no value, creating a tight sequence of value-creating steps to flow smoothly towards the customer, developing systems to downstream customers, pulling value from the next upstream activity, and undertaking continuous improvement to a state of perfection [13]-[15]. Since its birth, the powerful concept of lean thinking has been widely extended from its original application in auto manufacturing industries into other economic sectors, including healthcare, education, food service, construction, electronics, printing industry, and even home projects. For example, in healthcare, the application of lean principles reflected in: reducing medications errors in Norwegian hospitals [16]; improving the efficiency and effectiveness of a clinical or an administrative process inside an Irish hospital [17]; tackling cultural performance and operational issues at a macro-level to improve patient care while simultaneously reducing costs [18]; focusing on eliminating waste as perceived by patients hence maximizing quality and safety for the patients [19]; increasing teamwork, creating user friendly work areas and processes, changing management styles and expectations, increasing staff empowerment and involvement, and streaming the supply chain within the perioperative area [20]; achieving a significant reduction in both the number of hospitalization days and the number of patients affected by healthcare-associated infections in the general surgery departments [21]; and improving technique to build an Adult Sickle Cell Disease Medical Home [22].

Furthermore, in education, the effect of lean implementation has introduced a pathway: to map the student value stream starting with kindergarten and look at all activities, in the classroom and beyond, to analyze where value is added and where waste is created [23]; to achieve rational and optimal educational system by providing quality in the learning experience, method of delivery, and student outcomes [24]-[26]; to improve administrative processes through an innovative and engaging learning experience by involving undergraduate students instead of employees [27]; to elaborate different approaches to quality improvement in education by dealing with the problem of improving technology supports and services for instructional purposes in a US school district system [28]; to propose a waste management framework in the Brazilian Education System that allows universities to organize their activities and select tools or practices to optimize their efforts to create value for final users [29]; to bolster students' learning through problem solving in a virtual simulation environment [30]; and to mitigate six common types of mistakes made by teachers in a high-school environment [5].

Similarly, in food service, the adaptation of lean strategy provided a roadmap: to understand contextual factors and to analyze the benefits and barriers of improving operational efficiency and productivity among the European food Small and Medium-sized Enterprises (SMEs) [31], [32]; to lay the foundations for the Greek food sector to be competitive in a global scenario of an economic downturn [33]; to fill the gap on the lack of knowledge while developing a tailor-made framework for companies and practitioners in the European food SEMs to increase the production efficiency [34]; to bring out pertinent factors and useful insights from a Norwegian dairy producer toward greater environmental sustainability in fresh-food supply chains [35]; to create valid and reliable multi-item measurement scales in identifying potential improvement opportunities to enhance food industry firms' performance and competitiveness [36]; and to reduce the waste costs such as time and food, increase employee utilization while lowering labor costs, operate at a more efficient inventory level, and improve the profitability in the selected American restaurant categories [7].

Similarly, in other economic sectors, the lean approach has offered an implicit strategy: to increase efficiency without compromising occupational health and safety of workers in manufacturing industries [37]; to develop sustainable building rating systems while reducing the environmental impact of construction [38]; to create a more reliable and quick work flow in design and construction of a capital project while delivering value to the customer [39]; to remove the bottlenecks affecting quality and productivity in electronics repair industries [40]; to assess the impact of lean and green practices on an organization's effectiveness and competitive advantage [41], [42] in manufacturing and service sectors [43] as well as ecological and social objectives [44]; and to address key organizational factors to successfully implement strategies and systematically adopt change in printing industries [45].

As part of the lean philosophy, RCA establishes a practical and structured approach to identify causal factors and to resolve underlying problems. RCA offers a management practice tool to enable companies to achieve organizational improvement and operational efficiency with the goal to unravel the underlying problematic issues and to propose a resolution to tackle or eliminate them altogether [46]. The process involves with the data collection, cause charting, root cause identification, mitigation strategy, and active implementation [47]. RCA includes a variety of analytical methods and techniques such as the Six Sigma methodology, the 5 Whys technique and the Failure Modes and Effect Analysis (FMEA).

John and Kadadevaramath [48] applied the Six Sigma methodology, including brainstorming and principal components analysis techniques, to improve the resolution time performance in software development. In the insurance industry, whereas Sarker et al. [49] employed Six Sigma, including process mapping, stepwise regression, and FMEA to analyze insurance claim submissions in order to reduce claim processing cycle times, John and Parikh [50] applied the Six Sigma methodology, including statistical analysis, dynamic regression, and integer programming to reduce the daily backlog percentage of accident and injury claims of an outsourcing insurance claim processing process. Lean models and frameworks were established to integrate lean tools with ISO 9001: 2015 requirements [51]; for improving the reliability of lean systems [52], and for understanding change in a lean environment [53].

Whereas lean thinking and lean principles have been successfully applied in manufacturing and service industries with great impact; its application in home projects has been less visible. This research explores the effect of lean thinking and strategies to diagnose problems encountered with home projects by using appropriate methodologies such as 5 Whys and an innovative lean PFMEA technique

The 5 Whys technique offers an interrogative scientific approach to explore the cause-and-effect relationships underlying a particular problem by repeating why 5 times to unravel the nature of the problem while suggesting a potential resolution to such problem [46]. The 5 Whys is attributed to Sakichi Toyoda, the founder of the Toyota Motor Corporation, for developing the technique during the evolution its manufacturing methodologies and processes. Ohno [12], the architect of the Toyota Production System, described the 5 Whys method as the basis of Toyota's scientific approach by repeating why five times until a countermeasure becomes apparent to offer resolutions and to mitigate the risk and prevent the issue from recurring [54]. The tool has widespread use beyond automobile industry; and if performed effectively, it can help transform a reactive culture or one that moves from one crisis to the next into a forward-looking culture or an organization that solves problems before they may even occur or escalate into a full-blown crisis [55].

Similarly, the FMEA provides a powerful design tool yet subjective analysis for the systematic identification of possible Root Causes and Failure Modes and the estimation of their relative risks [56]. FMEA is applicable at the various levels of system decomposition from the highest level of block diagram down to the functions of discrete components [57]. The FMEA team determines the effect of each failure, identifies the single crucial failure points, ranks each failure according to the criticality of the failure effect and its probability of occurring. The causes of failure are said to be Root Causes and may be defined as mechanisms that lead to the occurrence of a failure. The timing of the FMEA analysis is essential; and the main goal is to ideally determine and then limit or avoid potential risks within the early stages of product design or process development [58]. Conducting an FMEA on the existing products or processes may also yield benefits to ultimately achieve higher reliability, higher quality, and enhanced safety. Recent modifications to traditional FMEA include implementing an ordinal rating scale for risk analysis in radiation oncology [59]; combining FMEA analysis with identification of key determinants in building industry supply chains [60]; applying a new Doubly Technique for Order of Preference by Similarity to the Ideal Solution (DTOPSIS) approach within a fuzzy PFMEA for rankings of listed failure causes in the milk process industry [61]; and integrating lean with PFMEA in the automotive industry [62].

3 | Applying Lean at Home

Example 1 (5 Whys). Air conditioning unit runs continuously, but does not cool home.

In this example, a new Honeywell thermostat (*Fig. 1*) was installed to replace the original thermostat that was approximately 20 years old.

Fig. 1. New Honeywell thermostat.

The homeowner followed the wiring installation instructions in the owner's manual (*Fig. 2*).

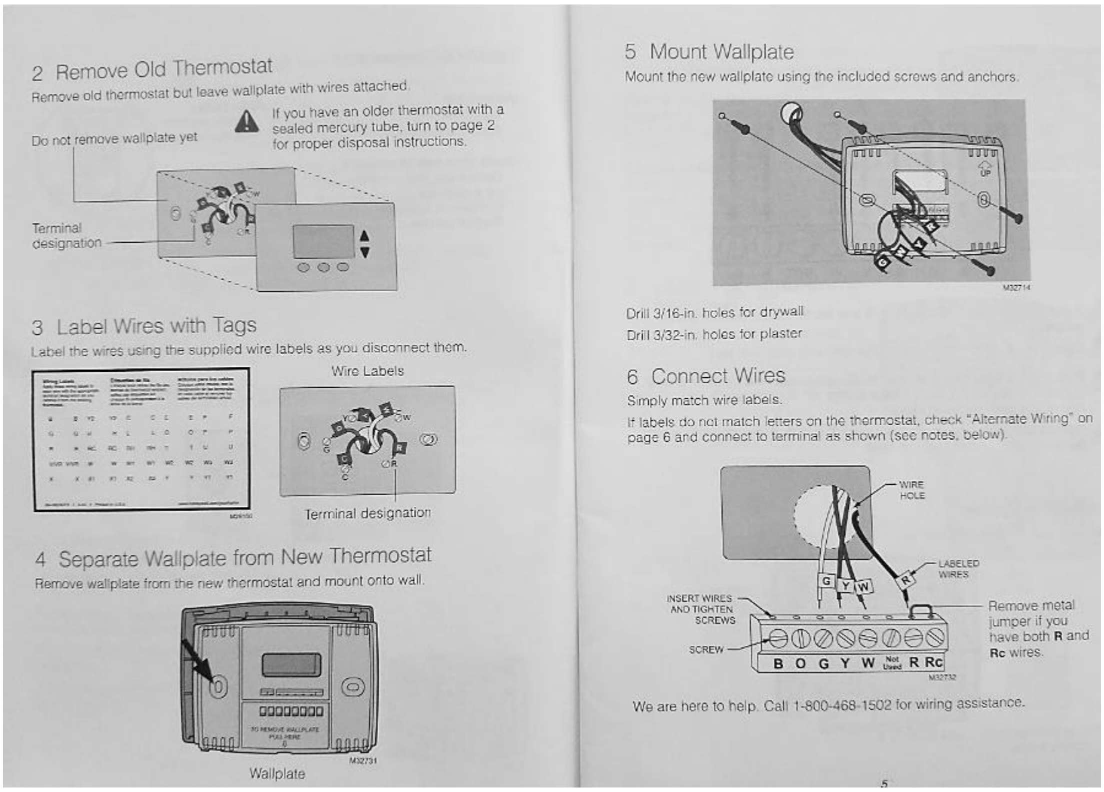


Fig. 2. Wiring instructions for Honeywell thermostat.

After installation, the thermostat ran continuously, did not cool the home, and would rarely turn off. The 5 Whys technique in *Fig. 3* was used to diagnose this issue.

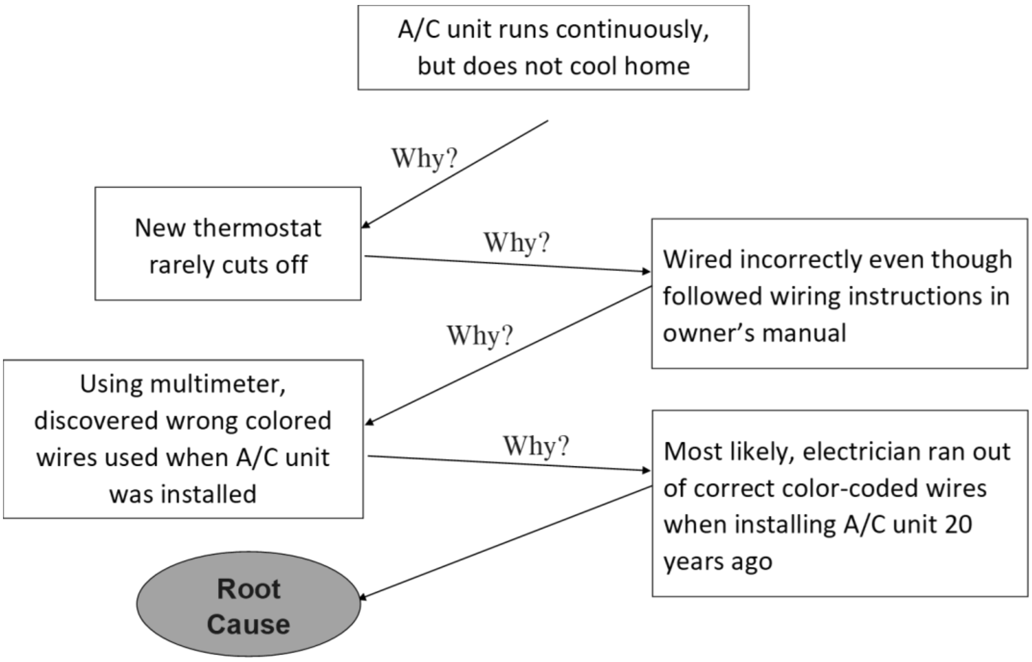


Fig. 3. 5 Whys diagram for diagnosing issue.

Solution

Once a multimeter was used to test voltage, current, and resistance, the root cause became readily apparent. Matching the color-coded wires to the contacts shown in the wiring instructions proved futile because two of the color-coded wires used in the original installation were incorrect. A plausible reason is that the electrician may have run out of the correct color-coded wires to properly install the original thermostat. By matching the wrong color-coded wires to the proper contacts corrected this issue. As a result, both the thermostat and HVAC system now run as intended.

Example 2 (Lean PFMEA). Why John Deere X304 Mower starts, then stop.

An innovative lean Process Failure Modes and Effects Analysis (PFMEA) allows for a structured way to diagnose a problem by identifying key process steps, then listing all failure modes, their effects, and the possible causes of those effects according to the severity and likelihoods of occurrence and detection, along with current controls in place, and then recommended lean actions.

In this example, a John Deere X304 riding mower would start every time, then stop. This provided great consternation to the homeowner, so a lean PFMEA was employed to determine all possible root causes as shown in *Table 1*.

The troubleshooting focus of this problem was abated to; 1) starting the mower, 2) the fuel system, as indicated in the process step column on the lean PFMEA chart. The mower started each time using a standard work procedure of inserting and turning the ignition key while the right foot simultaneously depressed the brake pedal, so that was not the problem. Preventive maintenance measures were applied as; 1) the date code on the battery indicated the battery was determined to still have a long useful life, 2) both sparkplugs were removed to check the sparkplug gap, and the sparkplug head was then cleaned on a wire wheel before re-installing. The problem-solving focus then centered on the fuel system, since this category had the highest RPN scores.

An overall 6s inspection of the mower exterior followed, particularly around the engine area. Upon removal of the fuel filter, grass shavings were observed inside the fuel filter, as shown in *Fig. 4*.

Fig. 4. Grass shavings visible inside fuel filter.

The result of a Kaizen event required replacing the fuel filter rather than cleaning it since the fuel filter was a sealed unit. But how did grass shavings get into a sealed fuel filter? Further visual inspection revealed that the fuel filter is connected by a fuel line on both the top and bottom ports of the fuel filter. The fuel line from the bottom port of the fuel filter is connected to the fuel tank. The fuel line from the top port of the fuel filter is connected to the carburetor.

Table 1. Lean PFMEA on why John Deere mower would start, then stop.

Procrss Step/ Function	Potential Failure Mode	Potential Failure Effects		Potential Causrs of Failure		Current Process Controls: Prevention			Recommended Action(S)	Responsibility	Action Taken				
What Is the Process Step, Change of Feature under Investigation?	In What Ways Could the Step, Change or Feature Go Wrong?	What is the Impact on the Customer If this Failure is not Prevented or Corrected?	Severity (1-10)	What Causes the Step, Change or Feature to Go Wrong? (How could it Occur?)	Occurrence (1-10)	What Controls Exist that Either Prevent or Detect the Failure?	Detection (1-10)	RPN	What Are the Recommended Actions for Reducing the Occurrence of the Cause or Improving Detection?	Who is Responsible for Making Sure the Actions are Completed?	What Actions were Completed (and When) with Respect to the RPN?	Severity (1-10)	Occurrence (1-10)	Detection (1-10)	RPN
Start mower	Missing key	Mower wont start	8	Forgot where key is located	10	Store key in same location each time	1	80	6s	Operator	Designated top right tray in rolling toolbox	8	1	2	16
	Broken key	Mower wont start	8	Turned old key too hard	10	Replace old key	1	80	jidoka	Operator	Replace worn key	8	1	2	16
	battery	Mower wont start	8	Batterys useful life ended	10	Check date on battery label	3	240	Preventive maintenance	Operator	Replace battery every 3 years	3	2	2	12
	Brake pedal not depressed	Mower wont start	5	Forgot to depress brake pedal when turning key	3	Familiarization with start-up procedure	3	45	Standard work	Operator	Depress brake pedal when starting mower	1	1	1	1
	Sparkplugs	Mower wont start	6	Dirty sparkplugs	8	Visual inspection	3	144	Preventive maintenance	Operator	Annual replacement during tune-up	2	1	2	4
	Old gas	Mower starts, then dies	8	Gas stored too long in gas container	3	Visual inspection	8	192	Poka-yoke	Operator	Put label on gas container with data purchased	3	1	2	6
	Gas tank clogged	Mower wont start	8	Debris entering gas tank	8	Non- sealed system	10	640	kaizen	Operator	Visual inspection of clean gas poured in gas tank	8	1	2	16
	Fuel line clogged	Mower starts, then dies	8	Debris entering fuel line	9	Non- sealed system	10	720	kaizen	Operator	Inspect fuel line before each use	8	2	4	64
	Fuel filter dirty	Mower starts, then dies	8	Debris entering fuel filter	10	Visual lnspection	6	480	Visual management	Operator	Inspect fuel filter before each use	8	1	1	8
	Carburetor clogged	Mower starts, then dies	8	Debris entering carburetor	10	Non- sealed system	10	800	kaizen	Operator	Visual inspection of carburetor before each use	8	2	7	112
						Total		3421			Tatal				255

Solution

This Kaizen event allowed for a structured approach to the homeowner in tracing the origin of grass shavings into the fuel filter via the fuel line back to the fuel tank. It was suspected that grass shavings entered the fuel tank most likely from grass debris around a moist spout on the gas container when filling the gas tank for the most recent mowing. The grass shavings worked their way down in the gas tank to the port entering the fuel line which allowed grass shavings to travel to and enter the fuel filter, and then continued out of the top port of the fuel filter through the fuel line into the carburetor. Since the John Deere mower was already 14 years old, the homeowner decided to replace the carburetor along with the fuel lines and fuel filter (see, *Fig. 5*). The gas tank was removed and thoroughly cleaned. The gas container spout was also cleaned of moisture and debris. By taking these action steps, the RPN was significantly reduced from 3, 421 to 255 and the John Deere X304 riding mower runs like new again.

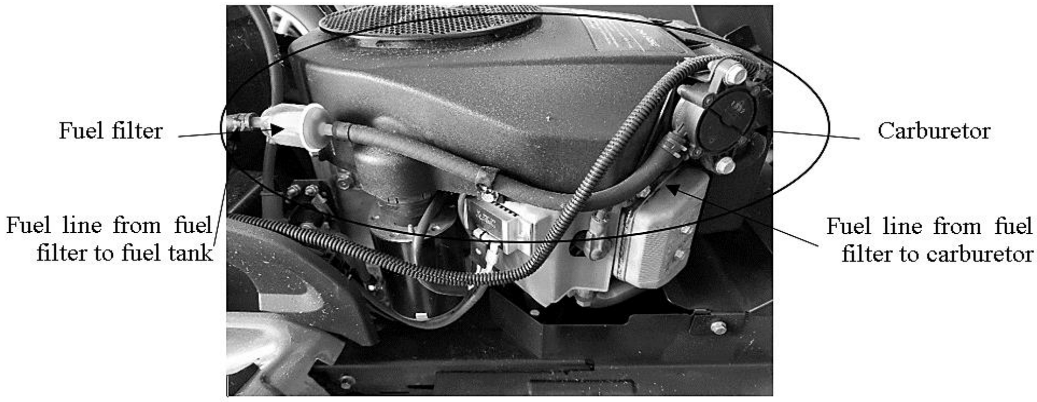


Fig. 5. Replaced fuel filter, fuel line, and carburetor.

4 | Conclusion

The 5 Whys technique was used to diagnose why a newly installed thermostat ran continuously and yet did not cool the home. The use of a multimeter led to the discovery that the wrong color-coded wires were used to connect the outside air conditioning unit to the wall thermostat during the initial installation of the air conditioning unit some 20 years ago. Lean PFMEA was used to determine why a John Deere X304 riding mower would start, then stop. The severity of potential failure effects, the likelihood of occurrence of the potential causes of those effects, and the likelihood of detecting those causes based on current controls in place was analyzed. A total RPN of 3,421 was calculated for the current state in which the focus abated to; 1) starting the mower, 2) the fuel system. A Kaizen event led to discovery of causes and corrective actions, such as cleaning the gas tank, replacing the fuel lines, fuel filter, and carburetor, reducing the overall RPN to 255. With regard to the fuel system, the RPN was significantly reduced from 2,832 to 206. Corrective actions restored the mower back to an operable state.

5 Whys and lean PFMEA are practical methodological techniques that can help diagnose root causes in malfunctioning equipment. This, in turn, can lead to a more rapid repair and restoration of homeowner equipment previously deemed inoperative.

5 | Areas for Future Study

Lean applications can be easily adopted for home use. Areas for future study include applying 6s and visual management techniques to organize a home interior/exterior, garage, or workshop; mistake-proofing common tasks conducted around the home by the homeowner, such as using ladders; creating checklists to perform intermittent activities, such as changing the oil in a car, proper tire rotation on a vehicle, and starting equipment that requires a sequence of steps; applying setup time reduction techniques for washing the car, doing the laundry, installing fixtures (such as lights or ceiling fans, etc.), painting, assembling a bicycle, or doing yard work.

Funding

No funding supported this research.

Conflicts of Interest

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

[1] Womack, J. P., Jones, D. T., & Roos, D. (2007). *The machine that changed the world: the story of lean production--Toyota's secret weapon in the global car wars that is now revolutionizing world industry*. Simon and Schuster.
[2] Aherne, J., & Whelton, J. (Eds.). (2010). *Applying lean in healthcare: a collection of international case studies*. CRC Press.
[3] Bowerman, J. (2007). *Introducing Lean in healthcare*. Bradford: Emerald Publishing Limited.
[4] Hadfield, D. (2006). *Lean healthcare: implementing 5s in lean or Six SIGMA projects: 5 keys for improving the workplace environment*. MCS Media. https://www.amazon.com/Lean-Healthcare-improving-workplace-environment-ebook/dp/B001DA3H06
[5] Keyser, R. S. (2022). Lean in education: mistake-proofing methods used by teachers at a magnet high school. *Journal of management & engineering integration*, *12*(1), 49-58.
[6] Martínez Sanahuja, S. (2020). Towards lean teaching: non-value-added issues in education. *Education sciences*, *10*(6), 160.
[7] Keyser, R. S., Marella, V. K., & Clay, K. (2017). Lean restaurants: improving the dining experience. *Journal of higher education theory and practice*, *17*(7), 67-79.
[8] Keyser, R. S. (2021). Lean at home: applying RCA techniques to home projects. *Journal of applied research on industrial engineering*, *8*(2), 104-115.
[9] Martinez, F. (2019). Lean home services in Czech Republic. *International journal of Lean Six Sigma*, *10*(3), 784-802.
[10] Brown, K. A., Schmitt, T. G., Schonberger, R. J., & Dennis, S. (2004). Quadrant Homes applies lean concepts in a project environment. *Interfaces*, *34*(6), 442-450.
[11] Krafcik, J. F. (1988). Triumph of the lean production system. *Sloan management review*, *30*(1), 41-52.
[12] Ohno, T. (1988). *Toyota production system: beyond large scale production*. Productivity Press, New York.
[13] Womack, J. P., & Jones, D. T. (1997). Lean thinking—banish waste and create wealth in your corporation. *Journal of the operational research society*, *48*(11), 1148-1148. https://doi.org/10.1057/palgrave.jors.2600967
[14] Chay, T., Xu, Y., Tiwari, A., & Chay, F. (2015). Towards lean transformation: the analysis of lean implementation frameworks. *Journal of manufacturing technology management*, *26*(7), 1031-1052.
[15] Swank, C. K. (2003). The lean service machine. *Harvard business review*, *81*(10), 123-130.
[16] Antony, J., Forthun, S. C., Trakulsunti, Y., Farrington, T., McFarlane, J., Brennan, A., & Dempsey, M. (2019). An exploratory study into the use of Lean Six Sigma to reduce medication errors in the Norwegian public healthcare context. *Leadership in health services*, *32*(4), 509-524.
[17] Laureani, A., Brady, M., & Antony, J. (2013). Applications of lean six sigma in an Irish hospital. *Leadership in health services*, *26*(4), 322-337.
[18] Matthias, O., & Brown, S. (2016). Implementing operations strategy through Lean processes within health care: the example of NHS in the UK. *International journal of operations & production management*, *36*(11), 1435-1457.
[19] Cohen, R. I. (2018). Lean methodology in health care. *Chest*, *154*(6), 1448-1454.
[20] Kimsey, D. B. (2010). Lean methodology in health care. *AORN journal*, *92*(1), 53-60.

[21] Montella, E., Di Cicco, M. V., Ferraro, A., Centobelli, P., Raiola, E., Triassi, M., & Improta, G. (2017). The application of Lean Six Sigma methodology to reduce the risk of healthcare–associated infections in surgery departments. *Journal of evaluation in clinical practice*, *23*(3), 530-539.

[22] Johnson, S., Hartigan, S., Holt, E., Sop, D., McHenry, C., Lipato, T., ... & Smith, W. R. (2019). Using Lean Six Sigma to develop a patient centered medical home for adults with sickle cell disease. *Blood*, *134*, 3408.

[23] Alves, A. C., Flumerfelt, S., & Kahlen, F. J. (Eds.). (2016). *Lean education: an overview of current issues*. Springer. https://doi.org/10.1007/978-3-319-45830-4

[24] Balzer, W. K. (2020). *Lean higher education: Increasing the value and performance of university processes*. CRC Press.

[25] Antony, J., Krishan, N., Cullen, D., & Kumar, M. (2012). Lean Six Sigma for higher education institutions (HEIs): challenges, barriers, success factors, tools/techniques. *International journal of productivity and performance management*, *61*(8), 940-948.

[26] Vukadinovic, S., Djapan, M., & Macuzic, I. (2017). Education for lean & lean for education: a literature review. *International journal for quality research*, *11*(1), 35.

[27] Doman, M. S. (2011). A new lean paradigm in higher education: a case study. *Quality assurance in education*, *19*(3), 248-262.

[28] LeMahieu, P. G., Nordstrum, L. E., & Greco, P. (2017). Lean for education. *Quality assurance in education, 25*(1), 74-90.

[29] Klein, L. L., Tonetto, M. S., Avila, L. V., & Moreira, R. (2021). Management of lean waste in a public higher education institution. *Journal of cleaner production*, *286*, 125386.

[30] Gadre, A., Cudney, E., & Corns, S. (2011). Model development of a virtual learning environment to enhance lean education. *Procedia computer science*, *6*, 100-105.

[31] Dora, M., & Gellynck, X. (2015). Lean Six Sigma implementation in a food processing SME: a case study. *Quality and reliability engineering international*, *31*(7), 1151-1159.

[32] Dora, M., Van Goubergen, D., Kumar, M., Molnar, A., & Gellynck, X. (2013). Application of lean practices in small and medium-sized food enterprises. *British food journal*, *116*(1), 125-141.

[33] Psomas, E., Antony, J., & Bouranta, N. (2018). Assessing lean adoption in food SMEs: evidence from Greece. *International journal of quality & reliability management*, *35*(1), 64-81.

[34] Van Goubergen, D., Dora, M., Molnar, A., Gellynck, X., & Kumar, M. (2011). Lean application among European food SMEs: findings from empirical research. *IIE annual conference Proceedings* (p. 1). Institute of Industrial and Systems Engineers (IISE).

[35] Powell, D., Lundeby, S., Chabada, L., & Dreyer, H. (2017). Lean Six Sigma and environmental sustainability: the case of a Norwegian dairy producer. *International journal of lean six sigma*, *8*(1), 53-64.

[36] Costa, L. B. M., Godinho Filho, M., Fredendall, L. D., & Ganga, G. M. D. (2021). Lean six sigma in the food industry: construct development and measurement validation. *International journal of production economics*, *231*, 107843.

[37] Singh, C., Singh, D., & Khamba, J. S. (2021). Exploring an alignment of lean practices on the health and safety of workers in manufacturing industries. *Materials today: proceedings*, *47*, 6696-6700.

[38] Aristizábal-Monsalve, P., Vásquez-Hernández, A., & Botero, L. F. B. (2022). Perceptions on the processes of sustainable rating systems and their combined application with Lean construction. *Journal of building engineering*, *46*, 103627.

[39] Tsao, C. C., Tommelein, I. D., Swanlund, E. S., & Howell, G. A. (2004). Work structuring to achieve integrated product–process design. *Journal of construction engineering and management*, *130*(6), 780-789.

[40] Samuel, R., Rajesh, M., Rajanna, S., & Franklin, E. (2021). Implementation of lean manufacturing with the notion of quality improvement in electronics repair industry. *Materials today: proceedings*, *47*, 2253-2257.

[41] Teixeira, P., Coelho, A., Fontoura, P., Sá, J. C., Silva, F. J., Santos, G., & Ferreira, L. P. (2022). Combining lean and green practices to achieve a superior performance: the contribution for a sustainable development and competitiveness—an empirical study on the Portuguese context. *Corporate social responsibility and environmental management*, *29*(4), 887-903.

[42] Udokporo, C. K., Anosike, A., Lim, M., Nadeem, S. P., Garza-Reyes, J. A., & Ogbuka, C. P. (2020). Impact of lean, agile and green (LAG) on business competitiveness: an empirical study of fast moving consumer goods businesses. *Resources, conservation and recycling*, *156*, 104714. https://doi.org/10.1016/j.resconrec.2020.104714

[43] Silva, S., Sá, J. C., Silva, F. J., Ferreira, L. P., & Santos, G. (2020). Lean green—the importance of integrating environment into lean philosophy—a case study. *Proceedings of the 6th European lean educator conference: ELEC 2019 6* (pp. 211-219). Springer International Publishing. https://doi.org/10.1007/978-3-030-41429-0_21

[44] Das, K. (2018). Integrating lean systems in the design of a sustainable supply chain model. *International journal of production economics*, *198*, 177-190.

[45] Ainul Azyan, Z. H., Pulakanam, V., & Pons, D. (2017). Success factors and barriers to implementing lean in the printing industry: a case study and theoretical framework. *Journal of manufacturing technology management*, *28*(4), 458-484.

[46] Gangidi, P. (2018). A systematic approach to root cause analysis using 3× 5 why's technique. *International journal of lean six sigma*, *10*(1), 295-310.

[47] Rooney, J. J., & Heuvel, L. N. V. (2004). Root cause analysis for beginners. *Quality progress*, *37*(7), 45-56.

[48] John, B., & Kadadevaramath, R. S. (2020). Improving the resolution time performance of an application support process using Six Sigma methodology. *International journal of lean six sigma*, *11*(4), 663-686.

[49] Ashok Sarkar, S., Ranjan Mukhopadhyay, A., & Ghosh, S. K. (2013). Improvement of claim processing cycle time through Lean Six Sigma methodology. *International journal of lean six sigma*, *4*(2), 171-183.

[50] John, B., & Parikh, P. (2020). Improving the insurance claim processing process using Six Sigma methodology. *International journal of six sigma and competitive advantage*, *12*(4), 348-368.

[51] Sá, J. C., Vaz, S. Carvalho, O., Lima, V., Morgado, L., Fonseca, L., Doiro, M., & Santos, G. (2022). A model of integration ISO 9001 with Lean six sigma and main benefits achieved. *Total quality management & business excellence*, *33*(1-2), 218-242. https://doi.org/10.1080/14783363.2020.1829969

[52] Keyser, R. S., & Sawhney, R. S. (2013). Reliability in lean systems. *International journal of quality & reliability management*, *30*(3), 223-238.

[53] Keyser, R. S., Sawhney, R. S., & Marella, L. (2016). A management framework for understanding change in a lean environment. *Tékhne*, *14*(1), 31-44.

[54] Serrat, O., & Serrat, O. (2017). The five whys technique. In *Knowledge solutions: tools, methods, and approaches to drive organizational performance* (pp. 307-310). Springer.

[55] King, R. H., & Baum, N. (2018). Problem solving in the medical practice using the five whys. *Journal of medical practice management: MPM*, *34*(3), 177-179.

[56] Arabian-Hoseynabadi, H., Oraee, H., & Tavner, P. J. (2010). Failure modes and effects analysis (FMEA) for wind turbines. *International journal of electrical power & energy systems*, *32*(7), 817-824.

[57] Cassanelli, G., Mura, G., Fantini, F., Vanzi, M., & Plano, B. (2006). Failure analysis-assisted FMEA. *Microelectronics reliability*, *46*(9-11), 1795-1799.

[58] Lipol, L. S., & Haq, J. (2011). Risk analysis method: FMEA/FMECA in the organizations. *International journal of basic & applied sciences*, *11*(5), 74-82.

[59] Baehr, A., Oertel, M., Kröger, K., Eich, H. T., & Haverkamp, U. (2020). Implementing a new scale for failure mode and effects analysis (FMEA) for risk analysis in a radiation oncology department. *Strahlentherapie und Onkologie*, *196*, 1128-1134. https://doi.org/10.1007

[60] Kulińska, E., Masłowski, D., & Dendera-Gruszka, M. (2021). New application of FMEA analysis in the heavy industry supply chain. *European research studies journal*, *24*(2B), 600-616.

[61] Das, I., Panchal, D., & Tyagi, M. (2022). A novel PFMEA-Doubly TOPSIS approach-based decision support system for risk analysis in milk process industry. *International journal of quality & reliability management*, *39*(1), 1-29.

[62] Banduka, N., Veza, I., & Bilić, B. (2016). An integrated lean approach to process failure mode and effect analysis (PFMEA): a case study from automotive industry. *Advances in production engineering and management*, *11*(4), 355-365. https://scidar.kg.ac.rs/handle/123456789/8986

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 56–83.

Paper Type: Research Paper

Integration of Two-Stage Assembly Flow Shop Scheduling and Vehicle Routing Using Improved Whale Optimization Algorithm

Vahid Bahmani[1], Mohammad Amin Adibi[1,*], Esmaeil Mehdizadeh[1]

[1] Department of Industrial Engineering, Faculty of Industrial and Mechanical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran; v.bahmani@ymail.com; maadibi@aut.ac.ir; emqiau@yahoo.com.

Citation:

Bahmani, V., Adibi, M. A., & Mehdizadeh, E. (2023). Integration of two-stage assembly flow shop scheduling and vehicle routing using improved whale optimization algorithm. *Journal of applied research on industrial engineering, 10*(1), 56-83.

Received: 16/02/2022 Reviewed: 20/03/2022 Revised: 09/05/2022 Accepted: 01/06/2022

Abstract

This paper provides an integrated model for a two-stage assembly flow shop scheduling problem and distribution through vehicle routing in a soft time window. So, a Mixed-Integer Linear Programming (MILP) model has been proposed with the objective of minimizing the total cost of distribution, holding of products, and penalties of violating delivery time windows. To solve this problem, an improved meta-heuristic algorithm based on Whale Optimization Algorithm (WOA) has been developed. The main innovations in the study include considering soft time window, sequence-dependent setup time, delivery time window, heterogeneous vehicles, holding costs of final products, and unrelated assembly machines. A comparison of the integrated and non-integrated model in a case study of industrial gearboxes production shows that the integrated model compared to the non-integrated model has saved 15.6% and 13.6% in terms of delay time and total costs, respectively. Computational experiments also indicate the efficiency of improved WOA in converging to optimal solution and achieve better solution in comparison to the Genetic Algorithm (GA). The results show that increasing the setup time can lead to an increase in total costs. It can be said that the increase of setup time increases the completion of time jobs. Also, the costs increased with decreasing the transport fleet capacity to -20%. The reason for this is that by reducing the capacity of vehicles, the model has to use more vehicles.

Keywords: Sentence Two-stage assembly, Vehicle routing, Whale optimization algorithm, Genetic algorithm.

1 | Introduction

In today's competitive environment, manufacturing companies are seeking to minimize costs to succeed in supply chains. Integrating supply chain decisions is one of the cost-saving methods. Efficient supply chain design involves complex issues such as inventory management, production scheduling, and order distribution [1]. On the other hand, flexibility and timely delivery are essential to customer satisfaction [2]. In the non-integrated approach, the manufacturing section prefers fewer preparations to reduce production costs, and more transportation travels to reduce delivery time, while the distribution section prefers less transportation travels with larger volumes to reduce distribution costs. In addition to the conflict between these two, this approach causes an increase in the costs of both sections and the entire system. On the other hand, integrating production and distribution schedules can reduce system costs. Production and distribution are complex problems,

Corresponding Author: maadibi@aut.ac.ir

https://doi.org/10.22105/jarie.2022.329251.1450

and therefore their integration will increase the complexity of the problem. However, it is expected that total costs be reduced by integrating the production and distribution processes [3]-[7].

The objective of this study is to determine order production and distribution schedule and routes that minimize total company costs including fixed costs of using vehicles, variable travel costs, earliness and tardiness penalties, as well as holding costs. In this study, after modeling the integrated production and distribution problem, it will be compared with the non-integrated model. The Whale Optimization Algorithm (WOA) will be also used as a new and efficient algorithm for production and distribution. The other sections of this study are as follows:

Section 2 will present relevant literature. Section 3 will describe the problem and define the mathematical model. In Section 4, the used meta-heuristic algorithms will be introduced and developed. Section 5 of the study contains some experiments and a case study to evaluate the performance of the proposed model and comparison of the performance of the proposed algorithms. Finally, the conclusion and suggestions for the future will be presented in the last section.

2 | Literature Review

The two-stage assembly flow shop is one of the most widely used scheduling problems in manufacturing industries. Potts et al. [8] well demonstrated the utility of this problem. By studying the two-stage assembly scheduling problem with the makespan objective function, they proved that the problem is NP-hard in the first stage due to the existence of two machines. Lee et al. [9] used this problem to assemble a fire engine. They used a branch and bound algorithm for the two-stage assembly scheduling problem with the makespan objective function and analyzed their boundary error by presenting three heuristic approaches. In another study, Allahverdi and Al-Anzi [10] investigated queries scheduling in a distributed database system. In this regard, many real-world problems can be modeled using the two-stage assembly scheduling problem. Therefore, the two-stage assembly problem can be considered as a general case of the flow shop problem. Tozkapan et al. [11] described this problem with the performance measure of total weighted flow time. With creating a lower bound, they obtained good results for problems of logical scale and proposed a heuristic algorithm for large scale problems. Allahverdi and Al-Anzi [10] studied the problem presented by Tozkapan et al. [11] using total completion time. They developed three algorithms, including Tabu Search (TS), Hybrid Tabu Search (HTS), and Simulated Annealing (SA). Computational experiments showed that although the CPU time of all three algorithms was approximately the same, and HTS improved the error rate by 60 and 90 percent compared to TS and SA, respectively. Sung and Kim [13] addressed the problem of scheduling a two-stage multi-machine assembly flow shop. They proposed a branch and bound algorithm and an efficient heuristic algorithm to minimize the sum of completion times. Allahverdi and Al-Anzi [10] also considered the problem with bi-criteria of makespan and maximum tardiness, and proposed three heuristic algorithms, including Particle Swarm Optimization (PSO), TS and Self-adaptive Differential Evolution (SDE). The analyses showed that both SDE and PSO had better performance than TS, but PSO performance was better than SDE. Koulamas and Kyparisis [14] generalized the two-stage problem to the three-stage assembly scheduling problem regarding collection and transportation with the goal of minimizing the makespan. They also proposed several heuristic algorithms and evaluated the worst-case ratio bound of those algorithms.

Zhang and Tang [15] integrated Preventive Maintenance (PM) with the problem of two-step assembly scheduling by presenting a Mixed Integer Linear Programming (MILP) model. To solve the problem, they proposed an iterative greedy algorithm based on PM and two heuristics Mixed Constrained Machine Preventive Maintenance (MCMTPM) and Net Economic Hybrid Preventive Maintenance (NEHPM). Numerical results showed that the proposed Iterated Greedy Preventive Maintenance (IGPM) embedded with NEHPM and local reference search performed better than the other 9 intelligent methods. Pourhejazy et al. [16] addressed the issue of distributed assembly timing with distributed start-up time sequences (DTSAFSP-SDSTs) with the aim of minimizing makespan. They used the Iterated Greedy algorithm to solve this problem effectively. Numerical experiments showed that the Improved Iterated Greedy algorithm offers the best solutions in most cases.

Hatami et al. [17] generalized the model proposed by Koulamas and Kyparisis [14], by considering sequence-dependent set-up times. Their mathematical model aimed to minimize the weighted sum of the mean flow time and maximum tardiness. They proposed two algorithms, SA and TS, to solve the problem and compared their performance. Andrés and Hatami [18] formulated two mathematical models to solve the three-stage assembly flow shop problem. Their objective was to minimize the total completion time by considering the sequence-dependent setup time in the first and third stages. Their results showed that they could find optimal solutions for problems with n = 15 (number of jobs) and m = 4 (number of parts). Maleki-Darounkolaei et al. [19] investigated this problem by considering the sequence-dependent setup time in the first stage as well as the blocking times between successive stages. They proposed a SA algorithm to solve the problem with the objective function of minimizing the weighted sum of the two objectives of the makespan and the mean completion time. Dalfard et al. [20] studied the problem by considering the sequence-dependent setup time and transportation times. Their objective function included minimizing the sum of total weighted squared earliness, total weighted squared tardiness, number of tardy jobs and makespan. For solving the problem, they used a hybrid Genetic Algorithm (GA) and concluded that for jobs more than 10, the results were not comparable between Lingo 8 and the hybrid GA.

Mozdgir et al. [21] considered the problem of two-stage assembly scheduling with non-identical assembly machines with the objective function of minimizing the weighted sum of the two criteria of mean completion time and makespan. They proposed a hybrid variable neighborhood search heuristic to solve the MILP model. Tian et al. [22] considered the problem with two criteria of mean completion time and makespan, and proposed a Discrete Particle Swarm Optimization (DPSO) algorithm to solve the problem. The results of that study indicated the efficiency of DPSO. Allahverdi et al. [23] studied the problem by assuming setup times as zero and the objective function of minimizing total tardiness. They proposed two versions of the SA algorithm, two versions of cloud theory-based SA, an insertion algorithm, and a GA to solve the problem. The results showed that one version of the SA combined with PIA had a better performance than the other algorithms. Allahverdi and Aydilek [12] proposed two new algorithms for a two-stage assembly scheduling problem considering separate setup times and compared four existing algorithms. The results of the analysis indicated that the error of the best algorithm is less than other algorithms by 54%-98%. Navaei et al. [24], also addressed the problemby considering several non-identical assembly machines and the makespan objective function. They developed a MILP model and proposed a hybrid SA algorithm to solve the problem.

Shoaardebili and Fattahi [25] studied the three-stage assembly flow shop scheduling problem simultaneously with machine availability constraints and two objective functions of minimizing the sum of weighted tardiness and earliness and minimizing total weighted completion times. Analyses indicated that of the two NSGA II and MOSA algorithms proposed, NSGA II had a better performance. Komaki and Kayvanfar [26] studied the two-stage assembly scheduling problem with identical assembly machines and the release date of jobs. The objective function of their model was makespan. They proposed a Grey Wolf Optimizer (GWO) algorithm to solve the problem. The analyses showed that the GWO algorithm exhibited better performance than the other more well-known algorithms. The two-stage assembly scheduling problem has developed in many ways over time. Allahverdi and Al-Anzi [10] considered the problem with m machines in the first stage and the objective function of the total completion time. They developed three heuristic approaches, consisting of an HTS, an SDE, and Novel Self-adaptive Differential Evolution (NSDE) algorithm to solve the problem. Computational experiments showed that NSDE had better performance than the other two algorithms.

Wang et al. [27] considered the two-stage assembly flow shop scheduling problem with batch delivery to one customer. For solving the problem with the objective function of minimizing the weighted sum of average arrival time at the customer and the total delivery cost, they proposed two heuristic methods

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

based on SPT and LPT and a new hybrid meta-heuristic (HGA-OVNS). The computational results showed the superiority of the HGA-OVNS meta-heuristic algorithm. Kazemi et al. [28] studied the problem of two-stage assembly flow shop scheduling with identical assembly machines and batch delivery. Their objective was to schedule jobs considering batches that would minimize the sum of tardiness and delivery costs. They proposed the Imperialist Competition Algorithm (ICA) and Hybrid Imperialist Competition Algorithm (HICA) to solve the MILP model. Their computational results indicated the superiority of the HICA algorithm in terms of the objective function but the ICA algorithm needed a comparably less time for implementation. Jung et al. [29] considered the two-stage assembly flow shop scheduling problem to assemble products with dynamic components-sizes and makespan objective function. In their MILP model, they considered the setup time to process the components of a new product. They proposed three GAs with different chromosome representations to solve large-scale problems. Wu et al. [30] addressed the two-stage flow shop scheduling problem using three machines and with learning phenomenon. Their objective was to minimize the total completion time and used the branch and bound algorithm to solve small-scale problems. In addition, six versions of the hybrid PSO algorithm were proposed for small- and large-scale problems and three different data types. In addition, ANOVA was used to evaluate the performance of the proposed algorithms. Goli and Davoodi [31] presented a coordinated model of production and distribution with a constant rate of demand in the supply chain. They developed two algorithms including SA refrigeration simulation and Biography-Based Optimization (BBO) algorithm. Numerical results showed better performance of BBO algorithm.

Basir et al. [32] considered the problem of two-stage assembly with batch delivery. They presented a Mixed-Integer Linear Programming (MILP) model with the aim of minimizing the weighted number of tardy jobs and the sum of delivery costs. They proposed a two-stage Improved Genetic Algorithm (IGA) with a hierarchical decision-making approach. Yavari and Isvandi [33] integrated the two-stage assembly scheduling problem by ordering the parts needed to process the components. They developed a MILP model to minimize the sum of the total weighted completion time, parts ordering, and holding cost. They proposed a GA and the computational results showed that the integrated approach improved the supply chain performance up to 8.16%. Luo et al. [34] investigated a two-stage assembly scheduling problem to minimize makespan considering separate setup times. They proposed a hybrid branch and bound algorithm. Their computational results showed that the algorithm performed better than one of the available methods. Talens et al. [35] addressed the problem of two-stage assembly scheduling with identical assembly machines and the objective function of minimizing total completion time. They presented two heuristic algorithms of CH_{MMA} and $BSCH_{MMA}$. The computational results showed that the proposed heuristic methods had better performance than the available heuristic ones. Lin and Chen [36], provided a dynamic scheduling of two-stage assembly flow shop to minimize the total tardiness as a Markov decision process. A Proximal Policy Optimization (PPO) algorithm was developed to efficiently train agent using production data. Hosseini et al. [37] investigated a two-stage production system consists of a fabrication stage followed by an assembly stage. Moreover, a heuristic algorithm and two proper lower bounds were introduced as references to evaluate the performance of the proposed heuristic algorithm.

In a study, Masruroh et al. [38] addressed the issue of multi-product production planning and integrated distribution in the product supply chain network. Their proposed model has two stages, the first of which is determined in order to maximize the profit of production, delivery and inventory planning for each product. Then, in the second stage, the exact production schedule is optimized to minimize the total start-up costs. Numerical results showed a significant reduction in costs and an increase in annual profits. Based on literature review section the literature review table is as follows.

Table 1. Summary of literature review.

No	Author	Soft Time Window	Sequence-Dependent Setup Time	Delivery Time Window	Heterogeneous Vehicles	Holding Costs of Final Products	Unrelated Assembly Machines	Integration of Flow Shop Scheduling And Vehicle Routing	Metaheuristic/Heuristic
1	Zhang and Tang [15]			*			*		*
2	Pourhejazy et al. [16]		*				*		*
3	Lin and Chen [36]	*		*					*
4	Masruroh et al. [38]	*							-
5	Hosseini et al. [37]				*				*
6	Talens et al. [35]	*					*		*
7	Luo et al. [34]					*	*		*
8	Yavari and Isvandi [33]			*		*			
9	Basir et al. [32]			*					*
10	Goli and Davoodi [31]			*					*
11	Wu et al. [30]	*							*
12	Kazemi et al. [28]	*					*		*
13	Wang et al. [27]				*		*		*
14	Allahverdi et al.[23]		*				*		*
15	Allahverdi and Al-Anzi [10]	*					*		*
16	Komaki and Kayvanfar [26]			*					*
17	Shoaardebili and Fattahi [25]	*					*		*
18	Mozdgir et al. [21]	*					*		*
19	Tian et al. [22]	*					*		*
20	Navaei et al. [24]				*				*
21	Dalfard et al. [20]	*		*					*
22	Hatami et al. [17]	*	*						*
23	This Study	*	*	*	*	*	*		*

In none of the previous studies, the two-stage assembly flow shop problem has been integrated with the distribution of orders through vehicle routing. However, distribution stage decisions are highly dependent on production stage decisions, and their integration can lead to minimizing total system costs. Therefore, the main innovation of the current study is to provide a new mathematical model and solution for the problem of two-stage assembly flow shop and vehicle routing. Other innovations in the study include considering soft time window, sequence-dependent setup time, delivery time window, heterogeneous vehicles, holding costs of final products, and unrelated assembly machines. Since production, assembly, and vehicle routing decisions are all short-term and operational, it is not advisable to use exact solutions that need a long time to perform. Therefore, an improved and novel meta-heuristic methods will be used in this study. For this purpose, the Improved Whale Optimization Algorithm (IWOA) is proposed and applied for the first time to integrate a two-stage assembly flow shop and vehicle routing.

3 | Problem Description and Proposed Mathematical Model

In this section, the structure of the problem is first described in detail. In the next step, after introducing the symbols, integrated and non-integrated mathematical models are presented.

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

3.1 | Problem Description

The problem under study involves decision making in three related areas including production, assembly, and distribution. In the different production systems, these three areas are performed sequentially and successively and independent planning and optimization are performed for each area. Integrating production, assembly and distribution decisions is a novel approach that has been addressed in this study and has been introduced as the Integration of Production, Assembly and Distribution (IPAD) problem. First, the IPAD model is introduced, and then the non-integrated model is presented. In the integrated model, it is assumed that at moment 0, J jobs are present in the production system. Each job represents an order. Each order has an independent I components, each of which requires independent processing. Each process is performed on a specific machine. It is important to note that there is a sequence-dependent setup time for each job to be processed. Therefore, determining the sequence of jobs has a substantial effect on the completion time of different jobs.

Once all the components of each job are processed, they should be assembled on one of the l unrelated machines in the second stage. Since vehicles are available at the completion time of the last job (Cmax), each product is held in a temporary warehouse after completion of assembly. Therefore, the assembled products must be held in the warehouse until the start of loading and distribution. This holding is costly for the company and it is necessary to plan the production and assembly in such a way that the holding time and costs of the orders are minimized.

In the distribution stage, there are V vehicles with limited capacity in the system. Each vehicle has a fixed cost as well as a variable cost per 1 km distance traveled. At this stage, time is a crucial pillar. Each customer has a time window for delivering orders. Servicing outside this time window will impose stupendous costs on the company. Therefore, at this stage, attempts will be made to determine an optimal route for each vehicle that will lead to the least cost on the route and the least penalties of violating the time window. *Fig. 1* illustrates the structure of this problem schematically.

Other assumptions are as follows:

I. All orders are available in the system at time zero.
II. All production and assembly machines are available since time zero.
III. Preemption is not allowed on any machine.
IV. Machine idle time in the first stage is not allowed.
V. Processing time and machine setup time for all jobs are predetermined.
VI. The assembly time of each job is definite and predetermined.
VII. In the distribution stage, the delivery of orders to customers has a specific time.
VIII. Travel time is the same for all vehicles.
IX. Each customer has a soft time window whose violation imposes costs on the company.

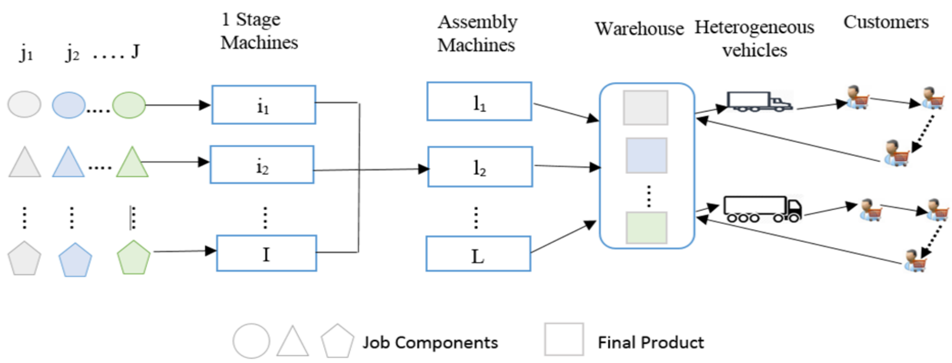


Fig. 1. Schematic representation of the integration of two-stage assembly flow shop and vehicle routing.

In this section, based on the considered assumptions, a distribution-production integrated mathematical model is presented.

3.2 | Notations

We used the following notations to formulate the problem.

Table 2. Model notations.

Notation	Description
Indices	
i, i'	Index of manufacturing machine or the component related to each order in the first stage, $i = 1,2,\dots, I$.
j, j'	Index of job(order) or manufactured product, $j = 1,2,\dots, J$
r	Index of position (rank) of orders in the sequence
l	Index of non-identical assembly machines in the second stage $l = 1,2,\dots, L$.
v	Index of vehicles
e, e', k	Index of nodes, 0 for origin node (production location-manufacturer) and one node for each customer, $e = \{0\} \cup \{1,\dots, K\} = 0,1,2,\dots,K$.
Parameters	
p_{ij}	Processing time of component i of order (job) j in the first stage.
$ST_{ij'ir}$	Setup time of machine i for job j after job j' at position r.
q_{jl}	Assembly time of order j on assembly machine l in the second stage.
ho_j	Cost of holding order j in warehouse after assembly.
av_{ev}	It is 1 if the node e can be serviced by the device v, otherwise 0.
Ca_v	Capacity of vehicle v.
pn_{jk}	Demand of customer k for order j.
$CO_{ee'v}$	Travel cost between node e and e' with vehicle v.
$ti_{vee'}$	Travel time between node e and e' with vehicle v.
S_k	Service time for customer k.
$[ed_k, ld_k]$	Service time window of customer k.
ep_k	Order delivery earliness penalty of customer k.
lp_k	Order delivery tardiness penalty of customer k.
Notation	**Description**
fc_v	Cost of using vehicle v
M	A large number
Decision Variables	
X_{ijr}	It is 1 if processing of job j starts on machine i at position r of the first stage, otherwise 0.
Z_{jrl}	It is 1 if the order j in the at position r of the first stage is assembled on the assembly machine l in the second stage, otherwise 0.
$W_{vee'}$	It is 1 if vehicle v travels to arc (e, e'), otherwise 0.
CT_{jr}	Completion time of job j on production machines at position r.
C_{ijr}	Completion time of the processing of the machine i of job j in the first stage on the production machines at position r.
$Start_{ijr}$	Start time of the processing of the component i of order j in the first stage on the production machines at position r.
C_{max}	Completion time of the assembly of last job in the second stage.
G_{rl}	Completion time of the assembly of order at position r on the assembly machine l in the second stage.
ET_{vk}	Earliness of vehicle v while arriving at customer k.
LT_{vk}	Tardiness of vehicle v while arriving at customer k.
ρ_{ve}	Arrival time of vehicle v to node e.

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

3.3 | Mathematical Model Formulation

According to the parameters and variables defined, a MILP model will be developed for the problem.

$$TC=\sum_{v=1}^{V}\sum_{e=0}^{K}\sum_{\substack{e'=0\\e\neq e'}}^{K}CO_{ee'v}\,W_{vee'}+\sum_{v=1}^{V}\sum_{e'=1}^{K}fc_v\,W_{v0e'}+\sum_{k=1}^{K}\sum_{v=1}^{V}ep_k\,ET_{vk}+\sum_{k=1}^{K}\sum_{v=1}^{V}lp_k\,LT_{vk}$$
$$+\sum_{j=1}^{n}\sum_{r=1}^{R}ho_j\left(Cmax-CT_{jr}-\sum_{l}Z_{jrl}q_{jl}\right). \tag{1}$$

$$\sum_{r}X_{ijr}=1 \ \text{ for all } \ i,j, \tag{2}$$

$$\sum_{j}X_{ijr}=1 \ \text{ for all } \ i,r, \tag{3}$$

$$X_{ijr}=X_{i-1jr} \ \text{ for all } \ i,j,r, \tag{4}$$

$$\sum_{l}Z_{jrl}=X_{ijr} \ \text{ for all } \ i,j,r, \tag{5}$$

$$Start_{ijr}\geq\sum_{j'\neq j}C_{ij'r-1}-M\left(1-X_{ijr}\right) \ \text{ for all } \ i,r, \tag{6}$$

$$C_{ijr}=Start_{ijr}+\sum_{j'}X_{ijr}X_{ij'r-1}ST_{jj'ir}+X_{ijr}*P_{ij} \ \text{ for all } i,j,r, \tag{7}$$

$$CT_{jr}=\max_i\left\{C_{ijr}\right\} \ \text{ for all } \ j,r, \tag{8}$$

$$CT_{jr}\leq MX_{ijr} \ \text{ for all } \ j,r, \tag{9}$$

$$G_{1l}=\sum_{j}Z_{j1l}\left(CT_{j1}+q_{jl}\right) \ \text{ for all } \ j,l, \tag{10}$$

$$G_{rl}=\max\left\{G_{(r-1)l},\sum_{j=1}^{J}Z_{jrl}CT_{jr}\right\}+\sum_{j}Z_{jrl}q_{jl} \ \text{ for all } r>1,l, \tag{11}$$

$$C_{max}\geq G_{rl} \ \text{ for all } \ r, \tag{12}$$

$$\sum_{v=1}^{V}\sum_{\substack{e=0\\e\neq k}}^{K}av_{ve}W_{vek}=1 \ \text{ for all } \ k, \tag{13}$$

$$\sum_{v=1}^{V}\sum_{e,k\in S}W_{vek}\leq|S|-1,S\subseteq\{1,2,\ldots,K\};2\leq|S|\leq K-1. \tag{14}$$

$$\sum_{\substack{e=0\\e\neq k}}^{K}W_{vek}-\sum_{\substack{e=0\\e\neq k}}^{K}W_{vke}=0 \ \ \forall v \ \&k \ \&av_{vk}=1, \tag{15}$$

$$\sum_{e}\left(\rho_{ve}+ti_{vek}\right)\leq\rho_{vk}+\sum_{e}\left(1-W_{vek}\right)M \ \text{ for all } \ v,k, \tag{16}$$

$$\rho_{v0}=C_{max} \ \text{ for all } \ v, \tag{17}$$

$$\sum_{k=1}^{K}W_{v0k}\leq 1 \ \text{ for all } \ v, \tag{18}$$

$$\sum_{e=0}^{K}\sum_{\substack{k=1\\e\neq k}}^{K}\sum_{j=1}^{J}pn_{jk}W_{vek}\leq Ca_v \ \text{ for all } \ v. \tag{19}$$

According to the parameters and variables defined, a MILP model will be developed for the problem.

$$ET_{vk}\geq ed_k\sum_{e=0}^{K}W_{vek}-\rho_{vk} \ \text{ for all } k\,\&v. \tag{20}$$

$$LT_{vk} \geq \rho_{vk} - ld_k \quad \text{for all } k \,\&\, v. \tag{21}$$

$$X_{ijr}, Z_{jrl}, W_{vee'} \text{ are Binary Variables} \quad \text{for all } i, j, r, l, e, v. \tag{22}$$

$$Cmax, CT_{jr}, Start_{ijr}, C_{ijr}, ET_{vk}, LT_{vk}, G_{rl}, \rho_{vk} \geq 0 \quad \text{for all } i, j, r, k, l, v. \tag{23}$$

The *Objective Function (1)* consists of five terms; the first and second terms of which are to minimize the total costs of routing and vehicle usage, respectively. The third and fourth terms calculate the total penalties for earliness and tardiness of the delivery time window, respectively. The fifth term minimizes the cost of holding orders in the warehouse of final products. *Constraints (2)* and *(3)* indicate that each job in each processing is processed only in a specific priority and vice versa; each position contains only 1 job. *Constraint (4)* states that the position of each component of a given job must be the same in all machines. In other words, a unique position is assigned to each job, which is the same in processing all of its components. Based on *Constraint (5)*, a job should only be in one position of the assembly machine sequence. In addition, the index of job j in the first stage (r) is transferred to the second stage according to this relation. *Constraint (6)* states that the start time of processing job j on machine i at position r is higher than the completion time of the job preceding job j on machine j (job at position r-1) if Xijr=1. In other words, the processing of job j on the machine i at position r can begin immediately after the job at position r-1 is completed. In *Constraint (7)*, the completion time of job j on machine i is calculated. Accordingly, the completion time of each job on each machine equals the sum of the start time, setup time and processing time. *Constraint (8)* indicates that the completion time of job j at the production stage is equal to the completion time of the last component of job j. Also, *Constraint (9)* states that the decision variable CTjr takes value only if Xijr=1. Based on *Constraint (10)*, the completion time of assembly of the order placed at the first position of the assembly machine is equal to the completion time of production its components plus its assembly time. Based on *Constraint (11)*, the completion time of the assembly of jobs at the second position and thereafter equals the maximum completion time of the previous job and the completion time of the production of the components of the job at that position plus the completion time of the job at the position r. *Constraint (12)* calculates the completion time of the assembly of the final order. Based on *Constraint (13)*, the customer e is served only by one of the authorized vehicles. Based on *Constraint (14)*, it is not possible to create a sub tour. Based on *Constraint (15)*, if arrival at the node of customer e takes place by vehicle v, the exit from it must take place only by vehicle v. *Constraint (16)* computes the arrival time of the vehicle at each node based on the time traveled by the vehicle and the origin node. The start time of the vehicle's transportation is equal to the completion time of the last order. This Constraint is shown in *Eq. (17)*. According to *Constraint (18)*, each vehicle can remain in the origin node or, if necessary, can depart from the origin node to at most one of the customer nodes. The capacity constraint of each vehicle is shown in *Constraint 19*. *Constraints (20)* and *(21)* calculate the earliness and tardiness of each vehicle, respectively. *Constraints (22)* and *(23)* represent the type of decision variables.

3.4 | Linearization

Given the nonlinearity of the *Constraints (7)*, *(8)* and *(11)*, the following describes how to linearize these Constraints and consider them in the model.

The following variables are used to linearize nonlinear constraints:

$xx_{ijj'r}$: Binary variable for linearizing *Constraint (7)*.

a_{ijr} : Binary variable for linearizing *Constraint (8)*.

β_{jr} : The variable required to linearize *Constraint (11)*.

The non-linearizing factor in *Constraint (7)* is $xx_{ijj'r} = X_{ijr} X_{ij'r-1}$. *Constraint (7)* is replaced by the following relations.

$$xx_{ijj'r} \le X_{ijr} \quad \text{for all } i,j,j',r. \tag{24}$$

$$xx_{ijj'r} \le X_{ij'r-1} \quad \text{for all } i\ ,j,j',r. \tag{25}$$

$$xx_{ijj'r} \ge X_{ijr} + X_{ij'r-1} - 1 \quad \text{for all } i,j,j',r. \tag{26}$$

$$C_{ijr} = Start_{ijr} + \sum_{j'} xx_{ijj'r} ST_{jj'ir} + X_{ijr} * P_{ij} \quad \text{for all } i,j,r. \tag{27}$$

According to the relation $CT_{jr} = \max_i \{C_{ijr}\}$, *Constraint (8)* is a nonlinear equation. For linearization, this relation is replaced by the following Constraints;

$$CT_{jr} - C_{ijr} \ge -M(1-\alpha_{ijr}) \quad \text{for all } i,j,r. \tag{28}$$

$$CT_{jr} - C_{ijr} \le \ M(1-\alpha_{ijr}) \quad \text{for all } i,j,r. \tag{29}$$

$$C_{i'jr} C_{ijr} + M(1-\alpha_{ijr}) \quad \text{for all } j,i,i',r; i \ne i'. \tag{30}$$

$$\sum_{i}^{I} = \alpha_{ijr} = 1 \quad \text{for all } j,r. \tag{31}$$

The non-linearizing factor of the *Constraint (11)* is $\beta_{jrl} = Z_{jrl} CT_{jr}$. They can be linearized by considering the following relations:

$$\beta_{jrl} \le Z_{jrl}.M \quad \text{for all } j\ ,r,l, \tag{32}$$

$$\beta_{jrl} \le CT_{jr} \quad \text{for all } j\ ,r,l, \tag{33}$$

$$\beta_{jrl} \ge CT_{jr} - (1-Z_{jrl}).M \quad \text{for all } j\ ,r,l. \tag{34}$$

3.5 | Non-Integrated Model

This section presents a non-integrated model. In this approach, first the mathematical model of production and assembly (first stage) is optimized. Then the optimal solution, as a parameter, is transferred to the mathematical model of the distribution (second stage). By optimizing the second stage model, the best distribution of orders among customers is determined. The first stage mathematical model consists of the following relations:

$$\text{Min } Z1 = \sum_{j=1}^{n} \sum_{r=1}^{R} ho_j \left(Cmax - CT_{jr} - \sum_{l} Z_{jrl} q_{jl} \right). \tag{35}$$

Constraints (2)-(6), Constraints (9) to *(10), Constraint (12),* and *Constraints (24)-(34),* which $CTjr$ and $Zjrl$ are decision variables as described above.

After optimizing the first stage model, the $Cmax$ value is transferred to the second stage model as a parameter. This model includes the following constraints:

$$\text{Min } Z2 = \sum_{v=1}^{V} \sum_{e=0}^{K} \sum_{\substack{e'=0 \\ e \ne e'}}^{K} CO_{ee'} W_{vee'} + \sum_{v=1}^{V} \sum_{e'=1}^{K} fc_v W_{v0e'} + \sum_{k=1}^{K} \sum_{v=1}^{V} ep_k ET_{vk} + \sum_{k=1}^{K} \sum_{v=1}^{V} lp_k LT_{vk}. \tag{36}$$

Constraints (13)-(21), which $Wvee', Wvoe'$ and ET_{vk} are decision variables as described above.

3.6 | Solution Methods

In this section, the proposed solutions method which are WOA, improved WOA, and GA are presented.

3.7 | Whale Optimization Algorithm (WOA)

Whales are cetaceans with a long tail. The interesting thing about whale life, which is the inspiration for this algorithm, is the way of feeding and hunting in humpback whales, known as bubble netting. In this way, each whale releases air bubbles beneath the sea and creates walls of rising air in the water. The fishes that are inside the air wall move toward the center of the circular bubble territory due to fear, and immediately the whale swallows many of them while rising from the water. This algorithm was proposed by Mirjalili and Lewis [39]. According to the WOA, humpback whales are able to detect and surround the prey's position. Because the optimal position in the search space is unclear, the WOA assumes that the best available solution is the target prey or a point near to it. Once this point is determined, the search for other optimal points and the updating of position is continued. This behavior is represented by the following equations.

$$D = | c.X^*(t) - X(t) | . \tag{37}$$

$$(t+1) = X^*(t) - A.D. \tag{38}$$

In the above equations, t represents the iteration of the algorithm, C and A the coefficients vectors, $X*$ the best-obtained position, and X the current Whale position. It should be noted that the value of $X*$ is updated in each iteration. The following equations are used to determine A and C values:

$$A = 2a.r - a. \tag{39}$$

$$C = 2.r. \tag{40}$$

Where, a is a vector that controls the variation of the solutions and its initial value is 2 that may reduce to 0 in different iterations. r is also a random vector ranging from 0 to 1. To implement the WOA, it is necessary to define the position of the whales based on a justified solution to the problem. In this study, the position of each whale is defined as a set consisting of 4 different vectors. These vectors are defined as follows. The first vector is a vector containing J cells that shows the sequence of jobs. All these cells take numbers ranging from 0 to 1. Ordering the numbers from largest to smallest can determine the sequence of jobs. The second vector is a vector of J cells ranging from 0 to 1. This vector shows the assignment of jobs to assembly machines. To describe it more clearly, suppose there are three assembly machines. Cells with values ranging from 0 to 0.334 (1 divided by 3) are assigned to assembly machine 1. Cells ranging from 0.334 to 0.664 (2 divided by 3) are assigned to assembly machine 2, and finally those ranging from 0.664 to 1 are assigned to assembly machine 3. The third vector is a vector with E cells ranging from 0 to 1. Each cell, depending on its value, represents the assignment of the customer to a vehicle. The way customers are assigned to machines is exactly in accordance with the procedure described for assigning jobs to machines. For example, if there are two machines, cells with a value of less than 0.5 are assigned to machine 1, and cells with a value greater than 0.5 are assigned to machine 2. The fourth vector has a length of E that shows the priority of customers in visits. Customer priority is determined by ordering the numbers of this vector from largest to smallest. For example, if the number of orders is 3, the number of assembly machines is 2, the number of vehicles is 2, and the number of customers is 3, four vectors relating to a hypothetical solution are according to *Table 3*.

Table 3. An example of a justified solution in the WOA.

Vector 1	0.26	0.48	0.64
Vector 2	0.35	0.27	0.95
Vector3	0.17	0.84	0.35
Vector 4	0.53	0.37	0.18

According to the first row, the sequence of jobs on different production machines is 3-2-1. In the assembly stage, according to the second row, orders 1 and 2 are assigned to assembly machine 1 and order 3 to assembly machine 2. In the distribution stage, according to the third row, customers 1 and 3 are assigned to the first machine and customer 2 to the second machine. To determine the order of customer visits by vehicle 1 (visit order of customers 1 and 3), see the fourth row. According to this row, the number related to customer 1 is greater than that related to customer 3 ($0.18 < 0.53$), so customer 1 is first visited and then customer 3 is visited. For vehicle 2, because it visits only one customer, the customer priority is not applicable.

3.8 | Improved Whale Optimization Algorithm

Despite the suitability of the WOA search algorithm, this algorithm converges quickly in some cases. Also, this algorithm greedily seeks to improve the existing set of solutions, which causes the algorithm to trap in local optima. Due to the problems mentioned for the WOA search algorithm, the necessary changes will be made in the following to develop this algorithm and increase its efficiency. This development is inspired by the study of Alinaghian and Goli [40].

I. In the step of generating a new solution in the WOA algorithm, each cell of the solution vector is randomly generated from the values of the best existing solution, according to *Eqs. (37)* and *(38)*. But in the new method, first a vector is generated from the combination of all existing solutions and then new solutions will be generated with its help and based on a random process.

II. The procedure for setting parameter a exits from the uniform mode-from 2 to 0-and will be set in such a way that the components of the new solution move towards the best possible solution.

III. In order to update the set of solutions, first the answers are sorted according to the value of the objective function. Then the solutions that have the following two characteristics are removed from the set of solutions: 1) have a weak objective function, 2) have similar solutions in the set of solutions.

In order to better understand the developed meta-heuristic algorithm, the IWOA algorithm is described step by step considering these changes.

Step 1. Determining the problem parameters and the algorithm.

Step 2. Generating random initial values for the set of problem solutions and calculating the value of the objective function of each of them (F(xj)).

Step 3. Generating new solutions:

$$FDR_i = \max_{j,l}\{\frac{f(x_i)-f(x_j)}{|x_{jl}-x_{il}|}\}. \tag{41}$$

$$D = \left|c.X_{i^*}(t) - X(t)\right|. \tag{42}$$

$$X(t+1) = X_{i^*}(t) - A.D. \tag{43}$$

In this relation, $f(x_i)$ and x_{il} are respectively the value of the objective function and the value of the l^{th} component of the new solution, also $f(x_j)$ and x_{jl} are respectively the value of the objective function and the value of the l^{th} component of the j^{th} existing solution:

$$x' = f(worst) - f(best). \tag{44}$$

$$p = \frac{1}{\sqrt{2\pi}} e^{-\frac{x'^2}{2}}. \tag{45}$$

$$a' = 0.1 \times (1-p). \tag{46}$$

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

In the above equations, f (best) and f (worst) are the values of the objective function and the worst solution in the set of existing solutions, respectively. Finally, parameter a’ is placed in equations 39 and 40 to complete the process of generating new solutions. *Fig. 2* shows the flowchart of the IWOA algorithm. Moreover, the pseudocode of IWOA is illustrated as *Algorithm 1*.

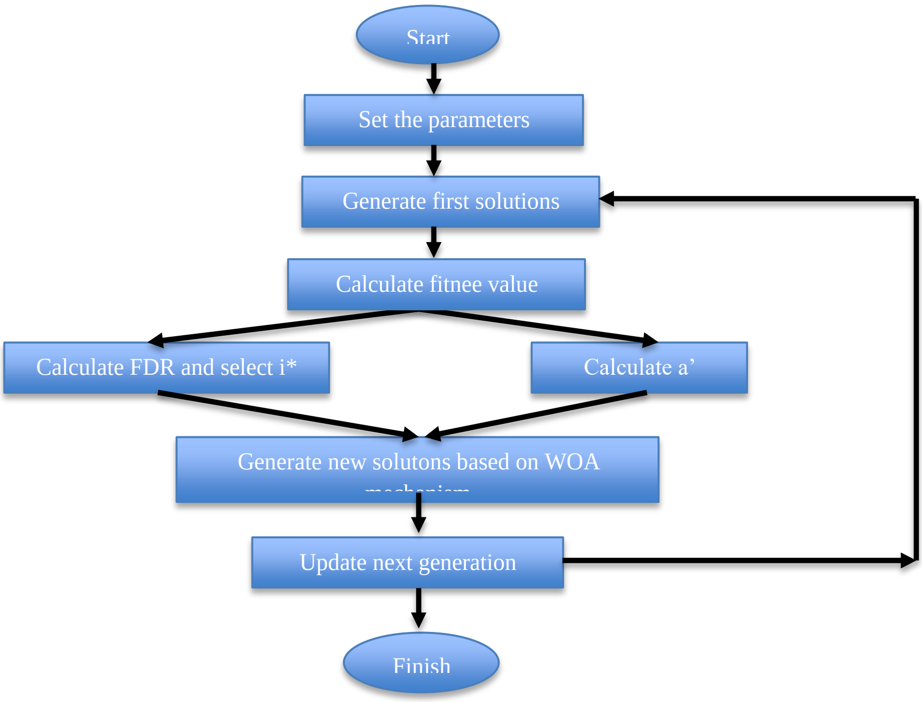


Fig. 2. The flowchart of proposed IWOA.

Algorithm. 1. The pseudocode of proposed IWOA.

```
Randomly initialize tha whale population (X)
Evaluate the fitness value of each whale
Specify the best search agent X*
While t<Max_iter
Calculate a’
Calculate FDR and select i*
For each search agent
Generate new solution using Eqs. (37)-(40) and save it as X(t+1).
End for
Evaluate X(t+1)
Update X*
End while
Report X*
```

3.9 | Genetic Algorithm

The most important characteristic of Genetic Algorithm (GA) is simplicity. The steps of the GA are illustrated in Algorithm 3. First, the solution to the problem is defined in the form of a chromosomal structure (coding). By introducing the fitness function, the quality of solutions in each chromosome is expressed as a number. Then, a specific number of chromosomes is generated randomly (or quasi-randomly). These chromosomes are known as the primary population and are evaluated based on the fitness function. Now, two chromosomes are selected for reproduction and using these two chromosomes, a new chromosome is generated (mating). With a specific probability, a number of genes of some chromosomes are changed. Performing the selection, mating and mutation steps creates a new population (generation) of chromosomes. If the chromosomes converge to the optimal response, the reproduction operation will be stopped. Otherwise, each generation will be produced from the previous generation until the desired solution is obtained or the stop criterion of the algorithm is applied [31], [41], [42].

First, a coding system should be defined. This coding system is called the chromosome. The chromosome used in this study is a chromosome with real numbers between 0 and 1. The structure of this chromosome is exactly the same as that proposed for the WOA.

Next, it is necessary to create the initial population. This initial population is randomly generated in the range between 0 and 1. Then the fitness value of each solution is calculated. The fitness value exactly equals the total cost of production, assembly, and distribution. Then, a set of solutions is selected as the parent using the roulette wheel method. Of the parents, two new solutions (children) are generated from each two parents (P1 and P2) using linear combination. For this purpose, one parameter α is randomly generated in the range of [-sigma, + sigma] (the sigma is the control parameter that should be set). Then, new solutions (SP1 and SP2) are generated by the *Eqs. (47)* to *(48).*

$$SP1 = \alpha P1 + (1-\alpha) P2. \tag{47}$$

$$SP2 = \alpha P2 + (1-\alpha) P1. \tag{48}$$

In such circumstances, it can be assured that the solutions are the perfect combination of the parent's solutions. Finally, it is checked that the cells of SP1 and SP2 take a value between 0 and 1. If a cell takes a value less than 0, it is changed to 0. If a cell takes a value greater than 1, it is changed to 1. This operation is performed with the Probability of Crossover (PC) in each iteration. Then the mutation is performed. Therefore, a cell is selected from a chromosome, and then its value is replaced by a random value. This operation is performed with the probability of mutation (pm) in each iteration. In the next stage, a number of solutions will be selected equal to the number of Pop size from the set of available solutions (parents, Crossover solutions, mutation solutions) and included in the next iteration. The flowchart of proposed GA is as *Fig. 3*.

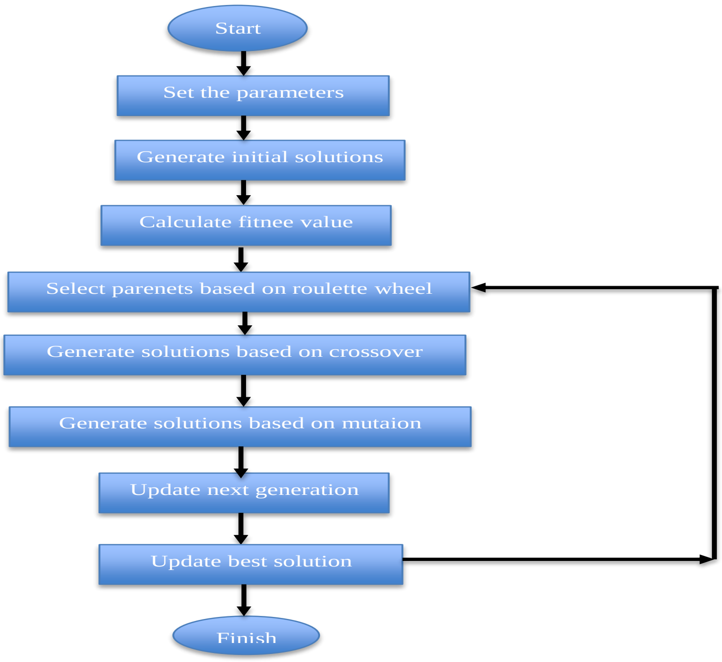


Fig. 3. The general structure of the GA.

3.10 | Computational Results

The aim of this section is fourfold: 1) to validate small-scale mathematical models, 2) to investigate the efficiency of proposed solution methods, 3) to present a case study, and 4) to conduct and report a sensitivity analysis.

3.11 | Validation of Mathematical Models

To validate the integrated mathematical model and the meta-heuristic algorithms, a small-scale numerical example is implemented in this section and the results are presented. The data in this numerical example has been designed in a way that the optimal solution can be clearly understood. The goal is to produce three types of gearboxes, shown by the symbols A, B, and C. The factory produces and sends for two customers in cities C1 and C2. *Table 4* shows data on customer demand, production and assembly times, and holding costs for each product. It should be noted that each time unit is considered to be 5 minutes.

Table 4. Data on processing and assembly times and holding cost of final products.

Job	A	B	C
The demand for customer 1 for each products	2	7	0
The demand of customer 2 for each products	0	1	4
Processing time of component 1 on machine 1 in the first stage	1	1	1
Processing time of component 2 on machine 2 in the first stage	2	2	2
Assembly time on the first assembly machine in the second stage	3	3.2	2.2
Assembly time on the second assembly machine in the second stage	2.7	2.6	1.4
The holding cost of each unit of product in the warehouse	1	1	1

In the production stage, there is a setup time for the production of each component of the different gearboxes. Sequence-dependent setup time for all jobs is defined as one unit of time. This company uses one type of vehicle to deliver products to customers. This vehicle has a capacity of 100 units and a fixed cost of 50 units. The variable costs of applying vehicles are also shown in *Table 5*. The data in *Table 5* shows the distance between the factory and customer location. In this case, 1 monetary unit of cost is considered for each unit of travel time.

Table 5. Data on travel time between different nodes.

	Factory	E1	E2
Factory	0	40	85
E1	40	0	133
E2	85	133	0

In addition, the delivery time windows and other data on customer delivery are shown in *Table 6*.

Table 6. Data on delivery time window, service time, and earliness and tardiness penalties of customer delivery.

Customer	E1	E2
Delivery time window	[50 , 60]	[20 , 60]
Service time	1	0.6
Earliness penalty	5	5
Tardiness penalty	15	17

We used a PC with Core I5 CPU processor under the windows 8.1 operating system with 4GB of RAM. The mathematical MIP model of the problem was implemented in GAMS software and solved with CPLEX 24.0.1 solver. The proposed algorithms were coded by MATLAB 2017 R1 software. The above data was entered the GAMS optimization software as the parameters of the integrated mathematical model, and then the model was optimized. The optimal solution to the problem is equal to 2468.1 monetary units. The Cmax value is also 9.4. In the production stage, the optimal sequence of jobs is defined as 1-2-3. *Fig. 4* shows the scheduling of jobs on different machines.

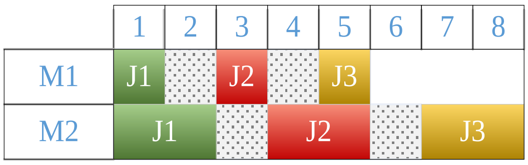


Fig. 4. Scheduling of jobs in the production stage in the integrated model.

In the assembly stage and in the optimal solution, job 2 is assigned to assembly machine 1, and jobs 1 and 3 to the assembly machine 2. *Fig. 5* illustrates the scheduling of these jobs during the assembly stage.

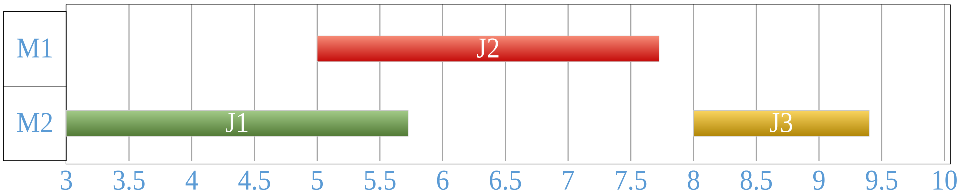


Fig. 5. Assembly scheduling in integrated model.

As *Fig. 3* illustrates, each job is assembled immediately after its production is completed, which confirms the accuracy of results. According to manual calculations, the Cmax value is 9.4, which confirms the validity of the results. In the following, the routing of vehicles between customers is specified. *Fig. 6* illustrates the created tour.

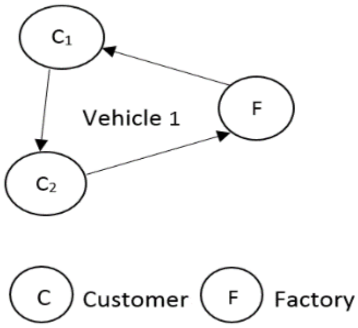


Fig. 6. The optimal solution of product distribution in the integrated model.

After comparing the arrival time of vehicle to each customer with its time window, it is revealed that the earliness and Tardiness delivery of orders to customer 1 and 2 are 0.6 and 122.4, respectively.

Data analysis shows the accuracy of the outputs, and the study of similar solutions shows that the lowest possible cost is 2468.1. Therefore, the mathematical model and its results can be considered as reliable. In the next stage, it will be necessary to investigate the validity of the WOA and GA. So, the problem defined is optimized by each of WOA, IWOA and GA in this section. *Fig. 7* illustrates the convergence of the WOA. Also *Fig. 8* illustrates the convergence of IWOA and finally *Fig. 9* shows the convergence of GA.

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

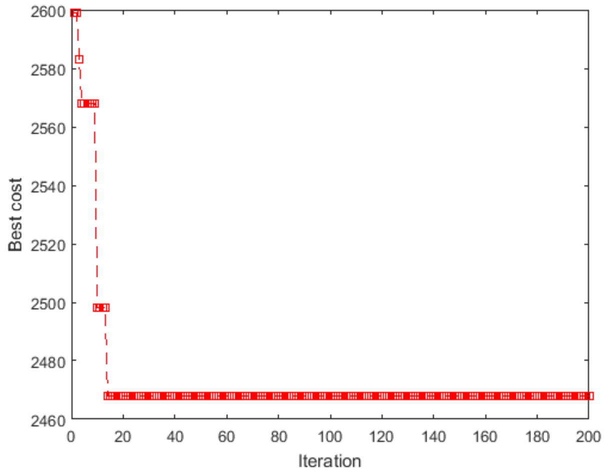


Fig. 7. The convergence of WOA.

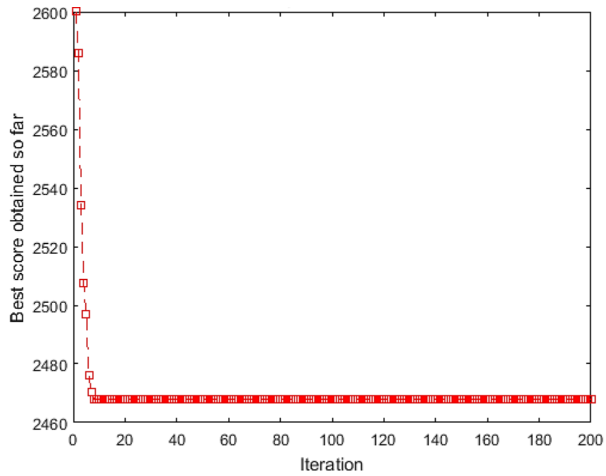


Fig. 8. The convergence of IWOA.

As illustrated in *Fig. 8*, WOA obtained the objective function of 2600 in its first iteration. After various iterations, this value reached 2468.1 in iteration 13. Moreover, in *Fig. 9*, IWOA is converged to the optimal solution after 8 iterations. This suggests that WOA and IWOA can converge to the optimal solution of the problem after a small number of iterations, and therefore its results will be valid. In addition, the proposed improvement for WOA leads to speed up the convergence in about 60%.

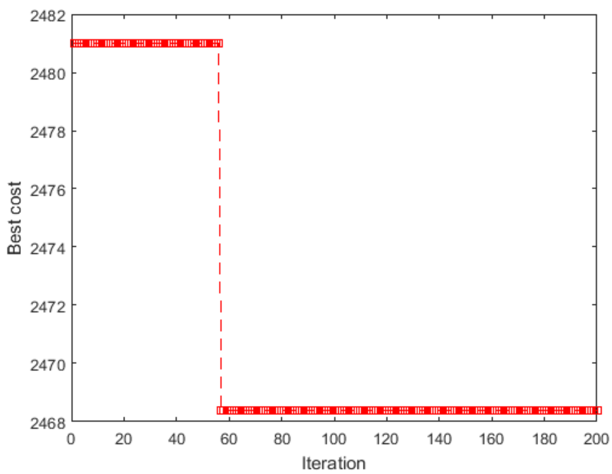


Fig. 9. The convergence of GA.

As can be seen in *Fig. 9*, the GA was able to converge to the overall optimal solution of the problem. Therefore, the results of this algorithm will be valid. The difference between the WOA, IWOA and GA is that the WOA and IWOA converged to the optimal solution after 13 and 8 iterations respectively, but the GA reached this convergence after 57 iterations. Therefore, the IWOA was strongest in this regard. While the WOA and IWOA started at the range of 2600, the GA started at the range of 2481. Therefore, the convergence rate of the WOA and IWOA was higher. In other words, the improvement that the WOA and IWOA is equal to 2600-2468.1 = 118.9. By dividing this phrase by 13 iterations, the convergence rate of the WOA is obtained as 9.1 and dividing by 8 iterations, the convergence of IWOA is obtained as 14.8. In other words, an average cost reduction of 14.8 units occurred per iteration before convergence in IWOA. This index is equal to 2481-2468.1 = 12.9 for the GA and the convergence rate is 0.22. In other words, the GA produced an average improvement of 0.22 units per iteration before convergence.

3.12 | Investigating the Efficiency of Proposed Solution Methods

In this section of the computational results, the performance of WOA, IWOA and GA will be investigated. In this regard, initially 40 instances in small, medium and large scales are generated. Data on these problems is shown in *Table 7*. The values of the parameters for each of the instances are randomly generated from continuous uniform distribution. The lower limit and the upper limit of each parameter are shown in *Table 8*.

These problems were optimized in the GAMS environment for exact solution optimization, and in the MATLAB environment with WOA, IWOA and GA. The time limit for optimization was considered as 3600 seconds. The results are presented in *Table 9*. In this table, Z represents the value of the objective function, T (in seconds) is the solution time, and Gap (%) is the relative error compared to exact method (GAMS).

Table 7. Scale of generated instances.

Size	Samples	I	J	E	V	L
Small scale	P1	2	3	3	1	2
	P2	3	3	3	1	2
	P3	3	4	3	1	2
	P4	3	4	4	1	2
	P5	3	4	5	2	3
Median scale	P6	5	5	5	3	3
	P7	4	5	5	2	3
	P8	5	6	6	2	3
	P9	5	7	7	3	4
	P10	6	8	9	3	3
Large scale	P11	10	10	10	4	5
	P12	12	15	11	5	6
	P13	14	17	15	6	7
	P14	16	20	17	7	9
	P15	18	23	19	9	10
	P16	20	25	20	10	11
	P17	25	27	25	13	13
	P18	30	30	27	15	14
	P19	40	40	35	20	20
	P20	50	50	40	30	25
	P21	60	60	50	30	30
	P22	70	70	60	40	35
	P23	80	80	70	40	40
	P24	90	90	80	50	45
	P26	100	100	90	50	50
	P26	110	100	100	60	55
	P27	120	120	110	60	60
	P28	130	120	120	70	65
	P29	140	140	130	70	70
	P30	150	140	140	80	75
	P31	160	160	150	80	80
	P32	170	160	160	90	85
	P33	180	180	170	90	90
	P34	190	180	180	100	95
	P35	200	200	190	100	100
	P36	210	200	200	110	105

Table 7. Continued.

Size	Samples	I	J	E	V	L
	P37	220	210	210	110	110
	P38	230	210	220	120	115
	P39	240	220	230	120	120
	P40	250	220	240	130	125

Table 8. Upper and lower limits of parameter values.

Parameters	Lower Bound	Upper Bound
Transportation time	20	140
Service time	0.5	1
Assembly time	1.4	3.4
Demand	0	5
Production time	1	4
Holding cost	1	4
Earliness penalty	3	5
Tardiness penalty	10	20
The cost of travel	30	135
Vehicle fixed cost	50	60
Vehicle capacity	1000	1200
Lower limit of the time window	20	90
Upper limit of the time window	90	140
Setup time	1	2

Table 9. Results of optimization of various problems.

Samples	GAMS		WOA			IWOA			GA		
	Z	T	Z	T	Gap	Z	T	Gap	Z	T	Gap
P1	2468.1	1.02	2468.1	0.69	0.00%	2468.10	0.74	0.00%	2468.1	3.43	0.00%
P2	2527.6	2.83	2527.6	4.88	0.00%	2527.60	4.88	0.00%	2527.6	6.06	0.00%
P3	2768.6	6.47	2811.6	5.11	1.53%	2811.60	5.53	1.53%	2847.8	7.34	2.78%
P4	1802.8	20.52	1867.8	5.31	3.48%	1860.20	5.82	3.09%	1956.7	7.53	7.87%
P5	3176.2	176.41	3257.7	5.59	2.50%	3203.40	5.77	0.85%	3261.6	7.81	2.62%
P6	2469.6	870.1	2505.7	6.67	1.44%	2493.10	7.31	0.94%	2539.1	8.82	2.74%
P7	2494.2	3600	2594.6	7.11	3.87%	2590.90	7.16	3.73%	2666.1	8.91	6.45%
P8	-	-	5435.2	6.77	0.05%	5432.48	6.86	0.00%	5450	8.44	0.32%
P9	-	-	5925.9	7.68	0.24%	5911.71	8.12	0.00%	5991.4	9.15	1.33%
P10	-	-	10449.3	8.24	0.30%	10417.97	8.38	0.00%	10572.8	9.94	1.46%
P11	-	-	15559.4	10.23	0.66%	15456.43	11.17	0.00%	16330.5	12.02	5.35%
P12	-	-	40140.9	13.31	0.83%	39809.42	15.73	0.00%	41197.2	15.72	3.37%
P13	-	-	63812.5	15.47	3.40%	63646.83	17.77	3.14%	61645.4	18.37	0.00%
P14	-	-	112338.9	18.3	3.02%	111234.88	18.71	2.05%	108951.5	21.94	0.00%
P15	-	-	160383.9	21.86	0.40%	159745.69	23.37	0.00%	167182	25.47	4.45%
P16	-	-	203856.4	24.36	0.52%	202792.18	28.22	0.00%	206792.5	28.74	1.93%
P17	-	-	224476.3	29.36	0.81%	222647.77	30.25	0.00%	225363.8	34.39	1.21%
P18	-	-	308709	32.02	0.93%	305827.44	37.96	0.00%	310627.1	38.1	1.55%
P19	-	-	707211.9	47.04	0.99%	700228.22	50.46	0.00%	722761.3	56.14	3.12%
P20	-	-	1393914.8	69.07	0.71%	1384042.51	79.83	0.00%	1441985.4	78.4	4.02%
P21	-	-	1474033.6	73.96109	1.95%	1473993.55	86.22	1.95%	1445261.2	85.58	0.00%
P22	-	-	1514966.5	77.51197	0.78%	1507566.97	90.98	0.30%	1503110.9	100.37	0.00%
P23	-	-	1586944.4	83.80926	0.00%	1586944.45	85.67	0.00%	1690395.3	104.30	6.12%
P24	-	-	1648002.2	89.64728	0.43%	1640941.26	95.57	0.00%	1879600.5	104.74	12.70%
P25	-	-	1759013.3	90.45446	0.00%	1759013.34	96.16	0.00%	2097065.9	121.79	16.12%
P26	-	-	1859640.8	97.89633	0.00%	1859640.77	99.07	0.00%	2334516.7	132.68	20.34%
P27	-	-	1898552.4	103.3518	0.00%	1898552.37	118.97	0.00%	2446831.0	149.12	22.41%
P28	-	-	1958180.3	108.4347	0.82%	1942096.92	114.81	0.00%	2449723.4	161.90	20.72%
P29	-	-	2068810.9	123.0877	0.00%	2069128.64	140.90	0.02%	2657742.2	191.73	22.16%
P30	-	-	2084668.5	131.1225	0.00%	2097761.34	142.28	0.62%	2834937.8	207.59	26.47%
P31	-	-	2120083.1	143.3223	0.00%	2130588.50	143.42	0.49%	2908072.7	230.02	27.10%
P32	-	-	2271524.2	153.0421	0.36%	2263271.44	180.52	0.00%	3178261.6	232.05	28.79%

Bahmani et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

Table 9. Continued.

Samples	GAMS Z	GAMS T	WOA Z	WOA T	WOA Gap	IWOA Z	IWOA T	IWOA Gap	GA Z	GA T	GA Gap
P33	-	-	2280345.1	171.9941	0.29%	2273815.41	188.03	0.00%	3563090.8	262.49	36.18%
P34	-	-	2294721.4	206.2778	0.06%	2293410.52	234.82	0.00%	3722680.6	290.42	38.39%
P35	-	-	2432103.7	214.6102	0.00%	2442126.30	246.32	0.41%	4022717.4	338.17	39.54%
P36	-	-	2538269.2	220.4266	0.25%	2531966.67	252.29	0.00%	4112688.6	359.08	38.44%
P37	-	-	2548623.6	251.7006	0.54%	2534978.60	266.94	0.00%	4621513.2	384.91	45.15%
P38	-	-	2696561.3	294.6928	0.86%	2673321.19	328.58	0.00%	4715799.6	411.56	43.31%
P39	-	-	2941324.5	318.2205	0.02%	2940708.34	367.36	0.00%	5162264.5	465.35	43.03%
P40	-	-	2990058.0	329.2949	0.61%	2971686.19	384.68	0.00%	5183784.6	541.62	42.67%
Average			1155917	90.54822	0.82%	1153416.531	100.941	0.48%	1646829.41	132.055	14.50%

As seen, GAMS was only able to optimize 7 problems. Moreover, the average WOA solution time is 90.54 seconds and its average error is 0.82%. These indices are 100.94 and 0.48% in the IWOA. The first comparisons demonstrate that the proposed improvement leads to find better solutions with lower costs. The average GA solution time is 132.05 seconds and its average error is 14.23%. Therefore, the IWOA performed better in terms of both solution time and error in comparison to GA too. It should be noted that the gap% obtained is based on the comparison of each approach with the exact solution method (GAMS).

Fig. 10 illustrates the comparison of solution times of GAMS and meta-heuristic methods for seven GAMS-solvable problems. *Fig. 11* illustrates the comparison of solution times of meta-heuristic methods for the 40 solved problems, and *Fig. 12* shows the comparison of the values of the objective functions for the two meta-heuristic methods.

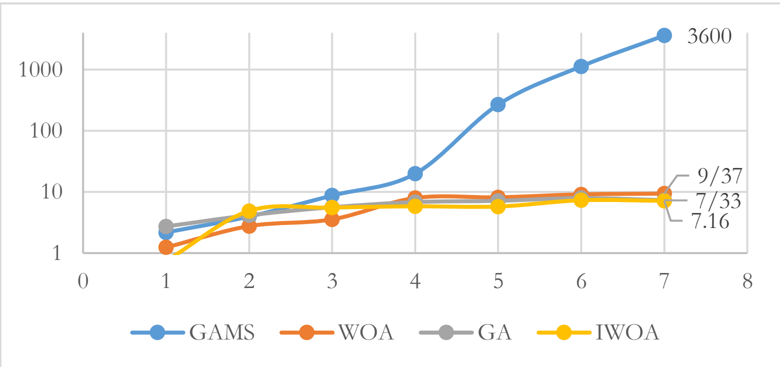


Fig. 10. Comparison of solution times of GAMS and meta-heuristic algorithms in small scale.

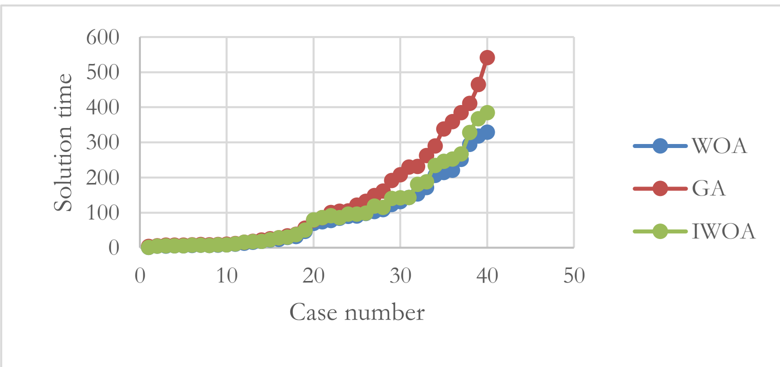


Fig. 11. Comparison of solution times between meta-heuristic algorithms in a test problem.

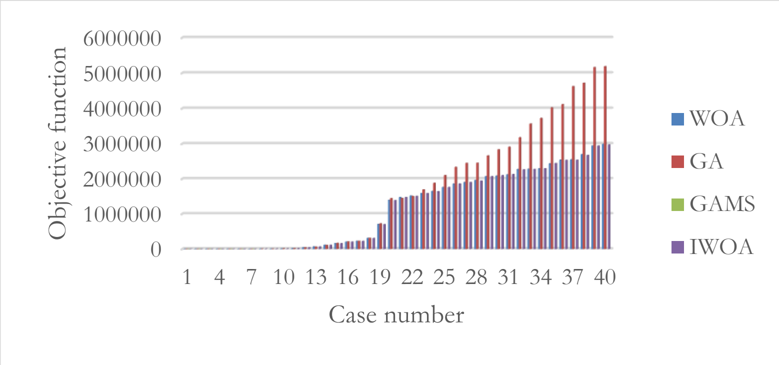


Fig. 12. Comparison of the value of the objective function.

As can be seen, for all solved instances, GA's solution time was longer than the WOA and IWOA. Moreover, for solving large-scale problems, the GA exhibited better performance than WOA only in problems 13 and 14, and in other problems, WOA provided a better objective function. On the other hand, IWOA spend more time than WOA but its convergence faster and achieve more quality solutions. Therefore, the efficiency of IWOA compared to WOA and GA could be well explained.

4 | Case Study

To conduct validation, the integrated and non-integrated mathematical models presented were implemented in an industrial gearbox manufacturing factory in Iran. The results will be presented in this section. This industrial gearbox factory is located in Ar city and produces 7 types of gearboxes. These gearboxes are shown with the symbols A, B, C, D, E, F and G. The factory produces and sends products to six customers in the cities of As, Ta, Kh, Sh, Sm and Tr. *Table 10* presents customers demand data, production and assembly times, and holding costs for each product. It should be noted that each unit of time is considered to be 5 minutes.

In the production stage, there is a setup time for the production of each component of the different gearboxes. *Table 11* shows the setup time of each job on each machine. According to this table, the setup time of each job is determined dependent on the previous job in the sequence. For example, if job A is processed immediately after job B, the setup time for job A is two time units. These times are equal for all machines.

Table 10. Data on processing and assembly times, and holding costs of final products.

Job	A	B	C	D	E	F	G
Demand of customer 1 of products	2	7	0	3	1	2	2
Demand of customer 2 of products	0	1	4	2	1	1	1
Demand of customer 3 of products	4	0	2	3	2	2	2
Demand of customer 4 of products	5	2	2	0	3	3	3
Demand of customer 5 of products	3	1	3	2	3	1	1
Demand of customer 6 of products	2	3	1	2	2	1	1
Processing time of component 1 on machine 1 in the first stage	2	3	2	3	2	1	1
Processing time of component 2 on machine 2 in the first stage	1	1	2	2	4	2	2
Processing time of component 3 on machine 3 in the first stage	2	3	3	3	2	2	2
Processing time of component 4 on machine 4 in the first stage	1	3	3	2	4	3	3
Processing time of component 5 on machine 5 in the first stage	3	3	2	2	3	1	1
Processing time of component 6 on machine 6 in the first stage	3	3	3	1	1	2	3
Processing time of component 7 on machine 7 in the first stage	4	1	2	3	1	2	2
Assembly time on the assembly machine in the second stage	3	3	2	3	3	1	1
The holding cost of each unit of product in the warehouse	10	10	11	9	10	10	11

This company uses 3 types of vehicles to send products to customers. The capacity and fixed cost of using these vehicles are shown in *Table 12*. The variable costs of using vehicles are according to *Table 13*. The data in *Table 14* show the distance between the factory and the customer locations. In this case,

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

for each unit of time (5 minutes), a distance of 5 km is travelled. One monetary unit of cost is considered per one monetary unit.

Table 11. Data on sequence-dependent setup time.

	A	B	C	D	E	F	G
A	1	2	1	2	2	1	1
B	1	2	1	2	1	1	1
C	1	1	1	2	2	1	1
D	2	1	1	1	2	1	1
E	1	1	2	1	1	1	2
F	1	1	1	1	1	1	2
G	1	2	2	1	1	1	2

Table 12. Data on capacity and fixed cost of vehicles.

Vehicle Type	1	2	3
Capacity	100	150	320
Fixed Cost	500	600	1000

Table 13. Data on travel time between different nodes.

	Ar	As	Ta	Kh	Sh	Sm	Tr
Ar	0	40	85	65	33	16	20
As	40	0	133	114	7	30	15
Ta	85	133	0	165	121	45	90
Kh	65	114	165	0	105	40	25
Sh	33	7	121	105	0	30	60
Sm	90	30	20	80	30	0	90
Tr	35	20	60	55	33	20	0

The above data was included in the GAMS optimization software as the parameters of the integrated mathematical model and the model was optimized. The optimal solution of the problem was obtained as 22462 monetary units. Moreover, the Cmax value was obtained as 28. In the production stage, the optimal sequence of jobs was determined as 6-5-2-7-3-4-1. *Fig. 13* illustrates the scheduling of jobs on different machines.

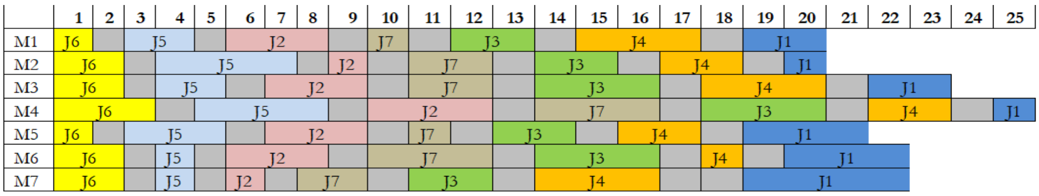


Fig. 13. Scheduling jobs in the production stage of the integrated model.

Based on the obtained results in *Fig. 13* in the optimal solution, different machines have idle times. The reason for this is that the integrated model seeks to complete the processing of the various components. The important point in *Fig. 13* is that the sequence of jobs is defined in a way that all setup times are equal to its minimum value *Fig. 14* illustrates the scheduling of these jobs during the assembly stage.

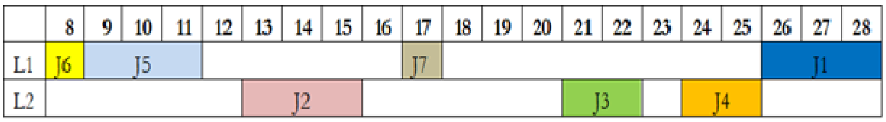


Fig. 14. Assembly scheduling in integrated model.

As illustrated in *Fig. 14*, each job is placed in the assembly stage immediately after production, which confirms the accuracy of the results. In addition, based on manual calculations, the Cmax value is equal to 28, which shows the validity of the results. In the following, the routing of vehicles between customers will be determined. For this purpose, vehicles 1 and 3 were used. The tours of each vehicle are illustrated in

Fig. 15. Based on the results, it is clear that all customers were visited once and each vehicle formed a tour which started and finished in the factory. This also confirms the accuracy of the results.

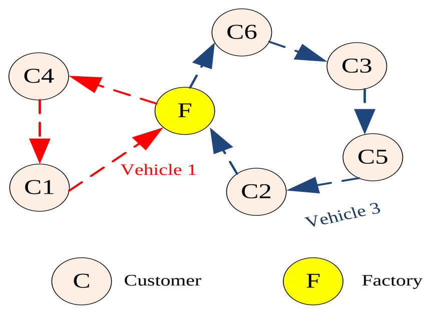


Fig. 15. The optimal solution of product distribution in the integrated model.

To evaluate the results of the integrated model, a two-stage model was optimized with data provided in *Tables 11-15*. In the first stage model, after optimization, the sequence of jobs was obtained as 6-5-2-7-4-3-1 and the Cmax value was equal to 28. In other words, the results of the first stage model are exactly equal to the results of the integrated model. Then the second stage model was optimized. *Fig. 16* illustrates the optimal routes of the second stage model.

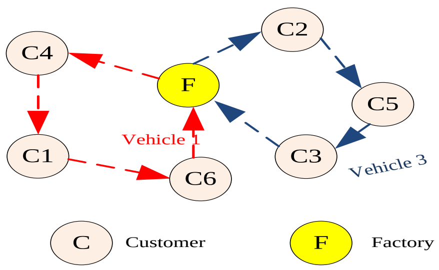


Fig. 16. Optimal routes in the non-integrated model.

The results of this section show that the routes formed in the non-integrated model are completely different from those in the integrated model. Therefore, earliness and tardiness can serve as a criterion to determine the superiority of either of the integrated and non-integrated approaches relative to the other one. *Table 14* shows comparison of earliness and tardiness as well as production and distribution costs in the two models. *Table 14* shows comparison between the different costs of the two models.

As can be seen, both models performed equally well. The reason for this is that the transportation of products started at 28. Moreover, the lower limit of the time window of all customers is less than this time. Therefore, it was not possible to create earliness in any of the models. In total, the integrated model had a tardiness of 142 time units, while the non-integrated model had a tardiness of 186. In other words, the integrated model saves 23.65% in delay time. An examination of the costs of the two models shows that the integrated model has a higher fixed cost than the non-integrated model. However, with expending higher fixed costs, 2.34% of variable costs will be saved. Regarding delays, the cost of tardiness is 2698 in the integrated model and 3361 in the non-integrated model, which the integrated model is 19.70% better than the non-integrated model. Overall, the total cost of distributing products among customers is 20196 monetary units in the integrated model and approximately 21760 monetary units in the non-integrated model, which about 7% financial savings is obtained in the integrated model. Therefore, a comprehensive study of the above models shows that the performance of the integrated model is much better than the non-integrated model; and simultaneous decision-making on production, assembly and distribution can significantly minimize system costs, and also increase customer satisfaction by reducing delays.

Table 14. Earliness and tardiness in integrated and non-integrated models.

	Customer	C1	C2	C3	C4	C5	C6
Integrated model	Earliness	-	-	-	-	-	-
	Tardiness	24	44	21	-	49	4
Non-integrated models	Earliness	-	-	-	-	-	-
	Tardiness	24	19	18	-	57	68

Table 15. Distribution costs in integrated and non-integrated models.

	Earliness Total	Total Tardiness	Distribution Fixed Cost	Distribution Variable Cost	Delays Cost	Total Distribution Cost
Integrated model	0	142	1500	17968	2698	20196
Non-integrated model	0	186	1500	18399	3361	21760
Percentage of Integrated model superiority	-	23.65%	0%	2.34%	19.70%	7.31%

4.1 | Sensitivity Analysis

The purpose of sensitivity analysis is to investigate the effect of fluctuations in important parameters of the mathematical model on the value of the objective function. This effect will be examined independently. In other words, by assuming other parameters constant, the effect of one parameter on the value of the objective function will be analyzed. In terms of cost factors, this analysis is quite clear. Fluctuations in any type of cost have a direct effect on the value of the objective function.

However, some parameters do not have a significant effect on the objective function. These parameters include setup time and capacity of the transport fleet. The effect of these parameters on the objective function will be examined below.

4.2 | Sensitivity Analysis of Setup Time

In the validation problem, the values of the setup time were presented. In this section, the values of this parameter fluctuate between -20% and +20% (*Table 16*, *Fig. 17*).

Table 16. Sensitivity analysis of setup time.

Percentage change in the parameter	-20%	-10%	0%	10%	20%
Objective function value	6050.8	6053.9	6053.9	6054.1	6057.9

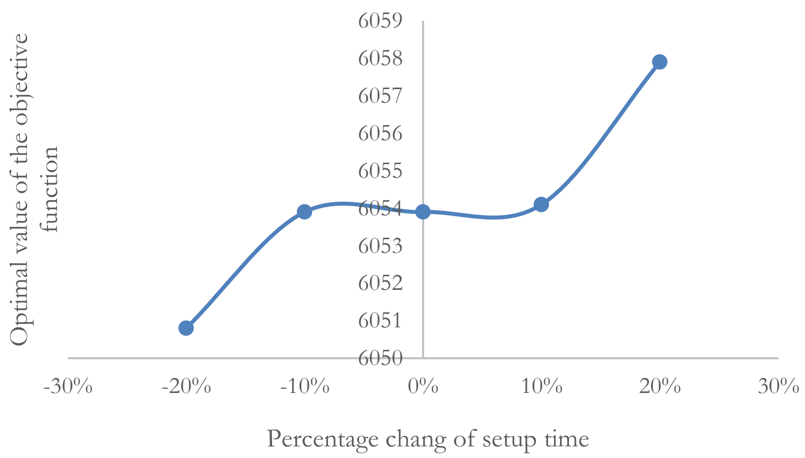


Fig. 17. Objective function values relative to setup time changes.

The results in *Fig. 17* and *Table 16* show that increasing the setup time can lead to an increase in total costs. It can be said that the increase of setup time increases the completion of time jobs. As a result, the distribution stage starts later and eventually the customers receive their orders with more delay. Therefore, the cost of delay in customer delivery increases. It should be noted that minor changes in this parameter did not influence the value of the objective function and only a severe increase of up to +20% leads to changes.

4.3 | Sensitivity Analysis of Transport Fleet Capacity

In the validation problem, the capacity values of the transport fleet were presented. In this section, the values of this parameter fluctuate between -20% and +20%, and in *Table 17* and *Fig. 18*, the results of the sensitivity analysis of this parameter are presented.

Table 17. Sensitivity analysis of transport fleet capacity.

Percentage change in the parameter	-20%	-10%	0%	10%	20%
Objective function value	6050.8	6053.9	6053.9	6054.1	6057.9

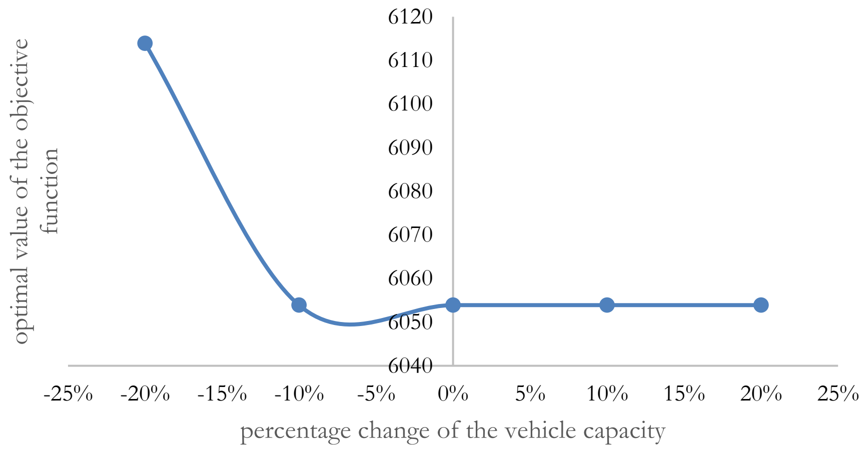


Fig. 18. The value of objective function relative to variations in transport fleet capacity.

As shown in *Fig. 18* and *Table 17*, no change occurs in the value of the objective function from -10% up to +20% fluctuations. The costs increased with decreasing the transport fleet capacity to -20%. The reason for this is that by reducing the capacity of vehicles, the model has to use more vehicles; and as a result, more fixed costs are imposed on the model, and therefore the optimal value of the objective function increases.

5 | Conclusion

In this study, the two-stage assembly flow shop problem and transport fleet routing were studied. The key innovation of this research was the integration of two-stage assembly scheduling and vehicle routing decisions. In this study, the costs of holding, routing, and penalties for violating the time window were minimized. Therefore, an integrated model and a non-integrated model were presented. An improved version of WOA is proposed to optimize the studied problem. Comparison of the integrated model and the two-stage model showed that the integrated model saved 23.65% of delay time, the integrated model showed better performance than the two-stage model by 13.6% in terms of total costs.

An examination and comparison of the applied solutions showed that in all solved problems, the WOA's solution time was less than the GA's and IWOA's. However, the proposed IWOA leads to about 60% improvement in cost reduction in comparison to WOA. This is while WOA performs better than GA in 80% of large-scale problems, and GA only provides a better objective function value in 20% of the problems (problems 13 and 14). Therefore, the efficiency of the IWOA, compared to the WOA and GA, can be well approved.

Analysis of the results on the algorithms studied in this study shows that new meta-heuristic algorithms such as WOA can perform much better and more powerfully than conventional and old algorithms and provide higher speed and higher quality, and this algorithm Have the ability to replace the old algorithms well. Also, due to the newness of the WOA algorithm, various improvements can be made in it, which can implement the search process in the WOA algorithm better and more powerfully.

In this study, IWOA algorithm have led to improvements in both the neighborhood creation structure and the choice of answers for subsequent iterations, resulting in a 60% improvement in cost savings over the WOA. Therefore, it is generally concluded that the IWOA algorithm can be introduced as a new and efficient algorithm both in terms of optimization speed and quality of the results found, and other researchers in this field are suggested to focus more on Rely on this algorithm and use its advantages over other meta-heuristic algorithms.

The management-related achievements of this research show that in factories, integration of decisions related to production, assembly and distribution can help managers in controlling costs and creating coordination between production and distribution units. The need for this integration will be intensified when there are multiple customer orders. In such a situation, rough-cut planning cannot provide the necessary coordination among production, assembly, and distribution, and it is necessary to use up-to-date scientific tools. This research can be a comprehensive decision making for managers of manufacturing organizations. The limitations of the research are as follows:

I. The meta-heuristics algorithms require access to a computer system equipped with features such as high RAM and CPU.
II. As there was no official database for some parts of cost elements, the driver's estimations were asked to help. The questions about the travel costs for each route have been categorized and the estimated costs have been entered into the mathematical model.

To develop this research, it is suggested that uncertainty in customer demand be considered as possibilistic programming or robust optimization. It is also recommended to include planning on product waste and supply of raw materials in the problem. Regarding the solution methods, it is suggested that a heuristic algorithm be developed for the problem, and also other meta-heuristic algorithms such as Runner-Root Algorithm (RRA) be used to solve this problem.

Acknowledgments

The authors are tankful from Deputy for Research and Technology of Islamic Azad University, Qazvin Branch with the code 21440311962004.

Funding

This study is extracted from the Ph.D dissertation in which is supported by Islamic Azad University, Qazvin Branch with the code 21440311962004.

Conflicts of Interest

There is no financial interest to report.

Reference

[1] Negri, M., Cagno, E., Colicchia, C., & Sarkis, J. (2021). Integrating sustainability and resilience in the supply chain: a systematic literature review and a research agenda. *Business strategy and the environment*, *30*(7), 2858-2886.

[2] Burggräf, P., Wagner, J., Lück, K., & Adlon, T. (2017). Cost-benefit analysis for disruption prevention in low-volume assembly. *Production engineering*, *11*, 331-342.

[3] Mohammadi, S., Al-e-Hashem, S. M., & Rekik, Y. (2020). An integrated production scheduling and delivery route planning with multi-purpose machines: a case study from a furniture manufacturing company. *International journal of production economics*, *219*, 347-359.

[4] Wu, C. C., Gupta, J. N., Cheng, S. R., Lin, B. M., Yip, S. H., & Lin, W. C. (2021). Robust scheduling for a two-stage assembly shop with scenario-dependent processing times. *International journal of production research*, *59*(17), 5372-5387.

[5] Wu, C. C., Zhang, X., Azzouz, A., Shen, W. L., Cheng, S. R., Hsu, P. H., & Lin, W. C. (2021). Metaheuristics for two-stage flow-shop assembly problem with a truncation learning function. *Engineering optimization*, *53*(5), 843-866.

[6] Abdolazimi, O., Shishebori, D., Goodarzian, F., Ghasemi, P., & Appolloni, A. (2021). Designing a new mathematical model based on ABC analysis for inventory control problem: a real case study. *RAIRO-operations research*, *55*(4), 2309-2335.

[7] Khanchehzarrin, S., Shahmizad, M., Mahdavi, I., Mahdavi-Amiri, N., & Ghasemi, P. (2021). A model for the time dependent vehicle routing problem with time windows under traffic conditions with intelligent travel times. *RAIRO-Operations research*, *55*(4), 2203-2222.

[8] Potts, C. N., Sevast'Janov, S. V., Strusevich, V. A., Van Wassenhove, L. N., & Zwaneveld, C. M. (1995). The two-stage assembly scheduling problem: complexity and approximation. *Operations research*, *43*(2), 346-355.

[9] Lee, C. Y., Cheng, T. C. E., & Lin, B. M. T. (1993). Minimizing the makespan in the 3-machine assembly-type flowshop scheduling problem. *Management science*, *39*(5), 616-625.

[10] Allahverdi, A., & Al-Anzi, F. S. (2006). Evolutionary heuristics and an algorithm for the two-stage assembly scheduling problem to minimize makespan with setup times. *International journal of production research*, *44*(22), 4713-4735.

[11] Tozkapan, A., Kırca, Ö., & Chung, C. S. (2003). A branch and bound algorithm to minimize the total weighted flowtime for the two-stage assembly scheduling problem. *Computers & operations research*, *30*(2), 309-320.

[12] Allahverdi, A., & Aydilek, H. (2015). The two stage assembly flowshop scheduling problem to minimize total tardiness. *Journal of intelligent manufacturing*, *26*, 225-237.

[13] Sung, C. S., & Kim, H. A. (2008). A two-stage multiple-machine assembly scheduling problem for minimizing sum of completion times. *International journal of production economics*, *113*(2), 1038-1048.

[14] Koulamas, C., & Kyparisis, G. J. (2001). The three-stage assembly flowshop scheduling problem. *Computers & operations research*, *28*(7), 689-704.

[15] Zhang, Z., & Tang, Q. (2021). Integrating flexible preventive maintenance activities into two-stage assembly flow shop scheduling with multiple assembly machines. *Computers & industrial engineering*, *159*, 107493. https://doi.org/10.1016/j.cie.2021.107493

[16] Pourhejazy, P., Cheng, C. Y., Ying, K. C., & Lin, S. Y. (2021). Supply chain-oriented two-stage assembly flowshops with sequence-dependent setup times. *Journal of manufacturing systems*, *61*, 139-154.

[17] Hatami, S., Ebrahimnejad, S., Tavakkoli-Moghaddam, R., & Maboudian, Y. (2010). Two meta-heuristics for three-stage assembly flowshop scheduling with sequence-dependent setup times. *International journal of advanced manufacturing technology*, *50*(9-12), 1153-1164.

[18] Andrés, C., & Hatami, S. (2011, September). The three stage assembly permutation flowshop scheduling problem. *V international conference on industrial engineering and industrial management* (pp. 867-875). http://adingor.es/congresos/web/uploads/cio/cio2011/metodos_cuantitativos/867-875.pdf

[19] Maleki-Darounkolaei, A., Modiri, M., Tavakkoli-Moghaddam, R., & Seyyedi, I. (2012). A three-stage assembly flow shop scheduling problem with blocking and sequence-dependent set up times. *Journal of industrial engineering international*, *8*, 1-7.

[20] Dalfard, V. M., Ardakani, A., & Banihashemi, T. N. (2011). Hybrid genetic algorithm for assembly flow-shop scheduling problem with sequence-dependent setup and transportation times. *Technical gazette*, *18*(4), 497-504.

Bahmani et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 56-83

[21] Mozdgir, A., Fatemi Ghomi, S. M. T., Jolai, F., & Navaei, J. (2013). Two-stage assembly flow-shop scheduling problem with non-identical assembly machines considering setup times. *International journal of production research*, *51*(12), 3625-3642.

[22] Tian, Y., Liu, D., Yuan, D., & Wang, K. (2013). A discrete PSO for two-stage assembly scheduling problem. *The international journal of advanced manufacturing technology*, *66*, 481-499.

[23] Allahverdi, A., Aydilek, H., & Aydilek, A. (2016). Two-stage assembly scheduling problem for minimizing total tardiness with setup times. *Applied mathematical modelling*, *40*(17-18), 7796-7815.

[24] Navaei, J., Ghomi, S. F., Jolai, F., Shiraqai, M. E., & Hidaji, H. (2013). Two-stage flow-shop scheduling problem with non-identical second stage assembly machines. *The international journal of advanced manufacturing technology*, *69*, 2215-2226.

[25] Shoaardebili, N., & Fattahi, P. (2015). Multi-objective meta-heuristics to solve three-stage assembly flow shop scheduling problem with machine availability constraints. *International journal of production research*, *53*(3), 944-968.

[26] Komaki, G. M., & Kayvanfar, V. (2015). Grey Wolf Optimizer algorithm for the two-stage assembly flow shop scheduling problem with release time. *Journal of computational science*, *8*, 109-120.

[27] Wang, K., Ma, W. Q., Luo, H., & Qin, H. (2016). Coordinated scheduling of production and transportation in a two-stage assembly flowshop. *International journal of production research*, *54*(22), 6891-6911.

[28] Kazemi, H., Mazdeh, M. M., & Rostami, M. (2017). The two stage assembly flow-shop scheduling problem with batching and delivery. *Engineering applications of artificial intelligence*, *63*, 98-107.

[29] Jung, S., Woo, Y. B., & Kim, B. S. (2017). Two-stage assembly scheduling problem for processing products with dynamic component-sizes and a setup time. *Computers & industrial engineering*, *104*, 98-113.

[30] Wu, C. C., Chen, J. Y., Lin, W. C., Lai, K., Liu, S. C., & Yu, P. W. (2018). A two-stage three-machine assembly flow shop scheduling with learning consideration to minimize the flowtime by six hybrids of particle swarm optimization. *Swarm and evolutionary computation*, *41*, 97-110.

[31] Goli, A., & Davoodi, S. M. R. (2018). Coordination policy for production and delivery scheduling in the closed loop supply chain. *Production engineering*, *12*(5), 621-631.

[32] Basir, S. A., Mazdeh, M. M., & Namakshenas, M. (2018). Bi-level genetic algorithms for a two-stage assembly flow-shop scheduling problem with batch delivery system. *Computers & industrial engineering*, *126*, 217-231.

[33] Yavari, M., & Isvandi, S. (2018). Integrated decision making for parts ordering and scheduling of jobs on two-stage assembly problem in three level supply chain. *Journal of manufacturing systems*, *46*, 137-151.

[34] Luo, J., Liu, Z., & Xing, K. (2019). Hybrid branch and bound algorithms for the two-stage assembly scheduling problem with separated setup times. *International journal of production research*, *57*(5), 1398-1412.

[35] Talens, C., Fernandez-Viagas, V., Perez-Gonzalez, P., & Framinan, J. M. (2020). New efficient constructive heuristics for the two-stage multi-machine assembly scheduling problem. *Computers & industrial engineering*, *140*, 106223. https://doi.org/10.1016/j.cie.2019.106223

[36] Lin, X., & Chen, J. (2021). Deep reinforcement learning for dynamic scheduling of two-stage assembly flowshop. *Advances in swarm intelligence: international conference on Swarm Intelligence, ICSI 2021* (pp. 263-271). Springer International Publishing. https://doi.org/10.1007/978-3-030-78811-7_25

[37] Hosseini, S. M. H., Sana, S. S., & Rostami, M. (2022). Assembly flow shop scheduling problem considering machine eligibility restrictions and auxiliary resource constraints. *International journal of systems science: operations & logistics*, *9*(4), 512-528.

[38] Masruroh, N. A., Fauziah, H. A., & Sulistyo, S. R. (2020). Integrated production scheduling and distribution allocation for multi-products considering sequence-dependent setups: a practical application. *Production engineering*, *14*, 191-206.

[39] Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*, 51-67.

[40] Alinaghian, M., & Goli, A. (2017). Location, allocation and routing of temporary health centers in rural areas in crisis, solved by improved harmony search algorithm. *International journal of computational intelligence systems*, *10*(1), 894-913.

[41] Tirkolaee, E. B., Goli, A., Hematian, M., Sangaiah, A. K., & Han, T. (2019). Multi-objective multi-mode resource constrained project scheduling problem using Pareto-based algorithms. *Computing*, *101*, 547-570.

[42] Rashidi Komijan, A., Ghasemi, P., Khalili-Damghani, K., & HashemiYazdi, F. (2021). A new school bus routing problem considering gender separation, special students and mix loading: a genetic algorithm approach. *Journal of optimization in industrial engineering*, *14*(2), 23-39.

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 84–97.

Paper Type: Research Paper

Cost Efficiency in the Presence of Time Dependent Prices

Jafar Pourmahmoud[1,*], Naser Kaheh[1], Farhad Hosseinzadeh Lotfi[2]

[1] Department of Applied Mathematic, Azarbaijan Shahid Madani University, Tabriz, Iran; pourmahmoud@azaruniv.ac.ir; n.kaheh@azaruniv.ac.ir.

[2] Department of Mathematical, Science and Research Branch, Islamic Azad University, Tehran, Iran; farhad@hosseinzadeh.ir.

Citation:

Pourmahmoud, J., Kaheh, N., & Hosseinzadeh Lotfi, F. (2023). Cost efficiency in the presence of time dependent prices. *Journal of applied research on industrial engineering, 10*(1), 84-97.

Received: 28/05/2022 Reviewed: 02/07/2022 Revised: 05/08/2022 Accepted: 01/09/2022

Abstract

In traditional cost-efficiency models, inputs and outputs, as well as input prices were known as constant values for each decision-making unit In our daily applications, however, market entry prices vary at different times. In other words, input prices for Decision-Making Units (DMUs) are time dependent. Traditional methods cannot calculate the cost efficiency of DMUs with time-dependent prices. This paper proposes a new method to calculate the cost efficiency of DMUs in the presence of time-dependent prices. The proposed model is a parametric programming problem model depending on time. In the presented model, the inputs and outputs are functions in terms of time, which is not present in the models introduced by other researchers. New definitions for time-dependent cost efficiency have also been introduced. The cost efficiency of DMUs is measured over a given time and the units are ranked according to the time obtained. Finally, a numerical example has been presented to illustrate the proposed method.

Keywords: Data envelopment analysis, Cost efficiency, Time Dependent prices, Ranking.

1 | Introduction

Data Envelopment Analysis (DEA), first introduced by Charnes et al. [4], is a non-parametric method for evaluating the performance of Decision-Making Units (DMUs). With the help of DEA, the efficiency of the DMUs can be calculated and the DMUs can be ranked. Many studies have been conducted in this area, including Banker et al. [1], Edalatpanah [7], Khodabakhshi and Cheraghali [17], and Maghbouli and Moradi [20] Managers, in addition to evaluating DMUs, are always looking for all ways of manufacturing their products with minimal costs. In practice, DMUs can be valued in terms of costs, profits, or revenues if input and output prices are available. In fact, in cost efficiency models, the ability to yield current outputs is evaluated at minimum cost.

The concept of cost efficiency was first introduced by Farrell [11] and then by Färe et al. [9] used the linear programming model to develop cost efficiency. They defined the cost-effectiveness of a DMU as the ratio between the minimum production cost and the actual observed cost. Tone [29] improved

Corresponding Author: pourmahmoud@azaruniv.ac.ir

https://doi.org/10.22105/jarie.2022.344543.1475

the Färe et al. [9] by evaluating DMUs using cost-based production possibility rates instead of traditional production possibility rates. Other studies in this area include Kholmuminov and Wright [18], Haralayya and Aithal [12], [13].

Traditional methods (such as Färe et al. [9] and Tone [29]) of the cost efficiency of different DMUs have been calculated when all data, including inputs, outputs and prices of inputs, are determined precisely. However, in practice, the values for the outputs or inputs or their prices may be uncertain.

Recently, various uncertain data have been introduced, including fuzzy data, interval data, and stochastic data. In addition, many studies have been conducted to calculate the efficiency of DMUs, as well as the cost efficiency, profit efficiency, and revenue efficiency of DMUs in the presence of uncertain data. Some of these studies are discussed here.

In recent years, many researchers have studied the use of fuzzy theory for cost efficiency. Jahanshahloo et al. [16] examined for the first time the assessment of cost efficiency considering the fuzzy DEA. Several studies have subsequently been carried out in this area, including Puri and Yadav [25] Pourmahmoud and Sharak [23], [24]. Camanho and Dyson [3], and Hosseinzadeh Lotfi et al. [15] are the first ones to calculate DMUs cost efficiency in the presence of interval data. Camanho and Dyson [3] calculated the cost efficiency of DMUs while the inputs and outputs values at each DMU were certain and input prices were interval. They calculated the cost efficiency of DMU in which the input values and output values of each DMU as well as the input prices were certain. Sun et al. [27] and Dyvak et al. [6] are the most recent studies in the field of interval data.

There are many models and methods associated with time in DEA, but most of them attempt to study DMU performance over different time periods. These methods can be divided into three categories including productivity index, window analysis and dynamic systems. In the network structure, the Malmquist Productivity Index (MPI), first introduced by the Swedish economist Malmquist [21], is used to measure changes in productivity over time. Dynamic systems are repetitions of the single-period systems which are connected by carryovers, where a single-period system can have any particular structure. Färe and Grosskopf [10] studied this topic earlier. Li et al. [19] studied dynamic prediction of financial distress using Malmquist DEA. Pourmahmoud [22] introduced a new model for ranking DMU based on Dynamic DEA. Data window analysis method was first introduced by Charnes and Cooper [5], as a window analysis. The window analysis method evaluates the performance of each DMU as if it had a different identity at any point in time. In this method, each window consists of a specific number of studies years, beginning with the base year and continuing for the duration of the window. The efficiency values of each DMU are calculated each year, taking into account that the average of the efficiency calculated in this window is the efficiency value of that DMU in this window. By moving the window to a new period (deleting the base year and adding a year at the end of the window), the efficiency values in the new window are calculated for DMUs. Finally, the performance of each DMU is evaluated by comparing each window's efficiency scores to other DMUs over the period.

Another time-related problem is calculating the efficiency of DMUs with time-varying data. Another type of data, called time-dependent data, was first reported by Taeb et al. [28]. They calculated the efficiency of DMUs where the input and output values were time-dependent. In fact, the price of gold, oil, stocks, etc. depends on time. Sometimes the price fluctuations are strong, so the change in cost efficiency over a period of time is very significant. A rapid calculation is required to have a real and updated value for cost efficiency. In this study is to calculate the cost efficiency of DMUs in which market prices for inputs are time dependent and the values of inputs and outputs are certain. Recently, studies have been conducted on cost efficiency, including Fallahnejad et al. [8], Soleimani-Chamkhorami and Ghobadi [26], and Hatami-Marbini and Arabmaldar [14], which none of them is time-dependent data. In this study, we consider inputs and outputs as a function of time. This feature is not present in any of the previous models. This is the superiority of the model presented in this study compared to previous models. Literature gap is given in the table below.

Table 1. Literature gap.

Researches	Method Used
Previous Researches [3], [15], [16], [28]	Calculating the efficiency of DMUs or their cost efficiency in the presence of inputs and outputs time-dependent or fuzzy or interval.
Current study	Calculation of DMUs cost efficiency in the presence of time-dependent prices.
Comparison	In none of the previous models, cost efficiency has been calculated in the presence of time-dependent input prices. In this study, for the first time, input prices are considered as a function of time.

This paper is organized as follows. Section 2 includes an introduction to the traditional cost efficiency from the viewpoint of Färe et al. [9]. Section 3 presents the proposed time-dependent cost efficiency model. In Section 4, a numerical example illustrating the proposed approach is provided. Finally, conclusion is given in Section 5.

2 | Introduction to the Old-Style Cost Efficiency

Suppose DMU_k produces the output $y_k=(y_{1k},y_{2k},...,y_{sk})$ using the input $x_k=(x_{1k},x_{2k},...,x_{mk})$ while all input and output components are nonnegative and $x_k\neq 0$, $y_k\neq 0$. Suppose also that the input price vector for DMU_k is $w_k=(w_{1k},w_{2k},...,w_{mk})$. Therefore, the observed cost by DMU_k is:

$$C_k=\sum_{i=1}^{m} w_{ik}x_{ik}.$$

DMU_k is cost efficient when the observed cost by DMU_k is the lowest cost that can produce y_k. Suppose there are a set of n observations on the DMUs each of which produces s number of outputs by using m number of inputs. Banker et al. [1] defined the production feasibility set for these data in Variable Return to Scale (VRS) mode as follows:

$$T=\{(x,y): x\geq\sum_{j=1}^{n}\lambda_j x_j, y\leq\sum_{j=1}^{n}\lambda_j y_j, \sum_{j=1}^{n}\lambda_j=1, \lambda_j\geq 0, j=1,2,...,n\}.$$

Färe et al. [9] introduced cost minimization model as follows:

$$\begin{aligned} C_k^* = {} & \min\ w_k x \\ & \text{s.t. } (x,y_k)\in T. \end{aligned} \tag{1}$$

Model (1) in VRS mode is written as follows:

$$\begin{aligned} C_k^* = {} & \min\sum_{i=1}^{m} w_{ik}x_i \\ & \text{s.t. } \sum_{j=1}^{n}\lambda_j x_{ij}\leq x_i, i=1,2,...,m, \\ & \sum_{j=1}^{n}\lambda_j y_{rj}\geq y_{rk}, r=1,2,...,s, \\ & \sum_{j=1}^{n}\lambda_j=1, \\ & \lambda_j\geq 0, j=1,2,...,n. \end{aligned} \tag{2}$$

If $x^*=(x_1^*,x_2^*,...,x_m^*)$ is the optimal solution for *Model (2)*, the minimum cost for producing y_k is

$$C_k^*=\sum_{i=1}^{m} w_{ik}x_i^*.$$

Pourmahmoud et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 84-96

Since $(x_k, y_k) \in T$, and C_k is the actually observed cost by DMU_k and C_k^* is also the minimum possible cost for producing the output of DMU_k, so $C_k^* \leq C_k$. DMU_k is cost efficient whenever $C_k^* = C_k$. Färe et al. [9] introduced the cost efficiency of DMU_k as follows:

$$\mathrm{CE}_k = \frac{\mathrm{C}_k^*}{\mathrm{C}_k}.$$

Evidently, $CE_k \leq 1$; and the closer CE_k gets one, the more efficient DMU_k becomes in terms of cost.

3 | Proposed Model for Time Dependent Cost Efficiency

Since in real market ices of gold, oil, and stock etc. often fluctuate over time in some cases, cost efficiency evaluation of DMUs is considered over a period of time. Therefore, the cost efficiency of DMUs that use these inputs would vary by time. Hence, in such a case the estimation would be time dependent. Based on the Färe model, suppose market prices at time t for input i of DMU_k equals $w_{ik}(t)$. $w_{ik}(t)$ is a time dependent function that may be a constant function. Being constant means that the price of the i-th input of DMU_k over a period of time is constant. Therefore the observed cost by DMU_k (at the moment t) is

$$\mathrm{C}_k(\mathrm{t}) = \sum_{i=1}^{m} \mathrm{w}_{ik}(\mathrm{t})\mathrm{x}_{ik}. \tag{3}$$

DMU_k is cost efficient at moment t when the observed cost by DMU_k is the lowest cost that can produc y_k.

If market prices of the inputs are time dependent, the minimum cost for producing y_k will also be time dependent. Therefore, *Model (2)* can be rewritten as follows to estimate the cost of DMU_k in VRS mode over t∈ [a, b]:

$$\begin{aligned} \mathrm{C}_k^*(\mathrm{t}) = \min & \sum_{i=1}^{m} \mathrm{w}_{ik}(\mathrm{t})\mathrm{x}_i \\ \text{s.t.} & \sum_{j=1}^{n} \lambda_j \mathrm{x}_{ij} \leq \mathrm{x}_i, \mathrm{i} = 1,2,\ldots,\mathrm{m}, \\ & \sum_{j=1}^{n} \lambda_j \mathrm{y}_{rj} \geq \mathrm{y}_{rk}, \mathrm{r} = 1,2,\ldots,\mathrm{s}, \\ & \sum_{j=1}^{n} \lambda_j = 1, \\ & \lambda_j \geq 0, \mathrm{j} = 1,2,\ldots,\mathrm{n}. \end{aligned} \tag{4}$$

Definition 1. A function of Time Dependent Cost Efficiency (TDCE) for DMU_k, $CE_k(t)$ is defined as follows:

$$\mathrm{CE}_k(\mathrm{t}) = \frac{\mathrm{C}_k^*(\mathrm{t})}{\mathrm{C}_k(\mathrm{t})}. \tag{5}$$

Obviously, in *Relation (5)*, $C_k(t) \neq 0$ for all t since the observed cost by a DMU would never be zero. Also with respect to *Model (3)* for all t, $CE_k(t) \leq 1$. If for DMU_k at t_i, $CE_k(t_i) = 1$ then this DMU will be the cost efficient at t_i consequently, DMU_k may be efficient in one moment and inefficient in another moment. Hence, for DMU_k there will be three states. The following definition shows TDCE state of $DMUs$.

Definition 2. In terms of TDCE, the following three states can be considered for DMU_k:

I. If $\text{for all } t \in [a,b];\ CE_k(t)=1,$ then DMU_k is named global-efficient based on TDCE.

II. If $\text{there exists } t_1,t_2 \in [a,b]; CE_k(t_1)=1, CE_k(t_2)<1,$ then DMU_k is named local-efficient based on TDCE.

III. If $\text{for all } t \in [a,b];\ CE_k(t)<1,$ then DMU_k is named none-efficient based on TDCE.

The following is a graph of the cost efficiency in terms of time for DMUs in all three states:

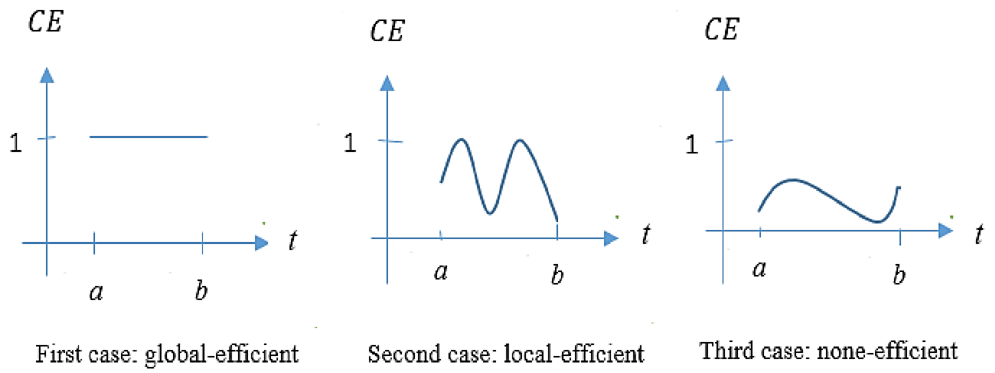


Fig. 1. Graph of the TDCE for DMUs.

As mentioned above, a more comprehensive definition of TDCE can be introduced and ranking DMUs can also be done using this definition.

Definition 3. The value for the TDCE of DMU_k over $T=[t_1,t_2]\,(t_1<t_2)$ which is shown by the symbol CE_k^T can be calculated as follows:

$$\mathrm{CE}_k^T=\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\mathrm{CE}_k(t)dt. \tag{6}$$

Theorem 1. For every time interval $T=[t_1,t_2]$ $(t_1<t_2)$ we have $CE_k^T \le 1$.

Proof: given that for all t∈$T=[t_1,t_2], CE_k(t)\le 1$, then:

$$\int_{t_1}^{t_2}\mathrm{CE}_k(t)dt \le \int_{t_1}^{t_2}dt \Rightarrow \int_{t_1}^{t_2}\mathrm{CE}_k(t)dt \le t_2-t_1 \Rightarrow \frac{1}{t_2-t_1}\int_{t_1}^{t_2}\mathrm{CE}_k(t)dt \le 1 \Rightarrow \mathrm{CE}_k^T \le 1.$$

The geometric proof of the theorem can be stated using the data in *Fig. 2* as follows. In *Definition 3*, the value of $\int_{t_1}^{t_2} CE_k(t)dt$ is the area under the graph of the function $CE_k(t)$ shown by the colored part in *Fig. 2*. t_2-t_1 is also the value of $\int_{t_1}^{t_2} dt$, that is, the area under the graph of the constant function $CE_k(t)=1$. So the value for CE_k^T is always less than or equal to one.

Proposition 1. In TDCE evaluation of DMUs, at any given moment, there is always a DMU whose cost efficiency is one.

Proof: In the calculation of DMUs cost efficiency, in the traditional *Model (2)*, there is always at least a DMU whose cost efficiency equals one. Since time is fixed, when DMUs in the given $t=\alpha$ is evaluated, *Model (4)* becomes the traditional *Model (2)*. So at this moment, there is a DMU whose cost efficiency equals one.

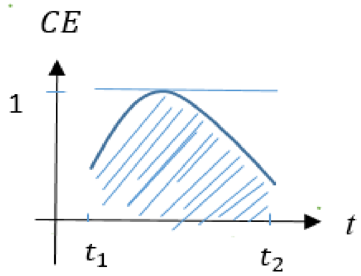


Fig. 2. The cost efficiency versus time.

Corollary 1. In TDCE evaluation of DMUs, there may not be a DMU that is global-efficient, but there is always a DMU that is local-efficient. Therefore, it can be said that in this type of evaluation, at any time interval, for n DMU under evaluation, there is always a DMU that is not totally none-efficient.

Definition 4. In the time interval I = [a, b] for DMU_k, the *Definition 3* is rewritten as follows:

$$CE_k^I = \frac{1}{b-a}\int_a^b CE_k(t)dt. \tag{7}$$

In the specific case $I_0=[0,1]$, the TDCE value of DMU_k is equal to:

$$CE_k^{I_0} = \int_0^1 CE_k(t)dt. \tag{8}$$

Definition 5. If we have in $T=[t_1,t_2]$ for DMU_k and $DMU_q : CE_q^T < CE_k^T$, then efficiency of DMU_k is higher than efficiency of DMU_q and as a result DMU_k is more efficient than DMU_q. Hence, using this definition, DMUs can be ranked according to their performance.

Solving algorithm for Model (4)

The price of the inputs in the objective function of *Model (4)* is a function of the variable t, which changes continuously. Given the definition of the parametric programming problem, *Model (4)* is also a parametric programming problem. The model can be solved by a parametric problem solving algorithm. The model resolution algorithm in *Model (4)* is proposed by Bazaraa et al. [2]:

1. For t = a, solve the model with the simplex algorithm.
2. Replace the alterations ed by changes in the objective function cost vector using the sensitivity analysis in optimal table extracted from Step 1. In other words, calculate the objective function values and $z_j - c_j$ of the non-basic variables by considering the cost vector and replace them in the row of the optimal table objective function extracted from Step 1. If the final table fails to be unified after the effect of parameter t, unify it.
3. Find the permissible range of parameter t by setting $z_j - c_j \le 0$ to keep the table optimized. Then increase t until the table loses optimality and select the first available non-basic variable $z_j - c_j > 0$ as the basic input variable.
4. Repeat Step 3 until for every t and for all non-basic variables $z_j - c_j \le 0$. Then the optimal solution is obtained.

Pourmahmoud et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 84-96

4 | Numerical Examples

Three numerical examples in different modes are presented in this section to illustrate the proposed method. All three examples measure TDCE functions and TDCE values for five two-input, one-output DMUs. The input values for the DMUs have not changed in all three examples. In the first example, the output of all DMUs is one. In the second example, we assume that the input prices are different for each DMU. Finally, in the third example, we not only look at input prices differently, but also at different outputs. The proposed method had good results for the three cases.

Example 1. Suppose that there are five DMUs with two inputs and one output. *Table 2* shows the crispt inputs of DMUs. Additioanly, assume that the output of all DMUs is the same value as $y=1$ and the vector of time dependent prices for inputs is the same for all DMUs as $(1+5t, t+1)$.

Table 2. The inputs for 5 DMUs.

DMUs / Inputs	A	B	C	D	E
Input 1	5	3	4	2	2
Input 2	1	2	5	4	6

The DMUs in *Table 2* are shown in *Fig. 3* in the input space. *Fig. 3* shows that A, B and D are strongly efficient, E is weakly efficient and C is inefficient.

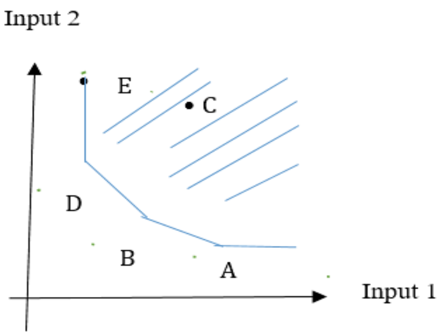


Fig. 3. Show DMUs in inputs space.

Since the output of all DMUs is assumed to be equal to 1, the minimum cost to produce this output must be the same. *Model (4)* is applied for DMUs in *Table 2* and the following results are obtained;

$$C_A^*(t)=C_B^*(t)=C_C^*(t)=C_D^*(t)=C_E^*(t)=\begin{cases}17t+5;0\le t\le 1/3,\\ 14t+6;t>1/3.\end{cases}$$

By considering the inputs of each DMU and their costs and using *Eq. (3)*, the observed cost by each DMUs is obtained as follows:

$$C_A(t)=26t+6,$$

$$C_B(t)=17t+5,$$

$$C_C(t)=25t+9\ ,$$

$$C_D(t)=14t+6\ ,$$

$$C_E(t)=16t+8.$$

By applying *Eq. (5)* for each DMUs, the TDCE function of each DMU is:

$$CE_A(t)=\begin{cases}\dfrac{17t+5}{26t+6}, & 0\le t\le 1/3,\\[2ex] \dfrac{14t+6}{26t+6}, & t>1/3.\end{cases}$$

$$CE_B(t)=\begin{cases}1, & 0\le t\le 1/3,\\ \dfrac{14t+6}{17t+5}, & t>1/3.\end{cases}$$

$$CE_C(t)=\begin{cases}\dfrac{17t+5}{25t+9}, & 0\le t\le 1/3,\\[2ex] \dfrac{14t+6}{25t+9}, & t>1/3.\end{cases}$$

$$CE_D(t)=\begin{cases}\dfrac{17t+5}{14t+6}, & 0\le t\le 1/3,\\[2ex] 1, & t>1/3.\end{cases}$$

$$CE_E(t)=\begin{cases}\dfrac{17t+5}{16t+8}, & 0\le t\le 1/3,\\[2ex] \dfrac{14t+6}{16t+8}, & t>1/3.\end{cases}$$

Cost efficiency of DMUs can be easily calculated at any time. As an example, in $t=0.1$ cost efficiency values with four decimal places for A, B, C, D and E are 0.9139, 1.0000, 0.5547, 0.9054 and 0.5833, respectively.

Suppose that DMUs are evaluated in interval $I_0=[0,1]$. In this case, the TDCE value of DMUs in I_0 is calculated by applying *Eq. (8)* to *4* decimal places are shown in *Table 3*.

Table 3. The TDCE value of DMUs over I_0 with the same time dependent price.

DMU	A	B	C	D	E
$CE_k^{I_0},\ k=A,B,C,D,E$	0.6970	0.9625	0.5965	0.9775	0.7890

According to *Table 3*, none of the DMUs are global-efficient and the DMUs were fully ranked. The ranking of these DMUs is as follows: D, B, E, A, C.

Considering TDCE functions of each DMU, D and B are local-efficient and A, C and E are non-efficient. Although it is not necessary to have a DMU whose TDCE value is equal to one in evaluating DMUs in terms of TDCE, this may also happen. If we consider the time interval under evaluation, instead of $I_0=[0,1]$, the time interval [0, 1/3], B will be global-efficient. According to the results of *Table 3*, the ranking of the DMUs in terms of TDCE may change compared to their traditional ranking. For example, in the traditional model and by considering the same outputs $y=1$, DMU_A is strong efficient and DMU_E is weakly efficient, while according to *Table 3*, DMU_E rank is better than DMU_A rank. Also, by to the results in *Table 3*, TDCE value of DMU_D during I_0 is greater than TDCE value of DMU_B. However, in the time interval [0, 1/3], their ranking be opposite. Since during this time interval, TDCE value of DMU_B is equal to one, but TDCE value of DMU_D is less than one. So the ranking of DMUs in terms of TDCE would be changed if the time intervals were changed.

In this example, the input prices for all DMUs were assumed to be the same, although most of the time this is not the case. Since the DMUs offer their inputs from different malls and at different prices, the input prices may not be the same for all DMUs. This case is examined in the following example.

Example 2. Suppose that there are five DMUs with two inputs and one output. The inputs values and time dependent input prices for these DMUs are presented in *Table 4*. Moreover, assume thate output of all DMUs is the same value as $y=1$.

Table 4. The input values and time dependent input prices for 5 DMUs.

DMU	Inputs		Time Dependent Input Prices	
	Input 1	Input 2	Input Price 1	Input Price 2
A	5	1	$9t^2+2$	$2t^2+3$
B	3	2	$1+6t$	$1+2t$
C	4	5	t^2+1	$2t^2+1$
D	2	4	$1+5t$	$1+t$
E	2	6	$7t+2$	$3t+3$

Model (4) is applied to the data in *Table 4* and the following result is obtained:

$$C_A^*(t)=\begin{cases}31t^2+12, & 0\le t\le 2\sqrt{5}/5,\\ 26t^2+16, & t>2\sqrt{5}/5.\end{cases}$$

$$C_B^*(t)=\begin{cases}22t+5, & 0\le t\le 1/2,\\ 20t+6, & t>1/2.\end{cases}$$

$$C_C^*(t)=7t^2+5,$$

$$C_D^*(t)=\begin{cases}17t+5, & 0\le t\le 1/3,\\ 14t+6, & t>1/3.\end{cases}$$

$$C_E^*(t)=\begin{cases}27t+12, & 0\le t\le 4,\\ 26t+16, & t>4.\end{cases}$$

By considering the inputs of each DMU and their cost and using *Eq. (3)*, the observed cost by each DMU is obtained as follows:

$$C_A(t)=47t^2+13,$$

$$C_B(t)=22t+5,$$

$$C_C(t)=14t^2+9,$$

$$C_D(t)=14t+6,$$

$$C_E(t)=32t+22.$$

By applying *Eq. (5)* for each $DMUs$, the TDCE function of each DMU is as following:

Pourmahmoud et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 84-96

$$CE_A(t)=\begin{cases}\dfrac{31t^2+12}{47t^2+13}, & 0\le t\le 2\sqrt{5}/5,\\[2ex] \dfrac{26t^2+16}{47t^2+13}, & t>2\sqrt{5}/5.\end{cases}$$

$$CE_B(t)=\begin{cases}1, & 0\le t\le 1/2,\\ \dfrac{20t+6}{22t+5}, & t>1/2.\end{cases}$$

$$CE_C(t)=\frac{7t^2+5}{14t^2+9},$$

$$CE_D(t)=\begin{cases}\dfrac{17t+5}{14t+6}, & 0\le t\le 1/3,\\[2ex] 1, & t>1/3.\end{cases}$$

$$CE_E(t)=\begin{cases}\dfrac{27t+12}{32t+22}, & 0\le t\le 4,\\[2ex] \dfrac{26t+16}{32t+22}, & t>4.\end{cases}$$

According to the above functions, none of DMUs are global-efficient. B and D are local-efficient and A, C and E are none-efficient. However, if we take the time interval [0, 1/2], B will be global-efficient. It is good to mention that in this example, similar to example 1, there is at least one DMU at each time whose cost efficiency value is equal to 1. In other words, if $0\le t<1/3$, $CE_B(t)=1$ and B is cost efficient, if $1/3\le t\le 1/2$, $CE_B(t)=1$ and $CE_D(t)=1$ and if $1/2<t$, $CE_D(t)=1$.

Suppose DMUs are evaluated in the interval $I_0=[0,1]$. In this case, TDCE value of A, C, D, and E are unchanged compared to the previous example, and $CE_B^{I_0}=.9891$. These values are shown in *Table 5*.

Table 5. The TDCE value of DMUs over time interval I_0.

DMU	A	B	C	D	E
$CE_k^{I_0}$, $k=A,B,C,D,E$	0.8093	0.9891	0.5399	0.9775	0.6596

According to *Table 5*, DMUs were fully ranked. The ranking of these $DMUs$ is as follows: B, D, A, E, C.

By comparing *Tables* 3 and *5*, when the input prices are changed, the rankings of the $DMUs$ may also change. In Example 1, DMU_A initially was ranked fourth, but it was ranked third when input prices were considered differentlly for DMUs. In addition, the ranking of DMU_B and DMU_D also changed compared to Example 1.

In the previous two examples, we assumed the output of $DMUs$ to be one, but in real world, the outputs cannot always be the same. In the next example, we will examine this case.

Example 3. Suppose that there are five DMUs with two inputs and one output. The inputs and output values and time dependent input prices for these DMUs are presented in *Table 6*.

Table 6. The input and output values and time dependent input prices for 5 DMUs.

DMU	Inputs		Output	Time Dependent Input Prices	
	Input 1	Input 2		Input Price 1	Input Price 2
A	5	1	2	$9t^2+2$	$2t^2+3$
B	3	2	3	$1+6t$	$1+2t$
C	4	5	1	t^2+1	$2t^2+1$
D	2	4	2	$1+5t$	$1+t$
E	2	6	2	$7t+2$	$3t+3$

Model (4) is applied to the data in *Table 6* and the following result is obtained:

$$C_A^*(t)=\begin{cases}31t^2+12, & 0\le t\le 2\sqrt{5}/5,\\ 26t^2+16, & t>2\sqrt{5}/5.\end{cases}$$

$$C_B^*(t)=22t+5,$$

$$C_C^*(t)=7t^2+5,$$

$$C_D^*(t)=\begin{cases}17t+5, & 0\le t\le 1/3,\\ 14t+6, & t>1/3.\end{cases}$$

$$C_E^*(t)=\begin{cases}27t+12, & 0\le t\le 4,\\ 26t+16, & t>4.\end{cases}$$

By considering the inputs of each DMU and their costs and using *Eq. (3)*, the observed cost by each DMU is obtained as follows:

$$C_A(t)=47t^2+13,$$

$$C_B(t)=22t+5,$$

$$C_C(t)=14t^2+9,$$

$$C_D(t)=14t+6,$$

$$C_E(t)=32t+22.$$

By applying *Eq. (5)* for each DMUs, TDCE function of each DMU is as follows:

$$CE_A(t)=\begin{cases}\dfrac{31t^2+12}{47t^2+13}, & 0\le t\le 2\sqrt{5}/5,\\[2ex] \dfrac{26t^2+16}{47t^2+13}, & t>2\sqrt{5}/5.\end{cases}$$

$$CE_B(t)=1,$$

$$CE_C(t)=\frac{7t^2+5}{14t^2+9},$$

$$CE_D(t)=\begin{cases}\dfrac{17t+5}{14t+6}, & 0\le t\le 1/3,\\[2ex] 1, & t>1/3.\end{cases}$$

$$CE_E(t)=\begin{cases}\dfrac{27t+12}{32t+22}, & 0\le t\le 4,\\[2ex] \dfrac{26t+16}{32t+22}, & t>4.\end{cases}$$

Pourmahmoud et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 84-96

According to the above functions, in each interval under evaluation, B is global-efficient, D is local-efficient and A, C and E are none-efficient.

Suppose DMUs are evaluated in interval $I_0 = [0,1]$. In this case, TDCE value of DMUs in time interval I_0 is calculated by applying the *Eq. (8)* to 4 decimal places and the following results are obtained:

Table 7. The TDCE value of DMUs over time interval I_0.

DMU	A	B	C	D	E
$CE_k^{I_0}$, $k = A,B,C,D,E$	0.8093	1.0000	0.5399	0.9775	0.6596

According to *Table 7*, DMUs were fully ranked. The ranking of these DMUs is as follows: B, D, A, E, C. In this example, B is global, efficient in each interval under evaluation. This point did not exist in previous examples.

As these examples illustrate, the recommended method can be used in a number of ways. If the input prices are time-dependent and the same for all inputs or different for inputs, the proposed model is applicable. When evaluating DMUs, given time-dependent input prices, at any point in time there is at least one DMU whose cost-efficiency value is equal to 1. However, here may not be a DMU whose TDCE value in the interval under evaluation is equal to one. In another word, there may not be a globally efficient DMU, but there is at least one locally efficient DMU. In addition, this method can be a relatively powerful tool for classifying DMUs since, according to the formula defined for TDCE, it is very unlikely that the value of TDCE will be the same for DMUs.

4 | Conclusion

The study of DMUs provides a number of results, one of which is the cost efficiency of DMUs that introduced by DEA models. Cost efficiency evaluates a DMU's ability to produce current outputs at the lowest possible cost. In traditional cost efficiency models, the cost efficiency of the DMUs is calculated in a fixed time. Therefore, the prices are considered fixed and certain. While for most practical problems, input prices fluctuate over time. In other words, input prices are time dependent. In this study, a new model was introduced to calculate cost efficiency in the presence of time dependent input prices. In fact, in the model presented in this study, the inputs and outputs were considered a function of time. This feature is the superiority of this model over previous models. This model is a parametric programming problem and was solved with a parametric programming problem algorithm. Furthermore, the definitions of TDCE were provided. Finally, in the examples, the TDCE of DMUs was measured over a given time interval, and the DMUs were ranked. The researchers strongly recommend further research in the following areas. First, in this study, a numerical example was used to verify the proposed model. Other researchers can use this model to solve real problems. Second, the TDCE model proposed in this study was based on that proposed by Färe et al. [9], that introduced cost efficiency model. TDCE for DMUs can also be calculated based on Tone's cost efficiency model. Third, the model proposed in this study was in VRS mode. A model for constant return to scale (CRS model) can also be proposed.

Funding

Funding is not applicable.

Conflicts of Interest

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

[1] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078-1092.

[2] Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. (2011). *Linear programming and network flows*. John Wiley & Sons.

[3] Camanho, A. S., & Dyson, R. G. (2005). Cost efficiency measurement with price uncertainty: a DEA application to bank branch assessments. *European journal of operational research*, *161*(2), 432-446.

[4] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444.

[5] Charnes, A., & Cooper, W. W. (1984). Preface to topics in data envelopment analysis. *Annals of operations research*, *2*(1), 59-94.

[6] Dyvak, M., Melnyk, A., Rot, A., Hernes, M., & Pukas, A. (2022). Ontology of mathematical modeling based on interval data. *Complexity*, *2022*. https://doi.org/10.1155/2022/8062969

[7] Edalatpanah, S. A. (2018). Neutrosophic perspective on DEA. *Journal of applied research on industrial engineering*, *5*(4), 339-345.

[8] Fallahnejad, R., Asadi Rahmati, S., & Moradipour, K. (In Press). Decomposition of cost efficiency, given the set of price and cost production possibilities in data envelopment analysis. *Journal of applied research on industrial engineering*. http://www.journal-aprie.com/article_148535_6a05ccefae17e255435630d6d1731659.pdf

[9] Färe, R., Grosskopf, S., & Lovell, C. K. (1985). *The measurement of efficiency of production* (Vol. 6). Springer Science & Business Media.

[10] Färe, R., & Grosskopf, S. (1997). Intertemporal production frontiers: with dynamic DEA. *Journal of the operational research society*, *48*(6), 656-656.

[11] Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the royal statistical society: series A (General)*, *120*(3), 253-281.

[12] Haralayya, B., & Aithal, P. S. (2021). Inter bank analysis of cost efficiency using mean. *International journal of innovative research in science, engineering and technology (IJIRSET)*, *10*(6), 6391-6397.

[13] Haralayya, B., & Aithal, P. S. (2021). Study on cost efficiency in Indian and other countries experience. *Journal of advanced research in hr and organizational management*, *8*(1&2), 23-30.

[14] Hatami-Marbini, A., & Arabmaldar, A. (2021). Robustness of Farrell cost efficiency measurement under data perturbations: evidence from a US manufacturing application. *European journal of operational research*, *295*(2), 604-620.

[15] Hosseinzadeh Lotfi, F., Jahanshahloo, G. R., Shahverdi, R., & Rostamy-Malkhalifeh, M. (2007). Cost efficiency and cost Malmquist productivity index with interval data. *International mathematical forum*, 2(9), 441-453.

[16] Jahanshahloo, G. R., Hosseinzadeh Lotfi, F., Alimardani Jondabeh, M., Banihashemi, S., & Lakzaie, L. (2008). Cost efficiency measurement with certain price on fuzzy data and application in insurance organization. *Appl. math. sci*, *2*(1), 1-18.

[17] Khodabakhshi, M., & Cheraghali, Z. (2022). Ranking of Iranian executive agencies using audit court budget split indexes and data envelopment analysis. *Journal of applied research on industrial engineering*, *9*(3), 312-322.

[18] Kholmuminov, S., & Wright, R. E. (2017). *Cost efficiency analysis of public higher education institutions in Uzbekistan*. Retrieved from https://pureportal.strath.ac.uk/en/publications/cost-efficiency-analysis-of-public-higher-education-institutions-

[19] Li, Z., Crook, J., & Andreeva, G. (2017). Dynamic prediction of financial distress using Malmquist DEA. *Expert systems with applications*, *80*, 94-106.

[20] Maghbouli, M., & Moradi, F. (2021). The relation between emotional intelligence (EQ) and mathematics performance: a DEA-based approach. *Journal of applied research on industrial engineering*, *8*(Special Issue), 1-14.

[21] Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de estadística*, *4*(2), 209-242.

[22] Pourmahmoud, J. (2015). New model for ranking DMUs in DDEA as a special case. *International journal industrial mathematics*, *7*(2), 187-192.

[23] Pourmahmoud, J., & Sharak, N. B. (2018). Measuring cost efficiency with new fuzzy DEA models. *International journal of fuzzy systems*, *20*(1), 155-162.

[24] Pourmahmoud, J., & Sharak, N. B. (2020). Evaluating cost efficiency using fuzzy data envelopment analysis method. *Iranian journal of operations research*, *11*(1), 25-42.

[25] Puri, J., & Yadav, S. P. (2016). A fully fuzzy DEA approach for cost and revenue efficiency measurements in the presence of undesirable outputs and its application to the banking sector in India. *International journal of fuzzy systems*, *18*, 212-226.

[26] Soleimani-Chamkhorami, K., & Ghobadi, S. (2021). Cost-efficiency under inter-temporal dependence. *Annals of operations research*, *302*(1), 289-312.

[27] Sun, J., Chen, M., Fu, Y., & Luo, H. (2021). Allocating fixed resources for DMUs with interval data. *RAIRO-operations research*, *55*(2), 505-520.

[28] Taeb, Z., Hosseinzadeh Lotfi, F., & Abbasbandy, S. (2017). Determine the efficiency of time depended units by using data envelopment analysis. *International journal of research in industrial engineering*, *6*(3), 193-201.

[29] Tone, K. (2002). A strange case of the cost and allocative efficiencies in DEA. *Journal of the operational research society*, *53*(11), 1225-1231.

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 97–112.

Paper Type: Research Paper

A Data-Driven Design for Gas Turbine Exit Temperature Spread Condition Monitoring System

Nastaran Hajarian[1], Farzad Movahedi Sobhani[1,*], Seyed Jafar Sadjadi[1]

[1] Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran; n.hajarian@srbiau.ac.ir; fmovahedi@iau.ac.ir.

[2] Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran; sjsadjadi@iust.ac.ir.

Citation:

Hajarian, N., & Movahedi Sobhani, F., & Sadjadi, S. J. (2023). A data-driven design for gas turbine exit temperature spread condition monitoring system. *Journal of applied research on industrial engineering, 10*(1), 97-112.

Received: 27/07/2021 Reviewed: 01/08/2021 Revised: 09/10/2021 Accepted: 01/12/2021

Abstract

One of the most complex and costly systems in the industry is the Gas Turbine (GT). Because of the complexity of these assets, various indicators have been used to monitor the health condition of different parts of the GT. Turbine Exit Temperature (TET) spread is one of the significant indicators that help monitor and detect faults such as overall engine deterioration and burner fault. The goal of this article is to use data-driven approaches to monitor TET data to detect faults early, as fault detection can have a significant impact on GT reliability and availability. In this study, the TET data of v94.2 GT is measured by six temperature transmitters to show a detailed profile. According to the statistical tests, TET data are high dimensional and time-dependent in the real world industry. Hence, three distinctive methods in the field of the GT are proposed in this study for early fault detection. Conventional Principal Component Analysis (PCA), Moving Window Principal Component Analysis (MWPCA), and Incremental Principal Component Analysis (IPCA) were implemented on TET data. According to the results, the conventional PCA model is a non-adaptive method, and the false alarm rate is high due to the incompatibility of this approach and the process. The MWPCA based on V-step-ahead and IPCA approaches overcame the non-stationary problem and reduced the false alarm rate. In fact, these approaches can distinguish between the normal time-varying and slow ramp fault processes. The results showed that IPCA could detect fault situations faster than MWPCA based on V-step-ahead in this study.

Keywords: Early fault detection, Data-driven, Gas turbine exit temperature, Time-varying, PCA model, MWPCA model, IPCA model.

1 | Introduction

As the demand for systems safety and product quality increases, the role of process monitoring in industrial procedures becomes more prominent [1]. One of the most complex and costly systems in the industry is the Gas Turbine (GT). There is increasing attention to condition monitoring of GTs in power plants operations and maintenance. The main reasons for this consideration are to enhance the reliability and availability of these valuable assets [2]. As a result, employing the monitoring strategy appears to be important to ensure that faulty conditions are detected early and that a forced plant shutdown is avoided [3]. Because of the complexity of these assets, various indicators have been used to monitor the health condition of different parts of GT, such as coast down-time, vibration, performance, maximum turbine outlet temperature at minimum fuel flow,

Corresponding Author: fmovahedi@iau.ac.ir

https://doi.org/10.22105/jarie.2021.296470.1361

Turbine Exit Temperature (TET) spread, Bearing temperature during coast down, etc. [4]. “TET spread” is one of the essential indicators that helps monitor and detect the presence of faults such as overall engine deterioration and burner fault.

There are several conceptual approaches to detecting and diagnosing failure, as well as distinct categories presented by different researchers. None of these classifications are comprehensive, and each researcher has his or her own viewpoint. Zhang et al. [5] argued that fault detection methods could be either model-based or data-based methods (*Fig. 1*). However, Chiang et al. [6] believed that three fault detecting and diagnosing methods are data-driven, analytical-based, and knowledge-based methods.

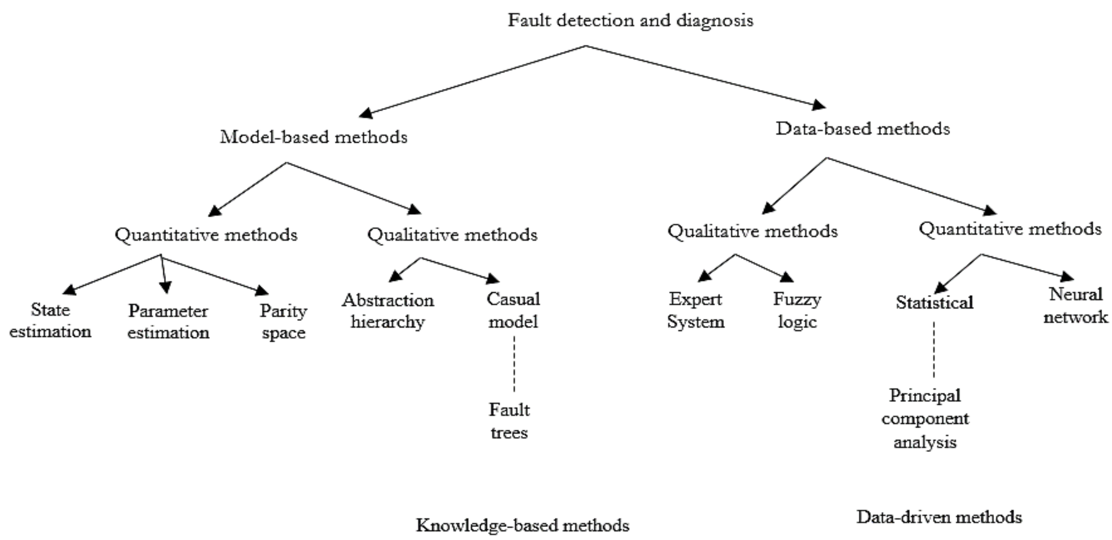


Fig. 1. Classification of fault detection and diagnosis method [5].

The analytical approach utilizes first principles to construct mathematical models of the system [7]. The analytical model cannot be applied for large-scale and complex systems. According to [8], the knowledge-based methods are mostly rule-based expert systems. The failure cases and engineers' experience are used to formulate the rules. When the detailed mathematical model cannot be reached, and when the number of inputs, outputs, and states of a system is logically small, the knowledge-based method will yield the best results [8]. The data-driven methods use product life-cycle data to detect faults and they do not depend on the first-principles models [9], [10].

Hence, data-driven methods can be used for large-scale and complex systems which are inexpensive and high accuracy [7], [11]. Upon the availability of the data of product or system, data-driven methods are chosen, but the system model is not [12]. As *Fig. 1* shows, data-driven methods consist of statistical and non-statistical methods [5]. One of the data-driven statistical tools is Principal Component Analysis (PCA) models. Instead of using base-models, we introduced PCA models for monitoring and identifying early failures in a GT using TET spread indicators in this study. Because the basic principles for constructing mathematical models of GTs are not readily available, and such a model is typically difficult to obtain due to the complexity and high dimensionality of a GT, there appears to be sufficient historical data for employing a statistical method.

Statistical Process Monitoring (SPM) developed the quality control charts to detect a system deviation from the normal behavior. When the number of variables or dimensions of a problem is beyond one value, it is necessary to use a multivariate statistical approach [13].

Monitoring, in a power plant context, means assessing the measured data in the simplest form to distinguish data of the normal operation from abnormal data. Nowadays, GTs are equipped with a lot of sensors, of which the collected data are used for monitoring purposes [14].

The idea of monitoring the TET spread monitoring has been proposed for more than two decades, and different methods have been suggested until now [4], [15]-[21]. Various research has been done in TET spread monitoring since 1992 [4], [15]-[21]. All of them mentioned the importance of TET monitoring to prevent catastrophic GT damages. In an ideal condition, the measurements of all thermocouples in exhaust must be the same. Nevertheless, in the real world, it is not. Knowles [17] described all possible reasons for this fact and suggested monitoring temperature patterns instead of temperature spread. In this method, a pattern is extracted from the healthy condition of the GT to compare the current TET pattern. Tsalavoutas et al. [21] presented a statistical method to evaluate the changes in temperature patterns. None of these researchers considered the influence of operation status and ambient conditions on their models.

Some researchers applied TET to diagnose a specific failure. Medina et al. [19] focused on combustion chambers failure detection with the aid of TET monitoring. They developed a TET model based on the basic principles of a GT, which estimates the TET, then compared the actual temperature measurement with the TET estimation to detect each combustor chamber failure. Korczewski [18] delivered a method to detect the failures of the automatic engine control system according to the alteration in TET during transient conditions such as start-up and acceleration processes. Besides, he presented diagnostic tolerances based on the statistical quality control methods. Jinfu et al. [16] presented an early fault detection method of the hot component. They introduced an indicator to detect the early faults of the hot component in the GT. Kenyon et al. [22] proposed an anomaly detection in exhaust gas temperature system by data mining algorithm to monitor the related parameters when anomalies are identified. Navi et al. [23] proposed partial kernel PCA for sensor fault detection of an industrial GT. Palmé et al. [20] used Auto Associative Kernel Regression method to simulate the TET pattern. In this method, the value of each thermocouple measurement would be estimated according to the previous records of exhaust temperature. Literature review shows that previous studies were often in the field of fault detection of the GT with an analytical approach, and there was no special attention to the important point that the trend of TET measured data is high dimensional and non-stationary. They are affected by operation status and ambient conditions.

According to our findings, no research has been conducted on early fault detection by monitoring TET spread using a statistical approach with non-stationary and high-dimensional data assumptions. This study aims to find an appropriate approach for early fault detection of turbine gas using TET spread, with a focus on data that will change over time, namely non-stationary data. In this paper, a low and straightforward calculation cost method will be proposed. It will help us better understand the condition of the TET pattern during the encounters of any fault generation and propagation. This study presents two approaches to cope with high-dimensional and time-dependent features. The Moving Window Principal Component Analysis (MWPCA) and Incremental Principal Component Analysis (IPCA) are two new approaches that can solve the dimensionality and time-dependent problems.

This paper also looks into the basic PCA, MWPCA, and IPCA in Section 2. In addition, Section 3 presents the details of the results of implementation and fault detection methods based on the conventional PCA, MWPCA, and IPCA. Finally, the conclusion is given in Section 4.

2 | The Multivariate Statistical Approach based on PCA, MWPCA, IPCA

Researchers believe that among the monitoring of multivariate processes, PCA, partial least squares, canonical correlation analysis and factor analysis are the suitable monitoring approaches [24].
The mentioned approaches and their expansion indicate the capability of conducting a high-dimensional data process. All turn the high-dimensional process into a lower-dimensional subspace and control the process behavior accordingly [25]. In the PCA, the original space variables, usually correlated, are transformed linearly into a new space of variables; these new variables are uncorrelated or orthogonal to each other [26]. If $X \in R^{n\times m}$ is the standardized data matrix with zero mean and unit variance with the scale parameter vectors x and S as the mean and variance vectors, respectively, where n denotes the sample

number and m represents the variable number. Defining the covariance matrix R is the first step to obtain PCA:

$$R = \frac{1}{n-1} X^T X. \tag{1}$$

Then accomplishing Singular-Value Decomposition (SVD) decomposition on R is as follows:

$$R = V \Lambda V^T, \tag{2}$$

where Λ represents a diagonal matrix consisting of the non-negative real eigenvalues as decreasing ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$). Columns of matrix V are the eigenvectors of R. Based on r principal eigenvalues, the transformation matrix $P \in R^{m \times r}$ is generated by choosing r eigenvectors or columns of V. The space of the measured variables is turned into the reduced dimension space by matrix P:

$$T = XP. \tag{3}$$

The columns of P are called loading and items of T are called scores. Scores are the values of the original measured variables that have been converted into the reduced dimension space [27]. According to *Eq. (3)*, the scores can also be turned into the original space as follows

$$\widehat{X} = TP^T. \tag{4}$$

The residual matrix E is calculated as:

$$E = X - \widehat{X}. \tag{5}$$

Finally, the original data space can be computed as

$$X = TP^T + E. \tag{6}$$

The important part of the implementation of PCA is choosing the number of principal components, which should be done carefully because TP^T illustrates the main source of variability and E represents the variability known as noise [28]. Several methods are proposed for choosing the number of principal components. SCREE procedure and Cumulative Percent Variance (CPV) approach are the most popular methods.

A plot of the eigenvalues is built in descending order by the SCREE procedure [28], which is a graphical approach that detects the knee in the curve. The number of knees on the plot indicates the count of principal components. CPV is the other method [29], [30] which determines the percent variance ($CPV(r \geq \%90$) calculated by the first r principal components as follows [31]:

$$CPV\ r) = \frac{\sum_{i=1}^{r} \lambda_i}{trace(R)}. \tag{7}$$

2.1 | Multivariate Statistical Process Control based on PCA

After constructing a PCA model based on the historical data collected, it is necessary to have an instrument that controls variation. It is possible to plot the multivariate control charts using the Hotelling T^2 and Square Prediction Error (SPE) or Q to detect the fault. Determining two orthogonal subspaces of the original space can decrease the monitoring of these two variables (T^2 and Q).

The significant variation and the random noise in the data can be controlled by T^2 and Q, respectively. The T^2 statistic can be calculated for each new observation x by:

$$T^2 = x^T P \Lambda_r^{-1} P^T x, \tag{8}$$

where Λ_r is the squared matrix constructed by the first r rows and columns of Λ and, as previously mentioned, P is r eigenvectors or columns of V. The upper confidence limit for T^2 is acquired using the F-distribution:

Hajarian et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

$$T_\alpha^2 = \frac{r(n-1)}{n-r} F_{\alpha,r,n-r}, \tag{9}$$

where r is the number of the principal components and n is the number of samples in the data and α is the level of significance. A violation of the threshold would mean that variations of the system are out of control. Another statistic is the SPE or Q that can monitor the portion of the measurement space related to the lowest m-r eigenvalues. In fact, the Q statistic is calculated as the sum of squares of residuals.

$$Q = x^T\left(I - PP^T\right)x, \tag{10}$$

where I is the identity matrix.

The upper confidence limit for the Q can be calculated from its approximate distribution:

$$Q_\alpha = \theta_1\left[\frac{C_\alpha\sqrt{2\,\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0\ h_0 - 1)}{\theta_1^2} + 1\right]^{1/h_0}, \quad h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}, \theta_i = \sum_{j=r+1}^{m} \lambda_j^i, \tag{11}$$

where C_α is the value of the normal distribution with the α level significance. A violation of the threshold would indicate that an unusual event has occurred that had changed the covariance structure of the model [28].

Monitoring and fault detection based on the PCA model considers two steps:

1. Off-line: Acquire training data, collected under normal operation, this matrix must be normalized to zero mean and unit variance with the scale parameter vectors x and S as the mean and variance vectors, respectively. Then, we should implement PCA algorithm, find eigenvalue and eigenvector and determine the number of principal components and finally the upper control limits for T^2 and Q statistics.
2. Online:

   – *Get a new instance and scale it using the scale parameter vectors x and s.*
   – *Calculate $T^2$and Q statistics using the result of PCA model.*
   – *When the value of statistics is compared to thresholds, the violation is interpreted as an alarm.*
   – *Repeat from step A.*

The monitoring process with PCA is indicated in *Fig. 2*.

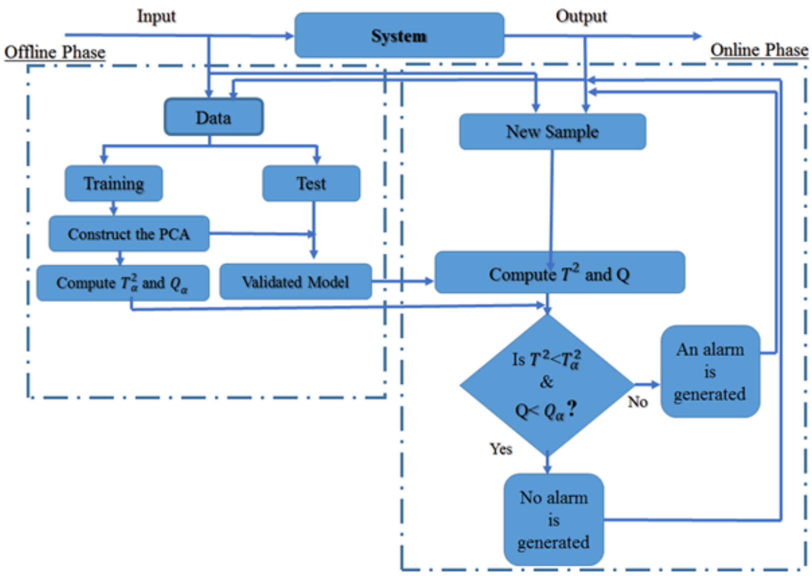


Fig. 2. Monitoring process with PCA method diagram [32].

2.2 | Multivariate Statistical Process Control Based on MWPCA

The process parameters, such as the mean or covariance, change with time, making the process non-stationary [33]. Several complementary MSPM methods have been introduced to tackle the time-varying issue. Three classes of approaches of Recursive Principal Component Analysis (RPCA), MWPCA and IPCA were applied to develop PCA methods to address non-stationary data [34].

Ketelaere et al. [35] investigated PCA-based statistical process-monitoring methods in terms of time-dependent and high-dimensional data, including MWPCA and RPCA [35]. RPCA techniques update the model for ever-increasing data consisting of new samples without discarding the old ones. Although RPCA is theoretically simple, it has been successfully employed for process monitoring. However, its implementation might not be easy for two main reasons: the ever-growing data set on which the model is updated, which eventually slows down the speed of adaptation as the data size increases. RPCA also consists of older data that are unrepresentative of the time-varying process. The forgetting factor cannot be easily selected without a priori knowledge of likely fault conditions when given to down-weight older samples [36]. The MWPCA method can tackle some of the limitations mentioned above by gathering a sufficient number of data points in the time-window to help build an adaptive process. Specifically, MWPCA removes older samples to choose the new samples representing the current operation process. Hence, for window size H, the data matrix at time t is $X_t=(x_{t-H+1}, x_{t-H+2}, \ldots, x_t)'$ and, at time t + 1, it is $X_{t+1}=(x_{t-H+2}, x_{t-H+3}, \ldots, x_{t+1})'$. The observations in the new window can be used to obtain the updated $\bar{x}_{t+1}$ and s_{t+1} [37]. The window includes a number of samples to cover enough process variation for modeling and monitoring purposes. Thus, window size is important. If a high number of samples select for the window, the MWPCA computation speed reduces drastically. If the result data used to enhance the computational efficiency is in a smaller window size, the relationship between the process variables will be important. If the model's adaptability to the process changes rapidly, it will be difficult to notice abnormal behavior, and the narrow window will be risky. Chiang et al. [6] determined the window size needed to estimate the T^2-statistic accurately. This was done according to the convergence of the T^2 distribution to the F distribution suggesting that minimum window sizes should be greater than approximately 10 times the number of variables. The monitoring process with MWPCA is indicated in *Fig. 3*. It is worth noting that over-fitting might be observed in MWPCA and a slow ramp could not be detected. As a result of the introduction of V-step-ahead prediction, MWPCA is based on the application delay. This approach is implemented using a model estimated at time t to predict the system behavior at time t + V and observe the likely faults. This step is used to ensure that the model does not over-adapt to the data and that it can detect faults that develop over time and are identified as regular observations at each time point [36], [38].

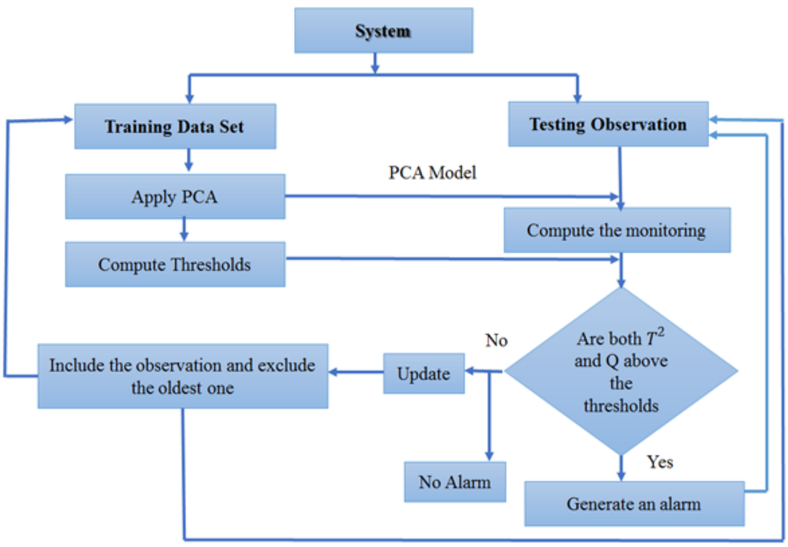


Fig. 3. Monitoring process with MWPCA method diagram [32].

Hajarian et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

2.3 | Multivariate Statistical Process Control based on IPCA

In adaptive methods, the monitoring models are updated by new sample and the normal time-varying information is added to the monitoring model to recognize between the normal time-varying and slow ramp fault processes. Another approach is IPCA.

Unlike the adaptive method, this method presents a novel approach in which the monitoring model is constant. When the PCA model implements time-varying process data, the PCs are also time-varying. This method introduces a new parameter as Incremental Principal Component (IPC). IPC calculate as follows:

$$\mathrm{IPC_i} = \mathrm{mean}\big(\mathrm{PC_i}\ \mathrm{k-L{:}\,k}\big) - \mathrm{PC_i}\ \mathrm{k-W-2L{:}\,k-W-L}). \quad (12)$$

The IPC is proposed to define the variations of the PCs. IPCs explain time-varying information. L indicates the number of PCs used to calculate the mean subtracted, and W indicates the interval of the two moments ($W > L$). A good introduction to IPCA can be found in [39]. IPCA model contains two steps as follows:

1. Offline:

   I. Acquire training data set A, which is collected under normal operation. This matrix must be normalized to zero mean and unit variance with the scale parameter vectors $\bar{x}$ and S as the mean and variance vectors, respectively.
   II. Construct the conventional PCA model with training data set, using the eigenvalue decomposition algorithm without reducing the dimension, and keep the loading matrix P and PCs.
   III. Compute the IPCs according to *Eq. (12)* in dataset A, and then compute the corresponding $IT^2$of each sample with the IPCs according to *Eq. (13)*.

$$\mathrm{IT_i^2} = \mathrm{IPC_i^T}\Lambda_{\mathrm{IPC}}^{-1}\mathrm{IPC_i}. \quad (13)$$

   IV. Use the Kernel Density Estimation (KDE) algorithm to define the control limit CL_j and CL of IT^2 as well as the threshold CL_j of IPC_j with the 99% confidence level.

2. Online step:

   I. Collect N_1 normal online samples, where $N_1 > W + 2 * L$. Normalize the samples through the means and variances of the training dataset A, and compute the PCs with the loading matrix P.
   II. Collect a new online sample. Normalize the sample with the means and variances of the training dataset A and calculate the PCs with the loading matrix P of the PCA model.
   III. Compute IPC_i of the i^{th} sample with the PCs of the forward W +2*L samples through *Eq. (12)*.
   IV. Compute the statistic $IT_i^2$of the i^{th}sample with the IPC_i of the i^{th} sample through *Eq. (13)*.
   V. If $IT_i^2 > CL$, then a fault may be present in the process. Otherwise, the sample is a normal one. In addition, when $IPC_{i,j}$>CL_j, the corresponding $PC_{i,j}$ is replaced by $PC_{i-i,j}$.
   VI. Repeat (II) – (III).

The monitoring process with IPCA is indicated in *Fig. 4*.

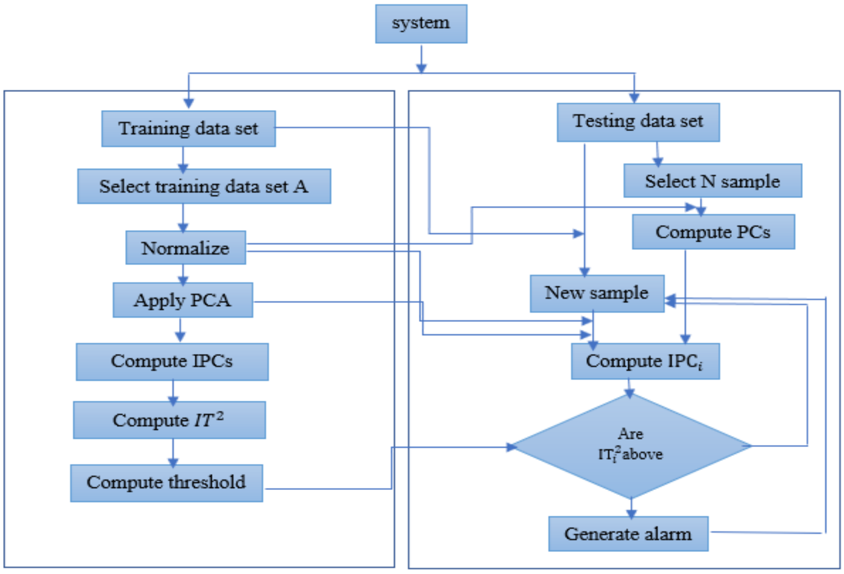


Fig. 4. Monitoring process IPCA method diagram.

3 | Static PCA, MWPCA Based on V-Step-Ahead and IPCA Applied to TET Spread

GTs are designed for many different purposes. In the industry, they are commonly used to drive compressors to transport gas through pipelines and generators that produce electrical power [40], [41]. In the past, the use of GTs has traditionally been limited to generating power during times of peak demand. Still, nowadays, they are being used in combined cycle power plants for baseload production [42]. Consequently, their availability, as well as reliability, play a significant role in these machines. The development of the GT in recent years has been facilitated most considerably by three factors:

I. Metallurgical developments that can be used to apply high temperatures in the combustor and turbine components.
II. Increased underlying knowledge of aerodynamics and thermodynamics.
III. Designing and simulating turbine airfoils and combustor and turbine blade cooling configurations by computer software.

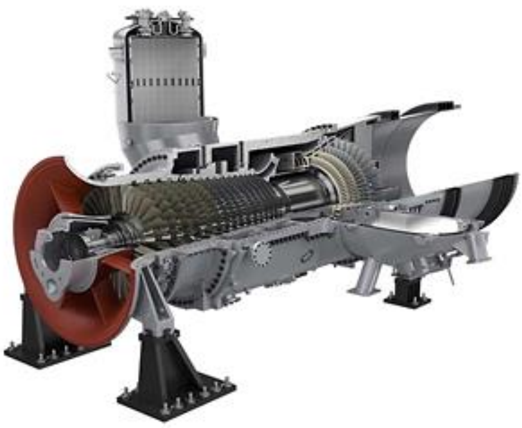


Fig. 5. Siemens V94.2 GT, from left to right: compressor, combustion chamber, turbine [43].

After compressing the air in the compressor, the fuel will be injected into it and combustion will increase the temperature of the gas. Turbine Inlet Temperature (TIT) is the average temperature of the flue gas that will face the first stage of turbine blades. During the expansion of the flue gas in the turbine, the pressure and temperature will decrease and the flue gas will leave the turbine with TET. The overview of a V94.2 GT manufactured by Siemens is shown in *Fig. 5*.

If TIT could be increased, GT's efficiency and specific power would improve. Nonetheless, there is a technological limit to designing and producing turbines that can withstand larger TIT. Since TIT is too hot to be measured directly, it is usually calculated by measuring TET. Both TET and TIT have a profile on their section due to the flue gas stream's rotation and turbulence. In v94.2 GTs, TET is being measured by six temperature transmitters to show a detailed profile. GT manufacturers use various methods to calculate the TIT regarding the measured TET. The operator keeps the GT in a protected condition by monitoring the TET.

The TET of an Iranian GT company is used to show the behavior of the methods throughout this study. As shown in *Fig. 6*, the data consists of six sensors, with each sensor representing the TET of V94-2 GT measured approximately with 1500-minute intervals.

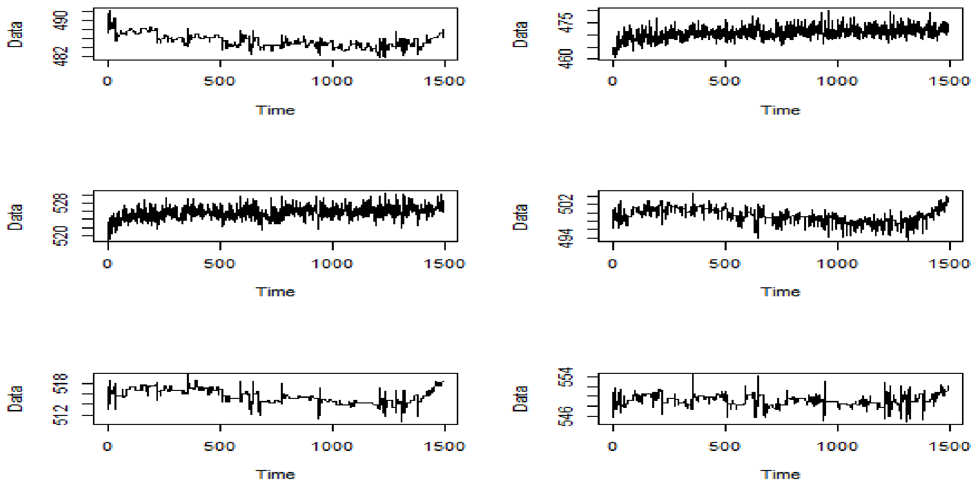


Fig. 6. Trends of TET data for normal process.

These data were collected when all parameters were under control. Statistical tests were implemented to detect the behavior of the data. These data are non-stationary following the test of Kwiatkowski–Phillips–Schmidt–Shin (KPSS). In this subsection, a static PCA model was applied on TET data for controlling the behavior of GT and early detecting fault.

PCA model was introduced into that data along with the results of the observations. Since all data was present, missing data methods were not used. Preprocessing is normally done in various fields. The type of preprocessing depends on the type of process. In the case of the TET data, no special preprocessing is necessary; standardizing the data was the only necessary procedure. After static PCA was applied to the TET data, three components were retained following the CPV criterion. The T^2- and Q-statistics were plotted. The first 400 observations were used to fit the underlying model. No considerable change in the vibration intensity of any of the sensors during this period was observed. The estimated model was evaluated against the well-behaved data observed before t = 400, and the confidence limit was set to 99%. Therefore, a distinction is made between phase I, which occurs when t > 400, and phase II. As shown in *Fig. 7*, conventional PCA cannot be a good instrument for TET data since the mean of data changes over time, and static PCA created a model with the first bunch of data and was unable to update the model over time. As a result, the model generates a false alarm, despite the fact that these changes are an essential element of the system. The first chart from the left is T^2 control chart and the second chart indicate monitoring by Q.

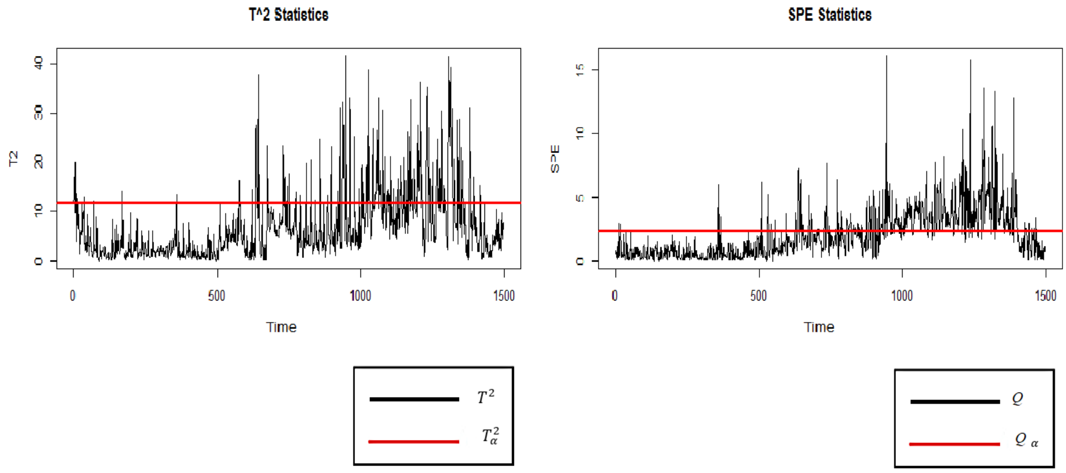


Fig. 7. Monitoring performance of conventional PCA for normal process.

As explained in the previous section, the MWPCA based on V-step-ahead prediction and IPCA were applied to the TET data to solve this problem. The T^2 and Q-statistics regarding MWPCA based on V-step-ahead were plotted in *Fig. 8*. Considering the points mentioned in the previous section, the opinion expert window size and delay size were 400 and 40, respectively. The confidence limit was set to 99%. The first chart from the left is T^2 control chart and the second chart indicate monitoring by Q.

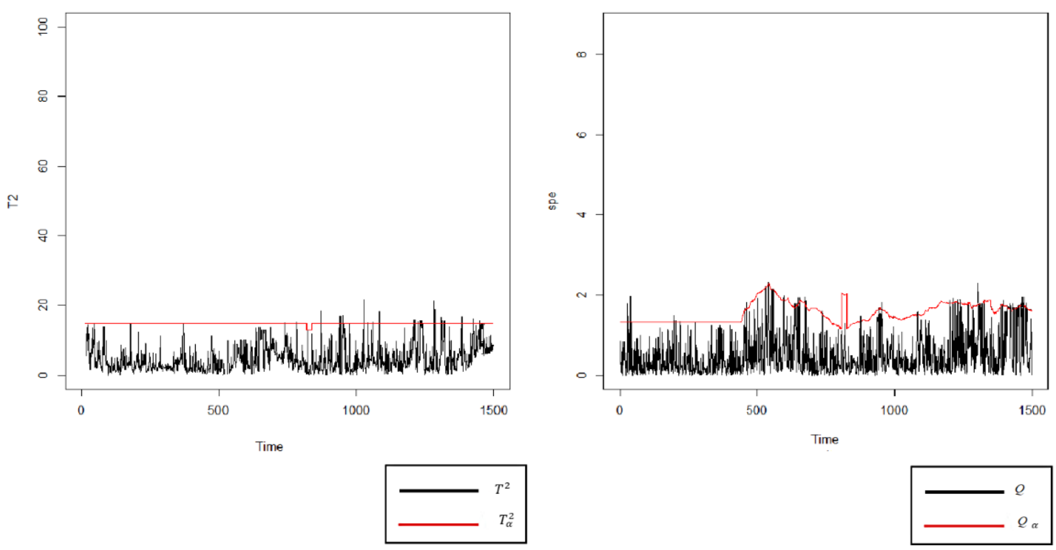


Fig. 8. Monitoring performance of MWPCA based on V-step-ahead for normal process.

Also, the IPCA method was applied to TET data. As shown in *Fig. 9*, the PCs are also non-stationary because the TET data is time-varying. The red lines represent the means of all PCs, and as can be seen, the PCs gradually deviate from the means. IPCs are computed through *Eq. (12)*, and the parameters L and W are set as 6 and 40, respectively. The results of IPCs were plotted in *Fig. 10*. This figure shows IPCs remained around mean and a statistic made by the IPCs will not increase slowly for the normal time-varying process.

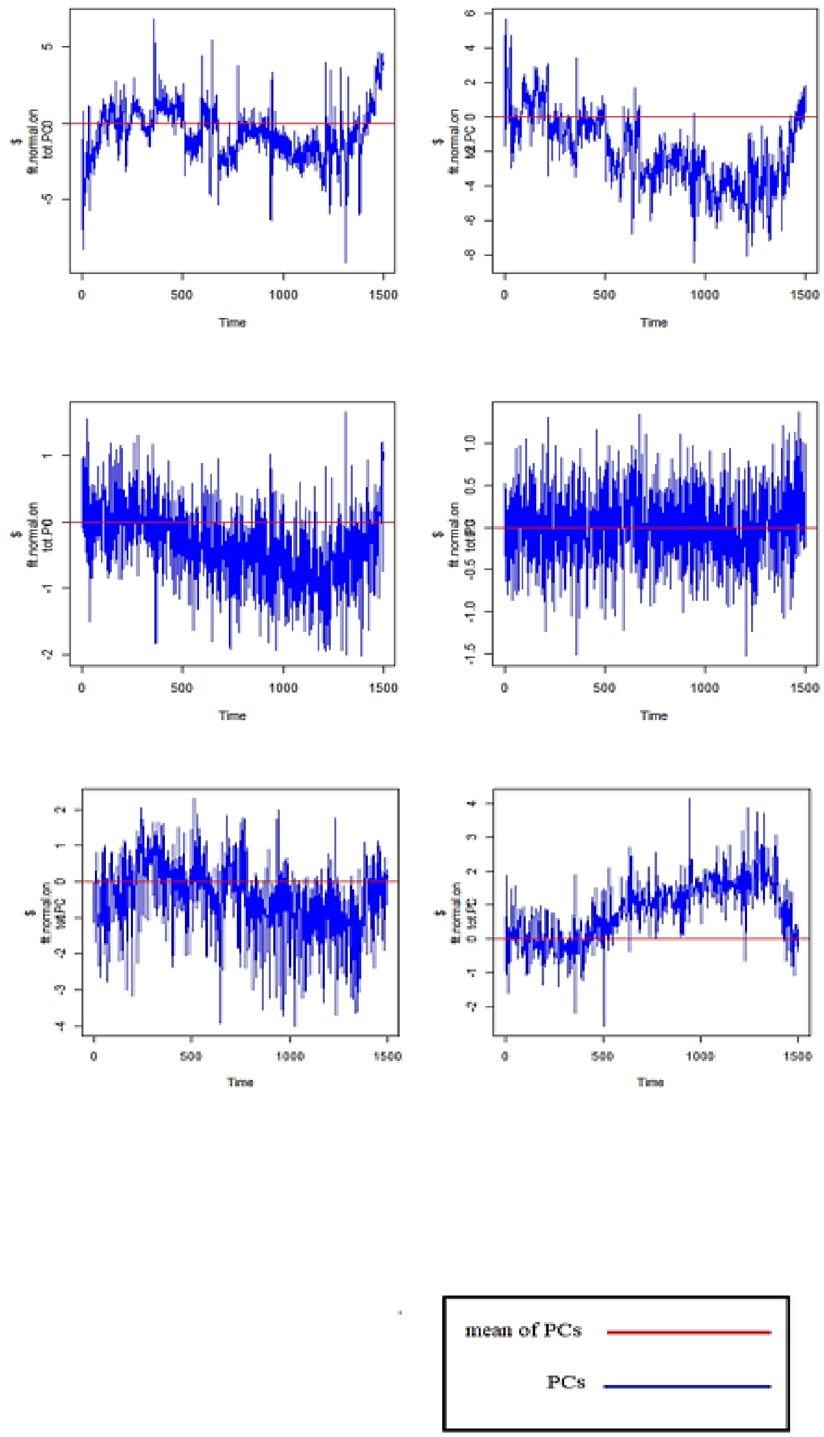


Fig. 9. Trends of PCs for normal process.

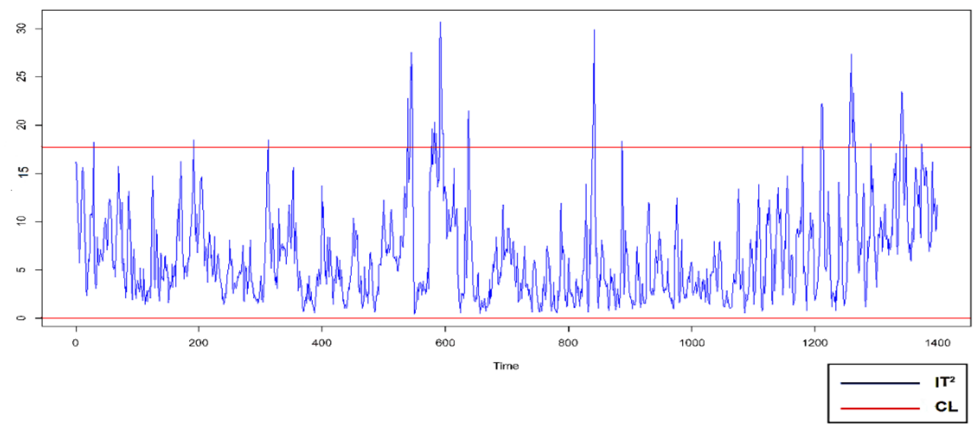


Fig. 10. Trends of PCs for normal process.

Hajarian et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

Then IT^2 was plotted in *Fig. 11*.

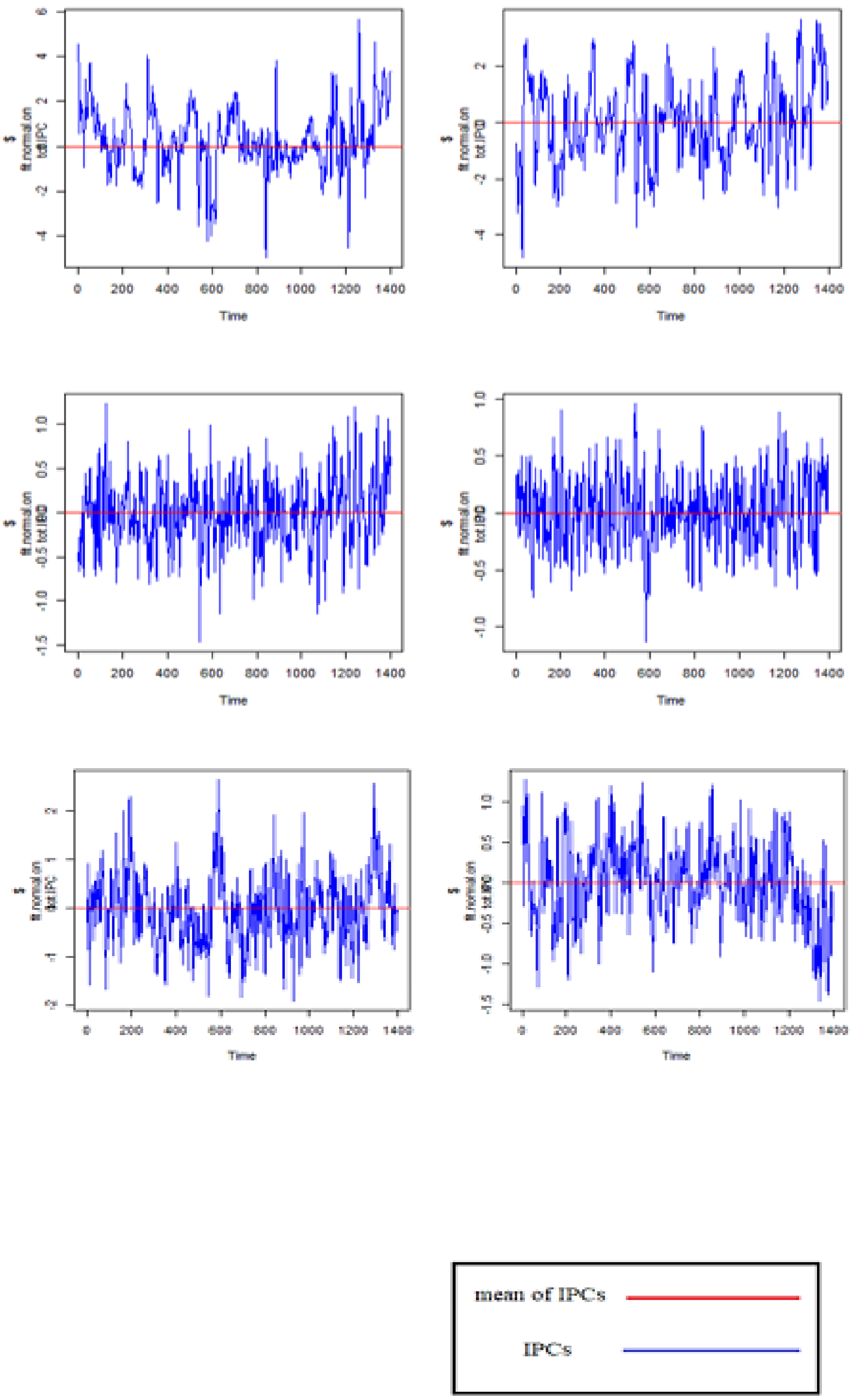


Fig. 11. Monitoring performance of IPCA process with fault 1.

A good fault detection technique should have robust against the training data set and react quickly in fault detection. The robustness is measured by computing the False Alarm Rate (FAR) upon fault free testing data set. Promptness fault detection is quantified by calculating Delay Time Detection (DTD) upon faulty testing data set. According to *Eq. (14)*, a FAR was calculated [44].

$$\text{FAR} = \frac{\text{Number of normal samples above the limits}}{\text{total number of normal sample}} * 100. \tag{14}$$

The results show that both MWPCA based on V-step-ahead and IPCA methods overcome non-stationary in this case. *Table 1* shows the percentage of false alarms rate and shows that IPCA performs better than MWPCA based on V-step-ahead.

Table 1. FAR percentage.

	MWPCA based on V-step-ahead	IPCA
T^2	5.3%	3.5%
Q	4%	

Two types of faults were added deliberately to check the power of MWPCA based on V-step-ahead prediction and IPCA in finding fault early.

Fault 1. Step change of second TET is introduced at the 1001st sample.

Fault 2. Linear ramp with 0.3 increments of second TET again is introduced beginning from the 1001st. As it was explained in the previous section, the T^2 statistic describes the useful information about system variation, and Q statistic represents model error or noise information. In case of a problem, the covariance structure of the model will be altered and the statistics can represent it. The results of MWPCA based on V-step-ahead are shown in *Figs. 12* and *13*.

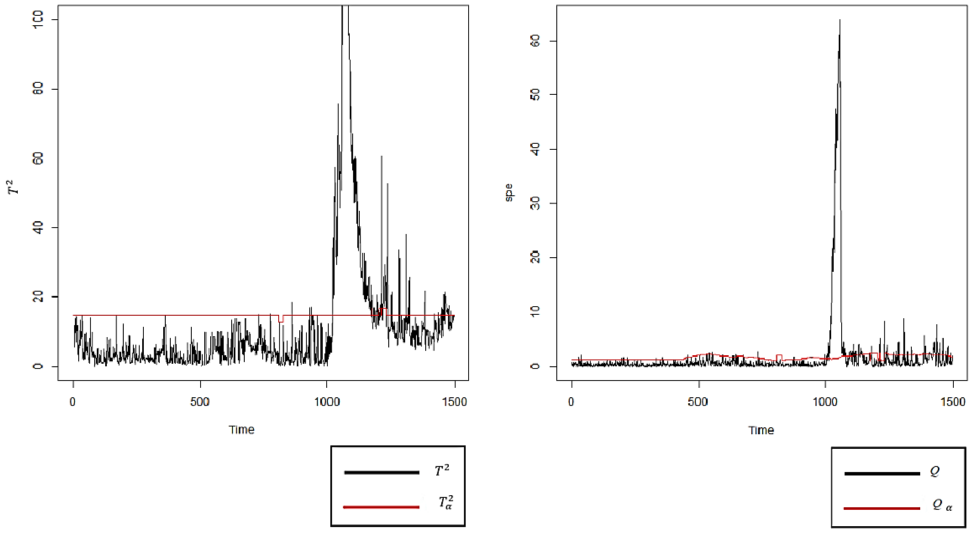


Fig. 12. Monitoring performance of MWPCA based on V-step-ahead for process with fault 1.

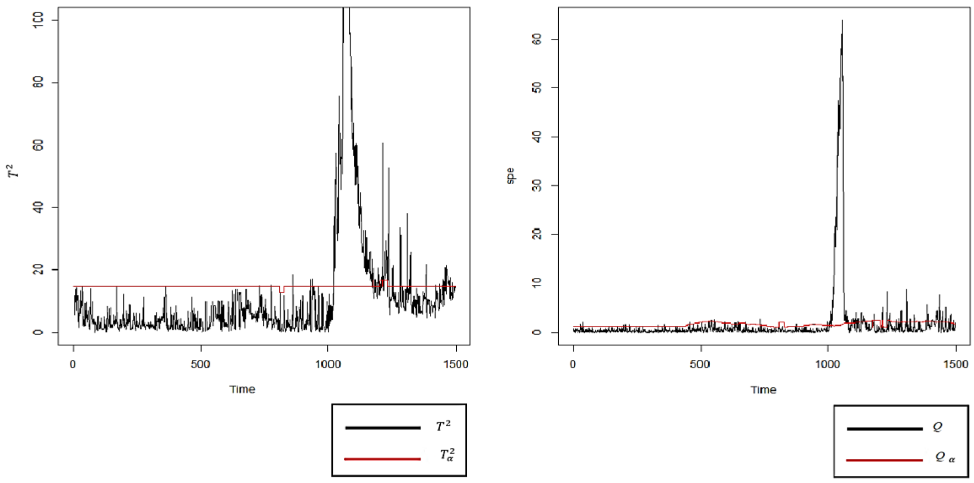


Fig. 13. Monitoring performance of MWPCA based on V-step-ahead process with fault 2.

The results of IPCA are presented in *Figs. 14* and *15*. The DTD index was used to compare these two methods:

$$\text{DTD} = \text{fault detection time} - \text{fault accorance time}. \tag{15}$$

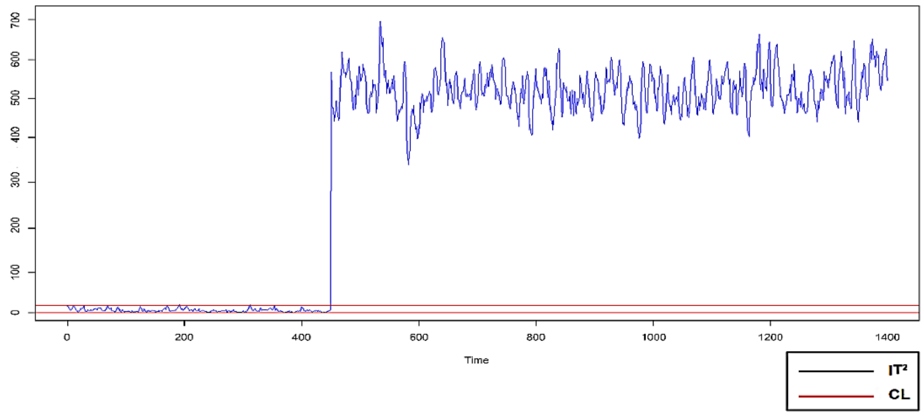


Fig. 14. Monitoring performance of IPCA process with fault 1.

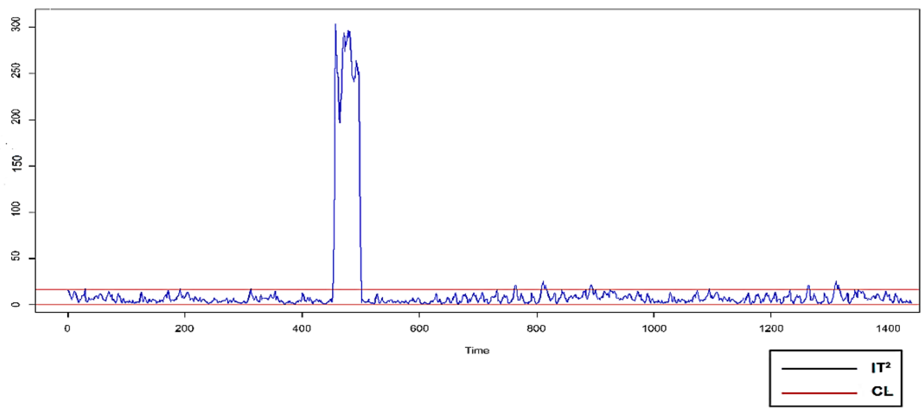


Fig. 15. Monitoring performance of IPCA process with fault 2.

As shown in *Table 2*, both of the methods easily detected the step-change process in time. Also, these charts can distinguish the normal time-varying and slow ramp fault processes.

But according to *Table 2*, slow ramp fault can be detected more quickly by the IPCA model in this case based on the TET data.

Table 2. Detection time delay.

	MWPCA based on V-Step-Ahead		IPCA
	T^2	Q	IT^2
Fault 1	0	0	0
Fault 2	18	11	9

4 | Conclusion

Early fault detection can play an important role in the reliability and availability of GTs. Therefore, providing a monitoring approach is necessary to guarantee early detection of faulty conditions before they lead to a forced plant shut down. The efficiency and specific power of GT would be improved if TIT could be increased. Since the TIT is too hot to be measured directly, it is usually calculated by measuring TET. This study used an appropriate data-driven approach for early fault detection of GTs. To detect faults, data-driven approaches rely on product life-cycle data rather than first-principles models. Hence, data-driven methods can be used for large-scale and complex systems; they are also cheap and inexpensive. PCA model is one of the data-driven methods. Still, unlike simulated data, actual data has ambient and system characteristics, and these parameters do not provide stationary data over

Hajarian et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

time. Finding an approach for early detection is important. In this study, the PCA model is implemented on six TET sensors, but it is not a suitable approach due to the non-stationary data. For this purpose, MWPCA based on V-step-ahead and IPCA are implemented on data. MWPCA based on V-step-ahead is a novel monitoring approach for non-stationary TET data. MWPCA based on V-step-ahead is an adaptable approach since it can update the monitoring model and control limit when the newly monitored sample is detected as a normal one. On the other hand, in the IPCA approach, the monitoring model remains unchanged and uses the new statistic called IT^2 to monitor data.

The results reveal that these approaches are data compatible and can detect a step-change fault in real time. However, when it comes to incremental ramp faults, the IPCA approach performs better. According to TET data behavior which is non-stationary and changes over time, a suitable approach has been recommended for early fault detection of turbine gas. In the future, there'll be more work to be done about how monitor high-dimensional, non-stationary, and autocorrelation data from GTs. It is also suggested to find the root of the faults by new fault isolation method and compare their results with classical methods.

References

[1] Sunadi, S., Purba, H. H., & Saroso, D. S. (2020). Statistical process control (SPC) method to improve the capability process of drop impact resistance: a case study at aluminum cans manufacturing industry in Indonesia. *Journal of applied research on industrial engineering*, *7*(1), 92-108.

[2] Molaei, S., & Cyrus, K. M. (2014). Robust design of maintenance scheduling considering engineering insurance using genetic algorithm. *International journal of research in industrial engineering*, *3*(1), 39-48.

[3] Liu, J., Liu, J., Yu, D., Kang, M., Yan, W., Wang, Z., & Pecht, M. G. (2018). Fault detection for GT hot components based on a convolutional neural network. *Energies*, *11*(8), 2149. https://doi.org/10.3390/en11082149

[4] Meher-Homji, C. B., & Bhargava, R. (1992). Condition monitoring and diagnostic aspects of gas turbine transient response. In *turbo expo: power for land, sea, and air* (78965, V004T11A006). American Society of Mechanical Engineers. https://doi.org/10.1115/92-GT-100

[5] Zhang, Y., & Jiang, J. (2008). Bibliographical review on reconfigurable fault-tolerant control systems. *Annual reviews in control*, *32*(2), 229-252.

[6] Chiang, L. H., Russell, E. L., & Braatz, R. D. (2000). *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media.

[7] Chiang, L. H., Russell, E. L., & Braatz, R. D. (2012). *Fault detection and diagnosis in industrial systems*. Springer London. https://doi.org/10.1007/978-1-4471-0347-9

[8] Fan, C. M., & Lu, Y. P. (2008). A bayesian framework to integrate knowledge-based and data-driven inference tools for reliable yield diagnoses. *2008 winter simulation conference* (pp. 2323-2329). IEEE.

[9] Alzghoul, A., Backe, B., Löfstrand, M., Byström, A., & Liljedahl, B. (2014). Comparing a knowledge-based and a data-driven method in querying data streams for system fault detection: a hydraulic drive system application. *Computers in industry*, *65*(8), 1126-1135.

[10] Yin, S. (2012). *Data-driven design of fault diagnosis systems* (Doctoral Dissertation, Duisburg, Essen, Universität Duisburg-Essen). https://core.ac.uk/download/pdf/33798712.pdf

[11] Meslameni, W., & Kamoun, T. (2021). Detection of an imbalance fault by vibration monitoring: case of a screw compressor. *Journal of applied research on industrial engineering*, *8*(1), 27-39.

[12] Sankavaram, C., Pattipati, B., Kodali, A., Pattipati, K., Azam, M., Kumar, S., & Pecht, M. (2009). Model-based and data-driven prognosis of automotive and electronic systems. *2009 IEEE international conference on automation science and engineering* (pp. 96-101). IEEE.

[13] Gol-Ahmadi, N., & Raissi, S. (2018). Residual lifetime prediction for multi-state system using control charts to monitor affecting noise factor on deterioration process. *Journal of applied research on industrial engineering*, *5*(1), 27-38.

[14] Losi, E., Venturini, M., Manservigi, L., Ceschini, G. F., & Bechini, G. (2019). Anomaly detection in gas turbine time series by means of bayesian hierarchical models. *Journal of engineering for gas turbines and power*, *141*(11), 111019. https://doi.org/10.1115/1.4044781

[15] Dundas, R. E., Sullivan, D. A., & Abegg, F. (1992). Performance monitoring of gas turbines for failure prevention. In *turbo expo: power for land, sea, and air* (78965, V004T10A013). American Society of Mechanical Engineers.

[16] Jinfu, L., Jiao, L., Jie, W., Zhongqi, W., & Daren, Y. (2017). Early fault detection of hot components in gas turbines. *Journal of engineering for gas turbines and power*, *139*(2). https://doi.org/10.1115/1.4034153

[17] Knowles, M. (1994). Gas turbine Temperature spread monitoring detection of combustion system deterioration. In *turbo expo: power for land, sea,* and air (78873, V005T17A002). American Society of Mechanical Engineers.

[18] Korczewski, Z. (2011). Exhaust gas temperature measurements in diagnostic examination of naval gas turbine engines. *Polish maritime research*, *18*(4), 49-53.

[19] Medina, P., Saez, D., & Roman, R. (2008). On line fault detection and isolation in gas turbine combustion chambers. In *turbo expo: power for land, sea, and air* (43123, pp. 315-324). https://doi.org/10.1115/GT2008-51316

[20] Palmé, T., Liard, F., & Therkorn, D. (2013). Similarity based modeling for turbine exit temperature spread monitoring on gas turbines. In *turbo expo: Power for land, sea, and air* (*55188*, p. V004T06A020). American Society of Mechanical Engineers.

[21] Tsalavoutas, A., Mathioudakis, K., & Smith, M. K. (1996). Processing of circumferential temperature distributions for the detection of gas turbine burner malfunctions. In *turbo expo: power for land, sea, and air* (78767, V005T15A010). American Society of Mechanical Engineers.

[22] Kenyon, A. D., Catterson, V. M., & McArthur, S. D. J. (2010). Development of an intelligent system for detection of exhaust gas temperature anomalies in gas turbines. *Insight-non-destructive testing and condition monitoring*, *52*(8), 419-423.

[23] Navi, M., Davoodi, M. R., & Meskin, N. (2015). Sensor fault detection and isolation of an industrial gas turbine using partial kernel PCA. *IFAC-papers online*, *48*(21), 1389-1396.

[24] Li, W., Peng, M., Liu, Y., Cheng, S., Jiang, N., & Wang, H. (2018). Condition monitoring of sensors in a NPP using optimized PCA. *Science and technology of nuclear installations*, *2018*, 1-16.

[25] Navi, M., Meskin, N., & Davoodi, M. (2018). Sensor fault detection and isolation of an industrial gas turbine using partial adaptive KPCA. *Journal of process control*, *64*, 37-48.

[26] Jolliffe, I. T. (2002). *Principal component analysis for special types of data* (pp. 338-372). Springer New York.

[27] Halligan, G. R., & Jagannathan, S. (2011). PCA-based fault isolation and prognosis with application to pump. *The international journal of advanced manufacturing technology*, *55*, 699-707.

[28] Jackson, J. E. (2005). *A user's guide to principal components*. John Wiley & Sons.

[29] Valle, S., Li, W., & Qin, S. J. (1999). Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods. *Industrial & engineering chemistry research*, *38*(11), 4389-4401.

[30] Jaffel, I., Taouali, O., Harkat, M. F., & Messaoud, H. (2015). Online process monitoring using a new PCMD index. *The international journal of advanced manufacturing technology*, *80*(5-8)), 947-957.

[31] Said, M., Fazai, R., Abdellafou, K. B., & Taouali, O. (2018). Decentralized fault detection and isolation using bond graph and PCA methods. *The international journal of advanced manufacturing technology*, *99*(1-4), 517-529.

[32] Li, W., Peng, M., & Wang, Q. (2018). False alarm reducing in PCA method for sensor fault detection in a nuclear power plant. *Annals of nuclear energy*, *118*, 131-139.

[33] Rato, T., Reis, M., Schmitt, E., Hubert, M., & De Ketelaere, B. (2016). A systematic comparison of PCA-based statistical process monitoring methods for high-dimensional, time-dependent processes. *AIChE journal*, *62*(5), 1478-1493.

[34] Gao, Y., Wang, X., Wang, Z., & Zhao, L. (2016). Fault detection in time-varying chemical process through incremental principal component analysis. *Chemometrics and intelligent laboratory systems*, *158*, 102-116.

[35] De Ketelaere, B., Hubert, M., & Schmitt, E. (2015). Overview of PCA-based statistical process-monitoring methods for time-dependent, high-dimensional data. *Journal of quality technology*, *47*(4), 318-335.

[36] Kruger, U., & Xie, L. (2012). *Statistical monitoring of complex multivatiate processes: with applications in industrial process control*. John Wiley & Sons.

[37] De Ketelaere, B., Rato, T., Schmitt, E., & Hubert, M. (2016). Statistical process monitoring of time-dependent data. *Quality engineering*, *28*(1), 127-142.

[38] Wang, X., Kruger, U., & Irwin, G. W. (2005). Process monitoring approach using fast moving window PCA. *Industrial & engineering chemistry research*, *44*(15), 5691-5702.

[39] Ding, S. X. (2014). *Data-driven design of fault diagnosis and fault-tolerant control systems*. Springer London.

[40] Langston, L. S., Opdyke, G., & Dykewood, E. (1997). Introduction to gas turbines for non-engineers. *Global gas turbine news*, *37*(2), 1-9.

[41] Ikpe, A. E., Iluobe, I. C., & Imonitie, D. I. E. (2020). Modelling and simulation of high pressure fogging air intake cooling unit of Omotosho phase II gas turbine power plant. *Journal of applied research on industrial engineering*, *7*(2), 121-136.

[42] International Energy Agency. (2006). *Energy technology perspectives*. https://www.iea.org/reports/energy-technology-perspectives-2006

[43] Siemens Energy. (2020). *Ingenious design to the core*. https://www.siemens-energy.com/global/en/offerings/power-generation/gas-turbines/sgt5-2000e.html

[44] Ammiche, M., Kouadri, A., & Bensmail, A. (2018). A modified moving window dynamic PCA with fuzzy logic filter and application to fault detection. *Chemometrics and intelligent laboratory systems*, *177*, 100-113.

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 113–124.

Paper Type: Research Paper

Estimating Production Losses from COVID-19 Pandemic Disruptions Based on Stock Returns

Obojobo Obukeajeta Donatus[1,*], **Chima Uzorh**[1]

[1] Department of Mechanical Engineering, Federal University of Technology Owerri, Nigeria; donatus4success@yahoo.com; a.cuzorh@yahoo.com.

Citation:

Donatus, O. O., & Uzorh, Ch. (2023). Estimating production losses from COVID-19 pandemic disruptions based on stock returns. *Journal of applied research on industrial engineering, 10*(1), 113-124.

Received: 29/09/2022 Reviewed: 01/11/2022 Revised: 09/12/2022 Accepted: 22/01/2023

Abstract

Economic or local disruptions that affect organizations' production activities often result in unexpected losses. An excellent example is the recent COVID-19 pandemic disruption which affected many economies globally. This study presents a deterministic model and uses simple regression analysis to estimate the average condition for production losses. Its corresponding components' input resources impact the overall estimates for selected organizations in Nigeria. It is anticipated that variability in economic activities is always accompanied by unconventional stock returns whose behavior indicates prevailing economic trends. Here we have looked at two organizations in the manufacturing sector as a case study Nigerian Breweries (NB) and Nestle Nigeria, whose stock prices [X] upon analysis reveal that at [X] ≤ N30 and [X] ≤ N821 are estimated conditions for zero net profit for both organizations respectively. Therefore, for NB, during the four quarters of the 2020 fiscal year, the following were assessed production losses, 3.47 billion Naira (Q_1), 4.17 billion Naira (Q_2), 3.72 billion Naira (Q_3) and 0.68 billion Naira (Q_4) with a total of 12.04 billion Naira annual estimated losses; with Costs of Goods Sold (COGS), Operating Expenses (OpEx) and Selling and Advertising Expenses (SAEX) having 39.6%, 44.5% and 15.9% impact on the estimates. Nestle Nigeria records estimated production losses of 5.8 billion Naira (Q_1), 6.4 billion Naira (Q_2), 4.2 billion Naira (Q_3), and -0.8 billion Naira (Q_4) (gain), resulting in a total 15.6 billion Naira annual estimated loss; and COGS, OpEx, and SAEX having 45.9%, 48.2% and 5.9% impact on the estimates respectively. This implies, SAEX had the most negligible percentage impact on overall estimated production losses for both organizations compared to COGS and OpEx. This study, therefore, reiterates the position of other economic reports describing the adverse effects of the pandemic in Nigeria; while also serving as an investment analysis guide to potential investors.

Keywords: Production losses, Estimation, Stock returns, Deterministic model, COVID-19 pandemic, Fiscal year.

1 | Introduction

Stock markets represent "systematic economic news", and their behavior is based on the outcomes of these news findings [6]. This news can be quantified in terms of a few driving variables like industrial production, risk premiums, inflation, and changes in inflation levels to analyze the behavior of stock prices [6]. The cost of a stock is the present value of cash flow that accrues to its owner [25]. Stock returns are used to measure the performance of a company stock [25]. Hence, stock returns are often viewed as economic health and performance indicators. For this reason, a company's financial performance for a given period can also be determined by examining stock returns occasioned by intense business disruptions, as in the case of the global COVID-19 pandemic disruption, which this research work is based on.

Corresponding Author: donatus4success@yahoo.com

https://doi.org/10.22105/jarie.2022.363902.1506

Many studies have attempted to examine the relationship between the financial ratios of companies and the performance of stock certificates. Upon review of relevant literature, studies based on profitability ratios can be grouped into two groups which are stock prices and stock return studies [25]. Although a wide range of studies examines the relationship between stock returns and a firm's profitability ratios, studies about the relationship between share prices and profitability rates are in more limited numbers [25]. These studies comprise the different research areas regarding methods and profitability ratios used, sectors, number of firms, and periods. An example is Cengiz and Püskül [27], who tried to identify profitability ratios related to stock returns. In one such research, earnings per share and book value per share were taken as independent variables and share prices as dependent variables. At the end of the regression data analysis, it was concluded that stocks prices movements are directly proportional to the profitability ratio. In other words, accounting information obtained from the company's balance sheet and income statements have a role in explaining stock price movement. They also examined and revealed the relationship between profitability and stock returns by identifying that an increase in equity and gross sales margin leads to a rise in stock returns.

In contrast, the increased operating profit margin results in decreased stock prices. In another study conducted in the Kuwaiti financial market between the years (2005-2014), the prediction power of 12 financial ratios was analyzed based on data set of 15 firms which took place in three different sectors; results show that the most effective ratios are Returns on Assets (ROA), Returns on Equity (ROE) and net profit ratio in the industry sector, while for service and investment sector they are ROA, ROE, Price-to-earnings ratio and earnings per share [25]. Although there is no perfect equation (relationship) that can predict exactly how the prices of stocks will move, there are market forces or factors that move a stock up or down. Still, ultimately, the price at any given moment is due to the supply and demand at that point in the market which is also subject to economic influence. However, these factors are grouped into fundamental, technical, and market sentiments. Fundamental factors drive stock prices based on the company's earnings and profitability from producing and selling goods and services. Technical factors related to a stock's price history in the market about chart patterns, momentum, and behavioral factors of traders and investors.

In contrast, the market sentiment is simply noise in the market. The investments made by traders based on sentiments, also known as "noise traders," is often compensated by the investment made due to mistaken judgments, thereby suggesting that the market is not affected by sudden noise but by informed investment decisions [22]. There for stock prices and returns can be evaluated using reliable estimation models to determine a company's financial performance. The financial performance of an organization and stock returns are somewhat related to each other [25]. However, the relationship between revenue growth and stock returns has been a puzzle in the corporate and academic environment. Revenue has consistently exhibited direct and significant effects on stock returns [3]. Many studies suggest a strong correlation between stock prices and the production output of organizations and industries. Fama [10] used regression models to justify this relationship between stock prices and real variables like production, cash flows, and Gross National Product (GDP) and growth rates of these variables.

Historically, the Nigerian stock market was established in 1960 but commenced active trading on June 5, 1961, first as the Lagos Stock Exchange; but later (in 1977) renamed the Nigerian Stock Exchange (NSE) [20]. Although trading commenced with few stocks like Nigerian Tobacco Company and Investment Company Limited, by December 2012, more than a hundred and ninety-six (196) equities were traded [20].

It is believed that business disruptions in any economic system greatly threaten organizations' success since incurred operational losses can be enormous. According to Pathak [1], disruption is a state of unbalance or disturbance that affects a system, economy, or nation, which can be an event or series of events causing damage to the normal functioning of these systems. It is a diversion from the state of the usual or expected. The disruptions have increased exponentially [8]-[12]. The average cost of natural disruptions has increased from $50 billion in the 1980s to $200 billion in recent years, and approximate

Obojobo and Uzorh | J. Appl. Res. Ind. Eng. 10(1) (2023) 113-124

losses worth $1.5 trillion were incurred between 2003 and 2013 [2]. In Nigeria's economic system report, recent data from the Nigeria Bureau of Statistics (NBS) indicates a decline in the performance of the industrial sector in the first quarter of 2020 on account of the impact of the COVID-19 pandemic, which led to a contraction in manufacturing activities and fall in both crude oil production and electricity generation. The estimated index of industrial production declined by 1.6 percent and 9.6 percent below the level in the preceding and corresponding quarter of 2019, respectively [5]. This decrease is attributed to a contraction in economic activities in all the subsectors during the review period [4]. Also, following the ease of previously imposed lockdown policy measures in some states of the federation to curtail the spread of the pandemic, depending on the firm and consumers' reactions, recovery was, therefore, slower or faster. Real GDP grew by 0.11 percent in 2020 Q4 compared with a contraction of 3.62 percent recorded in 2020 Q_3 and a growth of 2.55 percent in the corresponding quarter of 2019. Therefore, the annual GDP for 2020 contracted by 1.92 percent, compared with the growth of 2.27 percent recorded in 2019 [5]. This confirms the negative impact of the pandemic on the economy. Research has shown that economies often take time to return to business as usual. This, therefore, emphasizes the need to estimate production losses for the period under investigation accurately.

Undoubtedly, knowledge about how the global COVID-19 pandemic disruption affected production activities in Nigeria is imperative. This is necessary for loss estimation purposes and determining the contributions or impacts of various input resources on such estimates, which is a major challenge for organizations, governments, and stakeholders in every business environment. This study, therefore, aims to solve this problem for organizations to facilitate effective production planning and loss management during global or localized operational disruptions.

2 | Research Methodology

Research data for the following organizations were obtained from APT Securities and Funds Limited (member of the Nigeria Stock Exchange) at www.aptsecurities.com/nse-daily-price.php# and https://m.investing.com/equities/historical-data (Real-time NSE financial news provider) for analysis to estimate production losses of these organizations due to the global COVID-19 pandemic disruption using a deterministic model. Therefore, this model analyses the impact of the pandemic using stock price returns, which is believed to be a good representation of underlying economic and production activities in organizations during the affected period.

According to Pathak [1], a deterministic model is considered and used for effective analysis to predict the loss in production solely on the historical output data and input stock prices. This model was also used to estimate production losses from disruptions based on stock market returns as applicable to the 9/11 attacks, the Deepwater Horizon oil spill, and Hurricane Sandy. Many other studies like that of Fama [10], which made use of regression models to justify the relation between stock prices and real variables like production, cash flows, GDP, and the growth rates of these variables, also suggest a strong correlation of the stock prices with the production outputs of any industry. Therefore, consider a given organization (i) in an economy among (n) industrial organizations in a given sector and let the production output (net profit) for the organization (i) for the period (t) be denoted by (X_{it}). The production output (net profit) X_{it} is a linear function of the stock market index prices representing activities in (j) departments within an organization where (P_{jt}) is the stock market index price at time (t) for the organization (i) comprising of (j) departments or sections where J = 1, 2..., m. The linear coefficients relating the stock market index prices of the organization (i) to its production output (net profit) in any given economy is (a_{ij}), and (b_i) is the intercept. The production output (net profit) in the organization (i) at the period (t) is

$$X_{it} = \sum_{j=1}^{m} \left(a_{ij}P_{jt}\right) + b_i. \tag{1}$$

The regression coefficients (a_{ij}) and b_i will be calculated based on the average quarterly historical index stock prices and quarterly production output (net profit) for a given organization in the manufacturing sector under consideration for a given period. It is assumed that activities in these organizations reflect the larger economy.

Organization (i's) production may decline due to a disruptive event. We, therefore, assume that the average yearly production output (net profit) at the time step immediately before the disruptive event, t = 0, represents the average production output (net profit), and the production output (net profit) at any given time (t) afterward for the organization (i) as $X_{it.}$ L_{it} is the difference in production output at any given time, t, and reference pre-disruption time 0, which in this case represents the loss in actual production activities. This is based on the assumption that production would have been stable to a reasonable extent all through the year if not for the COVID-19 Pandemic, and changes in index stock prices accurately capture any fluctuation:

$$L_{it} = X_{io} - X_{it}. \tag{2}$$

Therefore, this formulation enables us to estimate production losses for organization (i) in terms of net profit losses for a given period (t). It is believed that for each period ranging from the first quarter of the year 2020, when the COVID-19 pandemic disruption impact was first noticed economically and later became intense, for Q_1, Q_2, $Q_{3,}$ and Q_4, $X_{it} < X_{io,}$ which indicates actual production losses expressed in net profit.

2.1 | Definition of Terms: COGS, OpEx, and Selling and Administration Expenses

Cost of Goods Sold (COGS)

Represent the direct cost related to the manufacturing of goods/services that are sold to customers. It does not include selling, interest, general, and administrative expenses [26].

Operating Expenses (OpEx)

These are expenses a business incurs through its normal business operations, such as rents, equipment, inventory costs, payroll, maintenance, and Research and Development (R&D) costs [26].

Selling and Advertising Expenses (SAEX)

Expenses related to the running of a business that is not directly included in the production of goods or delivery of services. Examples are utilities, insurance payments, marketing, advertising, and promotion expenses [26].

For a given business entity:

$$NP = \sum R - \sum E, \tag{3}$$

where NP = Net Profit, R = Revenue, E = Expenses.

$$\% \text{ Expense Impact on Estimated Losses} = \frac{RE}{FT} X100, \tag{4}$$

where R_E = Range of quarterly expenses, F_T = Total loss impact factor.

3 | Results and Discussion

Table 1. Showing quarterly stock prices of NB and Nestle Nigeria (2019).

Quarterly average stock prices in Naira (2019)	Q1	Q2	Q3	Q4
NB	73	62	51	53
Nestle Nigeria	1497	1453	1303	1347

Obojobo and Uzorh | J. Appl. Res. Ind. Eng. 10(1) (2023) 113-124

Table 2. Showing quarterly net profits for NB and Nestle Nigeria (2019).

Quarterly net profit in billions of Naira (2019)	Q1	Q2	Q3	Q4
NB	8.02	5.3	4.36	3.83
Nestle nigeria	12.8	13.4	10.6	8.8

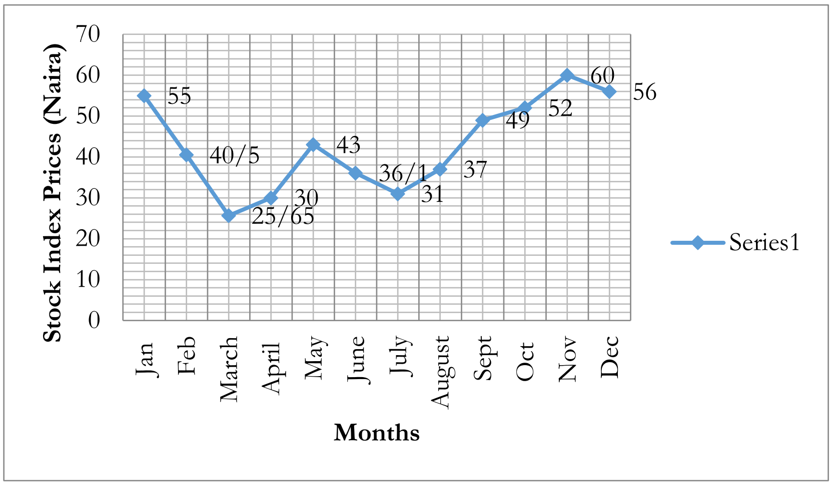


Fig. 1. NB stock prices pattern (2020).

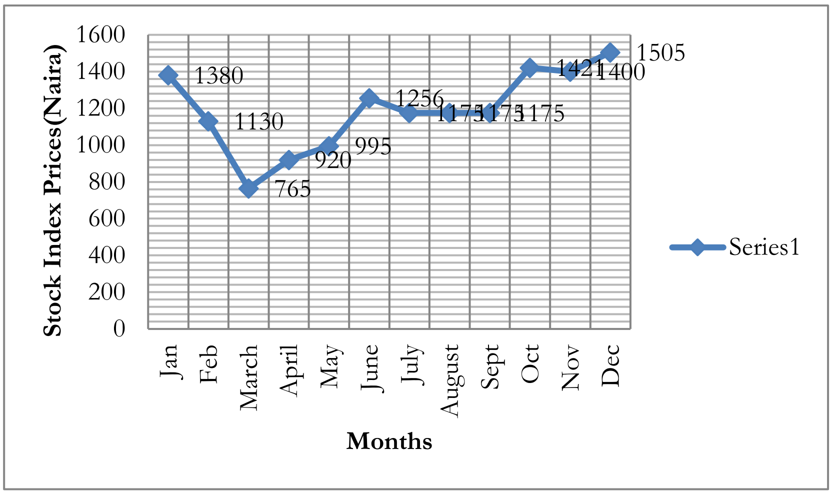


Fig. 2. Nestle Nigeria stock prices pattern (2020).

Fig. 1 and *2* above shows the monthly stock price movement patterns for both organizations during the 2020 COVID-19 pandemic disruption from January to December.

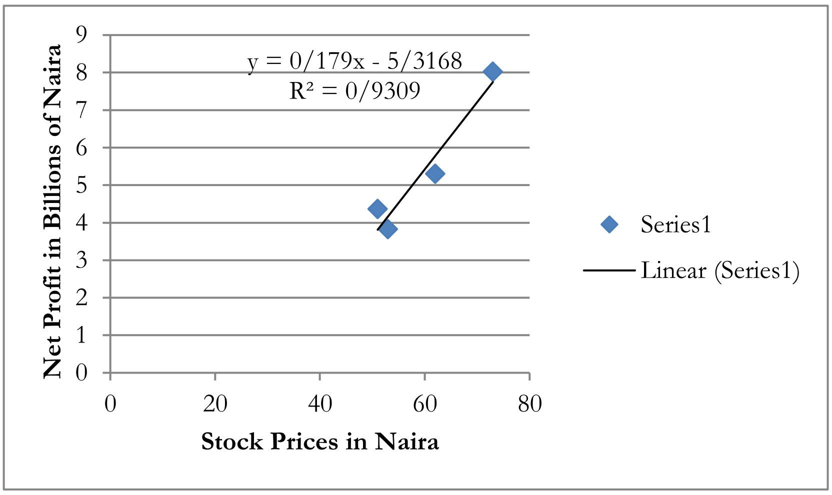


Fig. 3. Graph showing relationship between stock prices and Net Profit (NB) (2019).

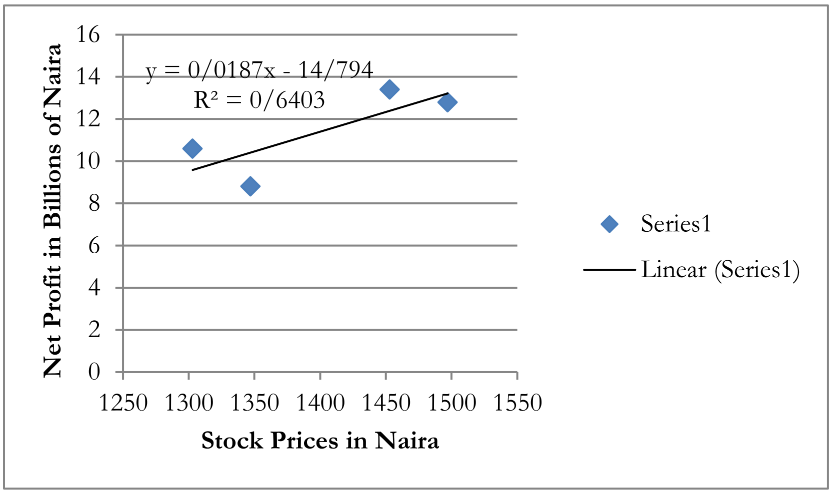


Fig. 4. Graph showing relationship between stock prices and net profit (Nestle Nigeria) (2019).

Figs. 3 and *4* Show the mathematical relationship between stock prices and net profits in the form of simple regression equations. For Nigerian Breweries (NB), we have $Y = 0.179X - 5.316$. Where Y is the dependent variable called net profit in billions of Naira, and X is the independent variable representing stock index prices. A correlation coefficient of $R^2 = 0.9$ indicates a significant relationship between the dependent and independent variables. Therefore, at $X \leq$ N30, the average estimated condition for zero net profit on production activities for the business year becomes a basis for estimating production losses for the year 2020. Similarly, for Nestle Nigeria, we have $Y = 0.018X - 14.79$, with $R^2 = 0.64$, which is also significant. Therefore, $X \leq 821$ is the average estimated condition for zero net profit on production activities for the same business year. These equations were used to estimate net profits for all quarters of 2020. These figures were compared to the average yearly net profit figures of the previous year (2019), which is assumed to be a disruption-free period to get the difference, thereby estimating their respective quarterly production losses as shown in *Tables 3* and *5* below. *Tables 4* and *6* below show the summary of estimated net profit losses and COGS, OpEx, and SAEX.

3.1 | Loss Estimation for 2020 COVID-19 Pandemic Disrupted Business Year

Table 3. Showing estimated quarterly production losses (NB).

Year/Quarter	Average Share Price (n), (2020)	Average Net Profit in Billions of Naira (2019)	Estimated Net Profit in Billions of Naira (2020)	Estimated Net Profit Losses in Billions of Naira
2020-Q1	40.38	5.38	1.39	3.47
2020-Q2	36.46	5.38	1.21	4.17
2020-Q3	39	5.38	1.66	3.72
2020-Q4	56	5.38	4.70	0.68

Note: We assume that COGS, OpEx, SAEX, and other negligible miscellaneous expenses make up the total production expenses for the business year. Generally, total expenses include labor, services, supplies, employee's salary, materials, inventory, depreciation, rents, insurance coverage, advertising, income tax, etc.

Table 4. Showing net profit losses (billions), cogs, OpEx, and saex (NB).

Year/Quarter	Net Profit Losses	COGS	OpEx	SAEX
2020-Q1	3.47	48.3	72.2	24.1
2020-Q2	4.17	44.4	64.5	20.3
2020-Q3	3.72	51.4	74.8	23.6
2020-Q4	0.68	74.3	95.9	21.9

Loss estimation for 2020 COVID-19 pandemic disrupted business year (Nestle Nigeria).

Obojobo and Uzorth | J. Appl. Res. Ind. Eng. 10(1) (2023) 113-124

Table 5. Showing estimated quarterly losses (NB).

Year/Quarter	Average Share Price (n), (2020)	Average Net Profit in Billions of Naira (2019)	Estimated Net Profit in Billions of Naira (2020)	Estimated Net Profit Losses in Billions of Naira
2020-Q1	1092	11.4	5.6	5.8
2020-Q2	1057	11.4	5.0	6.4
2020-Q3	1175	11.4	7.2	4.2
2020-Q4	1442	11.4	12.2	-0.8

Note: negative sign (-) indicates no loss and a potential excess net profit gain based on estimates.

Table 6. Showing net profit losses (Billions), COGS, OpEx, and SAEX (Nestle Nigeria).

Year/Quarter	Net Profit Losses	COGS	OpEx	SAEX
2020-Q1	5.8	38.7	52.8	14.1
2020-Q2	6.4	41.5	53.9	12.4
2020-Q3	4.2	42.5	55.8	13.3
2020-Q4	-0.8	45.2	60.1	15

3.2 | Relationships between Cogs, OpEx, SAEX, and Net Profit Losses

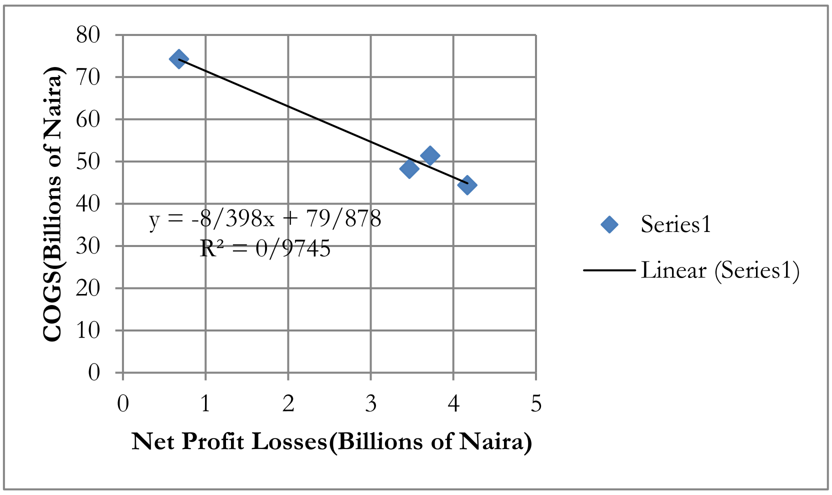


Fig. 5. Graph showing relationship between cogs and net profit losses (NB).

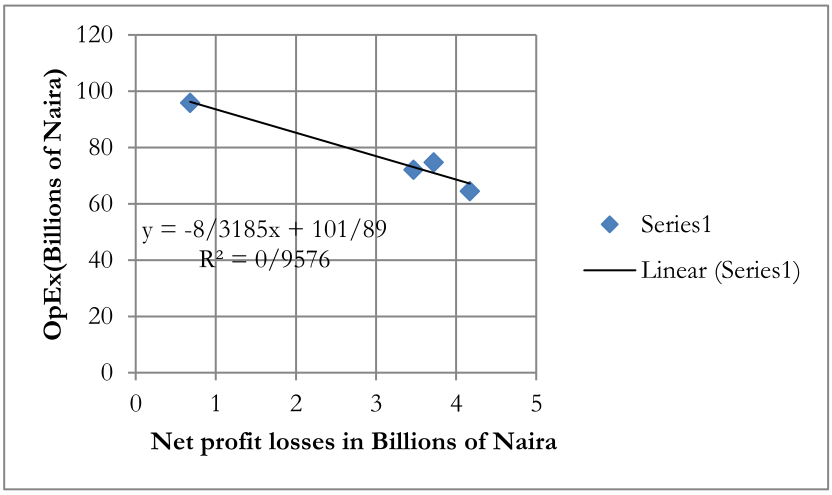


Fig. 6. Graph showing relationship between OpEx and net profit losses (NB).

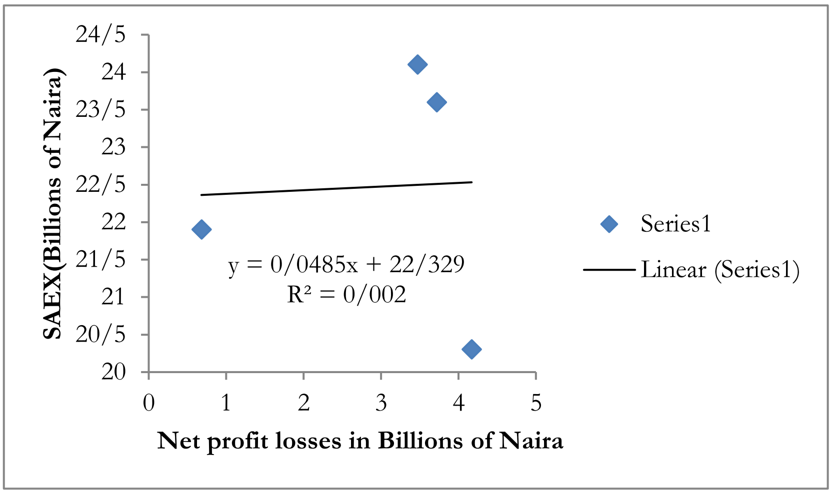


Fig. 7. Graph showing relationship between SAEX and net profit losses (NB).

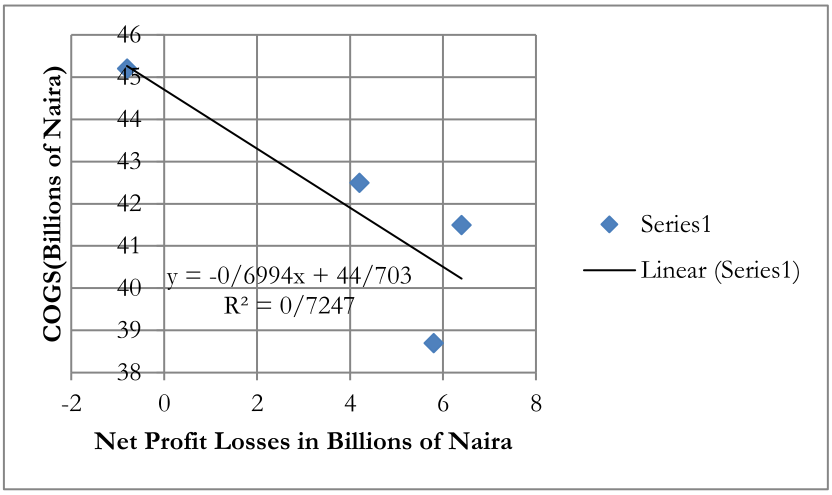


Fig. 8. Graph showing relationship between cogs and net profit losses (nestle).

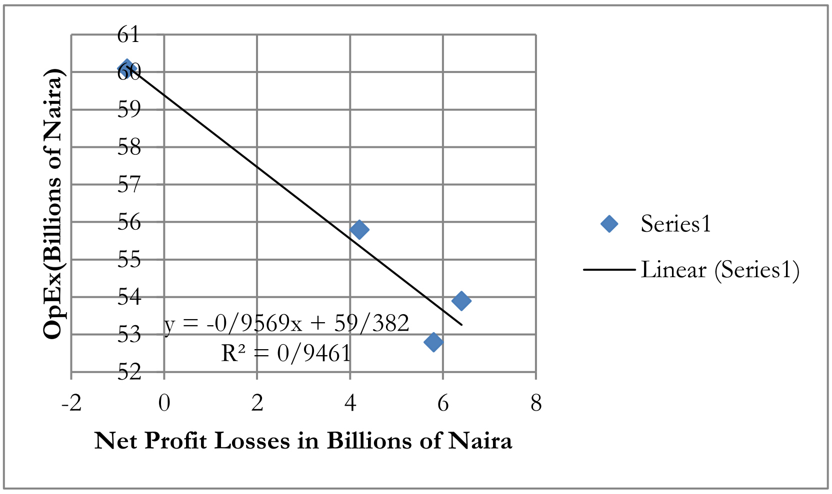


Fig. 9. Graph showing relationship between OpEx and net profit losses (nestle).

Obojobo and Uzorth | J. Appl. Res. Ind. Eng. 10(1) (2023) 113-124

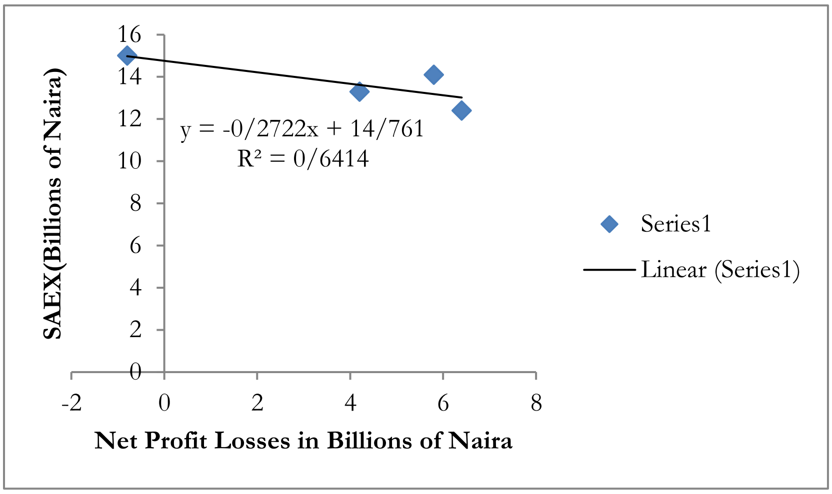


Fig. 10. Graph showing relationship between SAEX and net profit losses (nestle).

The above graph shows the relationship between the various production input resources and estimated net profit losses for both organizations.

3.3 | Percentage Expense Impact Factor on Estimated Productionlosses Determination

Table 6. Showing percentage expense impact on estimated production losses (NB).

Inputs (Billions of Naira)	Q1	Q2	Q3	Q4	Range	Estimated Loss Impact Factor (f)	%Expense Impact on Estimated Losses
COGS	38.7	41.5	42.5	45.2	6.5	6.5	39.6
OpEx	52.8	53.9	55.8	60.1	7.3	7.3	44.5
SAEX	14.1	12.4	13.3	15	2.6	2.6	15.9

Table 7. Showing percentage expense impact on estimated production losses (Nestle Nigeria).

Inputs (Billions of Naira)	Q1	Q2	Q3	Q4	Range	Estimated Loss Impact Factor (f)	%Expense Impact on Estimated Losses
COGS	48.3	44.4	51.4	74.3	29.9	29.9	45.9
OpEx	72.2	64.5	74.8	95.9	31.4	31.4	48.2
SAEX	24.1	20.3	23.6	21.9	3.8	3.8	5.9

From the above graphs (*Figs. 5-10*), COGS and OpEx show a strong negative correlation with net profit losses for both organizations, with correlation coefficients of 0.99 and 0.98 for NB and 0.85 and 0.97 for Nestle Nigeria, respectively. In contrast, SAEX shows no significant correlation with net profit losses for NB (0.04). However, that of Nestle Nigeria is slightly negatively significant at 0.64. Therefore, operational net profit losses during the COVID-19 pandemic for both organizations were caused primarily by unforeseen operational disruptions, which affected raw materials and other OpEx throughout the year. From *Table 6* and *7* above, the outcome of this analysis clearly shows the effect of OpEx on estimated production losses which is comparatively higher than others for both organizations, having a higher estimated loss impact factor (F), followed by COGS and SAEX. Nestle Nigeria, therefore, recorded the following percentage expense impact on estimated production losses: 44.5%, 39.6%, and 15.9% for OpEx, COGS, and SAEX, respectively, while NB has the following percentage expense impact on estimated production losses: 48.2%, 45.9% and 5.8% for OpEx, COGS, and SAEX respectively. Also, NB operational business performance reveals that at the stock price [X] ≤ N30 is the average estimated condition for zero net profit on production activities for the 2019 fiscal year, while that of Nestle Nigeria is at the stock price [X] ≤ 821. These estimates, no doubt, will serve as reliable guides to potential investors. Hence, we conclude that the COVID-19 pandemic's impact on OpEx led to a greater percentage of the estimated production losses in both organizations for the year. This, however, is possibly due to huge losses in operating person-hours occasioned by various lockdown policies to curtail the spread of the pandemic across the federation.

4 | Conclusion

From the preceding analysis, it is evident that the COVID-19 pandemic disruption led to significant production losses for both manufacturing organizations under consideration, even though the impact varies based on estimates as expected. A closer look at *Tables 1* and *2* in the Appendix shows the selected period's global Purchasing Managers' Index (PMI). This is an indicator of the global economic health of various economies during the COVID-19 pandemic. Between the last quarter of 2019 (Q_4) and the third quarter of 2020 (Q_3), the effect of the pandemic on economic activities was more severe compared to the last quarter of 2020 (Q_4) and beyond. From the tables, we can also see the quarterly GDP of selected countries for the period under review. GDP is a reflection of economic activities, especially in a nation's manufacturing and service sectors. All countries listed show a steady decline in GDP figures beginning from the last quarter of 2019 due to the impact of the pandemic, which also showed a slight improvement towards the end of the second quarter of 2020 (Q_2) for most countries, possibly due to the gradual lifting of COVID-19 restrictions and installation of containment measures to deaden the spread of the pandemic which led to a rise in GDP figures during the last quarter of the year. Accordingly, the average J. P Morgan Global Composite PMI accelerated in the fourth quarter of 2020, compared with the levels in the third quarter of 2020 and the corresponding quarter of 2019 [5]. Therefore, this study confirms the effectiveness of a loss estimation model that has also been used in previous research work with potential to significantly improve decision-making and enable industries and governments to expedite recovery actions by providing rich and informative data. Conclusively, the overall business output remained positive for the United States, United Kingdom, China, Germany, India, United Arab Emirates, and Italy but negative for Japan, South Africa and Nigeria [5]. Therefore, the overall economic impact of the COVID-19 pandemic on the selected organizations and the economy for the period considered remains negative compared to pre and post-pandemic eras.

Although we have successfully estimated the production losses and various inputs resources percentage impacts on estimates, it is important to note that the accuracy of such estimates largely depends on the correlation coefficients between the dependent and independent variables. The closer this value is to unity, the more accurate the results obtained.

Both NB and Nestle Nigerian should look into possible measures to optimize their business operations to enhance operational resilience. This can result in low OpEx, which accounted for the highest percentage contribution to estimated production losses since no business entity is entirely immune to unforeseen operational disruptions like in the case of the global COVID-19 pandemic disruption. By so doing, investment risks and operational losses will be minimal. Also, splitting input resources into specifics instead of grouping them will make possible a more detailed analysis and insights towards knowing the highest specific contributor to estimated production losses so that efforts can be channeled to reduce the negative impact of such individual input resources by being proactive. Also, other stock market indices can be analyzed for possible correlation with industrial production output variables which can also be used for loss estimation purposes with the aid of this model or other loss estimation models.

References

[1] Pathak, A. K. (2017). *Estimating production losses from disruptions based on stock market returns: applications to 9/11 attacks, the deepwater horizon oil spill, and hurricane sandy* (Master Thesis, Iowa State University). https://www.imse.iastate.edu/files/2014/03/PathakAditya-Kiran-thesis.pdf

[2] Al Kazimi, A., & Mackenzie, C. A. (2016, April). The economic costs of natural disasters, terrorist attacks, and other calamities: An analysis of economic models that quantify the losses caused by disruptions. *2016 IEEE systems and information engineering design symposium (SIEDS)* (pp. 32-37). IEEE

[3] Boesso, G., & Kumar, K. (2007). Drivers of corporate voluntary disclosure: a framework and empirical evidence from Italy and the united states. *Accounting, auditing & accountability journal, 20*(2), 269-296.

[4] Central Bank of Nigeria (CBN). (2020). *Economic report first quarter 2020*. https://www.cbn.gov.ng/out/2020/rsd/first%20quarter%202020%20cbn%20economic%20report.pdf

[5] Central Bank of Nigeria (CBN). (2020). *Economic report, fourth quarter*. https://www.cbn.gov.ng/Out/2021/RSD/Fourth%20Quarter%202020%20Economic%20Report_1.pdf

[6] Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of business*, *59*(3), 383-403.

[7] Choi, J. J., Hauser, S., & Kopecky, K. J. (1999). Does the stock market predict real activity? time series evidence from the G-7 countries. *Journal of banking & finance*, *23*(12), 1771-1792.

[8] Coleman, L. (2006). Frequency of man-made disasters in the 20th century. *Journal of contingencies and crisis management*, *14*(1), 3-11.

[9] Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. Wiley-Interscience.

[10] Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *The American economic review*, *71*(4), 545-565.

[11] Fama, E. F. (1990). Stock returns, expected returns, and real activity. *The journal of finance*, *45*(4), 1089-1108.

[12] Guha-Sapir, D., D'Aoust, O., Vos, F., & Hoyois, P. (2013). *The frequency and impact of natural disasters*. Oxford University Press.

[13] Haimes, Y. Y., Horowitz, B. M., Lambert, J. H., Santos, J., Crowther, K., & Lian, C. (2005). Inoperability input-output model for interdependent infrastructure sectors. II: case studies. *Journal of infrastructure systems*, *11*(2), 80-92.

[14] Hassan, M. R., & Nath, B. (2005). Stock market forecasting using hidden Markov model: a new approach. *5th international conference on intelligent systems design and applications (ISDA'05)* (pp. 192-196). IEEE.

[15] Osuizugbo, I. C. (2020). Disruptions and responses within Nigeria construction industry amid covid-19 threat. *Covenant journal in research & built environment (CJRBE)*, *8*(2), 37-48.

[16] Anyalechi, K. C., Ezeaku, H. C., Onwumere, J. U. J., & Okereke, E. J. (2019). Does oil price fluctuation affect stock market returns in Nigeria?. *International journal of energy economics and policy*, *9*(1), 194-199.

[17] Andam, K. S., Edeh, H., Oboh, V., Pauw, K., & Thurlow, J. (2020). *Estimating the economic costs of COVID-19 in Nigeria* (Vol. 63). Intl Food Policy Res Inst.

[18] Leontief, W. (Ed.). (1986). *Input-output economics*. Oxford University Press.

[19] Lian, C., & Haimes, Y. Y. (2006). Managing the risk of terrorism to interdependent infrastructure systems through the dynamic inoperability input–output model. *Systems engineering*, *9*(3), 241-258.

[20] Tumala, M. T., & Yaya, O. S. (2015). Estimating bull and bear betas for the Nigerian stock market using logistic smooth threshold model. *CBN Journal of applied statistics (JAS)*, *6*(1), 263-284.

[21] Miller, R. E., & Blair, P. D. (2009). *Input-output analysis: foundations and extensions*. Cambridge University Press.

[22] Morck, R., Shleifer, A., Vishny, R. W., Shapiro, M., & Poterba, J. M. (1990). The stock market and investment: is the market a sideshow?. *Brookings papers on economic activity*, *1990*(2), 157-215.

[23] Okuyama, Y. (2007). Economic modeling for disaster impact analysis: past, present, and future. *Economic systems research*, *19*(2), 115-124.

[24] Okuyama, Y. (2008). *Critical review of methodologies on disaster impact estimation*. https://onlineasdma.assam.gov.in/kmp/pdf/1491474441Okuyama_Critical_Review.pdf

[25] Natarajan, R., Sivakavitha, S., & Vasani, S. A. (2020). Relationship between stock return and firms' financial performance in BSE listed companies. *European journal of molecular & clinical medicine*, *7*(3), 4553-4559.

[26] Floyd, D. (2019). *26 Goldman Sachs Alumni who run the world (GS)*. https://www.investopedia.com/news/26-goldman-sachs-alumni-who-run-world-gs/

[27] Cengiz, H., & Püskül, A. S. Ö. (2016). Hisse Senedi Getirileri ve Karlılık Arasındaki İlişki: Borsa İstanbul Endeksinde İşlem Gören İşletmelerin Analizi. *Yalova sosyal bilimler dergisi*, *6*(12), 295-306. (**In Turkish**). https://dergipark.org.tr/en/download/article-file/272284

Appendix

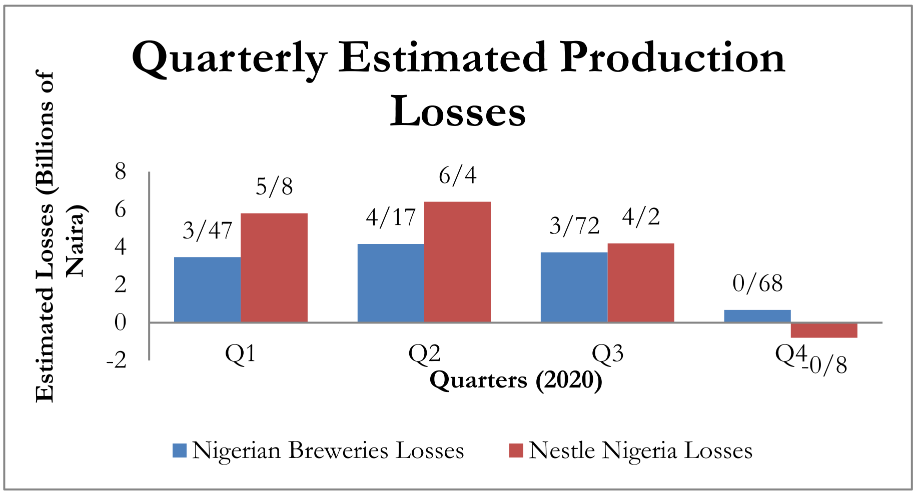


Fig. 1. Quarterly estimated production losses.

Table 1. Global purchasing managers index (PMI).

	2019 Q4	2020 Q3	2020 Q4
Composite	51.27	51.83	52.70
Manufacturing	50.07	51.57	53.53
Servics (business activity)	51.53	51.37	52.20
Employment level	50.93	49.33	51.40

Soursces: JP morgan, CBN staff compilation.

Table 2. Quarterly GDP in selected countries.

	Q4 2019	Q1 2020	Q2 2020	Q3 2020	Q4 2020
US	2.1	0.3	-9.0	-2.9	-2.4
UK	-1.7	-1.7	-21.5	-9.6	-7.8
CHN	6.0	-6.8	3.2	4.9	6.5
IND	4.7	3.1	-23.9	-7.5	0.4
GM	0.4	-2.1	-11.3	-4.0	-3.7
IT	0.1	-5.6	-17.7	-5.0	-6.6
JP	1.7	-1.8	-10.2	-5.0	-1.2
SA	-0.5	-0.1	-17.1	-6.0	Na
NG	2.6	1.9	-6.1	-3.6	0.11

Sources: Trading Economics/Various Country Websites, CBN Staff compilation.
Note: US, UK, CHN, IND, GM, IT, JP, SA and NG represent United States, United Kingdom, China, India, Germany, Italy, Japan, South Africa and Nigeria, respectively.

Obojobo and Uzorh | J. Appl. Res. Ind. Eng. 10(1) (2023) 113-124

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 125–140.

Paper Type: Research Paper

Open Innovation and SMEs: Providing a Model for Business Development (an Application on Iranian Industrial Park)

Younos Vakil Alroaia*

Department of Management, Entrepreneurship and Commercialization Research Center, Semnan Branch, Islamic Azad University, Semnan, Iran; y.vakil@semnaniau.ac.ir.

Citation:

Alroaia, Y. V. (2023). Open innovation and SMEs: providing a model for business development (an application on iranian industrial park). *Journal of applied research on industrial engineering, 10*(1), 125-140.

Received: 15/04/2021 Reviewed: 20/05/2021 Revised: 09/07/2021 Accepted: 01/09/2021

Abstract

The aim of this study is providing a developed model for Small and Mid-size Enterprises (SMEs) in Open Innovation (OI) activities. In this regard, an appropriate model was defined by studying the literature. Then, after selecting a sample of 60 SMEs the data were collected by a questionnaire and were analyzed with the smart PLS software. In the third stage, the relative importance of factors was tested from the perspective of 10 experts in the field of OI along with experienced managers of the SMEs with more than 15 years of work experience with the help of ANP and Promethee methods. The results showed that these factors include the parameters: product characteristics, inter-organizational factors, and environmental factors. In addition, the most important factors include product characteristics. Finally, several implications were made such as changing the degree of SMEs' participation in Open Innovation Activities (OIA), over time according to continuous monitoring of these moderators.

Keywords: Small and mid-size enterprises, Open innovation, Product characteristics, Inter-organizational factors, Environmental factors.

1 | Introduction

Policymakers have lunched many initiatives to stimulate the adoption of Open Innovation (OI). Given the reluctance of Small and Mid-size Enterprises (SMEs) to engage in OI, policy initiatives strive to encourage them to open up in order to faster technological progress and economic growth. Public institutions even provide consulting services to counter associated risks [37]. However, SMEs are key levers in both developed and developing economies. This sector not only contributes to more than 90% of the number of businesses and half of the world's employment, including micro businesses (businesses with less than 10 employees), but also accounts for 97% of businesses in Iran and are important sources of innovation [1]. For these reasons, governments are looking for ways to increase the productivity of SMEs.

Corresponding Author: y.vakil@semnaniau.ac.ir

https://doi.org/10.22105/jarie.2021.281399.1303

Several researchers have identified OI as an important strategy to overcome the weaknesses of SMEs, such as resource constraints, capacity constraints, and exposure to various types of risks [2], [3]. OI, defining as "the purposeful knowledge flow into or out of the organization in order to accelerate innovation and expand markets to exploit innovation outside the organization" [4], provides more efficient use of technology and knowledge for companies, which couldn't be accessed in any other way. Thanks to OI, companies can take advantage of innovative organizational talent instead of relying solely on internal research and development resources [5], [6]. Additionally, companies can reduce the risks of developing intellectual properties, because using OI can reduce fix costs (such as R&D personnel costs) in return for more variable costs associated with the purchase of tangible intellectual properties (such as the purchase of a patent). OI, with the help of the combination of internal innovation and systematic scanning of external knowledge, is a powerful lever for increasing the flexibility and efficiency of a firm's internal and external resources. In general, companies are creating financial and non-financial benefits for themselves by utilizing the benefits of Open Innovation Strategies (OIS) [7], [8].

Definition of OI shows that the concept of openness is in line with internal innovation and external exploitation of innovation. Despite the fact that OI in the past was conceptually contradictory with closed innovation, newer studies suggest the relationship between them as a spectrum of innovation [9]. In other words, Cornell [9] suggests that the superiority of choosing an OI approach over the closed one depends on the organization's conditions. A company may benefit from OIS in order to fulfill some of its projects, and at the same time may utilize closed innovation strategies for its other projects. Also, in some competitive environments, especially in the development of defensive technologies, companies rely exclusively on closed innovation strategies. In other words, a company, regardless of the industry it is operating in, should not use OIS [10] in all its periods of operation, since no matter how strongly the laws protect the intellectual properties of a company, some of its trade techniques and knowledge will be vulnerable to expose to other competitors. Therefore, companies should not merely be bound by the use of open or closed innovation strategies, since both approaches can be beneficial for them depending on circumstances [9]. Hence, the understanding of innovation as a spectrum, which gets closed in the one hand and becomes open on the other hand depending on the degree of firms' participation in Open Innovation Activities (OIA), can be useful. One of the growing trends in the world is about SMEs to take advantage of OIS [11]. Given the reluctance of SMEs to engage in OI policy initiatives strive to encourage them to open up in order to foster technological progress and economic growth. Public institutions even provide consulting services to counter associate risks [37]. However, to effectively increase collaboration for innovation in SMEs, not all OIS, policy mechanisms and frameworks may by equally appropriate. Taking into account the fact that some firms, especially smaller ones, face resource and capability constraints or cognitive barriers regarding the adoption of OI, major benefits such as reaching external knowledge and distant experts with unrestricted scope may be difficult to use [38]. Researches in this field have shown that the use of OIS varies according to the characteristics of the industry. Also, the growth of the use of these strategies is depicted as sudden and unpredictable jumps over time instead of a steady rate of growth [12]. On the other hand, scholars have shown that there is still very few number of SMEs benefiting from OIS, and many more can benefit from them even with limited participation in OIA [9]. So far, many studies have examined the utilization of OI in large companies, and there have not been enough researches on why SMEs, particularly in Iran, are not willing to participate in OIA. Therefore, this research seeks key factors that influence the decision of SMEs to participate in OI in the environment of Iran.

A review of OI literature in relation to Iranian SMEs shows that OI is not well-suited in this sector [13] and SMEs in Iran has problems such as lack of resources, lack of strong enough cash flow, lack of R&D department, innovation weakness, and required raw material supplement problems that cause poor performance. Paying more attention to the causes of SMEs' tendencies to participate in OIA can minimize their weaknesses and enhance this section's development [6]. Recent experiences in many countries all around the world have shown the effectiveness of the OIS in promoting the performance of SMEs [14]. As any innovation system's final output must ultimately produce new products and deliver new services, the considered function is essential in the innovation system [39]. Additionally, the knowledge that companies gain from cooperating with other organizations is very valuable to their success and gaining a competitive edge in the market [15]. What is evident is that the identification and the ranking of factors affecting the decision of SMEs to participate in OIA in order to assess the OI environment from the perspective of SMEs in Iran are the first prerequisites to increase the participation of this sector in OI environment in order to benefit from the advantages.

Due to the concept of OI, which was introduced in management literature in 2003, this research area still has a lot to study about [11]. There are several important research gaps in the study of OI literature [9]. This paper examines one of the most important research gaps proposed which is related to the critical need to identify key factors affecting the degree and the efficiency of SMEs' participation in OI environment [9], [16], [17]. In response to this need, the present study proposed a developed model of moderating factors affecting the decision of SMEs in Iran to exploit the benefits of OI or in other words, to identify the most important factors in accelerating or hindering the implementation of OIS in Iranian SMEs, which includes three broad categories such as product characteristics, inter-organizational factors and environmental factors. These factors are going to be explained in the literature review section. In the proposed model, the effect of factors such as technological turbulence and market turbulence has been investigated, which has not yet been sufficiently examined, especially in Iran.

The turbulent environment imposes various constraints on corporates' innovative performance that has not yet been examined as a factor influencing Iranian SMEs' decision to participate in OIA. Environmental changes include technological turbulence and market turbulence. Technological turbulence refers to the rate of changes in technology and the unpredictability of technology that rapidly abolishes the company's technical knowledge and requires the development of a new technology. Market turbulent refers to the rate of changes in customers' preferences and customers' demands that rapidly disrupt the current market knowledge of companies. Companies need to respond more quickly to unanticipated changes and supplant multiple alternatives to meet customer needs to maintain innovative market performance. In a highly turbulent market environment, the role of technological turbulence in OI should not be underestimated, because, in a rapidly changing market environment, companies do not know how to use these technologies in favor of the knowledge of market which is changing rapidly. A company that wants to prevent its technology from becoming obsolete should be able to rapidly commercialize it or develop it jointly [15]. In the field of OI, the effect of technological turbulence and market turbulence factors, which are less considered in the environment of Iranian SMEs, on these enterprises' tendency to take participate in OIA have been investigated. This model includes three general categories of product specifications, internal factors and environmental factors. Therefore, the main stimulus of the present research is to answer this main question: which factors moderate Iranian SMEs' degree of participation in OIA? In the following sections, the study includes a literature review, research methods, the empirical results, discussions, limitations, theoretical suggestions, and suggestions for future researches.

2 | Literature Review

Today, by increasing labor mobility, division of labor due to increasing globalization activities, increasing protection of intellectual properties due to the enactment of laws, and technical advances in remote cooperation, open source innovation has become a common strategy for businesses [5]. All of these factors facilitate and moderate the participation of businesses in the OI environment around the world. Other effective factors that influence (enhance / hinder) managers' decisions on the degree of participation in OIA, specifically SMEs, include three categories of factors: 1) product characteristics, 2) factors inter-organizational, and 3) environmental factors. Therefore, in line with the purpose of this study, which is to increase the participation of the SME sector in OIA, the following hypotheses are proposed:

2.1 | Product Characteristics Moderate the Degree of Iranian SMEs' Participation in OIA

OIA are more common in industries that produce more complicated products. Products with complex technology are more likely to require two or more innovative partners [10], [18], [19]. On the other hand, OI is more common in industries in which products have a shorter life cycle. The lifecycle of products often encourages SMEs to collaborate with other influential companies in the value chain in order to commercialize products more quickly. In other words, participation with other SMEs significantly reduces the innovative products' time of entering the market [9], [13]. Innovation are widely considered to be valuable capabilities associated with competitive advantage [40]. SMEs active in high-tech industries are seeking to take advantage of OI in line with their technological strategies [20]. It means in the initial and final stages of products maturity they tend to be more cooperative, especially with universities. After this stage, using more close innovation strategies, they enter the phase of innovation maturity, then again in the stage of innovation commercialization and marketing, they tend to participate more in OIA [9].

2.2 | Inter-Organizational Factors Moderate the Degree of Participation of Iranian SMEs in OIA

Cornell [9] believes that participation in OI needs an open business plan which helps an SME create a beneficial exchange of knowledge with the outside world. Business plan of a company affects its strategic management decisions in relation to participation in OIA [7]. For example, a SMEs may consist of several engineers and its business plan may be based on selling the patents of its inventions to other companies, while another SMEs products and commercializes its products through the acquisition of other companies' patents. On the other hand, the unique opportunity of a SMEs to gain access to resources and capabilities positively affect its ability to innovate and commercialize. A SMEs with more limited resources is looking for partnerships with other innovative partners. SMEs which do not have enough resources and capabilities needed to manage the whole process of exploration to exploitation of innovation are more willing to engage in OIA [21], [22]. Scholars have shown that there is a high correlation between the level of participation in OI and the need for more resources in SMEs. Accessing more resources can be seen in a variety of ways, such as increased funding, a more efficient supply chain, marketing capabilities, regulatory and legal resources, technical expertise, credit enhancement, and brand popularity and other assets and capabilities [3], [8]. Some scholars point out other vital elements that increase the effectiveness of participation in OIA for SMEs, including absorptive ability, search capability, cultural intelligence, effective organizational structure, and strong leadership [8], [9], [11], [21], [24], [25], [26], [27]. Absorptive capability is a dynamic and fundamental ability that is necessary for successful participation in OIA for SMEs and is related to the capability of a company in successful learning, analysis, and exploitation of external knowledge of the organization [23], [24], [28]. Such a capability requires the enhancement of an OI culture and sufficient experience to understand external innovation [3], [11], [25]-[28]. Also, search capability refers to the effective search in the environment in order to identify OI opportunities, and almost all SMEs are weak in searching due to their limited knowledge and limited ability to seek out externally-developed knowledge. Scholars have shown there is a high correlation between search capability and the SME's innovative performance [9], [24].

Cultural intelligence refers to the ability of an SME to create effective collaboration among members of diverse cultures. High cultural intelligence is an essential factor for working with foreign or local companies which have high cultural diversity [9], [13]. The use of organizational structures designed for OIS, such as organic structures, known for enhancing innovative capabilities in the organization, is necessary for SMEs willing to engage in OIA because such innovation structures encourage bottom-up innovation as well as the exchange of knowledge in the organization [8], [9]. On the other hand, strong leadership can positively affect the willingness of an SME to engage in OIA. They state that strong leadership is influential through resource management, performance management, and organizational motivation, the allocation of human and financial resources, the flow of knowledge, and the management of processes. Leaders can continuously strengthen the willingness and ability of SMEs to participate in OIA, through planning, monitoring, communicating and providing performance feedback along with the provision of internal and external rewards, [11], [19], [25], [27].

Excessive confrontation with multiple risks can undermine innovation, especially radical ones. Scholars have shown that SMEs are more willing to engage in OIA when they encounter high-risk innovation projects. The advantage of collaborating in OIA is that the risks and rewards of the outcome will be distributed among partners. On the other hand, the risks associated with R&D activities decrease for SMEs. Scholars have also shown that the reduction of non-systematic risks of R&D projects can strengthen the competitive position of SMEs in an industry [13], [19].

Additionally, participation in OIA, according to researchers, requires OI culture [3], [19], [25], [27]. Such organizations are potentially willing to sell or buy intellectual properties' rights to/from other organizations or partners in order to develop and commercialize innovation [3]. If an SME is willing to participate actively in OIA, it will promote OI culture continually. Such a goal is achievable through encouraging adaptive learning culture and promoting appropriate collaboration as well as knowledge management in the field of OI [3], [25], [29]. Previous experiences of an SME in contributing to OIA affect its future participation, as an SME builds its future strategies based on what it has learned from past experiences [2]. In other words, the poor performance of an OI strategy in the past can lead to a lack of participation of an SME in future OIA and vice versa [9].

2.3 | Environmental Factors Moderate the Level of Participation of Iranian SMEs in OIA

Characteristics of industry affect the degree of SMEs' participation in OIA [12]. For example, OIA can bring economies of scale to members of a collaboration network, especially for highly scale-intensive industries. Japanese companies active in the electronics industry rely on R&D collaborations to reach economy of scale. He also states that OIA in different industries have been adapted in various ways. For example, while outward OI is more common in high-tech industries, inward OI is more popular in low-tech industries [9]. In line with confirming the role of product characteristics in SMEs' tendency to adopt OI approach. In some industries, such as food industry, the dominant strategy is secrecy, which means the characteristics of the industry and the company's central strategy play an important role in the desire to participate in OIA [12], [25]. Development of joint inter-organizational innovation and commercialization activities are dependent on efficient network structures because these structures reflect and influence the extent to which SMEs participate in OIA [8], [13], [27], [30]. Companies, which use OI, are heavily dependent on the communications network system and that the communications network is changing over time [10]. They also claim that the most efficient network for the innovation exploration stage is usually a relatively open system between partners. When innovation becomes mature, it is necessary to make this network more closed and between a smaller number of partners with limited relationships [10]. This strategy generally increases the efficiency and the speed of the commercialization process of production [13]. The risk of participation in long-term relationships between organizations in the form of collaborative networks is an issue that has been warned by numerous studies because researchers believe that this can lead to network stagnation, which means collaboration reduction, new thinking avoidance, and resistance toward expanding the network of cooperation with new partners. It is mentioned that network structure conduciveness depends on partners' commitment, trust among the members, and the extent to which they participate in sharing resources, capabilities, experiences, and expertise. Such criteria along with the easiness of entering OI networks affect an SME's tendency to collaborate in the field of OI [2], [3], [8], [11], [14], [24], [26], [27], [31]. If the cost-benefit ratio increases for collaborating in OIA, SMEs will be less willing to engage in such activities [9] For example, the larger the network becomes, the more it is expensive and the more difficult it is to be managed [12], [29]. Costs of local and international cooperation are different because international cooperation can increase the costs of SMEs in certain circumstances, especially when talking about barriers of globalization such as different languages and logistical issues. SMEs often prefer to work in local markets rather than in international ones, in which they know customers, industry characteristics, and regulations better. In other words, developing activities in large geographical and dispersed markets increase their costs [9]. Despite the benefits of OI, SMEs are cautious about establishing such collaborations. They compare the risks of participation in OI to its benefits. For example, financial instability increases the potential partners' bankruptcy and the risk of establishing cooperative relationships. Therefore, selecting an appropriate partner for OIA is an important issue [2], [12]. On the other hand, establishing such collaborations with customers in all aspects of R&D process increases the risk of disclosing intellectual properties, according to some scholars [2], [18], [29].

The legal environment varies from country to country. These laws and regulations affect the ability of SMEs to implement OIS. Legal instability leads to undermining of OIA. Countries that have firm laws about intellectual property protection boost OIA among SMEs. Therefore, SMEs, before participating in any OIA with other partners, take into account the laws and regulations that monitor the protection of intellectual property, since many companies' motivation to implement OIS is to access other organizations' intellectual properties [18].

On the other hand, Cornell [9] states that restricting laws in some countries prevent the withdrawal of technology and knowledge from the country. This phenomenon is especially common when it comes to countries' defensive technology and knowledge [8], [9]. Government decisions, whether in the form of support or interference, affect the decision of SMEs to participate in OIA. For example, in a country where tax incentives are set up to attract foreign companies' cooperation in R&D projects are one of the government supportive levers to encourage SMEs to participate in OI networks. Many OECD countries are actively seeking to increase OIA between industry and education through governmental funding [8], [12], [32].

On the other hand, technological turbulence means rapid technological changes in products and production processes, which significantly reduces the life cycle of products and challenges SMEs. High technological turbulence makes SMEs more likely to use OIS because such strategies help them increase technology return rate when the technology is new. Companies, in rapidly changing environments, need to have access to new knowledge and new competitive advantages [7], [15], [33]. In addition to technological turbulence, turbulence in the market also influences the decision of SMEs to participate in OIA. Market turbulence refers to the rate of changes in the composition of customers, their preferences, competitive market conditions, and the degree of ambiguity and risk of business processes. SMEs that operate in turbulent environments need more changes in their products and services, due to changes in customer needs and desires. They are also more likely to participate in OIA. In stable markets, companies are less willing to align their products with the changing needs of customers, and therefore they feel less willing to participate in OIA [7], [29], [34] Explanations on the impact of technological turbulence and market turbulence on the degree of Iranian SMEs' participation in OIA lead to the following two hypotheses:

I. Technical turbulence moderates the degree of Iranian SMEs' participation in OIA.
II. Market turbulence moderates the degree of Iranian SMEs' participation in OIA.

Inspired by the results of previous literature, the importance of deciding to participate in OIA for SMEs is obvious. Despite the fact that previous researches have consistently emphasized the critical role of product characteristics, inter-organizational factors, and environmental factors, most of them, especially domestic scholars, have studied the role of these factors without considering technical turbulence and market turbulence in this issue.

Thus the present study, with the help of gathering many factors, has provided a developed model for factors moderating the decision of SMEs to participate in OIA as shown in *Fig. 1*.

Alroaia | J. Appl. Res. Ind. Eng. 10(1) (2023) 125-140

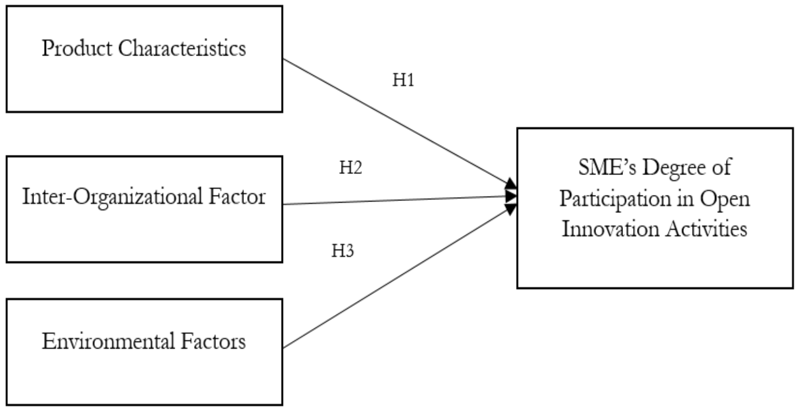


Fig. 1. The theoretical model.

3 | Research Methods

3.1 | Data Collection and Sample

In this section, in order to achieve the research's objectives, we examine theoretically the hypotheses. Therefore, the place of study, sampling method, data collection, method of measuring variables, research tool and data analysis methods have been described. By purpose, this cross-sectional research is a qualitative one and, by the method, this survey-based research is a descriptive one. To assess the hypotheses, data were collected from the managers of SMEs, with at least 5 years of professional experience, operating in industrial park by the use of convenient sampling method through two questionnaires. To collect data, the questionnaires were designed through the study of literature, dissertations, articles, and databases related to the research topic (both domestic and foreign ones). The first questionnaire was designed to test the theoretical model for moderating factors affecting the degree of SMEs' participation in OIA. It was arranged based on the Likert scale and assessed the perspective of 60 SMEs which have 10 to 49 employees. To determine the validity of the internal structure of the questionnaire, the convergent validity method was used with the Average Variance Extracted (AVE) of more than 0.6 and the divergent validity method was used by Fornell and Larcker methods. Since the partial least squares method was used to test the model with PLS 2 software, so to calculate the reliability of the questionnaire and the model from indicators such as Cronbach's alpha (more than 0.7), combined reliability (more than 0.7) and coefficients factor loads (more than 0.6) have been used [40]. The result is shown in *Table 1*.

Table 1. Convergent reliability and reliablity.

Latent Variables	Observable Variables	Convergent Validity AVE	Reliability Combined Reliability	Cronbach's Alpha
Product characteristics	Q_1-Q_3	0.619	0.866	0.794
Inter-organizational factors	Q_4-Q_9	0.759	0.925	0.885
Environmental factors	Q_{10}-Q_{19}	0.636	0.729	0.777

Subsequently, the second questionnaire, using the pairwise comparisons, determined the relative importance of the model's factors from the perspective of 10 experts in the field of OI along with experienced managers of the SMEs with more than 15 years of work experience. Therefore, in the present research, the field method is used to determine the key factors and confirm the theoretical model, the factors were weighed using the ANP method and were ranked using the Promethee method. In the following section, these methods are explained.

Alroaia | J. Appl. Res. Ind. Eng. 10(1) (2023) 125-140

3.2 | Research Instrument

The power of the relations between the latent variable and the observable variable is defined by the factor loading (λ). The value of factor loading is defined between zero and one. If the factor loading is less than 0.3, the relation between the variables is considered weak and is discarded. However, the acceptable value of factor loading is between 0.3-0.6 and values greater than 0.6 are desirable. In factor analysis, variables that measure a latent variable must have high factor loadings with that latent variable, and low factor loadings with other latent variables. The t-test is used to evaluate the significance of the relations between the variables with a significant level of 0.05. Therefore, if the t-value is calculated less than 1.96, the relation is not significant [35].

The Analytic Network Process (ANP) is a more general form of the Analytic Hierarchy Process (AHP) process used in Multi-Criteria Decision Making (MCDM). The ANP structures a decision problem into a network then use a system of pairwise comparisons to measure the weights of the components of the structure, and finally to rank the alternatives in the decision. The first step in the ANP method is creating a model and structure of the problem. The problem should be expressed clearly and be analyzed into a logical system like a network. Such a network structure can be achieved with the help of decision makers through brainstorming sessions or other appropriate methods. In the second step, pairwise comparisons are used to find out how the elements in the network are interacting [36]. The weights used at this stage are the input values to indicate the priorities of each criterion and sub-criterion in the Promethee method.

Promethee method is a multi-criteria decision-making methodology designed to discuss qualitative and discrete alternatives by Brans et al. [42]. This method is quite simple compared to other multi-criteria analysis methods, and it is especially appropriate for issues that require a limited number of actions to be ranked based on several criteria that are sometimes contradictory. This approach can be one of the most powerful decision-making methods that can help managers choose the best decision choices. The ranking of actions is performed by comparing the pair of actions in each criterion. The comparison is measured based on a predefined preference function with the domain [0, +1]. The preference function P is a function which compares two actions "a" and "b" in terms of the criterion j as follows:

$$\mathrm{Pj}\ \ \mathrm{a}\,,\mathrm{b}) = \mathrm{P}[\mathrm{dj}\ \ \mathrm{a}\,,\mathrm{b})]. \quad (1)$$

dj (a, b) = fj (a) - fj (b) denotes the difference in the size of two actions for criterion fj. This difference for criteria that have to be maximized will be significant only if fj (a)> fj (b). For the criteria that have to be minimized, the opposite condition is true. When a preference function has been associated with each criterion by the decision maker, all comparisons between all pairs of actions can be done for all criteria. A multi-criteria preference degree is then computed to globally compare every couple of actions:

$$\prod(\mathrm{a},\mathrm{b}) = \frac{\sum_{i=1}^{k}\pi_i P_i(a,b)}{\sum_{i=1}^{k}\pi_i}. \quad (2)$$

In order to position every action with respect to all the other actions, two scores are computed:

$$\varphi^{+}\ \mathrm{a}) = \sum\nolimits_{\mathrm{b}\in\mathrm{k}}\prod\ \mathrm{a},\mathrm{b}). \quad (3)$$

$$\varphi^{-}\ \mathrm{a}) = \sum\prod\ \mathrm{a},\mathrm{b}). \quad (4)$$

The positive preference flow quantifies how a given action is globally preferred to all the other actions while the negative preference flow quantifies how a given action is being globally preferred by all the other actions. The positive and negative preference flows are aggregated into the net preference flow:

$$\varphi\ \mathrm{a}) = \varphi^{+}\ \mathrm{a}) - \varphi^{-}\ \mathrm{a}). \quad (5)$$

This is the Promethee II complete ranking which is obtained by ordering the actions according to the decreasing values of the net flow scores. In other words, the criterion with the highest net flow has priority [42].

4 | Results

4.1 | The PLS Output

Table 2 shows the output of Smart PLS software which contains the factor loadings of the model's criteria. As mentioned, factor loadings above 0.3 are acceptable and criteria with factor loadings lower than 0.3 are excluded. According to the results, all model's factors have been confirmed. Also, the t-value was extracted as follows which are all greater than 1.96, thus, they're all acceptable.

Table 2. The smart PLS software output for t-values and factor loadings of the research model's criteria.

Latent Variables	Observable Variables	T-Value	Factor Loadig	Confirmed/not Confirmed
Product characteristics	More complicated products	274.139	0.978	✓
	Short product life cycle	317.906	0.959	✓
	Be in the earliest and latest stages of a product maturity (using more close innovation activities in between)	140.001	0.985	✓
Inter-organizational factors	High openness of firms' business models	266.000	0.977	✓
	Increase in need to access others' complementary resources and capabilities	52.338	0.936	✓
	Firms' high OI capabilities (absorptive capacity, search capabilities, cultural intelligence, conduciveness of organizational structure, strong leadership)	123.909	0.970	✓
	Increase in need to share project failure risks	81.671	0.963	✓
	Stronger OI culture	116.424	0.972	✓
	Increase in successful OI experiences	65.879	0.939	✓
Environmental factors	High scale-intensiveness an industry	476.452	0.991	✓
	High conduciveness of the network structure	98.137	0.968	✓
	High-tech and knowledge-intensive an industry	738.247	0.993	✓
	Low degree of industry reliance on secrecy	319.997	0.987	✓
	High easiness of entering OI networks	251.580	0.973	✓
	Low costs and risks of OI (costs associated with the size of the network, language diversity, cultural constraints along with the risks associated with the disclosure of intellectual property, reducing the cost-benefit ratio of OIA, partner insolvency)	99.954	0.974	✓
	High technological turbulence (high rate of rapid technological change in products and processes)	527.244	0.992	✓
	Low legal interventions and high legal stability and protection of intellectual property (local and international)	158.229	0.977	✓
	High market turbulence (high rate of rapid changes in the composition of customers, their preferences, and competitive market conditions)	367.924	0.989	✓
	High governmental support and funding	258.462	0.984	✓

4.2 | Weight Measurements and Ranking the Model's Criteria and Sub-Criteria Using ANP and Promethee Methods

After designing the ANP hierarchy model in super decision software and entering the data collected from the pairwise comparisons questionnaire, the weights of all criteria were calculated. In the third stage, in order to rank criteria, the value of the index Φ was calculated by Promethee software. Based on the outputs, we determined that the model's inconsistency is 0.08, which is less than 0.1, so the system's consistency is confirmed. Due to the system compatibility, we introduce the data of the paired comparisons into the

Super Decision software and calculate the weight of each of the factors. The output of this part is shown in *Table 3*.

Table 3. The model's main factors' weights and ranking.

Factor	Weight	Φ	Rank
Product characteristics	0.55	0.50	1
Inter-organizational factors	0.37	0.23	2
Environmental factors	0.08	-0.73	3

Given the output of Promethee software, the value of Φ for the first criterion, which is Product Characteristics, is 0.50 and more than other criteria. Therefore, it is the most prior criterion. Now, by assessing each sub-criterion of each factor, the weight and rank of each sub-criterion will be extracted. In the next step, we calculate the weight associated with the indicators for the product characteristics. By performing Super Decision software, the value of inconsistency for "product characteristics" criterion is equal to 0.00, which is less than 0.1, so the system's consistency is proven. In the following, we calculate the weight of the sub-criteria of this element using Super Decision and ranking them using the Promethee software. The result is shown in *Table 4*.

Table 4. The weight and ranking of sub-criteria relative to the factor of product characteristics.

Sub-Criterion	Weight	Φ	Rank
More complicated products	0.61	0.54	1
Short product life cycle	0.31	0.46	2
Be in the earliest and latest stages of a product maturity (using more close innovation activities in between)	0.08	-1.00	3

Given the output of Promethee software, the value of Φ for the first criterion, which is more complicated products, is 0.54 and more than other sub-criteria. Therefore, it is the most prior sub-criterion of Product Characteristics. By performing Super Decision software, the value of inconsistency for "inter-organizational factors" criterion is equal to 0.08, which is less than 0.1, so the system's consistency is proven. In the following, we calculate the weight of the sub-criteria of this element using Super Decision and ranking them using the Promethee software. The result is shown in *Table 5*.

Table 5. The weight and ranking of sub-criteria relative to the factor of inter-organizational factors.

Sub-Criterion	Weight	Φ	Rank
High openness of firms' business models	0.08	-0.15	4
Increase in need to access others' complementary resources and capabilities	0.03	-0.55	6
Firms' high OI capabilities (absorptive capacity, search capabilities, cultural intelligence, conduciveness of organizational structure, strong leadership)	0.13	0.29	2
Increase in need to share project failure risks	0.04	-0.48	5
Stronger OI culture	0.16	-0.07	3
Increase in successful OI experiences	0.56	0.96	1

Given the output of Promethee software, the value of Φ for the first criterion, which is increase in successful OI experiences, is 0.96 and more than other sub-criteria. Therefore, it is the most prior sub-criterion of inter-organizational factors. By performing Super Decision software, the value of inconsistency for "environmental factors" criterion is equal to 0.06, which is less than 0.1, so the system's consistency is proven. In the following, we calculate the weight of the sub-criteria of this element using super decision and ranking them using the Promethee software. The result is shown in *Table 6*.

Table 6. The weight and ranking of sub-criteria relative to the factor of environmental factors.

Sub-Criterion	Weight	Φ	Rank
High scale-intensiveness an industry	0.06	0.09	5
High conduciveness of the network structure	0.34	0.92	1
High-tech and knowledge-intensive an industry	0.13	0.16	4
Low degree of industry reliance on secrecy	0.02	-0.64	9
High easiness of entering OI networks	0.02	-0.66	10
Low costs and risks of OI (costs associated with the size of the network, language diversity, cultural constraints along with the risks associated with the disclosure of intellectual property, reducing the cost-benefit ratio of OIA, partner insolvency)	0.03	-0.43	8
High technological turbulence (high rate of rapid technological change in products and processes)	0.06	-0.19	7
Low legal interventions and high legal stability and protection of intellectual property (local and international)	0.17	0.63	2
High market turbulence (high rate of rapid changes in the composition of customers, their preferences, and competitive market conditions)	0.06	-0.15	6
High governmental support and funding	0.11	0.26	3

Given the output of Promethee software, the value of Φ for the first criterion, which is high conduciveness of the network structure, is 0.92 and more than other sub-criteria. Therefore, it is the most prior sub-criterion of Environmental Factors.

5 | Discussion

In this study, inspired by the research literature, the importance of increasing the participation of SMEs in OIA is identified. Despite the fact that literature continuously emphasizes the critical role of OIS in increasing the productivity of SMEs, most of the previous researches, especially domestic ones, have studied the role of product characteristics, inter-organizational and environmental factors. Therefore, the potential impact of technological turbulence and market turbulence, especially in Iran, have been less considered. The present study has been helpful in collecting a number of moderating factors that have contributed to the degree of SMEs' participation in OIA in Iran.

One of the categories of factors moderating Iranian SMEs' degree of participation in OIA is product characteristics. Product characteristics can enhance or hinder the probability of SMEs benefiting from OIA. These characteristics include the complexity of products, the product life cycle, and the product's earliest and latest stages of maturity. This research confirms the relative priority of the product characteristics in Iranian SMEs' degree of participation in OIA by analyzing the collected data. Due to the reasons given by experts, product characteristics include high complexity of products, short product life cycle, and being in the product's earliest and latest stages of maturity, increase the degree of Iranian SMEs' participation in OIA. In other words, these factors increase the likelihood of benefiting from participating in OIA by SMEs. Experts assert that the production process of technologically complicated products is often longer and it requires more financial resources. It is apparent, therefore, that SMEs with more limited financial resources and capabilities would be more willing to collaborate in OIA. Also, the risk of pursuing an advanced technology project is high. Innovation collaboration between companies enables them to divide this risk by allocating less funding by each partner. On the other hand, products with more sophisticated technology will expose the company to the risk of intellectual property disclosure, because complex products often consist of relevant products. Experts believe this reason is one of the main motivations for SMEs to participate in OIA. Such a result is in line with other studies [9], [10], [13], [19], [20]. Although it is rarely possible to find a study examining the environment of Iran. Therefore, the first hypothesis (H1), that is the product characteristics moderate the degree of Iranian SMEs' participation in OIA, is approved.

The research emphasizes the undeniable moderating role of inter-organizational factors in Iranian SMEs' degree of participation in OIA. These factors include openness of firms' business models, the need to access others' complementary resources and capabilities, SMEs' OI capabilities (absorptive capacity, search

Alroaia | J. Appl. Res. Ind. Eng. 10(1) (2023) 125-140

capabilities, cultural intelligence, conduciveness of organizational structure, and strong leadership), the need to share project failure risks, OI culture, and successful OI experiences. This conclusion confirms a large body of literature on the subject that inter-organizational factors have proven to have moderating impact on the degree of SMEs' participation in OIA [2], [3], [7]-[9], [11], [13], [19], [21]-[29]. Due to the importance of this issue, experts point out that the SMEs' managers' beliefs about the effectiveness and efficiency of participating in OIA depend heavily on the success or the failure of their past experiences and if their expectations don't meet the desired outcomes it may even lead to the formation of a cognitive bias against participating in such activities. They believe that acquiring successful experiences in the past will lead to optimism and willingness to participate in this field in the future. Therefore, the second hypothesis is confirmed. In other words, inter-organizational factors moderate the Iranian SMEs' degree of participation in OIA.

In addition, the findings of this study show the moderating effect of environmental factors on Iranian SMEs' degree of participation in OIA. Environmental factors contain scale-intensiveness of an industry, conduciveness of the network structure, high-technology and knowledge-intensiveness of an industry, degree of industry reliance on secrecy, the easiness of entering OI networks, the costs and the risks of OI (costs associated with the size of the network, language diversity, cultural constraints along with the risks associated with the disclosure of intellectual property, reducing the cost-benefit ratio of OIA, and partner insolvency), technological turbulence, legal interventions and high legal stability and protection of intellectual property (local and international), market turbulence, and governmental support and funding. Many studies confirm this conclusion [2], [3], [7]-[13], [15], [18], [19], [24]-[27], [30]-[34]. Particularly, experts point out the higher the team members of an OI network conduciveness, the higher the SMEs' willingness to participate in OIA. Higher conduciveness means committed partners in allocating resources, capabilities, experiences, trust, and expertise. Experts believe that conduciveness of the network structure also depends on the commitment of partners to actively participate in the collaboration network, acceptance of new ideas, and acceptance of new partners. Consequently, the third hypothesis (H3) is approved. In other words, environmental factors moderate Iranian SMEs' degree of participation in OIA.

Finally, despite the less-considered role of technological turbulence and market turbulence in scholars, particularly domestic ones, and despite their rankings as the sixth and seventh important factors, this research has strongly confirmed its positive impact on Iranian SMEs' degree of participation in innovation activities. Due to the strong moderating impact of these factors on Iranian SMEs' degree of participation in innovation activities, along with their relatively low priority, experts point out different reasons. They justify that when the SMEs' market environment becomes more turbulent and vague, under which circumstance the needs and preferences of customers frequently change, these organizations should strive to compete with other companies and stay ahead of the competition and need to acquire and exploit externally-developed knowledge and technology. The competitive environment is chaotic when there is a rapid change in technology that will create new expectations in the market, but because the costs of shifting to a new technology is high, due to the current economic conditions in Iran, and the type of industry in the assessment of this criterion is very influential, the role of these two factors, despite their high impact, is considered less important than other factors. By the way, the ambiguity in the environment (turbulent environment) and the low technologically developed SMEs (non-turbulent technology status) in Iran could have caused such a result, according to the experts. Therefore, two sub-hypotheses of this research were also confirmed. In other words, technical turbulence moderates Iranian SMEs' degree of participation in OIA (H31); and market turbulence moderates Iranian SMEs' degree of participation in OIA (H32).

6 | Theoretical Suggestions

As said, positive and negative findings regarding the impact of SMEs' participation in OIA in various researches show that, as many SMEs all over the world benefit from these types of activities, some of them may be reversely affected [28]. This indicates the existence of moderating factors affecting the degree and the effectiveness of SMEs' participation in OIA. Therefore, this research aims to collect the set of these factors and categorize them into three groups (product characteristics, inter-organizational factors, and environmental factors) considering the role of technological turbulence and market turbulence according to the environment of Iran as shown in a developed model in *Fig. 1*. In other words, these factors or conditions affect the probability of an SME to benefit from participation in OIA. Therefore, the following theoretical suggestions can be offered to SMEs' managers before they decide to engage in OI contributions:

I. SMEs should explicitly outline the degree and the scope of participation in joint activities in the field of OI.
II. SMEs need to evaluate the OI strategic options before deciding to take part in the OI area. According to the model presented in this study, there are moderating factors that affect the probability of benefiting from the strategic decisions in the field of OI that should be taken into consideration.
III. SMEs should continuously monitor the changes of the moderating factors, in order to adjust their degree of participation in OIA, and thus, maximize their profitability of participating in this field, besides minimizing the disadvantages of OIA such as costs and risks.

7 | Limitations and Directions for Future Research

The present research has been subject to some limitations in spite of the importance of its findings to identify the moderating factors that increase or hinder the participation of SMEs in OIA. This research is based on a convenient sampling method with the focus on managers and experts of SMEs in the industrial parks. Therefore, it is recommended that in order to generalize the results more accurately and achieve a comprehensive view in other parts of Iran and with a larger sample further researches are helpful. Additionally, few studies have been done to analyze the strategies that SMEs should implement confronting each moderating factor and the success rate of SMEs under such circumstances. For example, further research will be needed to assess the factors affecting SME participation in OIA and the effectiveness of SME performance including the role of OIS. Also, a better understanding of how SMEs can increase their success through OIA at the regional, national, and international levels is a matter that requires future scholars' attention.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Funding Source Declaration

Younos Vakil Alroaia performed conceptualization, methodology, software, and literature review and manuscript preparation. Vakil Alroaia performed data correction, writing original draft preparation, writing reviewing and editing references.

Author Agreement

I am submitting a manuscript for consideration of publication in JARIE. The manuscript is entitled "OI and SMEs: Providing a Model for Business Development". It has not been published elsewhere and that it has not been submitted simultaneously for publication elsewhere.

Acknowledgment

The author wish to thank Shima Mesaeli for assistance with particular technique and methodology. Also, the authors would like to thank the editor and the two anonymous reviewers for the constructive comments on improving an early version of this paper.

References

[1] UNIDO. (2003). *Strategy document to enhance the contribution of an efficient and competitive small and medium-sized enterprise sector to industrial and economic development in the Islamic Republic of Iran*. https://open.unido.org

[2] Shaikh, M., & Levina, N. (2019). Selecting an open innovation community as an alliance partner: looking for healthy communities and ecosystems. *Research policy*, *48*(8), 103766.

[3] Kratzer, J., Meissner, D., & Roud, V. (2017). Open innovation and company culture: internal openness makes the difference. *Technological forecasting and social change*, *119*, 128-138.

[4] Chesbrough, H., Vanhaverbeke, W., & West, J. (Eds.). (2006). *Open innovation: researching a new paradigm*. Oxford University Press on Demand. https://www.amazon.com/Open-Innovation-Researching-New-Paradigm/dp/0199226466

[5] Dahlander, L., & Gann, D. M. (2010). How open is innovation?. *Research policy*, *39*(6), 699-709.

[6] Mashdai, A., & Taghavi, Z. (2010). Open innovation: the most important competitive advantage of companies in the market. *The 10th conference of the Iranian aerospace society*, Tehran, Iran. Civilica. (**In Persian**). https://civilica.com/doc/134764

[7] Hung, K. P., & Chou, C. (2013). The impact of open innovation on firm performance: the moderating effects of internal R&D and environmental turbulence. *Technovation*, *33*(10-11), 368-380. https://doi.org/10.1016/j.technovation.2013.06.006

[8] Afrasiabi, R. (2014). The effect of internal and external factors on open innovation. *Journal of technology development management*, *2*(2), 29-46. (**In Persian**). DOI: 10.22104/jtdm.2015.135

[9] Cornell, B. T. (2012). *Open innovation strategies for overcoming competitive challenges facing small and mid-sized enterprises*. University of Maryland University College.

[10] Lyu, Y., He, B., Zhu, Y., & Li, L. (2019). Network embeddedness and inbound open innovation practice: the moderating role of technology cluster. *Technological forecasting and social change*, *144*, 12-24. https://doi.org/10.1016/j.techfore.2019.03.018

[11] Naqshbandi, M. M., & Tabche, I. (2018). The interplay of leadership, absorptive capacity, and organizational learning culture in open innovation: testing a moderated mediation model. *Technological forecasting and social change*, *133*, 156-167. https://doi.org/10.1016/j.techfore.2018.03.017

[12] Pervan, S., Al-Ansaari, Y., & Xu, J. (2015). Environmental determinants of open innovation in Dubai SMEs. *Industrial marketing management*, *50*, 60-68. https://doi.org/10.1016/j.indmarman.2015.05.029

[13] HashemiDehaghi, Z. (2019). Open innovation and its necessity in organizations. *Journal of industry and university*, *26*(25), 25. (**In Persian**). http://jiu.ir/article-1-92-fa.html

[14] Rahmanian, S., & Nasr, M. (2010). A network mediation model for open innovation in SMEs. *Fourth conference on technology management*, Tehran, Iran. Civilica. (**In Persian**). https://civilica.com/doc/94593/

[15] Ahmadi, Z., Shafiei, M., & Vakil, Y. (2018). The impact of open innovation on company innovative performance with the emphasis on moderating role of market turbulence (case study: small and medium-sized manufacturing companies in Semnan city). *Quarterly journal of industrial technology development*, *16*(32), 73-92. http://jtd.iranjournals.ir/article_30506_en.html

[16] Hossain, M., & Kauranen, I. (2016). Open innovation in SMEs: a systematic literature review. *Journal of strategy and management*, *9*(1), 58-73. https://doi.org/10.1108/JSMA-08-2014-0072

[17] West, J., & Bogers, M. (2017). Open innovation: current status and research opportunities. *Innovation*, *19*(1), 43-50. https://doi.org/10.1080/14479338.2016.1258995

[18] Frishammar, J., Ericsson, K., & Patel, P. C. (2015). The dark side of knowledge transfer: exploring knowledge leakage in joint R&D projects. *Technovation*, *41*, 75-88. https://doi.org/10.1016/j.technovation.2015.01.001

[19] Parhizkar, M. M., Jokar, A. A., & Darini, V. M. (2014). Identification of factors influencing on organizational innovation based on open innovation paradigm: case study, publication industry. *Industrial management studies*, *11*(31), 101-125. https://jims.atu.ac.ir/article_193_en.html

[20] Vaezi Nejad, M., & Sadeghi, J., (2008), Open innovation: the key strategy of internal technology-based companies for competition in internal and external markets. *Third international conference on strategic management*, Tehran, Iran. Civilica. (**In Persian**). https://civilica.com/doc/67151/

[21] Radziwon, A., & Bogers, M. (2019). Open innovation in SMEs: exploring inter-organizational relationships in an ecosystem. *Technological forecasting and social change*, *146*, 573-587. https://doi.org/10.1016/j.techfore.2018.04.021

[22] Lu, P., Yuan, S., & Wu, J. (2017). The interaction effect between intra-organizational and inter-organizational control on the project performance of new product development in open innovation. *International journal of project management*, *35*(8), 1627-1638. https://doi.org/10.1016/j.ijproman.2017.09.009

[23] Spithoven, A., Clarysse, B., & Knockaert, M. (2010). Building absorptive capacity to organise inbound open innovation in traditional industries. *Technovation*, *30*(2), 130-141. https://doi.org/10.1016/j.technovation.2009.08.004

[24] Brockman, P., Khurana, I. K., & Zhong, R. I. (2018). Societal trust and open innovation. *Research policy*, *47*(10), 2048-2065. https://doi.org/10.1016/j.respol.2018.07.010

[25] Sivam, A., Dieguez, T., Ferreira, L. P., & Silva, F. J. (2019). Key settings for successful open innovation arena. *Journal of computational design and engineering*, *6*(4), 507-515. https://doi.org/10.1016/j.jcde.2019.03.005

[26] Javid, Gh., & Bagherinejad, J. (2012). The relationship between open innovation process and the absorptive capacity of businesses. *Rushd technology magazine*, *31*(8), 51-61.

[27] Safdari Ranjbar, M., Manteghi, M., & Tavakkoli, Gh. (2014). Open innovation, a comprehensive view on concepts, approaches, trends and key success factors. *Journal roshd-e-fanavari*, *10*(40), 10-17. https://www.sid.ir/paper/145013/fa

[28] Gkypali, A., Arvanitis, S., & Tsekouras, K. (2018). Absorptive capacity, exporting activities, innovation openness and innovation performance: a SEM approach towards a unifying framework. *Technological forecasting and social change*, *132*, 143-155. https://doi.org/10.1016/j.techfore.2018.01.025

[29] Lopes, C. M., Scavarda, A., Hofmeister, L. F., Thomé, A. M. T., & Vaccaro, G. L. R. (2017). An analysis of the interplay between organizational sustainability, knowledge management, and open innovation. *Journal of cleaner production*, *142*, 476-488. https://doi.org/10.1016/j.jclepro.2016.10.083

[30] Mazzola, E., Perrone, G., & Kamuriwo, D. S. (2015). Network embeddedness and new product development in the biopharmaceutical industry: the moderating role of open innovation flow. *International journal of production economics*, *160*, 106-119. https://doi.org/10.1016/j.ijpe.2014.10.002

[31] Nestle, V., Täube, F. A., Heidenreich, S., & Bogers, M. (2019). Establishing open innovation culture in cluster initiatives: the role of trust and information asymmetry. *Technological forecasting and social change*, *146*, 563-572. https://doi.org/10.1016/j.techfore.2018.06.022

[32] Jugend, D., Jabbour, C. J. C., Scaliza, J. A. A., Rocha, R. S., Junior, J. A. G., Latan, H., & Salgado, M. H. (2018). Relationships among open innovation, innovative performance, government support and firm size: comparing Brazilian firms embracing different levels of radicalism in innovation. *Technovation*, *74*, 54-65. https://doi.org/10.1016/j.technovation.2018.02.004

[33] Chen, T., Li, F., Chen, X. P., & Ou, Z. (2018). Innovate or die: how should knowledge-worker teams respond to technological turbulence?. *Organizational behavior and human decision processes*, *149*, 1-16. https://doi.org/10.1016/j.obhdp.2018.08.008

[34] Ebrahimpour, M., Nopasand Asil, S. M., & Ahmadi Saravani, M. (2015). The impact of strategic flexibility on companies market performance: the mediating role of open innovation. *Quarterly journal of industrial technology development*, *13*(26), 5-16. http://jtd.iranjournals.ir/article_18081.html?lang=en

[35] Gorsuch, R. (1983). *Factor analysis* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

[36] Saaty, T. L. (1996). *Decision making with dependence and feedback: the analytic network process* (Vol. 4922, No. 2). Pittsburgh: RWS Publications.

[37] Carayannis, E. G., & Meissner, D. (2017). Glocal targeted open innovation: challenges, opportunities and implications for theory, policy and practice. *The journal of technology transfer*, *42*(2), 236-252. https://doi.org/10.1007/s10961-016-9497-0

[38] Chesbrough, H., Vanhaverbeke, W., & West, J. (Eds.). (2014). *New frontiers in open innovation*. Oup Oxford.

[39] Azami, A., Kazerooni, H., & Mohammadkhani Ghiasvand, H. (2021). Improvement solutions for production and service functions in the innovation system: prioritizing solutions using FMADM approach. *Journal of applied research on industrial engineering*, *8*(Spec. Issue), 1-23. DOI: 10.22105/jarie.2021.274952.1260

[40] Shariatmadari Serkani, E. (2015). Using DEMATEL–ANP hybrid algorithm approach to select the most effective dimensions of CRM on innovation capabilities. *Journal of applied research on industrial engineering*, *2*(2), 120-138.

[41] Hair Junior, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. SAGE Publications

[42] Brans, J. P., Vincke, P., & Mareschal, B. (1986). How to select and how to rank projects: the PROMETHEE method. *European journal of operational research*, *24*(2), 228-238.

Alroaia | J. Appl. Res. Ind. Eng. 10(1) (2023) 125-140

E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 141–154.

Paper Type: Research Paper

A Bi-Objective Optimal Task Scheduling Model for Two-Machine Robotic-Cell Subject to Probable Machine Failures

Bahareh Vaisi[1,*], Hiwa Farughi[2], Sadigh Raissi[3], Heibatolah Sadeghi[2]

[1] Young Researchers and Elite Club, South Tehran Branch, Islamic Azad University, Tehran, Iran; st_b_vaisi@azad.ac.ir.
[2] Department of Engineering, University of Kurdistan, Sanandaj, Iran; h.farughi@uok.ac.ir; h.sadeghi@uok.ac.ir.
[3] Department of Industrial Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran; Raissi@azad.ac.ir.

Citation:

Vaisi, B., Farughi, H., Raissi, S., & Sadeghi, H. (2023). A bi-objective optimal task scheduling model for two-machine robotic-cell subject to probable machine failures. *Journal of applied research on industrial engineering, 10*(1), 141-154.

Received: 09/04/2022 Reviewed: 12/05/2022 Revised: 17/07/2022 Accepted: 08/08/2022

Abstract

In this study, we model a stochastic scheduling problem for a robotic cell with two unreliable machines susceptible to breakdowns and subject to the probability of machine failure and machine repair. A single gripper robot facilitates the loading/unloading of parts and cell-internal movement. Since it is more complicated than the other cycles, the focus has been on the S_2 cycle as the most frequently employed robot movement cycle. Therefore, a multi-objective mathematical formulation is proposed to minimize cycle time and operational costs. The ε-constraint method is used to solve small-sized problems. Non-dominated Sorting Genetic Algorithm II (NSGA-II), is used to solve large-sized instances based on a set of randomly generated test problems. The results of several test problems were compared with those of the GAMS software to evaluate the algorithm's performance. The computational results indicate that the proposed algorithm performs well. Compared to GAMS software, the average results for maximum spread (D) and Non-Dominated Solutions (NDS) are 0.02 and 0.04, respectively.

Keywords: Breakdowns, Identical parts, NSGA-II, Probable failures, Robotic cell, Scheduling.

1 | Introduction

Flexible Manufacturing Systems (FMS) play an essential role in production systems and promptly respond to customer demands [1]. In such systems, robots are typically responsible for picking up products and loading/unloading machines, consequently, robots can facilitate the process and improve system productivity. A robotic cell is a type of FMS consisting of *m* Computerized Numerical Control (CNC) machines; some robotic cells also have an input and output buffer. In robotic cell problems, the primary focus of research is on scheduling robot tasks. Scheduling optimization, which improves the system's productivity when a manufacturing system must deal with uncertainty, is essential.

Scheduling is one of the most critical issues in all systems that optimize one or more objectives by considering the resource and operation constraints [2]. This issue's application in various domains, such as production and service systems, can assist the systems in achieving their desired performance

Corresponding Author: st_b_vaisi@azad.ac.ir

https://doi.org/10.22105/jarie.2022.336768.1465

goals [3]. As we can see in service systems such as hospitals, several factors can lead to a stop in operations, and proper scheduling that considers resources and constraints increases efficiency [4], [5]. Machine breakdowns and repair times have been relaxed in the scheduling optimization process of robotic cells so far. Therefore, the current study proposes scheduling a two-machine robotic cell that confronts breakdowns. The robot performs under the S_2 cycle as the most commonly used robot's movement cycle. We use the ε-constraint method to solve the small-sized and NSGA-II to solve the large-sized problems.

The review of the literature is summarized in the following section. Section 3 defines the problem and presents the mathematical model used to model it. The solution approach is presented in Section 4, followed by some numerical examples, sensitivity analysis of results, and discussion based on the model in Section 5. Finally, Section 6 reports on the paper's conclusions.

2 | Literature Review

In most previous studies on robotic manufacturing cells, the scheduling problem considers a single criterion. Their research's most important objective functions were minimizing cycle time and maximizing the cell's throughput. Such as the papers cited in [6]-[15]. Hoogeveen's [16] survey of multi-criteria scheduling was published. The problem of multi-objective scheduling in robotic cells has been studied by [17]-[30], among others.

Given the importance of completing different tasks on time, deterioration and delays between tasks incur enormous costs. Consequently, maintenance is a crucial aspect of industrial environments. Although there have been numerous studies on deterministic robotic cells, the issue of determining an unreliable robotic cell in both machine breakdowns distributed according to an exponential distribution and stochastic processing time (due to the probability of repair time) remains unsolved. Stochastic models incorporate uncertainty and utilize probability distributions in which the data are either known or can be estimated.

Recent studies have shown that stochastic factors, such as machine breakdown and uncertain repair time, significantly impact the scheduling in actual production environments. Considering a multilevel assembly system with multiple sublevel components, Sadeghi et al. [31] stated that it would be impossible to complete the items on time due to random machine breakdowns. They then proposed a mathematical model incorporating the uncertainty of lead time. Additionally, Sadeghi [32] stated that their operating costs might increase due to using tools and machines, thereby increasing the system's expenses. Utilizing preventive maintenance is proposed as a method for reducing operational costs. The history of the study of stochastic robotic cells is as follows.

Some previous studies focus on robotic cells with random processing time, which can be referred to [13], [33]-[35]. Shafiei-Monfared et al. [33] considered a robotic cell consisting of three machines and a robot in the center of the cell when a part processing time element is stochastic. Comparing the cycle times of a variety of scenarios in this robotic cell was an attempt to determine the productivity benefits of each. Geismar and Pinedo [34] presented the first analytic study of robotic cell operations where the process has a stochastic processing time, as is typical in the microlithography portion of semiconductor manufacturing. It was demonstrated how the proximity of the stochastic process to the bottleneck process influences throughput measurement in such cells. The robot's sequence time distribution function was identified and validated through simulation. In a different study, a robotic cell problem with variable processing times was formulated, and the effectiveness of heuristic and metaheuristic solution methods for optimizing output rate was demonstrated [35]. Tonke et al. [13] developed an online-offline scheduling approach based on the assumption of uncertain processing times to address real-world applications such as cluster tools in semiconductor manufacturing. Their research involved a dual-gripper robotic cell problem with pick-up constraints.

Vaisi et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 141-154

According to other studies like [28], [30], [36]-[39], robotic cells operate under a production system with machine failures and repairs. Savsar and Aldaihani [36] developed a model to analyze Performance Measures (PFM) of a Flexible Manufacturing Cell (FMC) consisting of two machines and a robot under various operational conditions, including machine failures and repairs. The model was based on Markov processes and determined closed-form probabilities of system states for calculating PFMs. In a separate study [37], fault-tolerant conditions were incorporated into the model, allowing the FMC to operate in a degraded state. A Markovian model was developed to determine the system's dependability and productivity under various operational conditions. In a similar study by [38], the Markov chain model was developed for both single- and dual-machine FMCs. The model was subsequently generalized to FMC with n machines. Researchers [28], [30], [39] investigated random failures in robotic cells with two and three machines. In recent researches conducted by [30], [39]-[42], the robotic cell produces a variety of parts in an uncertain environment.

This study addresses a stochastic issue for an unreliable two-machine robotic cell when it considers the probability of machine failure and its impact on cycle time uncertainty. A single gripper robot is utilized to load and unload identical parts. Here, the authors focused on the S_2 movement cycle, which is more complicated than the other cycles and is the most commonly used movement cycle for robots.

3 | Problem Definition and Modeling

One type of product flows over only one machine in the manufacturing cell, but there are two identical CNC machines, neither of which has operational priority. Typically, in a 2-machine cell, there are three possible robot cycle options for part displacement: S_1, S_2 and $S_{12}S_{21}$ [6], [40]. As mentioned previously, the scope of this paper is restricted to the S_2 cycle.

In the S_2, a robot initially takes place before the Input Buffer (IB). Then the following operations are followed sequentially by the robot: 1) the robot picks up a part, 2) moves to the first machine (M1), 3) loads M1, 4) the robot moves to the second machine (M2), 5) waits for the previous process to be completed on the part (if it is needed), 6) robot unloads the part from M2, then, 7) transfers the product to the Output Buffer (OB), 8) loads the OB, 9) robot returns to M1, 10) if it is needed the robot waits until the completion of the process, 11) unloads from M1, 12) transports the part to the M2, 13) loads the part on the M2 and finally 14) robot turns back to the IB [6]. $A_{01}\,A_{23}\,A_{12}$ infers the sequence of activities in the S_2 cycle as mentioned; A_{pq} is the robot's activity sequence from station p to station q for $p = 0, 1, 2$ & $q = 1, 2, 3$ (see [6], [40]). A typical linear two-machine robotic cell is shown in *Fig. 1*.

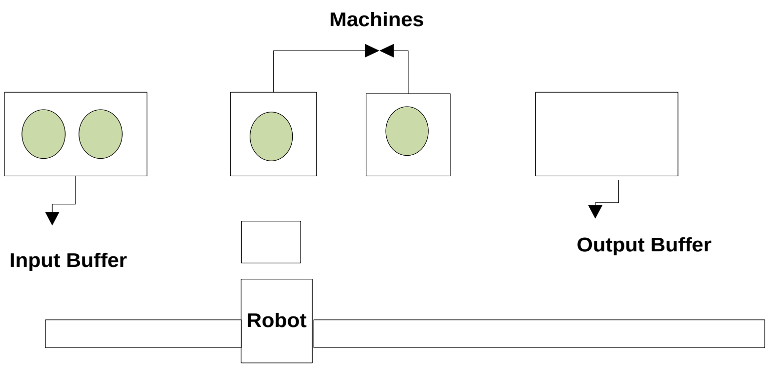


Fig. 1. Linear two machines robotic cell.

3.1 | Assumptions

In addition to the assumptions in [28], the basic assumptions for this study are:

I. Failure in the machines is probable.
II. Some repairs performed to the machine per part's entrance to the cell follow an exponential distribution. The time to repair each machine is equivalent to its processing time.
III. To produce each part, multiple operations are required. Some of the operations are done on the first machine and the rest is done on the second machine.
IV. Existence of probability distribution in repair time may face the processing time with uncertainty. So, it is assumed that; the machines are unreliable and there is a possibility of failure during the operation. It is also assumed that the number of repairs performed per machine for each part follows the Exponential distribution. By arriving the part into the cell, considering the probability of machine failure, there are uncertain processing times that complicate the analysis and modeling of the robotic cell. The following model is developed to analyze the desired robotic cell. The layout of the assumed robotic cell was based on *Fig. 1* and [21].

3.2 | Notations

The following parameters and variables are used in the proposed mathematical model:

${C_0}^k$	Machining cost per k[th] part (\$/min).
C_r	Repair cost for each breakdown (\$/min): without setup costs.
C_p	Preventive maintenance cost (\$/min).
C_{TOOL}	Cost of tool (\$/tool): tools replacement prohibited in an operating cycle.
λ	Failure rate: follows the Exponential distribution.
μ	Repair rate: follows the Exponential distribution.
t_i	Processing time of operation i (min).
$t_{r,j}$	Duration of a repair visit for machine j (min).
W_{R_j}	Duration of maintenance in machine j (min).
W^k	Robot's waiting time in a cycle for the k[th] part fed to the cell (min).
F^k	Total cost for the k[th] part fed to the cell (\$).
${T_{S_2}}^k$	The partial cycle time of S_2 for the k[th] part fed to the cell (min).
X_j^k	Number of repairs performed to the machine j for the k[th] part entered to the cell (integer random variable, exponential distribution (meaning $X_j^k \geq 1$).
a	The processing time for a part on the 1[st] machine.
b	The processing time for a part on the 2[nd] machine.
ϵ	Loading/unloading time.
δ	Time is taken by a robot to move between two adjacent stations.
$pr(TI_\Theta < TI < TI_{\Theta+1})$	The probability of breakdown occurrence for the Time Interval (TI) between TI_Θ and $TI_{\Theta+1}$.

$$O_{ij} = \begin{Bmatrix} 1, & if\ Operation\ i\ is\ allocated\ to\ machine\ j, \\ 0, & otherwise. \end{Bmatrix}$$

3.3 | Modeling

Literature reveals that various studies, such as [43], have focused on minimizing the total production costs of machining, tooling, and maintenance. In many instances, tooling cost is considered a constant value and has no impact on the optimization process. In the present study, tooling costs are also constant. It was added to the formula to complete the concept. Total production cost for the k^{th} part fed to the manufacturing cell was defined as the first objective function while minimizing partial cycle time for the k^{th} part was considered as the second objective function. *Eqs. (1)* and *(2)* represent the preferred objective functions.

Vaisi et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 141-154

In the proposed model, the constraints for robotic cell scheduling are *Eqs. (3)-(5)* and *(8)*. These equations are derived from the robot's move cycle definitions and redefined based on the assumed problem. The time between two consecutive repairs for the first machine is *a* and for the second machine, *b*. Each machine's maintenance time is represented by *Eq. (6)*. The failure rate per machine for each model run is calculated using an Exponential distribution. *Eqs. (7)* and *(9)* are decision variables related to the allocation of operations to machines. The formulation example follows:

$$\text{Min } F^k = \left(\sum_{j=1}^{2} \sum_{i=1}^{n} C_0{}^k \, t_i O_{ij} + C_r \, t_{r,j} \text{ pr } TI_\Theta < TI < TI_{\Theta+1}) \right) + C_{TOOL}, \tag{1}$$

$$\text{Min } T_{S_2}{}^k = 6 \in +8\delta + W^k, \tag{2}$$

s.t.

$$a = \textstyle\sum_{i=1}^{n} t_i O_{i1}) + W_{R_1}, \tag{3}$$

$$b = \textstyle\sum_{i=1}^{n} t_i O_{i2}) + W_{R_2}, \tag{4}$$

$$W^k = \text{Max} \begin{Bmatrix} 0, & aX_1{}^k - \left[2 \in +4\delta + W_{R_1}\right] \\ & bX_2{}^k - \left[2 \in +4\delta + W_{R_2}\right] \end{Bmatrix}, \tag{5}$$

$$W_{R_j} = \textstyle\sum_{i=1}^{n} t_{r,j} \text{ pr } TI_\Theta < TI < TI_{\Theta+1})O_{ij} \quad \text{for } j=1,2, \tag{6}$$

$$O_{i1} + O_{i2} = 1, \tag{7}$$

$$W^k \geq 0, \tag{8}$$

$$O_{ij} \in \{0,1\}. \tag{9}$$

4 | Solution Approach

Bi-objective optimization problems aim to identify a set of Pareto optimal solutions. This research uses the evolved ε-constraint method for small-scale problems. Recently, Vaisi [44] reviewed the application of optimization tools in robotic systems and revealed that the authors of nearly half of the research papers published between 2005 and September 2021 had used heuristic/metaheuristic algorithms to optimize robotic manufacturing system problems. Consequently, a well-known multi-objective meta-heuristic approach, NSGA-II, is utilized to solve the bi-objective model in the current study for large-scale problems. This algorithm is one of the most popular multi-objective optimization algorithms. After presenting the first version of this algorithm in 1995, its developers, the most significant among whom is Debb, presented the second version, NSGA-II, in 2002. *Algorithm 1* shows the Pseudo code of the NSGA-II.

Algorithm 1. The NSGA-II Pseudo code.

1. Create:
   Population size = P_t
   Repeat for a maximum number of iterations,
   $t = 0$
2. Generate child population = Q_t
   Apply:
   - Binary Tournament selection
   - One-point crossover and probable mutation
3. Combine P_t and Q_t to create a new population called Npop
4. Assign rank for each solution based on the non-domination sorting process
5. Create next-generation, (P_t) based on the lowest obtained ranks and highest Crowding Distance (CD)
6. Check the stopping criterion.
   While t < Maximum number of Iterations, do:
   If Yes (Go to step 7)/ If No $t = t + 1$ *and* (Go to step 2)
7. End of the algorithm.

4.1 | Crowding Distance Computation

The following two criteria determine the measures for better solutions:

Rank measure

The solution with the lower non-domination rank is preferred between two alternatives with different ranks. Alternatively, if both points belong to the same front, the point located in a region with fewer points is preferred [45].

Crowding distance

In instances where two selected particles occupy the same rank (both on the same side), the CD criterion is applied, as explained below:

For particles 2 to $n-1$, the CD, $I(d_k)$, is calculated based on *Eqs. (10)* and *(11)*.

$$\mathrm{CD_K} = \mathrm{I\ d_k)} + \cdots + \mathrm{I(d_k)_m}. \tag{10}$$

$$\mathrm{I\ d_k).m} = \frac{\mathrm{I\ K+1).M - I\ K-1).M}}{\mathrm{f_m^{max}}}. \tag{11}$$

The *Eq. (11)* represents the CD for the objective function m. Therefore, $I\ d_k)$ must be calculated and summed for all objective functions, as specified in *Eq. (10)*. After calculating the CD, the particle with the highest CD is selected, [46], [47].

4.2 | Solution Representation

The main objective is to determine the assignment of operations to machines and to plan the arrival of parts and their processes in the robotic cell. For setting parameters in the experimental design, this study uses a certain amount of algorithm repetition as a stopping criterion. The experiments were designed using the Taguchi method to adjust the algorithm's parameters. Therefore, a Taguchi-based experiment on ten randomly generated test problems was designed. Three experiment levels were chosen for each parameter, including crossover rate (p_c) and mutation rate (p_m), based on previous research and trial and error. The experiment levels of these parameters are displayed in *Table 1*.

Table 1. Levels of taguchi experiment.

Parameters	Levels		
	Low	Middle	High
p_c	0.6	0.8	0.9
p_m	0.1	0.15	0.2

As a result of the experiments and the fact that smaller response values are taken into account, the middle level in *Table 1* consists of appropriate combinations based on the average response factor:

Crossover = 80%, Mutation = 15%.

The population size and stopping criteria must be modified to implement the algorithm. By increasing the population size, the algorithm searches for more points in the space, and the quality and distribution of the results improve; however, if the number of population members becomes ten times greater, the time or required memory to solve the problem will be 100. Therefore, the population size for the proposed algorithm is 50. By increasing the number of algorithm replications, the model is given sufficient time to be solved, resulting in better results for larger values of this parameter. However, it should be noted that increasing the number of algorithm replications also increases the elapsed time. The suggested number of replications for the NSGA-II algorithm is 500, and the algorithm stops upon reaching 500 replications. Therefore, the maximum number of iterations is established as a stopping criterion, and 500 iterations are set as the stopping criterion.

A 2-string chromosome is employed to represent the assignment of processing times of operations to the machines. The first string represents the processing times assigned to the M1 and the second string is the processing times assigned to the second machine. There are i operations available for each part to be processed, so the probability density function of selecting operations is uniform. Consequently, in these two strings, a number between 2 and i shows the equivalent processing time allocated to the associated machine. $t_i(j)$ means assigning the i^{th} processing time to machine j. *Fig. 2* shows a chromosome with eight operations as an initial population sample.

$t_1(1)$	$t_2(2)$	$t_3(1)$	$t_4(1)$	$t_5(2)$	$t_6(1)$	$t_7(2)$	$t_8(2)$

Fig. 2. Forming an initial population.

Creating an initial population is the first step. In this study, the initial population is generated by generating 2-string chromosomes proportional to the size of the population. As stated previously, the 2-string chromosomes contain the processing times of operations required to produce i^{th} part. After forming the initial generation, individuals must be chosen to form the subsequent generation. Location of solutions in a Pareto front (lower fronts are superior) and CD are the selection criteria (in the same lower fronts). The new generation should be formed by altering specific characteristics of the parents.

In designing this algorithm, one-point Crossover and Probable Mutation have been used. *Fig. 3* represents the Crossover operation. The numerical values in *Fig. 3*, as an example, specify the value of processing times of operations to produce i^{th} part.

Forming the initial population for this problem (2-string chromosomes):

Parent 1	2.78	5.88	2.28	3.36	4.53	5.92	3.31	4.04
Parent 2	4.22	8.4	8.79	7.64	9.6	5.56	2.95	6.05

After One-point Crossover

Child 1	2.78	5.88	2.28	3.36	4.53	5.56	2.95	6.05
Child 2	4.22	8.4	8.79	7.64	9.6	5.92	3.31	4.04

Fig. 3. One-point Crossover.

5 | Results and Discussion

The proposed solution approach was tested on ten different test problems. These test problems are randomly generated by MATLAB R2016b and executed on an ASUS laptop with 8 GB of RAM and an Intel(R) Core(TM) i7-4500U processor running at 1.80 GHz 2.40 GHz. The designated test problems are listed in *Table 2*. *Table 3* describes the parameters and defined values for the considered robotic cell.

Table 2. Designated examples.

Test Problem	Number of Operations Per Part	Processing Times
1	8	(2,5)
2	8	(2,10)
3	8	(2,15)
4	8	(2,30)
5	20	(2,5)
6	20	(2,10)
7	20	(2,15)
8	30	(2,5)
9	30	(2,10)
10	30	(2,15)

Table 3. Characteristics of required parameters.

Parameters		
$C_{TOOL}=5$	$\epsilon=2$	$\mu=2$
$C_o=100$	$\delta=0.5$	$\lambda=3$
$C_r=20$	$C_p=12$	$t_r=[2,3]$

The results of running the NSGA-II algorithm on the test problems are presented in *Table 4*. It should be noted that for each test problem, the algorithm has been executed five times, and *Table 4* contains the best answers. Processing times have a uniform distribution within the specified range. In most test problems, the distance between the upper and lower bounds of the objective functions increases as the number of part operations increases.

Table 4. Characteristics.

Test Problem	The Lower Bound of the First Objective Function	The Upper Bound of the First Objective Function	The Lower Bound of the Second Objective Function	The Upper Bound of the Second Objective Function
1	2265.34	2346.18	91.01	91.21
2	2474.10	2721.64	110.72	117.54
3	2309.42	2369.03	102.96	104.51
4	5452.48	6275.21	428.44	1002.29
5	7652.46	8153.35	909.26	1092.75
6	6345.30	6353.20	731.17	777.68
7	6316.70	7321.20	725.25	731.91
8	11566.08	12492.32	2153.21	2307.21
9	15698.68	17440.95	3486.69	3747.26
10	19225.42	21234.51	4275.92	4483.20

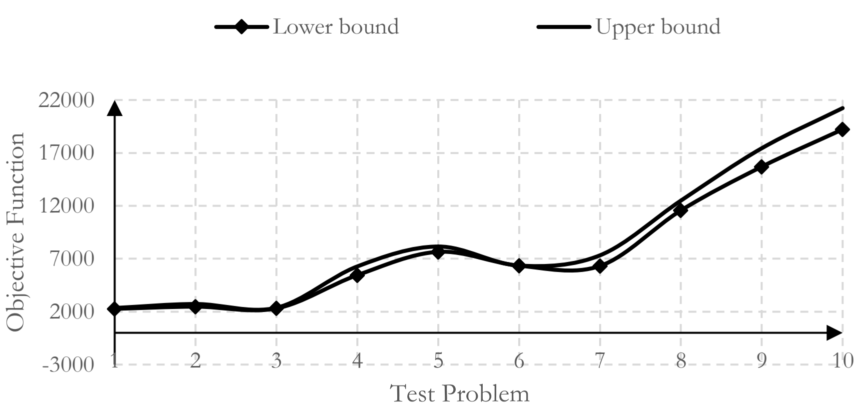


Fig. 4. Lower/upper bound changes of the first objective function in the test problems.

Figs. 4 and *5* demonstrate that the objective functions increase in value as the number of operations per part increases. In addition, the difference between the upper and lower bounds in the small-sized problems is not detectable, except for test problem number 4. As the processing time increases in this test problem, it affects the cycle time and total cost. The difference between the bounds is minor in instances with more than eight operations per part.

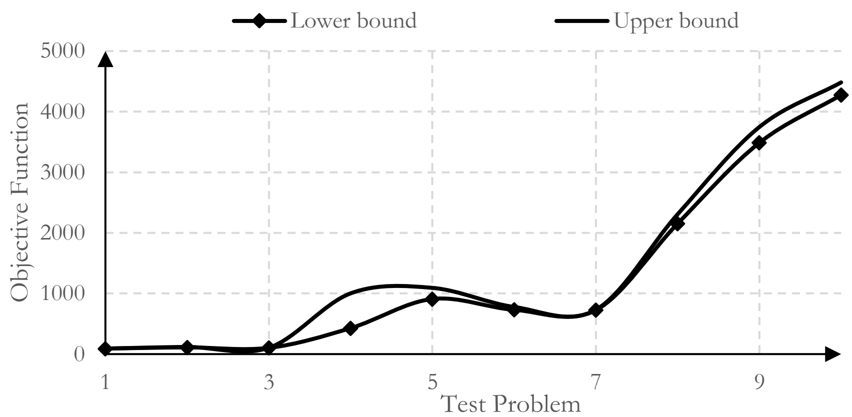


Fig. 5. Lower/upper bound changes of the second objective function in the test problems.

Changes in the first objective function are compared to alterations in the processing TI under the assumption of a constant number of operations. These variations are illustrated in *Figs. 6* to *8*. Consequently, increasing the processing TI causes the difference between the upper and lower total cost bounds to rise (first objective function).

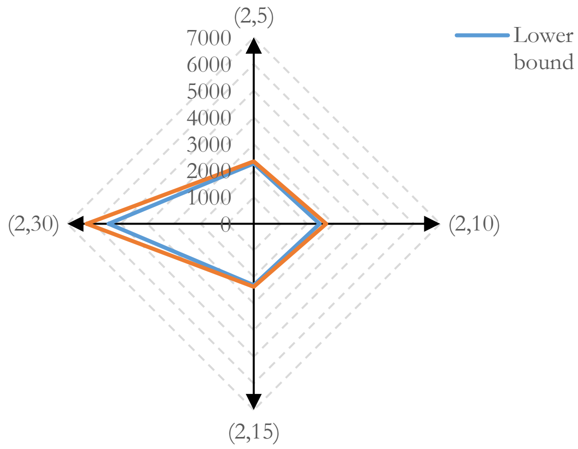


Fig. 6. Changes in the first objective function compared to changes in the processing TI for eight operations (problems 1 to 4).

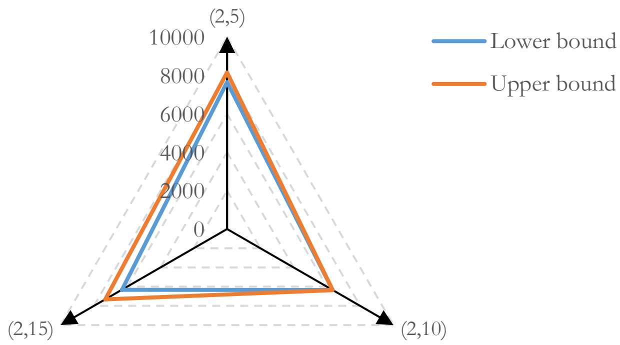


Fig. 7. Changes in the first objective function compared to changes in the processing TI for 20 operations (problems 5 to 7).

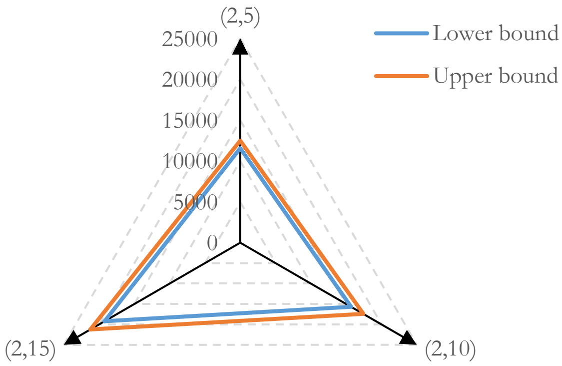


Fig. 8. Changes in the first objective function compared to changes in the processing TI for 30 operations (problems 8 to 10).

5.1 | Comparison of the Algorithms

Different test problems of varying sizes are executed to compare the proposed algorithms. The performance of the algorithms is then evaluated using three standard evaluation indices (criteria): computational time, maximum spread (D), and Non-Dominated Solutions (NDS). The size of the problem is the number of operations per part. The computational time is the average time required to provide a solution. Maximum spread evaluates the variety and distribution of Pareto front solutions using *Eq. (12)*. The performance of an algorithm with a higher maximum spread is superior. Finally, NDS displays the number of NDS obtained for each test problem [48], [49].

$$D = \sqrt{\sum_{m=1}^{M} \left(\max_i f_m^i - \min_i f_m^i\right)^2}. \tag{12}$$

Table 5 displays the mean values of the comparison metrics for each test problem based on the GAMS and NSGA-II results.

Table 5. The result of the NSGA-II algorithm and GAMS for the proposed programming.

Test Problem	Number of Operations Per Part	Processing Times	GAMS Results NDS	D	T(s)	NSGA-II Results NDS	D	T(s)
1	8	(2,5)	9	1913.10	539	9	1913.10	42
2	8	(2,10)	12	1284.20	932	12	1284.20	85
3	8	(2,15)	17	2455.20	1450	17	2455.20	259
4	8	(2,30)	24	2884.40	4356	23	2884.40	587
5	20	(2,5)	29	12871.30	5124	27	12256	218
6	20	(2,10)	45	23105.38	9879	45	22890	759
7	20	(2,15)	49	40972.12	15491	46	39681	1412
8	20	(2,30)	58	52321.72	35270	56	50227	3104
9	30	(2,5)	35	17981.00	22872	32	17683	1862
10	30	(2,10)	56	40201.00	42501	52	39681	2412
11	30	(2,15)	67	82812.23	61731	60	80685	2810
12	30	(2,30)	76	145238.45	71025	73	142256	3654
13	50	(2,5)	49	46208.39	48654	44	45773	3940
14	50	(2,10)	__	__	*	59	55764	4105
15	50	(2,15)	__	__	*	65	85282	6120
16	50	(2,30)	__	__	*	74	149283	7345

According to *Table 5*, the ε-constraint method can achieve the Pareto optimal solution set for small-sized problems. However, the NSGA-II algorithm could obtain the Pareto optimal solution set for the first four test problems; in the following four test problems, with 20 operations per part, the NSGA-II

Vaisi et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 141-154

results are very close to the results of the GAMS. Furthermore, the larger the size of the test problems, the more the proposed solutions have computational time. Therefore, to solve the large-sized problems, the NSGA-II algorithm was applied.

Relative Percentage Deviation (RPD) is calculated for 12 samples of the test problems using *Eq. (13)* to evaluate the results of NSGA-II. The average of RPD per index is determined for the algorithm, and according to the low value of RPD, this method applies to larger problems as well. *Table 6* displays the results.

$$\mathrm{RPD} = \left|\frac{\mathrm{Alg_{Sol}} - \mathrm{Best_{Sol}}}{\mathrm{Best_{Sol}}}\right| \times 100. \tag{13}$$

Table 6. The RPD of the NSGA-II algorithm for the test problems.

Test Problem	Number of Operations Per Part	Processing Time	RPD NDS	D	T(s)
1	8	(2,5)	0	0	0.92
2	8	(2,10)	0	0	0.91
3	8	(2,15)	0	0	0.82
4	8	(2,30)	0.04	0	0.86
5	20	(2,5)	0.07	0.05	0.96
6	20	(2,10)	0	0.01	0.92
7	20	(2,15)	0.06	0.03	0.91
8	20	(2,30)	0.03	0.04	0.91
9	30	(2,5)	0.08	0.02	0.92
10	30	(2,10)	0.07	0.01	0.94
11	30	(2,15)	0.10	0.03	0.95
12	30	(2,30)	0.04	0.02	0.95
Average	-	-	0.04	0.02	0.91

6 | Conclusion

The proposed model was primarily concerned with minimizing the production cost and S_2 cycle time in a two-machine, identical-parts robotic manufacturing cell subject to breakdowns such as machine failures and repairs. The problem was formulated, and a well-known metaheuristic algorithm, NSGA-II, was used to solve this bi-objective model. The solution approach was evaluated on some randomly generated problems, and the results were presented as the upper and lower bounds for the two objective functions. Due to the insignificant difference between the upper and lower bounds, the mean value can represent the real-valued amount of the objective functions. The results showed the robustness of the model and the algorithm. Expanding this problem and working on multiple parts of robotic cells for future research are recommended.

Conflicts of Interest

All co-authors have read and approved the manuscript, and there are no financial conflicts to disclose. We certify that the submission is original and is not currently being reviewed by another publication.

References

[1] Farughi, H., Dolatabadiaa, M., Moradi, V., Karbasi, V., & Mostafayi, S. (2017). Minimizing the number of tool switches in flexible manufacturing cells subject to tools reliability using genetic algorithm. *Journal of industrial and systems engineering*, *10*(Spec. Issue), 17-33. http://www.jise.ir/article_33655_b336d3237fefd7e86dedeaad450bb15b.pdf

[2] Zanjani, B., Amiri, M., Hanafizadeh, P., & Salahi, M. (2021). Robust multiobjective hybrid flow shop scheduling. *Journal of applied research on industrial engineering*, *8*(1), 40-55. https://dx.doi.org/10.22105/jarie.2021.252651.1202

[3] Rashidi, H., & Hassanpour, M. (2020). A deep-belief network approach for course scheduling. *Journal of applied research on industrial engineering*, 7(3), 221-237. https://dx.doi.org/10.22105/jarie.2020.243184.1185

[4] Hamid, M., & Tavakkoli-Moghaddam, R., Vahedi-Nouri, B., & Arbabi, H. (2020). A mathematical model for integrated operating room and surgical member scheduling considering lunch break. *International journal of research in industrial engineering*, *9*(4), 304-312.

[5] Khalili, N., & Shahnazari Shahrezaei, P., & Abri, A. G. (2020). A multiobjective optimization approach for a nurse scheduling problem considering the fatigue factor (case study: Labbafinejad hospital). *Journal of applied research on industrial engineering*, 7(4), 396-423. https://dx.doi.org/10.22105/jarie.2020.259483.1215

[6] Sethi, S. P., Sriskandarajah, C., Sorger, G., Blazewicz, J. & Kubiak, W. (1992). Sequencing of parts and robot moves in a robotic cell. *International journal of flexible manufacturing systems*, *4*, 331-358. https://doi.org/10.1007/BF01324886

[7] Fathian, M., Kamalabadi, I. N., Heydari, M., & Farughi, H. (2011). A Petri net model for part sequencing and robot moves sequence in a 2-machine robotic cell. *Journal of software engineering and applications*, *4*(11), 603. http://www.SciRP.org/journal/jsea

[8] Fathian, M., Nakhai Kamalabadi, I., Heydari, M., & Farughi, H. (2012). A petri net model for part sequencing and robot moves sequence in a 2-machine robotic cell. *Journal of software engineering and applications*, *4*, 603-608. https://www.scirp.org/html/8574.html

[9] Fathian, M., Nakhai Kamalabadi, I., Heydari, M., Farughi, H., & Naseri, F. (2013). Applying metaheuristic algorithms for output rate analysis in two-machine robotic manufacturing cells. *International journal of advanced robotic systems*, *10*(169), 1-11. https://doi.org/10.5772%2F56051

[10] Majumder, A., & Laha, D. (2016). A new cuckoo search algorithm for 2-machine robotic cell scheduling problem with sequence-dependent setup times. *Swarm and evolutionary computation*, *28*, 131-143. https://doi.org/10.1016/j.swevo.2016.02.001

[11] Wang, Z., Zhou, B., Trentesaux, D., & Bekrar, A. (2017). Approximate optimal method for cyclic solutions in multi-robotic cell with processing time window. *Robotics and autonomous systems*, *98*, 307-316. https://doi.org/10.1016/j.robot.2017.09.020

[12] Gultekin, H., Coban, B., & Akhlaghi, V. E. (2018). Cyclic scheduling of parts and robot moves in m-machine robotic cells. *Computers & operations research*, *90*, 161-172. https://doi.org/10.1016/j.cor.2017.09.018

[13] Tonke, D., Grunow, M., & Akkerman, R. (2019). Robotic-cell scheduling with pick-up constraints and uncertain processing times. *IISE transactions*, *51*(11), 1217-1235. https://doi.org/10.1080/24725854.2018.1555727

[14] Bożejko, W., Pempera, J., Smutnicki, C., & Wodecki, M. (2019). Cyclic scheduling in the manufacturing cell. In *Modelling and performance analysis of cyclic systems* (pp. 49-62). Cham: Springer International Publishing.

[15] Kim, H. J., & Lee, J. H. (2021). Scheduling of dual-gripper robotic cells with reinforcement learning. *IEEE transactions on automation science and engineering*, *19*(2), 1120-1136. https://doi.org/10.1109/TASE.2020.3047924

[16] Hoogeveen, H. (2005). Multicriteria scheduling. *European journal of operational research*, *167*(3), 592–623. https://doi.org/10.1016/j.ejor.2004.07.011

[17] Kayan, R. K. & Akturk, M. S. (2005). A new bounding mechanism for the cnc machine scheduling problems with controllable processing times. *European journal of operational research*, *167*(3), 624–643. https://doi.org/10.1016/j.ejor.2004.07.012

[18] Gurel, S. & Akturk, M. S. (2007). Considering manufacturing cost and scheduling performance on a CNC turning machine. *European journal of operational research*, *177*(1), 325-343. https://doi.org/10.1016/j.ejor.2005.11.029

[19] Turkcan, A., Akturk, M. S., & Storer, R. H. (2007). Due date and cost-based FMS loading, scheduling, and tool management. *International journal of production research*, *45*(5), 1183–1213. https://doi.org/10.1080/00207540600559955

[20] Restrepo, I. M., & Balakrishnan, S. (2008). Fuzzy-based methodology for multiobjective scheduling in a robot-centered flexible manufacturing cell. *Journal of intelligent manufacturing*, *19*, 421–432. https://doi.org/10.1007/s10845-008-0093-5

[21] Vaisi, B., Farughi, H., & Raissi, S. (2020). Multiobjective optimal model for task scheduling and allocation in a two machines robotic cell considering breakdowns. *WSEAS transactions on information science and applications*, 17, 1-8. https://www.wseas.com/journals/isa/2020/a025103-920.pdf

[22] Yildiz, S., Akturk, M. S., & Karasan, O. E. (2011). Bicriteria robotic cell scheduling with controllable processing times. *International journal of production research*, *49*(2), 569-583. https://doi.org/10.1080/00207540903491799

[23] Feng, J., Che, A., & Wang, N. (2014). Bi-objective cyclic scheduling in a robotic cell with processing time windows and non-Euclidean travel times. *International journal of production research*, *52*(9), 2505-2518. https://doi.org/10.1080/00207543.2013.849015

[24] Ma, K., Yan, P., & Dai, W. (2016). A hybrid discrete differential evolution algorithm for dynamic scheduling in robotic cells. *2016 13th international conference on service systems and service management (ICSSSM)* (pp. 1-6). IEEE.

[25] Abd, K., Abhary, K., & Marian, R. (2016). Multiobjective optimization of dynamic scheduling in robotic flexible assembly cells via fuzzy-based Taguchi approach. *Computers & industrial engineering*, *99*, 250-259. https://doi.org/10.1016/j.cie.2016.07.028

[26] Mansouri, S. A., Aktas, E., & Besikci, U. (2016). Green scheduling of a two-machine flow shop: trade-off between makespan and energy consumption. *European journal of operational research*, *248*(3), 772-788. https://doi.org/10.1016/j.ejor.2015.08.064

[27] Ghadiri Nejad, M., Shavarani, S. M., Vizvári, B., & Barenji, R. V. (2018). Trade-off between process scheduling and production cost in cyclic flexible robotic cells. *The international journal of advanced manufacturing technology*, *96*, 1081-1091. https://doi.org/10.1007/s00170-018-1577-x

[28] Vaisi, B., Farughi, H., & Raissi, S. (2018). Bi-criteria robotic cell scheduling and operation allocation in the presence of break-downs. *International journal of industrial engineering & production research*, *29*(3), 343-357. DOI: 10.22068/ijiepr.29.3.343

[29] Wu, X., Yuan, Q., & Wang, L. (2020). Multiobjective differential evolution algorithm for solving robotic cell scheduling problem with batch-processing machines. *IEEE transactions on automation science and engineering*, *18*(2), 757-775. DOI: 10.1109/TASE.2020.2969469

[30] Vaisi, B., Farughi, H., & Raissi, S. (2020). Schedule-allocate and robust sequencing in three-machine robotic cell under breakdowns. *Mathematical problems in engineering*, *2020*, 1-24. https://doi.org/10.1155/2020/4597827

[31] Sadeghi, H., Makui, A., & Heydari, M. (2016). Multilevel production systems with dependent demand with uncertainty of lead times. *Mathematical problems in engineering*, *2016*. https://doi.org/10.1155/2016/4967341

[32] Sadeghi, H. (2019). A forecasting system by considering product reliability, POQ policy, and periodic demand. *Journal of quality engineering and production optimization*, *4*(2), 133-148. DOI: 10.22070/JQEPO.2020.5087.1123

[33] Shafiei-Monfared, S., Salehi-Gilani, K., & Jenab, K. (2009). Productivity analysis in a robotic cell. *International journal of production research*, *47*(23), 6651-6662. https://doi.org/10.1080/00207540802372298

[34] Geismar, H. N., & Pinedo, M. (2010). Robotic cells with stochastic processing times. *IIE transactions*, *42*(12), 897-914. https://doi.org/10.1080/0740817X.2010.491505

[35] Al-Salem, M., & Kharbeche, M. (2017). Throughput optimization for the robotic cell problem with controllable processing times. *RAIRO-operations research*, *51*(3), 805-818. https://doi.org/10.1051/ro/2016064

[36] Savsar, M., & Aldaihani, M. (2008). Modeling of machine failures in a flexible manufacturing cell with two machines served by a robot. *Reliability engineering & system safety*, *93*(10), 1551-1562. https://doi.org/10.1016/j.ress.2007.06.002

[37] Savsar, M. (2010). Reliability modeling of a manufacturing cell operated under degraded mode. *Proceedings of the international conference on industrial engineering and operations management (IEOM 2010)* (PP. 374-385). IEOM. http://www.iieom.org/paper/211%20Savsar.pdf

[38] Hamasha, M. M., Alazzam, A., Hamasha, S., Aqlan, F., Almeanazel, O., & Khasawneh, M. T. (2015). Multimachine flexible manufacturing cell analysis using a Markov chain-based approach. *IEEE transactions on components, packaging and manufacturing technology*, *5*(3), 439-446. https://www.researchgate.net

[39] Vaisi, B., Farughi, H., & Raissi, S. (2021). Utilization of response surface methodology and goal programming based on simulation in a robotic cell to optimize sequencing. *Journal of quality engineering and management*, *10*(4), 327-338. https://dorl.net/dor/20.1001.1.23221305.1399.10.4.5.7

[40] Vaisi, B., Farughi, H., & Raissi, S. (2018). Two-machine robotic cell sequencing under different uncertainties. *International journal of simulation modelling*, *17*(2), 284-294. https://doi.org/10.2507/IJSIMM17(2)434

[41] Zahrouni, W., & Kamoun, H. (2021). Scheduling in robotic cells with time window constraints. *European journal of industrial engineering*, *15*(2), 206-225. https://doi.org/10.1504/EJIE.2021.114001

[42] Lee, J. H., & Kim, H. J. (2022). Reinforcement learning for robotic flow shop scheduling with processing time variations. *International journal of production research*, *60*(7), 2346-2368. https://doi.org/10.1080/00207543.2021.1887533

[43] Selim, M., & Gurel, A. S. (2007). Machining conditions-based preventive maintenance. *International journal of production research*, *45*(8), 1725-1743. http://dx.doi.org/10.1080/00207540600703587

[44] Vaisi, B. (2022). A review of optimization models and applications in robotic manufacturing systems: industry 4.0 and beyond. *Decision analytics journal*, 2, 100031. https://doi.org/10.1016/j.dajour.2022.100031

[45] Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *International conference on parallel problem solving from nature* (pp. 849-858). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45356-3_83

[46] Coello, C. A. C., Lamont, G. B., & Van Veldhuizen, D. A. (2007). *Evolutionary algorithms for solving multi-objective problems* (Vol. 5, pp. 79-104). New York: Springer. https://doi.org/10.1007/978-0-387-36797-2

[47] Kiyani Ghalehno, R., Niroomand, S., Didekhani, H., & Mahmoodirad, A. (2022). A multi-objective formulation for portfolio optimization of credit institutions branches: case study of Keshavarzi bank of Sistan and Baloochestan. *Journal of decisions and operations research*, *7*(2), 299-315. (**In Persian**). DOI: 10.22105/dmor.2021.257591.1260

[48] Sadeghi, H., & Mahmoodi, A. (2022). Multi-objective inventory model for material requirements planning with uncertain lead-time. *International journal of operational research*, *43*(4), 391-415. https://doi.org/10.1504/IJOR.2022.122811

[49] Shoaee, M., & Samouei, P. (2021). A cross-dock warehouse layout design using multi-objective gray wolf optimization algorithm. *Journal of decisions and operations research*, *7*(3). 453-465. (**In Persian**). DOI: 10.22105/dmor.2021.267139.1302

Publisher: Research Expansion Alliance (REA) on behalf of Ayandegan Institute of Higher Education, Iran.
Director- in-Charge: Seyyed Esmaeil Najafi
Editor-in-Chief: Nachiappan Subramanian
Sientific & Executive Manager: Seyyed Ahmad Edalatpanah

Journal of Applied Research on Industrial Engineering is an international scholarly open access, peer-reviewed, interdisciplinary and fully refereed journal. The mission of this journal is to provide a forum for industrial engineering educators, researchers, and practitioners to advance the practice and understanding of applied and theoretical aspects of industrial engineering and related areas.