RAIRO-Oper. Res. 56 (2022) 2775–2800
https://doi.org/10.1051/ro/2022095

RAIRO Operations Research
www.rairo-ro.org

A COMPREHENSIVE MODEL FOR DETERMINING TECHNOLOGICAL INNOVATION LEVEL IN SUPPLY CHAINS USING GREEN INVESTMENT, ECO-FRIENDLY DESIGN AND CUSTOMER COLLABORATIONS FACTORS

Razieh Beigizadeh[1], Aidin Delgoshaei[2,*], Mohd Khairol Anuar Ariffin[2], Sepehr Esmaeili Hanjani[3] and Ahad Ali[4]

Abstract. Technological innovations play a crucial role in designing an effective green supply chain. However, it is crucial to know the factors influencing technological innovation in a green supply chain. Some preconceptions show that technological innovation in a business can be affected by internal and external factors, and therefore there must be correlations between such factors to flourish the technological innovation and, subsequently, the green supply chain. Besides, predicting the technological innovation level in a supply chain can be vital and direct it to the Industry 5.0 goals. In this research, a 3-phased framework will be proposed to predict the Technological Innovation Level of Green Supply Chains. The scope of this research includes Green Investment, Eco-friendly Design and Customer Collaborations. In the 1st phase of the framework, dependent and independent factors considering the scope of the Research will be determined; and then, using statistical data analysis, the weight of factors, which reflects their impact on technological innovation (dependent factor), will be determined. Then, in the 2nd phase, a comprehensive model will be developed and trained. Using the data of supply chains that were gathered in the first phase, the train and test data would be selected. In continuation, the model will be trained and its performance will be evaluated using some metrics. Then, in the last phase (phase 3), the developed model will be used to predict the technological level of supply chains. The outcomes of this research can help top managers of supply chains to predict the level of technological innovation by investing a certain budget in improving the dependent variables. The outcomes demonstrated that Customer Collaboration (0.481), Eco-friendly design (0.419) and Green Investment (0.41) have significant impacts on technological innovation improvement in the studied cases, respectively. Besides, the results showed the superiority of the K-nearest Neighbor algorithm while using the Minkowski distance method and considering 5 neighbors. The findings indicated that the proposed framework could predict Technological Innovation with 0.751 accuracies. The outcomes of this research can be helpful for industry owners to predict the expected technological innovation level of their system by investing a certain budget in green investment, eco-friendly design and customer collaborations in their enterprises.

Mathematics Subject Classification. 90–10.

Received July 15, 2021. Accepted June 11, 2022.

Keywords. Green supply chain, technological innovation, design algorithm, green investment, eco-friendly design, customer collaborations.

[1] Department of Management, Islamic Azad University, Rasht Branch, Rasht, 41335 Guilan, Iran.

[2] Department of Mechanical and Manufacturing Engineering, Universiti Putra Malaysia, Serdang, 43400 Selangor, Malaysia.

[3] Department of Management, Islamic Azad University, North Tehran Branch, Tehran, 14515 Tehran, Iran.

[4] Department of Industrial Engineering, Lawrence Technological University, Southfield, MI 48075, USA.

*Corresponding author: delgoshaei.aidin@upm.edu.my

1. Introduction

Since the Industrial Revolution, manufacturing operations have wreaked havoc on the environment, disregarding nature, creating industrial waste, altering ecosystems, and depleting natural resources. Although there are hints that firms are becoming more environmentally conscious; nonetheless, some are still significant problems that happen due to such massive ignorance [15, 22].

During the last two decades, polemic legislation set to deter many companies in developed countries from harming the environment, there are still many barriers to overcoming such significant shortcomings in less developed industrial environments. Perhaps the reason behind such an unsuccessful result is twofold. Firstly, the existence of industry-specific barriers (*i.e.*, related to some of the broader joint external forces in the industry) and secondly, emerging organizational barriers (*i.e.*, related to internal barriers experienced in an organization) [15]. Perhaps, one solution for preventing such problems is green supply chain management. Green supply chain management that the Michigan State University first introduced Industrial Research Association in 1996 aims to protect the environment in supply chains throughout all stages of the life cycle of a product, from supplier selection to product use and recycling.

The drivers of an organization's move towards a green supply chain vary from suppliers, raw materials, manufacturing process, storage and transportation, and recycling. They strictly depend on the type of products and services provided by a supply chain. However, external factors such as laws and regulations and competitors influence the managerial decisions in a green supply chain [39].

While internal drivers come into consideration, the recent studies showed that a strong correlation existed between green supply chain management success and the level of applied technology innovation (T.I) [30]. Technological innovation causes more reliable production processing from the early stages of a supply chain, yielding less environmental damage [17]. Consequently, T.I results in increasing sustainability features. Recent studies show that technological innovation flourishing in a supply chain depends on many factors, including top management commitments, eco-friendly design, investment recovery, Green Purchasing, and customer collaboration [21].

In recent years, one of the major issues that manufacturing organizations usually face is how to improve their sustainability level to prevent legal fines by governments. T.I can be a solution. It can boost material usage, reduce wastage and improve the system's performance [32]. The synergies between environmental considerations and supply chain management allow supply chains to improve their environmental productivity, quality, and performance through the continuous flow of information. On the other hand, Yi and Xiao-li [40] showed that T.I improvement can lead to more sustainable developments.

Therefore, finding a way to predict the level of required T.I in a supply chain is vital. However, achieving a certain level of T.I requires a comprehensive understanding of various supply chain features to be considered a good strategy to switch from old technology and get relief from related troubles.

Therefore, this research tries to answer two crucial questions in green supply chains: "what are the effective drivers of a successful green supply chains management?" and "how much and where should top managers of supply chains should invest in improving their technological innovation level?" The present research tries to find the answer to the aforementioned questions by focusing on internal drivers in green supply chains.

Although the outcomes of this research can be generalized to manufacturing systems; however, the scope of this research is green supply chains. Besides, this study only focuses on the internal drivers of T.I in green supply chains, which can be considered the limitation of this research.

1.1. Motivations and novelties of the research

The motivation and novelties of this research are to propose a new design model to predict the technological innovation level in supply chains by considering top management commitments, eco-friendly design, and customer collaborations. This approach motivates top managers to determine the level of investments in T.I drive to achieve a certain level of technological innovation. For this purpose, a number of opted research studies will be investigated in continue.

2. Literature review

Flourishing in technological innovation depends on many internal and external factors. The level of innovation that can be expected from a business depends on potential capabilities in various supply chain areas. Lam *et al.* [19] discussed the existence of dynamic and multistage correlations between innovation and organizational factors. Consequently, by upgrading the technologies, recognizing and boosting the level of elements of a green supply chain will create a value chain that satisfies both customers' needs and environmental protection requirements [10]. Macchion *et al.* [24] and Macchion *et al.* [25] analyze the adoption of environmental sustainability practices and collaboration along the supply chain to achieve better innovation performance. Besides, technological innovation is a consequence of personal innovation. Delgoshaei *et al.* [8] discussed that personal and impersonal factors can play key roles in an expert's creativity level in a supply chain.

Scientists classified organizational innovation based on process, product, and technology level [29]. While improving technological innovation comes into consideration, *Green investments* supported strongly by top management play a vital role in providing a better environment to support the organization's future [3]. Many studies have shown that senior management support has played an important role in accepting and disseminating innovations in the organization [36]. On the other hand, the green investment will cause a flourishing economic status by using less energy and more productivity [12]. Zhang *et al.*, [41] showed that management innovation and T.I positively affect a system's performance.

Delgoshaei *et al.* [9] propose a fuzzy-based algorithm for determining a supply chain's rivals' strategy in uncertain market demand conditions. They showed that using strategic planning can help take correct action against rivals. Lee *et al.* [21] stated that setting up total quality management will cause increasing innovation speed, resulting in better market share. Senior management's role and other functional cooperation throughout a supply chain were proved before in novel field studies [20, 43]. Top management decisions can be considered an initial step in green supply chain management by taking accountable and efficient technological strategies that are in line with the green policy of a supply chain [2].

Eco-friendly design is another valuable and emerging tool for a green supply chain that addresses a product's capabilities while reducing the adverse environmental impacts on the universe [4, 23, 38]. To prevent these problems in the design, supply, production, and distribution of products, the environment and its protection should be considered [16,20]. Lee *et al.* [20] stated that using up-to-date technologies in manufacturing products that are less harmful to the environment at the same time makes them more competitive. However, in its initial step, Eco-friendly design mainly focuses on the technical upgrading of products and processes to reduce environmental costs [43]. Thus, Eco-friendly design can impact the level of technology that should be applied in a supply chain [20] by affecting material selection, processing technology, required machinery, transportation, and logistics, required human resources, material reuse, and recycling [6,7,27]. On the other hand, the environmental regulations and rules that prevent manufacturers from using harmful technologies are another reason for paying attention to the Eco-friendly design. For instance, since most electrical products sold in Europe are made in China, Chinese companies are responsible for following European rules, such as avoiding the use of certain prohibited substances [43]. Eltayeb and Zailani [11] A study in Malaysia examined the correlation between green supply chain schemes and performance outcomes (*i.e.*, environmental, economic, operational, and intangible) among 569 companies that were ISO14001 certified.

Customers are unignorably part of a chain value. The customers' needs will affect the design and functions of the products of a supply chain and all activities that will be done throughout the chain network. Companies are increasingly focusing on the needs of their customers and developing value-added products and services that can meet those needs. *Collaborating with customers* on environmental issues is defined as solutions in the supply chain cycle [42]. By working closely with customers for environmentally friendly design, cleaner production, and green packaging, companies can release new products while demonstrating environmentally-friendly innovation Chong and Ooi [5]. There is a wide range of characteristics and different behavioral aspects of green consumers, a comprehensive and innovative combination of numerous studies, and the effect of religious commitment (in terms of value) on green shopping behavior [32]. The terms Collaboration with customers in this study matches with the research of Zhu and Sarkis [42], Zhu *et al.* [43], and Lee *et al.* [20], where components such as customer collaboration for environmentally friendly design in a company, cooperation with the customer for cleaner production in the company, cooperation with the customer for green packaging in the company are taken into consideration. Therefore, stronger collaborations with customers to identify their needs and design, manufacture and distribute the products accordingly will affect the technology and innovation that a supply chain must use. As a result, customer collaboration will be considered another important factor influencing the technological innovation level in a green supply chain.

Using Green Materials for satisfying the environmental requirements in a supply chain is unavoidable [44]. On the other hand, most modern technologies require materials of higher quality, which leads to significant improvements in the environmental performance of companies while maintaining quality and cost objectives. Using green materials along with modern technology causes less wastage. Kapetanopoulou and Tagaras [18] showed that developing countries usually do not prioritize investment recovery systems because their waste management policies and crop recovery activities are not well established. Lee *et al.* [20] investigated the role of investment recovery on technological innovation.

Besides, Supervised algorithms showed excellent performance in forecasting and identifying level for various purposes in manufacturing systems and supply chains. There are novel references that show the successful application of supervised algorithms in risk assessment [1]; product distribution [13]; sustainable entrepreneurship [26, 28]; time forecasting [37] and bullwhip effect [33] in supply chains. Therefore, it can be ideal for predicting the technological innovation level of a green supply chain. Delgoshaei *et al.* [9] propose a fuzzy-based algorithm for determining a supply chain's rival's strategy in uncertain market demand conditions. They showed that using strategic planning can help take correct action against rivals.

An in-depth review of the literature shows that although invaluable research activities have been carried out on green investment, customer collaboration, and eco-friendly design, these variables' impact on the expected technological innovation level has not been addressed yet. Therefore, this research aims to develop a design-based model to find the impact of the mentioned factors on the technological level of a green supply chain.

The main purposes of this research are twofold. The first is to find the impacts of Customer Collaboration, Green Investment, and Eco-friendly Design on Technological Innovation Improvement in Green Supply Chains. The second purpose is to propose a model to predict the technological innovation level of a supply chain by investing a certain amount of budget in Customer Collaboration, Green Investment, and Eco-friendly Design. Such a framework can help stakeholders invest in the supply chain with more confidence while they can see the results of their investment in the technological innovation level of the supply chain.

The outcomes of this research can be helpful for industry owners to predict the expected technological innovation level of their system by investing a certain budget in Green Investment, Eco-friendly Design, and Customer Collaborations in their enterprises.

3. Research methodology

In this section, a 3-phase framework will be proposed for predicting the technological innovation level of supply chains. In the 1st phase of the framework, dependent and independent factors considering the research scope will be determined. Then, using statistical data analysis, the weight of factors, which reflects their impact on technological innovation (dependent factor), will be determined. Then, in the 2nd phase, a new model will be developed and trained. Using the data of supply chains gathered in the first phase, the train and test data would be selected. In continuation, the model will be trained, and its performance will be evaluated using some metrics. The performance of the used new algorithm will be evaluated by comparing it with other algorithms. Then, in the last phase (phase 3), the developed model will be used to predict the technological level of supply chains. Using the outcomes of the developed model can help top managers of supply chains to predict the level of technological innovation by investing a certain budget in improving the dependent variables. The flowchart of the proposed framework is presented in Figure 1.

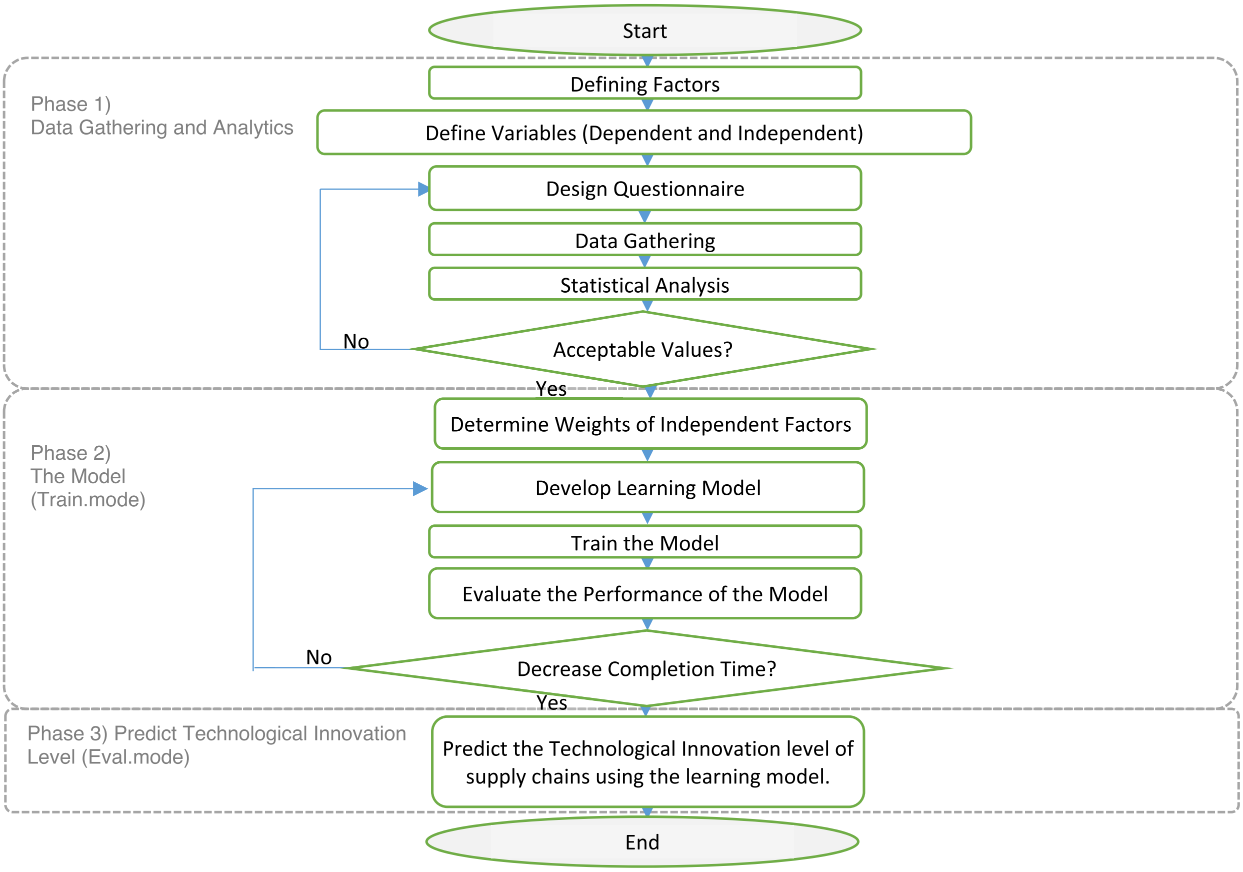

FIGURE 1. Flowchart of the proposed framework.

3.1. Novelties of the proposed method compared to the recent relevant references

Using the supervised methods for predicting T.I level in a green supply chain can be considered the main novelty of the Research. With the outcomes of the proposed method, the top managers can determine the best amount of investing in green materials, designing eco-friendly products, and expanding the correlations with their customers if they want to achieve a certain level of technological innovation.

Table 1 shows the contribution of the current research compared to the relevant references in the literature.

4. DATA GATHERING AND ANALYTICS (PHASE 1)

The first phase of the proposed framework is data gathering and statistical analysis. This section is designed to analyze the collected data. The first activity is to determine the required number of samples for statistical Analysis (Sect. 4.3). Since the domain of this study is in north Malaysia, the supply chains in this region are considered. Then, in Section 4.5, the validity of the questionnaire must be examined using Cronbach's Alpha to see if the questions are valid enough to be used for Statistical Analysis. In the next step, descriptive analyses will be done for each variable to get general information about them, such as the mean, standard deviation, and the minimum and maximum values they can get. This section can help determine any abnormal status in variables (such as considerable standard deviations). In the next step, the normality of each variable will be examined using Kolmogorov–Smirnov Normal Test (Sect. 4.2). In the next step, the statistical hypotheses of the

TABLE 1. Comparing the contribution of this research with the most relevant opted references.

Ref.	Scope	Contribution	Case study	Main findings
Yi and Xiao-li [40]	Technological Innovation	Sustainable Development	Regression Analysis	Regional Technological Innovation Improving can lead to sustainable developments
Zhang *et al.* [41]	Technological Innovation	Organization Performance Improvement	Empirical evidence collected from 304 Asian CEOs and top managers	The impacts on system performance by using a certain level of management and technical innovation
Zhu *et al.* [43]	Green Investment	Measuring the green supply chain implementation	341 Chinese manufacturers	A validated measurement scale can help evaluate the strengths and weaknesses in different facets of implementing GSCM practices in supply chains
Luo *et al.* [23]	Eco-friendly Design and Customer Collaboration	Role of different unit's cooperation in the success of Eco-friendly design in green supply chains	222 Chinese manufacturing organizations	Buyer-seller relationship and competitive environment influence green supply chain collaboration
Zhu and Sarkis [42]	Drivers of International wide GSC	Collaborating with customers on environmental issues is a crucial step in setting up a successful supply chain cycle	118 case studies from Chinese GSC	They found Chinese Manufacturing systems have a different environment that influences the GSC drivers
The current research	Technological Innovation	Prediction of Expected Technological Innovation level of a green supply chain by focusing on green investment, eco-friendly design and customer collaboration	254 Malaysian Supply Chains	1. Customer Collaboration, Green Purchasing and Eco-friendly Design have a significant impact on Technological Innovation 2. The proposed new model can successfully predict the technological innovation level of green supply chains by investing a certain budget on green investment, eco-friendly design and customer collaboration elements

research will be defined and examined using the ANOVA test (Sect. 4.3). Then, while the correlations between technological innovation and independent variables of the research are proved, the weights of each variable will be determined using a regression equation (Sect. 4.4).

4.1. Dependent and independent variables

When a body of information is gathered for research, it is first necessary to organize and summarize it in a meaningfully understandable and relevant way. Descriptive statistical methods are used for this purpose. A summary of the characteristics of descriptive statistics related to the variables used in this research is given in Tables 4–6. The reported statistics include central indices and criteria, including mean, minimum, maximum, and scatter indices of the standard deviation of the variables used in this study, extracted using SPSS 18 software.

4.2. Statistical population

A statistical population is a set of individuals or units with at least one common attribute. Usually, in any research, the study population is a statistical population that the researcher wants to study about the attribute(s) of the variables of its units. The statistical population of this research is medium and small supply chains, as mentioned in the next section (Sect. 4.3).

4.3. Sample size

In order to determine the statistical population size for the proposed method, the Cochran formula is used. The Cochran formula was firstly used in 1963 for determining sample size [34]. For this purpose, the medium- and large-scale supply chains in industrial zones in different North Malaysia states have been considered (the year 2020). According to the statistical values available in the county, there are 752 medium and large-scale factories in the food and wooden sectors that are available and active that can be considered. Since the statistical population is known, Cochran's formula for limited communities was used to determine the sample size, which is as follows.

$$n = \frac{N \times z^2_{\frac{\alpha}{2}} \times \sigma^2}{\varepsilon^2(N-1) + Z^2_{\frac{\alpha}{2}} \times \sigma^2} \tag{4.1}$$

$$n = \frac{752 \times (1.96)^2 \times (0.249)}{0.05^2(752-1) + (1.96)^2 \times (0.249)} = 254 \tag{4.2}$$

where σ^2 is the variance of the sample size; N is the size of the total experts in construction companies (we assume 1000 experts in K.L); n is the size of the statistical society; ε is the tolerable error (0.05) and $z_{\frac{\alpha}{2}}$ is the standard function value (1.96). According to Cochran's formula, the statistical sample size of 254 companies was obtained where these values can reflect the reality of data with 95% reliability based on statistical formulas as presented in Section 4.

4.4. Data gathering method

In this research, the following methods have been used to collect data:

– Scientific books and dissertations.
– Use of internal and external articles.
– Using a questionnaire.

The method of collecting research data is the field study, and the primary tool used in this research is a questionnaire. A questionnaire survey collects data from individuals using a formal questionnaire or interview guide (Tab. 2).

The conceptual research model data is collected through a field survey of people in managerial positions familiar with environmental issues and operations in manufacturing companies in an industrial area. In the questionnaire that companies have completed, the items that measure the dimensions of the conceptual research model are modeled as follows from previous Research.

Three dimensions of green supply chain management methods (Green Investment and Green Customer Collaborations) were used in this research after a detailed review of the research of Zhu *et al.* [43] and Lee *et al.* [20] used a total of 9 items and were measured against three green supply chain management methods. Nine items were also used to evaluate corporate technology innovation. These items were adapted from the research of Prajogo and Sohal [31], Singh and Smith [35], and Chiou *et al.* [4].

All items are measured by answers on a 5-point Likert scale with statements including strongly disagree = 1 to agree = 5 strongly. This study uses multiple regression to investigate the factors affecting technological innovation in green supply chain management in this field. The research questionnaire to examine the factors affecting green supply chain management (case study: manufacturing companies in an industrial area) is shown in Table 2.

4.5. Reliability and validity

Research questionnaires are considered standard questionnaires in terms of repetition in multiple types of research and have the necessary validity. Reliability is concerned with how measuring instruments produce the same results under the same conditions. Alpha less than 6.0 usually indicates poor reliability; 6.0 to 8.0 is acceptable, and higher than 8.0 indicates high reliability. Obviously, the closer this number is to one, the better. In this study, to measure the reliability of the questionnaires, the formula for calculating Cronbach's alpha is as follows.

$$r_a = \frac{J}{J-1}\left[1 - \frac{\sum S_i^2}{S^2}\right] \tag{4.3}$$

J: Number of questionnaire subset questions.

TABLE 2. Categorizing the common problems in line with the defined variables.

Independent variables			Dependent variable
Green Investment	Eco-friendly design	Green Customer Collaborations	Technological Innovation
Improve investment (sales) or inventory/additional materials in the company is appropriate. Waste and materials used in the company are sold. Additional capital equipment is sold in the company.	There is a product design to reduce material/energy consumption in the company. Product design is suitable for the reuse, recycling, and recycling of materials, parts, and components. Product design is appropriate to prevent or reduce hazardous products and/or their production process in the company.	There is cooperation with the customer for environmentally friendly design in the company. Work with the customer to produce cleaner in the company. There is cooperation with the customer for green packaging in the company.	The company can produce products with novelty features. The company uses the latest technology to produce a new products. The speed of production of a new product is fast enough/competitive enough. The company has enough new products to introduce to the market. The company has new products that have a leading position in the market. The company is technologically competitive. The company uses up-to-date/new technologies in the process. The company quickly adopts a process with the latest technological innovations. Processes, techniques, and technology are changing rapidly in our company.

TABLE 3. Reliability results using Cronbach's test.

Variable	Related questions in the questionnaire	Cronbach's alpha
Green Investment in supporting green supply chain	1–2–3	0.885
Environmental design compatibility	4–5–6	0.900
Green Customer Collaborations	7–8–9	0.896

S_i^2: The variance of the scores of each subset.

S_i: Total variance.

Cronbach's alpha coefficient indicates this questionnaire's sufficient reliability (Tab. 3). The reliability coefficient of the questions related to the research variables separately (alpha coefficient of Green Investment is 0.885, alpha coefficient of Eco-friendly Design 0.900, alpha coefficient of Green Customer Collaborations 0.896, and alpha coefficient of Technology Innovation 0.964 indicate the reliability of the questions of these elements.

TABLE 4. Description of the component of variables of the research.

Variable	Quantity	Minimum	Maximum	Mean	Std. var.
Green Investments	254	1.33	5	3.36	0.576
Environmental design compatibility	254	2	5	3.43	0.548
Green Customer Collaborations	254	1	5	3.15	0.652

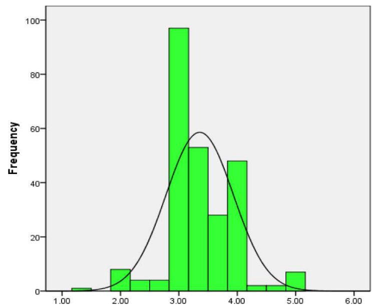

FIGURE 2. Histogram of the "Green Investments" variable.

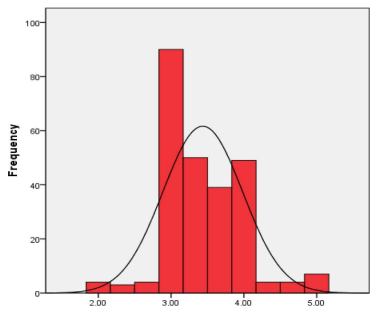

FIGURE 3. Histogram of the "Environmental design compatibility" variable.

4.6. Descriptive analysis of the variables

This section analyzes the descriptive Analysis to learn more about the variables and their status in the statistical society (Tab. 4 and Figs. 2–4).

4.7. Kolmogorov–Smirnov test (K-S)

This test is a distribution compatibility test for quantitative data. Suppose a researcher has a sample of metrics and wants to determine if it is a sample of a normally distributed society. The normality test of distribution is one of the most common tests for small samples that the researcher doubts are normal. For this purpose, the K-S test is a practical test. In SPSS software, this test matches four different distributions of normal, Poisson, exponential and uniform. This method is based on the difference between the relative cumulative frequency of observations and the expected value under the assumption of zero. Hypothesis zero states that the selected sample has a normal distribution (Poisson, exponential, or uniform). Kolmogorov–Smirnov test to match the distribution, compare the cumulative probabilities of the values in your data set with the cumulative probabilities of the same values in a particular theoretical distribution. If the difference is significant enough, this test will

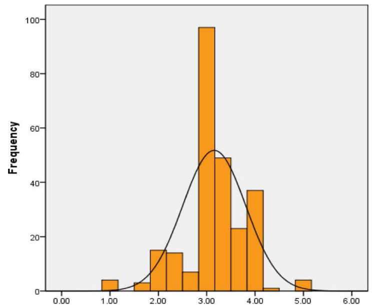

FIGURE 4. Histogram of the "Green Customer Collaborations" variable.

TABLE 5. The outcomes of the Kolmogorov–Smirnov test calculations for Green Purchasing.

One-sample Kolmogorov–Smirnov test					
		Green Investments	Eco-friendly Design	Green Customer Collaborations	Technological Innovation
N		30	30	30	30
Normal parameters^{a,b}	Mean	3.5889	3.6000	3.3667	3.2815
	Std. Deviation	0.58515	0.61526	0.57635	0.63577
Most extreme differences	Absolute	0.176	0.169	0.238	0.196
	Positive	0.176	0.169	0.238	0.171
	Negative	−0.159	−0.142	−0.196	−0.196
Kolmogorov–Smirnov Z		0.965	0.923	1.302	1.072
Asymp. Sig. (2-tailed)		0.309	0.361	0.067	0.201

Notes. [(a)]Test distribution is Normal. [(b)]Calculated from data.

show that the data does not match one of the theoretical distributions. In this test, if the decision criterion (p-value) is less than 5%, the null hypothesis is rejected, *i.e.*, the data cannot be from a specific distribution such as normal, Poisson, exponential, or uniform.

The null and void hypotheses in the Kolmogorov–Smirnov tests are as follows:

$$\begin{cases} H0: \text{The data for the variable follows the normal distribution} \\ H1: \text{The data for the variable does not follow the normal distribution.} \end{cases}$$

The results of the test of normality of research variables are as described in Table 5:

As shown in Table 5, it can be seen that the level of significance obtained is more than 0.05, so all research variables in the sample have a normal distribution.

The Cronbach's alpha results for the variables show that the questions are reliable enough to be used for statistical society. Then, the standard deviation values presented in Table 6 show that most responders have close ideas about questions. Besides, no outlier is found after data gathering, which represents the user data are real bile at the level of confidence of 95% ($1 - \alpha = 0.95$).

4.8. Statistical hypotheses

The hypothesis for the first variable is now defined as follows:

$$\begin{cases} H0: \text{Green Investments have a significant impact on Technological Innovations} \\ H1: \text{The Opposite side is true.} \end{cases}$$

Table 6. The impact of Green Investments on Technological Innovation.

Variable	Quantity	Correlation coefficient	Coefficient of determination	Significant level	Result
Green Investments	254	0.376	0.141	0.000	Approved

Table 7. Results of ANOVA test (Green Investments *versus* Technological Innovations).

ANOVA[b]						
Model		Sum of squares	*df*	Mean square	*F*	Sig.
1	Regression	14.165	1	14.165	41.506	0.000[a]
	Residual	86.001	252	0.341		
	Total	100.166	253			

Notes. [(a)]Predictors: (Constant), Green Investments. [(b)]Dependent Variable: Technological Innovation.

Table 8. Achieved coefficients for calculating regression formula (Green Investments *versus* Technological Innovations)

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.679	0.217		7.738	0.000
	Green Investments	0.410	0.064	0.376	6.442	0.000

Notes. [(a)]Dependent Variable: Technological Innovation.

As shown in Table 6, it can be seen that the level of significance obtained is less than 0.05, so this correlation is significant. The intensity of the correlation between the two variables of Green Investments and Technology Innovation is 0.376, which indicates a direct correlation between the two variables. Also, according to the beta weight, the regression equation is as follows (Tabs. 7 and 8):

$$E(Y) = 1.679 + 0.410X1. \tag{4.4}$$

On the other hand, the coefficient of determination between the two variables equals 0.141, which shows that the variable of investment improvement by 14.1% can predict technological innovation. Therefore, improving investment with technological innovation has a positive and significant effect. According to testing the research hypotheses with SPSS 18 software, the correlation coefficient and the probability obtained between the two variables of investment improvement with technological innovation are equal to 0.376 and 0.000, respectively. The results show that, at a 95% confidence level, there is a positive relationship between investment improvement and technological innovation. Furthermore, their relationship is significant according to the obtained significance levels ($p = 0.0000$, Sig. < 0.05).

On the other hand, the coefficient of determination between the two variables equals 0.141, which shows that the variable of investment improvement by 14.1% can predict technological innovation. Therefore, improving investment with technological innovation has a positive and significant effect, so that with increasing investment improvement, technological innovation also increases in the company. The results suggest that the firm's additional resources can be converted into revenue if the unused assets are resold, reducing the need to store and generate additional revenue. The study Kapetanopoulou and Tagaras [18] showed that developing countries

TABLE 9. The impact of Eco-friendly design on Technological Innovation.

Variable	Quantity	Correlation coefficient	Coefficient of determination	Significant level	Result
Eco-friendly design component	254	0.392	0.154	0.000	Approved

TABLE 10. Results of ANOVA test (Eco-friendly design *versus* Technological Innovations).

ANOVA[b]						
Model		Sum of squares	*df*	Mean square	*F*	Sig.
1	Regression	15.396	1	15.396	45.770	0.000[a]
	Residual	84.770	252	0.336		
	Total	100.166	253			

Notes. [(a)]Predictors: (Constant), Eco-friendly design. [(b)]Dependent Variable: Technological Innovations.

TABLE 11. Achieved coefficients for calculating regression formula (Eco-friendly design *versus* Technological Innovations).

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.621	0.215		7.532	0.000
	Eco-friendly design	0.419	0.062	0.392	6.765	0.000

Notes. [(a)]Dependent Variable: Technological Innovations.

do not prioritize investment recovery systems. As waste management policies and their recovery activities are not well established, companies are hesitant to leave on this trip. However, our results show that companies in Malaysia and other developing countries need to change their mindset and implement investment improvement strategies in their supply chain.

The second variable can be examined using the following hypothesis:

$$\begin{cases} H0: \text{Eco-friendly design has a positive and significant effect on technological innovation} \\ H1: \text{The Opposite side is true.} \end{cases}$$

As shown by Table 9, it can be seen that the level of significance obtained is less than 0.05, so this correlation is significant. The intensity of the correlation between the two Eco-friendly designs and technology innovation is 0.392, which indicates a direct correlation between the two variables. Also, according to the beta weight, the regression equation is as follows (Tabs. 10 and 11):

$$E(Y) = 1.621 + 0.419X1. \tag{4.5}$$

The coefficient of determination between the two variables is equal to 0.154, which shows that the Eco-friendly design variable of 15.4% can predict the variable of technological innovation. Therefore, Eco-friendly design is positively and significantly related to technological innovation.

Hypothesis 2: Eco-friendly design with technological innovation has a positive and significant effect.

TABLE 12. The impact of Green Customer Collaborations on Technological Innovation.

Variable	Quantity	Correlation coefficient	Coefficient of determination	Significant level	Result
Green Customer Collaborations	254	0.672	0.451	0.000	Approved

According to testing the research hypotheses with SPSS 18 software, the correlation coefficient and the probability obtained between the two Eco-friendly design variables with technological innovation are equal to 392.0 and 000.0, respectively. The results show that reliability level has a 95% positive correlation between Eco-friendly design and technological innovation. Furthermore, the correlation between them is significant ($p = 0.0000, \mathrm{Sig.} < 0.05$). On the other hand, the coefficient of determination between the two variables is equal to 154.0, which shows that the Eco-friendly design variable of 4.15% can predict the variable of technological innovation. Therefore, Eco-friendly design has a positive and significant impact on technological innovation, so with the increase of Eco-friendly design, technological innovation also increases in the company. The results indicate that by considering Eco-friendly design as a critical factor in green supply chain management, a company's environmental performance can be improved and reduce the environmental effects of the life cycle. Our study shows that products designed with environmental concerns are the primary target. For example, reducing materials or energy, reusing/recycling, avoiding hazardous chemicals that are significantly more likely to be innovative, with novelty features that come first in the market.

Therefore, the second hypothesis of the research that Eco-friendly design has a positive and significant effect on technological innovation is accepted. This hypothesis is consistent with the research of Zhu *et al.* [43] and Lee *et al.* [20].

Finally, the last impacts of the variable can be defined as below:

$$\begin{cases} H0: \text{Green Customer Collaborations have a positive and significant effect on technological innovation} \\ H1: \text{The Opposite side is true.} \end{cases}$$

As shown in Table 12, it can be seen that the level of significance obtained is less than 0.05, so this correlation is significant. The intensity of the correlation between the two variables of Green Customers Relations and Technology Innovation is 0.672, which indicates a direct correlation between the two variables. Also, according to the beta weight, the regression equation is as follows (Tabs. 13 and 14):

$$E(Y) = 1.425 + 0.481X1. \tag{4.6}$$

On the other hand, the coefficient of determination between the two variables equals 0.196, which shows that the Green Customer Collaborations variable of 45.1% can predict the technology innovation variable. Therefore, buying green with technological innovation has a positive and significant effect. According to testing the research hypotheses with SPSS 18 software, the correlation coefficient and the probability obtained between the two variables of green purchasing with technological innovation are equal to 0.672 and 0.000, respectively. The results show a 95% confidence that there is a positive relationship between customer collaboration and technological innovation. Moreover, their relationship is significant according to the obtained significance levels ($p = 0.0000, \mathrm{Sig.} < 0.05$).

On the other hand, the coefficient of determination between the two variables is equal to 0.451, which shows that the variable of cooperation with customers by 45.1% can predict technological innovation. Therefore, cooperation with customers with technological innovation has a positive and significant effect, so with increasing cooperation with customers, technological innovation also increases in the company. The results suggest that joint participation in environmental issues may improve the performance of companies. In other words, consumers prefer green products.

Table 13. Results of ANOVA test (Green Customer Collaborations *versus* Technological Innovations).

ANOVA[b]						
Model		Sum of squares	*df*	Mean square	*F*	Sig.
1	Regression	45.195	1	45.195	207.184	0.000[a]
	Residual	54.971	252	0.218		
	Total	100.166	253			

Notes. [(a)]Predictors: (Constant), Green Investments. [(b)]Dependent Variable: Technological Innovation.

Table 14. Achieved coefficients for calculating regression formula (Green Customer Collaborations *versus* Technological Innovations).

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.013	0.145		6.987	0.000
	Green Customer Collaborations	0.648	0.045	0.672	14.394	0.000

Notes. [(a)]Dependent Variable: Technological Innovation.

4.9. Result analysis

This section collected data through a field survey of manufacturing companies in an industrial area. Therefore, the statistical population of this study is medium and small manufacturing companies in an industrial area, the number of which is 750 companies. In this study, the available non-probability sampling method was used, and according to Cochran's formula, the statistical sample size was 254. In the questionnaire completed by the companies, the items that measure the dimensions of the conceptual research model are modeled as follows from previous research. Three dimensions of green supply chain management methods (Green Investments, Eco-friendly Design, and Green Collaborations with Customers) were used in this research after a detailed review of research by Zhu *et al.* [44] and Zhu *et al.* [43].

4.9.1. Outcomes of hypotheses

According to the obtained coefficients and significance level, the following results were obtained:

- The first hypothesis of the research is that the commitment of senior management to Top Management Commitments to technological innovation has a positive and significant effect.
- The second research hypothesis is accepted that Eco-friendly design has a positive and significant impact on technological innovation.
- The third research hypothesis is accepted that green shopping has a positive and significant effect on technological innovation.

5. The proposed model (phase 2)

After determining the effective factors on the technological innovation level of a supply chain, in this phase (phase 2), a new design-based model will be developed for predicting the technological innovation of the green supply chains using Python. Python is among the most powerful applications that provide a comprehensive and robust basis for design algorithms.

FIGURE 5. The block diagram of the proposed algorithm.

5.1. Libraries of Python

For this purpose, the following Libraries will be used:

- Numpy: for support for large-scale data and multidimensional arrays and matrices.
- Pandas: for using data sets and data frame.
- Scipy: for modules for optimization, linear algebra, integration, interpolation, special functions.
- Matplotlib: for drawing graphical views of the outcomes of the model.
- Sklearn: for using ML methods.

5.2. Jupyter as the platform of Python

Jupyter is one of the most powerful platforms for coding algorithms with Python. It is free and open-source and has been widely used throughout the past years. One dominant feature of Jupyter is testing the outcome of each script line exactly below it during the coding process. It helps algorithm developers to examine each line before going further.

5.3. Block diagram of the proposed algorithm

Figure 5 will show the process of developing the proposed algorithm.

5.4. Data in use

This section uses a CSV file containing 421 data, where 254 data from Section 4.3 is used, and the rest are generated by Python for training and evaluating the model. The dataset will be imported into Python (Fig. 6).

In [2]:
```
TIData=pd.read_csv('D://1. Papers//1- Manuscripts//3- Supply Chain Management//2- Under Review//7- Aidin//New//2//KNN//Data.csv',
TIData
```

Out[2]:

	Row	Green Investment	Eco-friendly Design	Customer Relations	TechnologicalInnovation
0	1	10	10	2	3
1	2	1	2	1	3
2	3	2	3	10	2
3	4	4	1	6	1
4	5	7	7	6	5
...	...	...	...	...	...
416	417	5	7	3	1
417	418	2	9	8	2
418	419	2	2	1	3
419	420	5	8	9	4
420	421	7	8	7	5

421 rows × 5 columns

FIGURE 6. Dataset in use.

In [39]:
```
TIData.describe()
```

Out[39]:

	Row	Green Investment	Eco-friendly Design	Customer Relations	TechnologicalInnovation
count	421.000000	421.000000	421.000000	421.000000	421.000000
mean	210.959620	5.375297	5.524941	5.572447	2.750594
std	121.743488	2.978775	2.890456	2.811981	1.646651
min	1.000000	1.000000	1.000000	1.000000	1.000000
25%	106.000000	3.000000	3.000000	3.000000	2.000000
50%	211.000000	5.000000	5.000000	6.000000	3.000000
75%	316.000000	8.000000	8.000000	8.000000	3.000000
max	421.000000	10.000000	10.000000	10.000000	10.000000

FIGURE 7. Descriptive analysis of the dataset.

5.5. Descriptive analysis

To continue and before further processing, the data must be described statistically. For this purpose, descriptive Analysis is done using the Pandas library, as shown in Figure 7.

As shown in Figure 7, all responders answered all features thoroughly. Therefore, there is no need for a data pre-processing step.

5.6. Test and train data selection

This research uses the "train_test_split" command of the *sklearn* library, the data is divided into two sections where 70% of the data set will be used for training purposes, and 30% will be considered for training purposes test data. The figure below shows the dimensions of the "Xtrain" and "Xtest" matrices (Fig. 8).

```
In [3]: X=TIData.iloc[:,[2,3]].values
        X

Out[3]: array([[10,  2],
               [ 2,  1],
               [ 3, 10],
               [ 1,  6],
               [ 7,  6],
               [ 2,  8],
               [ 7,  9],
               [ 3,  1],
               [ 9,  2],
               [ 9,  6],
               [10,  1],
               [ 5,  4],
               [ 1, 10],
               [ 2,  5],
               [ 9,  2],
               [ 2,  4],
               [ 7,  4],
               [ 7,  8],
               [ 4,  2],
```

```
In [4]: y=TIData.iloc[:,4].values
        y

Out[4]: array([ 3,  3,  2,  1,  5,  2,  4,  2,  1,  5,  1,  2,  1,  1,  1,  3,  2,
                3,  2,  3,  3,  1,  6,  4,  3,  3,  1,  3,  4,  2,  2,  5,  8,  5,
                2,  4,  2,  6,  2,  4,  1,  4,  3,  2, 10,  2,  4,  1,  3,  6,  3,
                2,  1,  1,  3,  3,  2,  1,  1,  1,  2,  1,  1,  4,  1,  3,  2,  3,
                1,  1,  2,  3,  4,  2,  2,  3,  1,  4,  7,  2,  2,  5,  2,  3,  1,
                2,  2,  9,  1,  1,  1,  2,  2,  2,  5,  9,  4,  1,  1,  3,  1,  6,
                1,  9,  2,  1,  6,  3,  2,  1,  2,  3,  3,  6,  2,  3,  3,  2,  3,
                2,  2,  3,  3,  2,  4,  2,  3,  3,  3,  4,  3,  3,  6,  3,  1,  1,
                1,  2,  3,  2,  1,  1,  2,  3,  2,  2,  1,  3,  1,  3,  2,  3,  1,
                2,  2,  1,  5,  3,  3,  3,  3,  3,  4,  3,  1,  2,  2,  5,  2,  8,
                2,  3,  5,  3,  3,  3,  3,  4,  5,  3,  6,  1,  2,  2,  1,  5,  1,
                7,  2,  3,  1,  3,  1,  1,  6,  3,  3,  2,  2,  1,  3,  2,  4,  3,
                3,  1,  1,  2,  2,  3,  4,  3,  3,  3,  6,  2,  3,  3,  3,  3,  2,
                2,  1,  3,  3,  3,  2,  1,  2,  2,  2,  3,  1,  1,  1,  1,  3,  3,
                3,  2,  3,  6,  3,  3,  3,  2,  2,  1,  2,  1,  4,  2,  3,  6,  1,
                1,  1,  3,  3,  2,  5,  3,  2,  2,  1,  3,  3,  1,  3,  2,  3,  1,
                2, 10,  4,  3,  3,  2,  2,  3,  2,  5,  1,  8,  2,  3,  3,  2,  3,
                3,  2,  2,  3,  1,  2,  2,  2,  2,  1,  6,  5,  3,  1,  3,  2,  1,
                2,  3,  2,  3,  1,  3,  1,  3,  2,  2,  2,  3,  3,  2,  3,  7,  6,
                3,  4,  2,  3,  4,  1,  1,  1,  3,  4,  6,  1,  3,  3,  1,  2,  1,
                3,  1,  4,  2,  3,  1,  1,  3,  3,  1,  7,  2,  6,  3,  3,  3,  1,
                3,  2,  3,  3,  7,  1,  1,  2,  5,  1,  3,  5,  2,  3,  2,  1,  3,
                4,  3,  3,  3,  8,  8,  7,  2,  2,  3,  2,  1,  1,  6,  3,  1,  3,
                3,  2,  1,  2,  1,  3,  3,  2,  4,  1,  3,  3,  2,  5,  3,  2,  3,
                3,  2,  3,  1,  3,  3,  2,  3,  1,  2,  3,  4,  5], dtype=int64)
```

FIGURE 8. Clustering dataset into train and test sets.

5.7. Fitting process

KNN learning phase works based on calculating the distances (Minkowski, Manhattan, etc.) between a point and all nearest data around it, depending on the number of predefined neighbors. Then, the label can be estimated using the most frequently observed labels of the K-nearest data around the point.

This strategy is handy, specifically while data clustering according to the features is meaningful. In other words, in a society where closer data have similar feature values, KNN works excellent.

5.8. Determine the best share of test and train split rate

Choosing the best amount of train and test share split is vital for boosting the accuracy of the proposed algorithm. It can also show a valuable hint for recognizing the over-fitting or under-fitting. For this purpose, using a "For" loop, the proposed KNN algorithm will be run ten times. In each iteration, the algorithm chooses a specific test and train share. Table 15 shows the test and train split value in each iteration. Figure 9 indicates that the algorithm's performance will be maximized when the split rate is 0.3.

TABLE 15. Test-Train split rate.

Iteration	Test and train split rate
1	0
2	0.05
3	0.1
4	0.2
5	0.25
6	0.3
7	0.35
8	0.4
9	0.45
10	0.5

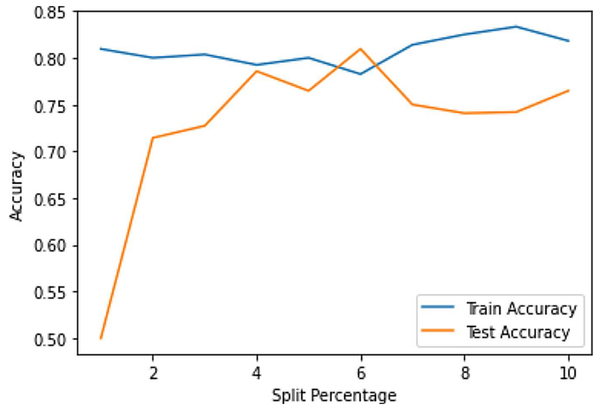

FIGURE 9. The accuracy of train and test data for the proposed algorithm.

5.9. K-nearest neighbors method

K-nearest Neighbor Algorithm (KNN) is one of the effective methods used for the classification and regression of new data using a set of data used for training. This method has a very high ability to classify different groups with high features and is a very efficient method.

The K-nearest Neighbor Algorithm method has been widely used in various engineering, medicine, and management fields. The basis of this method is in the classification of different regions by cross-support vectors. When drawing regional lines, it tries to select the most confident margins using various functions such as exponential, polynomial, and sigmoid kernels that help classify the items in a better way.

5.10. Choosing appropriate number of neighbors

Choosing the correct number of neighbors in the KNN method plays a crucial role in its accuracy. Figure 10 indicates the different levels identified by KNN, while the number of neighbors is considered 5 and 10, respectively. The model will provide a better levels in all paired features while the number of neighbors is considered 10.

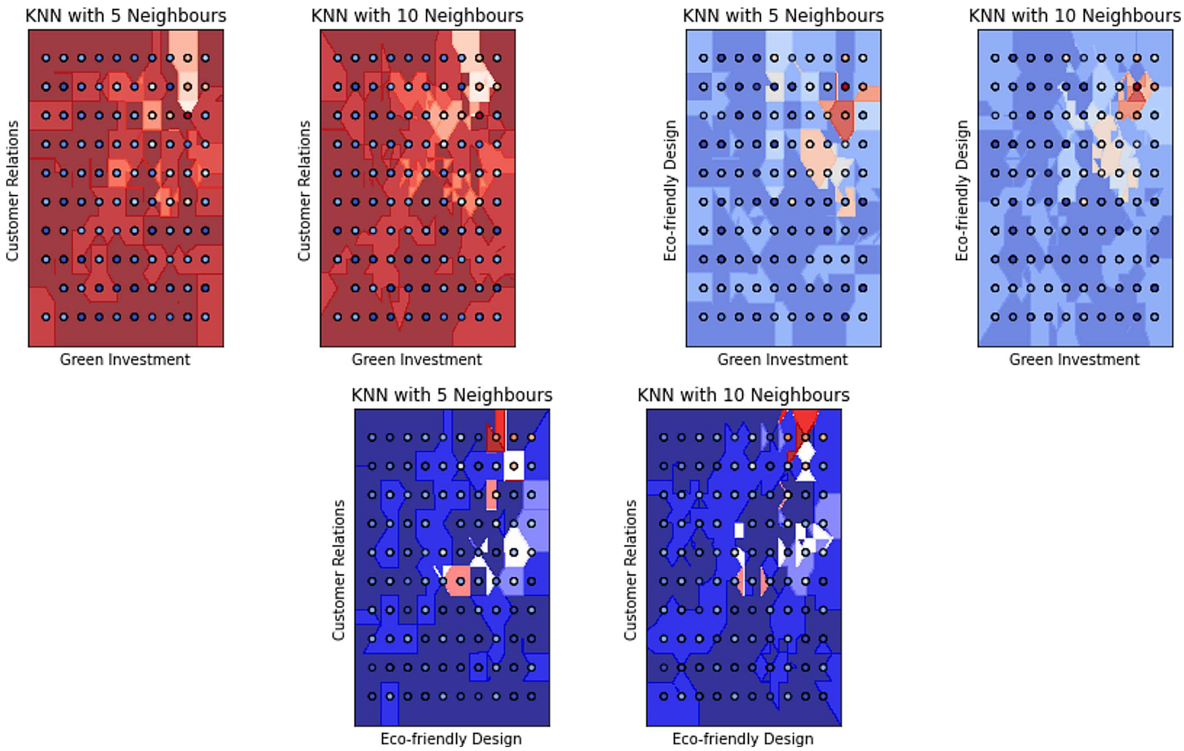

FIGURE 10. Comparing the regions while paired variables are taken into account.

5.11. Comparing Minkowski, Manhattan, Euclidean, and Cosine methods

KNN can be calculated using different distance methods. Although mostly they report almost the same results; but, it is essential to choose the best distance method. Figure 11 compares Minkowski, Manhattan, Euclidean, and cosine distance methods. It is shown that the difference between levels recognized by Minkowski, Manhattan, and Euclidean methods is neglectable. For solving the test data, the Minkowski method will be used.

One crucial step that must be taken after evaluating the performance of the proposed model; and before using actual data for predicting the Technological Innovation is to measure the model's accuracy by calculating the errors. For this purpose, Mean Absolute Error and Mean Squared Error is used (Tab. 18):

$$\mathrm{MAE} = \left(\frac{1}{n}\right) \sum |Y_{\text{test}} - Y_{\text{pred}}| \tag{5.1}$$

$$\mathrm{MSE} = \left(\frac{1}{n}\right) \sum (Y_{\text{test}} - Y_{\text{pred}})^2. \tag{5.2}$$

Using both MAE and MSE measures the model's accuracy without considering the direction of the values, and therefore some values may neutralize the effects of other values in the opposite direction. However, it can show the average magnitude of the errors in a set of predictions, and therefore MAE should not be ignored (Tab. 17).

5.12. Comparing the performance of different supervised methods for the developed model

In this section, the levels that KNN offer will be compared with a series of other methods, including Linear Regression, Logistic Regression, Random Forest (RF), Naive Base (Gaussian NB), Support Vector Machine

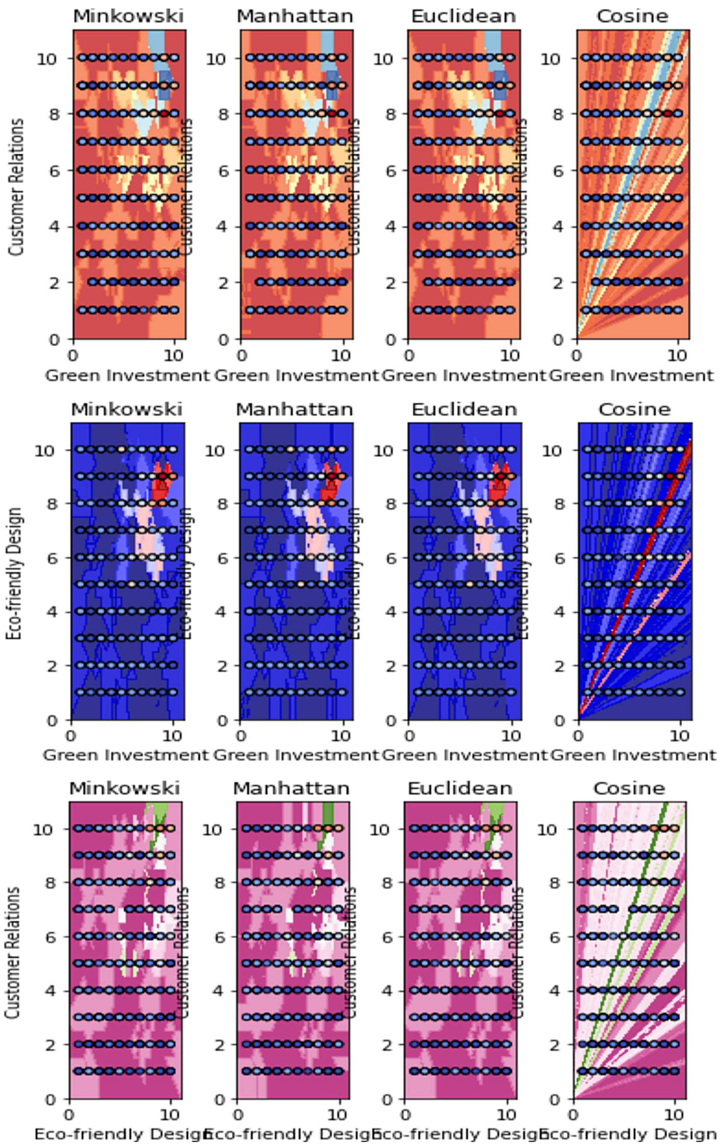

FIGURE 11. Results of comparing Minkowski, Manhattan, Euclidean, and Cosine distance methods.

(SVM), and Multi-Layer Perceptron (MLP) (Fig. 12). The settings of the classifiers are shown in Table 16. Using the settings of Table 16, the data is applied to the six well-known supervised methods to compare their performance with the KNN. The outcomes are presented in Table 17.

Comparing different methods indicated that Logistic Regression could provide a base with higher accuracy (0.525) and less Absolute Mean Error and Mean Squared Error simultaneously (0.86 and 1.4, respectively). The recognized regions of technological innovation by different methods are shown in Figure 12.

Figure 12 shows the clustering of data using different methods. In this figure, 2 points are noticeable. First is how strongly the method can identify the clusters, and the next is how it is sensitive to draw borderlines of a cluster. However, the more complex designs (lines in a figure) does not necessarily mean better performance as it may happen due to overfitting. In contrast, the very simple designs might happen due to under-fitting, which

Table 16. Settings of the methods that are used for the developed model.

Row	Algorithm	Used settings
1	Linear regression	clf1 = linear_model.LinearRegression()
2	Logistic regression	clf2 = LogisticRegression(random_state = 1, solver = “newton-cg”, multi_class = “multinomial”)
3	RF	clf3 = RandomForestClassifier(random_state = 1, n_estimators = 100)
4	NB	clf4 = GaussianNB()
5	SVM	clf5 = SVC(gamma = “auto”)
6	MLP	clf6 = MLPClassifier(solver = “lbfgs”, alpha = 1e-5, hidden_layer_sizes = $(8, 3)$, random_state = 1)
7	KNN	clf7 = KNeighborsClassifier(n_neighbors = 10, p = 2, metric = “minkowski”)

Table 17. Comparing the outcomes of different methods for the developed model.

Metric	Linear regression	Logistic regression	RF	NB	SVM	MLP	KNN
Accuracy score	0.339	0.525	0.356	0.456	0.347	0.198	0.751
Mean absolute error	1.2	0.86	1.13	1.0	1.03	0.882	0.9
Mean squared error	2.266	1.4	2.266	2.13	1.9	1.399	1.366

means the training process did not carry out well. In Figure 12, Linear Regression, Logistic Regression, Naive Bayes, and Multi-Layer Perceptron could provide more specific regions. However, these methods are underfitted and show less accuracy in clustering effective regions while Green Purchasing and Eco-Friendly Design are considered.

In contrast, Random Forest and KNN offer designs that are more accurate. Using the data used in this model, KNN Algorithm seems to draw the best clustering regions. It can help supply chain managers forecast

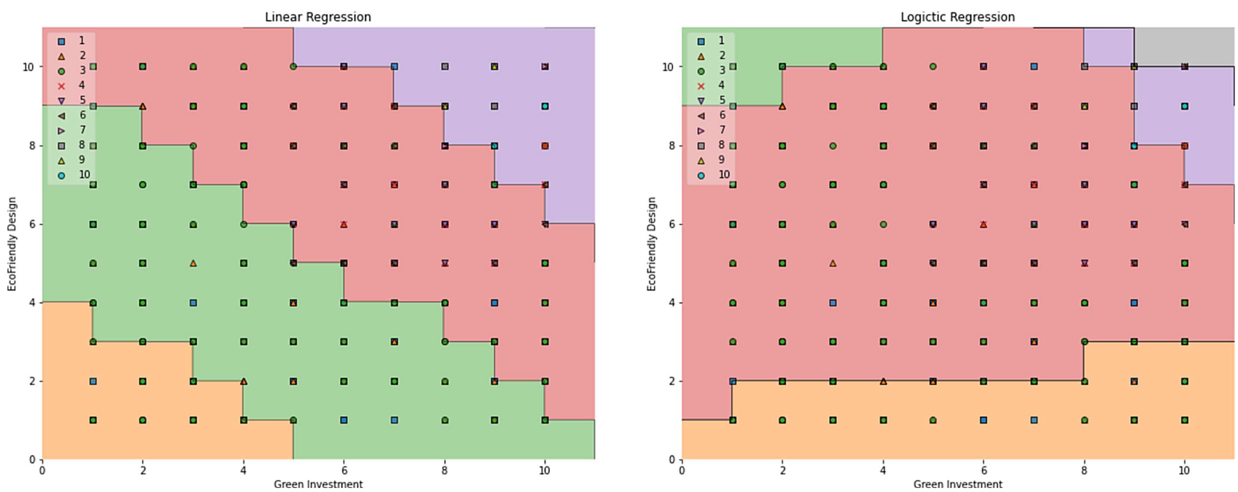

Figure 12. Comparing clustered regions by different methods.

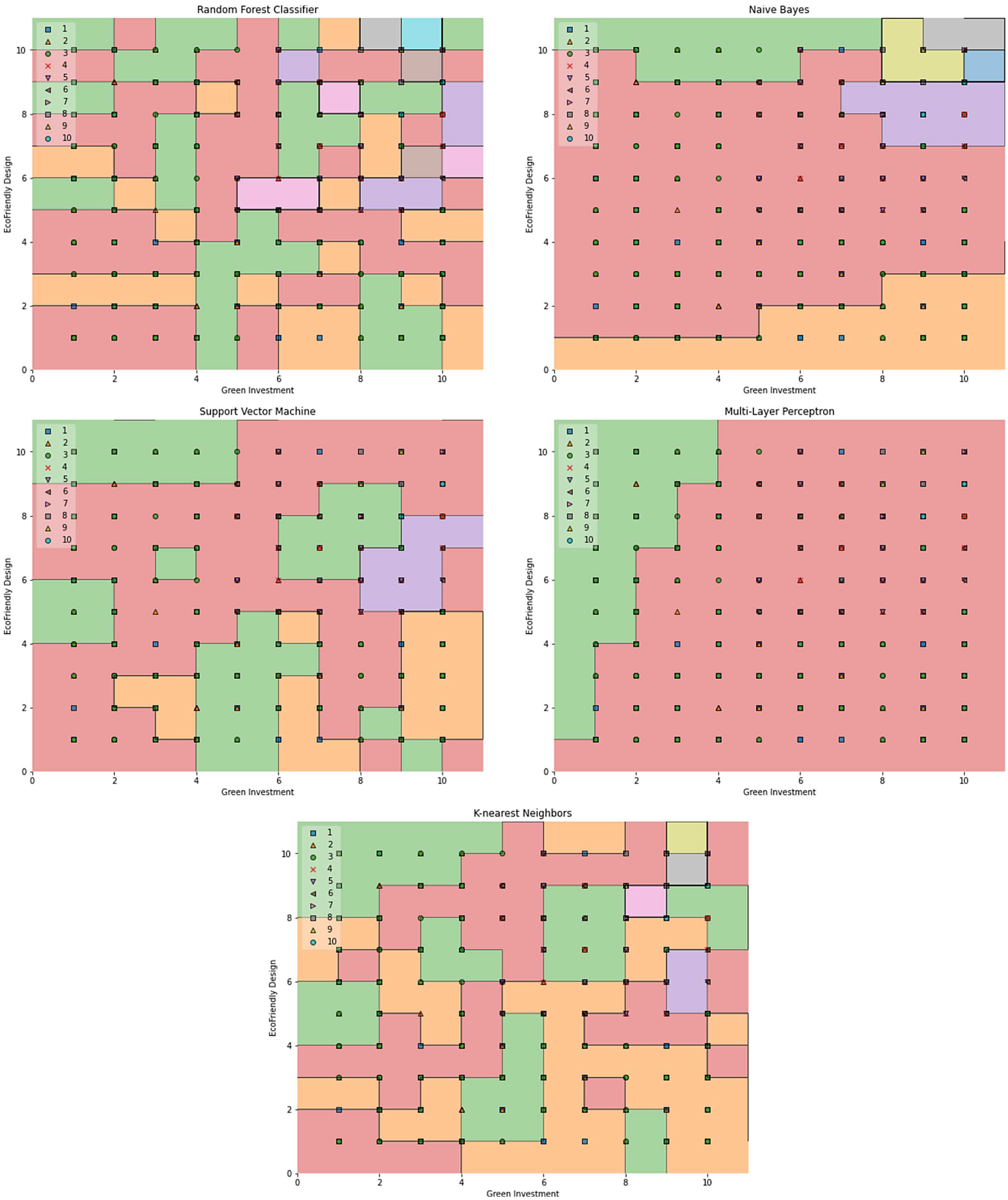

FIGURE 12. continued.

TABLE 18. Dataset for Validating the proposed KNN.

Row	Green Investment	Eco-Friendly Design	Customer Collaboration	Technological Innovation Label	Model Prediction	Whether SVM predicted Correct Label
1	6	6	6	4	4	✓
2	3	9	3	2	3	✓
3	9	5	1	2	2	✓
4	6	1	1	1	3	–
5	4	7	4	1	1	✓
6	1	3	8	2	2	✓
7	5	1	7	2	2	✓
8	5	4	2	1	1	✓
9	1	1	8	2	2	✓
10	8	6	3	1	1	✓
11	7	8	8	3	3	✓
12	9	3	7	2	2	✓
13	9	3	8	2	3	–
14	9	3	1	1	1	✓
15	8	1	3	2	2	✓
16	6	5	6	6	5	–
17	8	2	8	1	1	✓
18	8	3	6	1	1	✓
19	1	3	9	2	2	✓
20	1	1	1	1	1	✓
21	4	8	4	2	2	✓
22	2	7	4	2	2	✓
23	1	2	9	2	3	–
24	9	1	5	1	1	✓
25	9	2	4	1	1	✓
26	4	2	3	1	1	✓
27	5	5	1	1	1	✓
28	5	7	5	3	3	✓
29	8	4	6	2	3	–
30	9	6	9	5	4	–

technological innovation levels using a series of input data of green investment, eco-friendly design, and customer collaborations in the supply chain.

6. Predict technological innovation level (phase 3)

In this section, the trained ML model will be used to predict a supply chain's technological innovation level using the given data on Eco-friendly Design, Green Investment, and Customer Collaborations. For this purpose, a part of Section 5.4 that was considered for the evaluation process (Eval.data) is used. The outcomes are indicated in Table 8.

Table 18 indicates that the proposed method could successfully predict and classify the most validating data. Using 421 data that contain real data and generated data by Python for the training and evaluation of the model, the achievements are noticeable and worthy (precision = 0.8). Since this value is not less than the accuracy of training data (0.751), it is concluded that the ML model is not under-fitted and thus well trained. Moreover, it can be seen that while the model is confronted with new conditions, which were not used during the training process, it can still determine the class of the technological innovation level successfully.

7. Managerial insights

This research investigated the impact of different internal drivers on T.I. Comparing the key findings of this research with the recent references indicated that the (i) different levels of T.I may be seen by investing in internal drivers. (ii) Technological innovation can be flourished by improving the internal factors in a supply chain, such as top management commitments. This finding complies with [30]. In addition (iii), the outcomes showed that the T.I level can be predicted accurately (80%) by considering the green investment, eco-friendly design, and customer collaboration. This finding aligns with [18, 21, 32].

The outcomes of this research can help senior supply chain managers to determine the level of required investments in using green materials, designing eco-friendly products, and expanding the correlations with their customers to achieve a certain level of technology innovation that consequently leads to more system performance [41].

8. Conclusion

This research presents a 3-phase framework for predicting the technological innovation level of green supply chains by focusing on Green Investment, Customer Collaboration, and Eco-friendly Design Features. In the first phase of the framework, the model's independent variables were identified. In continuation, the positive correlations between drivers and T.I is investigated. Then, using statistical analysis on data of supply chains located in the northern states of Malaysia, the impact of drivers on T.I was determined, indicating that Customer Collaboration, Eco-friendly design, and Green Investment have noticeable impacts on T.I improvement in the studied cases (0.481, 0.419 and 0.41, respectively). These results align with [12, 14, 18, 21, 43].

In the second part of the Research, a comprehensive model is developed using the K-nearest Neighbor algorithm to predict the technological innovation level where Green Investment, Customer Collaboration, and Eco-friendly Design were considered as features of the model. The outcomes indicated that the T.I level could be effectively predicted in green supply chains using the proposed method (test score: 0.751). The findings complies with [30, 32]. The outcomes are also compared with Linear Regression, Logistic Regression, Naive Bayes, Multi-Layer Perceptron, Support Vector Machine, and Random Forest methods, demonstrating K-nearest Neighbor's superiority in outlining better designs for projecting the T.I level in supply chains.

In the third phase, the method is applied to a supply chain located in Kuala Lumpur. The outcomes showed that the developed model could be successfully used in practice and provide accurate results for 30 supply chains with less MAE and MSE values (0.9 and 1.36, respectively). Using the proposed method of this research, top managers of a supply chain can make financial decisions on investments on discussed technological innovation drivers (green materials, designing eco-friendly products, and expanding the correlations with their customers) with more confidence.

Future model development is suggested by developing a cloud-based application to use real-time data of drivers of T.I and predict T.I for coming periods. In addition, more expansion of the proposed method of this research can be carried out by proposing a fuzzy-based application to include uncertainty of the mentioned drivers of T.I.

Acknowledgements. Authors would like to thank the anonymous reviewers and the editor for their positive comments during the reviewing process.

References

[1] W.A. Abbasi, Z. Wang, Y. Zhou and S. Hassan, Research on measurement of supply chain finance credit risk based on Internet of Things. *Int. J. Distrib. Sens. Netw.* **15** (2019) 1550147719874002.

[2] A. Ahmadi-Javid and P. Hoseinpour, On a cooperative advertising model for a supply chain with one manufacturer and one retailer. *Eur. J. Oper. Res.* **219** (2012) 458–466.

[3] F.K. Chan, J.Y. Thong, V. Venkatesh, S.A. Brown, P.J. Hu and K.Y. Tam, Modeling citizen satisfaction with mandatory adoption of an e-government technology. *J. Assoc. Inf. Syst.* **11** (2010) 519–549.

[4] T.-Y. Chiou, H.K. Chan, F. Lettice and S.H. Chung, The influence of greening the suppliers and green innovation on environmental performance and competitive advantage in Taiwan. *Transp. Res. Part E: Logistics Transp. Rev.* **47** (2011) 822–836.

[5] A.Y.L. Chong and K.B. Ooi, Adoption of interorganizational system standards in supply chains. *Ind. Manage. Data Syst.* **108** (2008) 529–547.

[6] A. Delgoshaei and A. Ali, A hybrid genetic and simulated annealing algorithms for scheduling fashion goods supply chains. *Int. J. Adv. Heuristic Meta-heuristic Algorithms* **1** (2020) 30–37.

[7] A. Delgoshaei, A. Aram and A. Ali, A robust optimization approach for scheduling a supply chain system considering preventive maintenance and emergency services using a hybrid ant colony optimization and simulated annealing algorithm. *Uncertain Supply Chain Manage.* **7** (2019) 251–274.

[8] A. Delgoshaei, A.K. Aram and A.H. Nasiri, The effects of individual and organizational factors on creativity in sustainable supply chains. Paper presented at the International Conference on Logistics and Supply Chain Management (2020).

[9] A. Delgoshaei, M. MohammadAzari, S.E. Hanjani, F. Fard, R. Beigizadeh and A.K. Aram, A fuzzy logic-based machine-learning algorithm for product distribution in supply chains considering rival's strategic decisions. *Int. J. Ind. Eng.* **27** (2020) 933–958.

[10] H.A. Dogahe, H.R. Meydanghah and M.N. Imani, The effect of educational methods of supply chain management on conflict management ineducational environments. *Int. J. Supply Chain Manage.* **8** (2019) 18–26.

[11] T. Eltayeb and S. Zailani, Going green through green supply chain initiatives toward environmental sustainability. *Oper. Supply Chain Manage.: Int. J.* **2** (2014) 93–110.

[12] L. Eyraud, B. Clements and A. Wane, Green investment: trends and determinants. *Energy Policy* **60** (2013) 852–865.

[13] B. Fanoodi, B. Malmir and F.F. Jahantigh, Reducing demand uncertainty in the platelet supply chain through artificial neural networks and ARIMA models. *Comput. Biol. Med.* **113** (2019) 103415.

[14] C.A. Geffen and S. Rothenberg, Suppliers and environmental innovation. *Int. J. Oper. Prod. Manage.* **20** (2000) 166–186.

[15] P. González-Torre, M. Alvarez, J. Sarkis and B. Adenso-Díaz, Barriers to the implementation of environmentally oriented reverse logistics: evidence from the automotive industry sector. *Br. J. Manage.* **21** (2010) 889–904.

[16] S. Gupta and O.D. Palsule-Desai, Sustainable supply chain management: review and research opportunities. *IIMB Manage. Rev.* **23** (2011) 234–245.

[17] R. Isaksson, P. Johansson and K. Fischer, Detecting supply chain innovation potential for sustainable development. *J. Bus. Ethics* **97** (2010) 425–442.

[18] P. Kapetanopoulou and G. Tagaras, Drivers and obstacles of product recovery activities in the Greek industry. *Int. J. Oper. Prod. Manage.* **31** (2011) 148–166.

[19] S.Y. Lam, V.H. Lee, K.B. Ooi and K. Phusavat, A structural equation model of TQM, market orientation and service quality. *Manag. Serv. Qual.: Int. J.* **22** (2012) 281–309.

[20] V.-H. Lee, K.-B. Ooi, A.Y.-L. Chong and C. Seow, Creating technological innovation via green supply chain management: an empirical analysis. *Expert Syst. App.* **41** (2014) 6983–6994.

[21] V.-H. Lee, K.-B. Ooi, A.Y.-L. Chong and B. Lin, A structural analysis of greening the supplier, environmental performance and competitive advantage. *Prod. Plan. Control* **26** (2015) 116–130.

[22] C.-Y. Lin and Y.-H. Ho, Determinants of green practice adoption for logistics companies in China. *J. Bus. Ethics* **98** (2011) 67–83.

[23] J. Luo, A.Y.-L. Chong, E.W. Ngai and M.J. Liu, Reprint of "Green Supply Chain Collaboration implementation in China: the mediating role of guanxi". *Transp. Res. Part E: Logistics Transp. Rev.* **74** 37–49.

[24] L. Macchion, A. Moretto, F. Caniato, M. Caridi, P. Danese, G. Spina and A. Vinelli, Improving innovation performance through environmental practices in the fashion industry: the moderating effect of internationalisation and the influence of collaboration. *Prod. Plan. Control* **28** (2017) 190–201.

[25] L. Macchion, A.M. Moretto, F. Caniato, M. Caridi, P. Danese and A. Vinelli, International e-commerce for fashion products: what is the relationship with performance? *Int. J. Retail Distrib. Manage.* **45** (2017) 1011–1031.

[26] S. Mehrolia, S. Alagarsamy and V.M. Solaikutty, Customers response to online food delivery services during COVID-19 outbreak using binary logistic regression. *Int. J. Consum. Stud.* **45** (2021) 396–408.

[27] A. Molla and A. Abareshi, Green IT adoption: a motivational perspective. Paper presented at the PACIS (2011).

[28] C. Negrutiu, C. Vasiliu and C. Enache, Sustainable entrepreneurship in the transport and retail supply chain sector. *J. Risk Finan. Manage.* **13** (2020) 267.

[29] M.S. Nikabadi and A. Shahrokhnia, Multidimensional structure for the effect of innovation culture and knowledge sharing on the new product development process with emphasis on improving new product development performance. *Middle East J. Manage.* **6** (2019) 494–512.

[30] M. O'dwyer, A. Gilmore and D. Carson, Innovative marketing in SMEs. *Eur. J. Marketing* **43** (2009) 46–61.

[31] D.I. Prajogo and A.S. Sohal, The relationship between TQM practices, quality performance, and innovation performance. *Int. J. Qual. Reliab. Manage.* **20** (2003) 901–918.

[32] M. Rahbari, S.H.R. Hajiagha, M.R. Dehaghi, M. Moallem and F.R. Dorcheh, Modeling and solving a five-echelon location–inventory–routing problem for red meat supply chain: case study in Iran. *Kybernetes* (2020). DOI: 10.1108/K-10-2019-0652.

[33] M. Saberioon, P. Císař, L. Labbé, P. Souček, P. Pelissier and T. Kerneis, Comparative performance analysis of support vector machine, random forest, logistic regression and k-nearest neighbours in rainbow trout (Oncorhynchus mykiss) classification using image-based features. *Sensors* **18** (2018) 1027.

[34] A.S. Singh and M.B. Masuku, Fundamentals of applied research and sampling techniques. *Int. J. Med. Appl. Sci.* **2** (2013) 124–132.

[35] P.J. Singh and A.J. Smith, Relationship between TQM and innovation: an empirical study. *J. Manuf. Technol. Manage.* **15** (2004) 394–401.

[36] S.K. Singh, S. Gupta, D. Busso and S. Kamboj, Top management knowledge value, knowledge sharing practices, open innovation and organizational performance. *J. Bus. Res.*. **128** (2019). DOI: 10.1016/j.jbusres.2019.04.040.

[37] S. Taghiyeh, D.C. Lengacher and R.B. Handfield, A multi-phase approach for product hierarchy forecasting in supply chain management: application to MonarchFx Inc. Preprint arXiv:2006.08931 (2020).

[38] F. Testa and F. Iraldo, Shadows and lights of GSCM (Green Supply Chain Management): determinants and effects of these practices based on a multi-national study. *J. Cleaner Prod.* **18** (2010) 953–962.

[39] Y.-Y. Wang, Z. Hua, J.-C. Wang and F. Lai, Equilibrium analysis of markup pricing strategies under power imbalance and supply chain competition. *IEEE Trans. Eng. Manage.* **64** (2017) 464–475.

[40] S. Yi and A. Xiao-li, Application of threshold regression analysis to study the impact of regional technological innovation level on sustainable development. *Renew. Sustainable Energy Rev.* **89** (2018) 27–32.

[41] Y. Zhang, U. Khan, S. Lee and M. Salik, The influence of management innovation and technological innovation on organization performance. A mediating role of sustainability. *Sustainability* **11** (2019) 495.

[42] Q. Zhu and J. Sarkis, An inter-sectoral comparison of green supply chain management in China: drivers and practices. *J. Cleaner Prod.* **14** (2006) 472–486.

[43] Q. Zhu, J. Sarkis and K.-H. Lai, Confirmation of a measurement model for green supply chain management practices implementation. *Int. J. Prod. Econ.* **111** (2008) 261–273.

[44] Q. Zhu, J. Sarkis and K.-H. Lai, Green supply chain management implications for "closing the loop". *Transp. Res. Part E: Logistics Transp. Rev.* **44** (2008) 1–18.